

---

Proceedings of the  
22<sup>nd</sup> Amsterdam Colloquium

---



*22<sup>nd</sup>*  
Amsterdam Colloquium  
*2019*

Edited by Julian J. Schlöder, Dean McHugh & Floris Roelofsen

## Foreword

The papers contained in this collection were presented at the 22<sup>nd</sup> Amsterdam Colloquium, organised by the Institute for Logic, Language and Computation (ILLC) at the University of Amsterdam, December 18–20, 2019. The biennial Amsterdam Colloquia aim at bringing together linguists, philosophers, logicians, cognitive scientists and computer scientists who share an interest in the formal study of the semantics and pragmatics of natural and formal languages.

Besides the general programme, the 2019 edition featured two workshops on *Semantic Universals* and *Super Linguistics*, respectively, and the *Beth Lecture*, jointly organised with the E.W. Beth Foundation. The programme included 8 invited talks, 34 contributed talks, and a poster session with 23 contributed posters.

We would like to thank all programme committee members, listed below, and all reviewers for their efforts in selecting the contributed talks and posters. We would also like to thank Peter van Ormondt and Katrin Schulz for their practical help in organising the conference.

Lastly, we would like to thank the ILLC, the E.W. Beth Foundation, the Netherlands Organization for Scientific Research (NWO), and the European Research Council (ERC, grant nos 680220, 716230 and 758540) for financial support.

Julian J. Schlöder  
Dean McHugh  
Floris Roelofsen



## Programme Committees

### General programme

Maria Aloni	ILLC, University of Amsterdam
Paul Dekker	ILLC, University of Amsterdam
Luca Incurvati	ILLC, University of Amsterdam
Robert van Rooij (chair)	ILLC, University of Amsterdam

### Workshop on Semantic Universals

Floris Roelofsen	ILLC, University of Amsterdam
Jakub Szymanik	ILLC, University of Amsterdam
Wataru Uegaki	University of Edinburgh

### Workshop on Super Linguistics

Pritty Patel Grosz	University of Oslo
Philippe Schlenker	Ecole Normale Supérieure, Paris

# Table of Contents

## Invited

Conditionals in selection semantics .....	1
<i>Fabrizio Cariani</i>	
Tagging: Semantics at the iconic/symbolic interface .....	11
<i>Gabriel Greenberg</i>	

## General Programme

Neg-raising 'believe' and Maximize Presupposition .....	21
<i>Sam Alxatib and Ivana Durović</i>	
Abilities and success .....	30
<i>David Boylan</i>	
Attitudes, conditionals and margins for error .....	40
<i>David Boylan and Ginger Schultheis</i>	
Cognitive modeling for formal semantics: the organization of discourse representation structures in declarative Memory .....	50
<i>Adrian Brasoveanu and Jakub Dotlačil</i>	
Definiteness across languages: from German to Mandarin .....	60
<i>David Bremmers, Jianan Liu, Bert Le Bruyn and Martijn van der Klis</i>	
Keeping <i>dou</i> as a simple distributor .....	71
<i>Zhuo Chen</i>	
Attitude Semantics: embedding and probability for indicative conditionals .....	81
<i>Ivano Ciardelli</i>	
Testing formal pragmatics of questions through their ignorance inferences .....	91
<i>Alexandre Cremers and Zhuoye Zhao</i>	
Donkey anaphora in non-monotonic environments: an experimental investigation .....	101
<i>Milica Denić and Yasutada Sudo</i>	
Learning modal force: evidence from children's production and input .....	111
<i>Anouk Dieuleveut, Annemarie van Dooren, Ailis Cournane and Valentine Hacquard</i>	
String iconicity and granularity .....	122
<i>Tim Fernando</i>	
Time matters: the role of temporal boundaries in NPI licensing .....	132
<i>Julie Goncharov and Lavi Wolf</i>	
Double negation, excluded middle and accessibility in dynamic semantics .....	142
<i>Matthew Gotham</i>	
Dogwhistles, trust and ideology .....	152
<i>Robert Henderson and Elin McCready</i>	

<i>Only</i> reconstruction and backwards association .....	161
<i>Aron Hirsch and Michael Wagner</i>	
Loose talk, scale presuppositions and QUD .....	171
<i>Daniel Hoek</i>	
The anaphoric potential of indefinites under negation and disjunction .....	181
<i>Lisa Hofmann</i>	
Evidence acquisition time as belief-state change: a view from Mvskoke (Creek) .....	191
<i>Kimberly Johnson</i>	
Pair-list questions from dependent plurals .....	201
<i>William Johnston</i>	
The asymmetry of fake tense .....	211
<i>Hadil Karawani, Carina Kauf and Hedde Zeijlstra</i>	
Dynamic unioning plural logic .....	218
<i>Ezra Keshet</i>	
Coordinating complete answers: the case of <i>Tanto–Quanto</i> Conjunction .....	228
<i>Filipe Hisao Kobayashi and Enrico Flor</i>	
Hyperlogic: a system for talking about logics .....	238
<i>Alexander Kocurek</i>	
Indicative and subjunctive conditionals in commitment spaces .....	248
<i>Manfred Krifka</i>	
DPs and CPs in depiction complements .....	259
<i>Kristina Liefke</i>	
Not few but all quantifiers can be negated: towards a referentially transparent semantics of quantified noun phrases .....	269
<i>Andy Lücking and Jonathan Ginzburg</i>	
Crises of Identity .....	279
<i>Matthew Mandelkern</i>	
Negation and alternatives in conditional antecedents .....	289
<i>Dean McHugh and Alexandre Cremers</i>	
How prosody disambiguates between alternative and polar questions .....	299
<i>Erlinde Meertens</i>	
Combining neurophysiology and formal semantics and pragmatics: the case of the N400 and the P600 .....	309
<i>Ralf Naumann and Wiebke Petersen</i>	
On supererogation: one should go when going is good enough and not going is not .....	319
<i>David Yoshikazu Oshima</i>	
Speaker discourse roles and the discourse profile of reportative evidentials .....	327
<i>Roumyana Pancheva and Deniz Rudin</i>	

Just exhaustification. A ‘two stage’ theory of exclusives.....	337
<i>Daniele Panizza and Gennaro Chierchia</i>	
Fixing de Morgan’s law in counterfactual antecedents.....	347
<i>Jacopo Romoli, Paolo Santorio and Eva Wittenberg</i>	
Common ground: <i>in sensu composito</i> or <i>in sensu diviso</i> .....	357
<i>Merel Semeijn</i>	
Singular/plural contrasts: the case of informational object nouns .....	367
<i>Peter Sutton and Hana Filip</i>	
Existential semantics in equatives in Japanese and German .....	377
<i>Eri Tanaka, Kenta Mizutani and Stephanie Solt</i>	
The role of “fake” past tense in acquiring counterfactuals.....	387
<i>Maxime Tulling and Ailís Cournane</i>	
Asking between the lines: elicitation of evoked questions in text .....	397
<i>Matthijs Westera and Hannah Rohde</i>	
Negating conditionals in bilateral semantics .....	407
<i>Malte Willer</i>	
Two types of higher-order readings of WH-questions .....	417
<i>Yimei Xiang</i>	
Bridging distributivity and free choice: the case of Mandarin <i>dou</i> .....	427
<i>Zhuoye Zhao</i>	
On domain adjectives and the metaphors they modify .....	437
<i>Joost Zwarts</i>	

## **Workshop: Semantic Universals**

Monotonicity restored: “more” never means “purer” .....	445
<i>Alan Bale, Bernhard Schwarz and David Shanks</i>	
A subregular bound on the complexity of lexical quantifiers.....	455
<i>Thomas Graf</i>	
Semantic universals of intonation and particles .....	465
<i>Yurie Hara and Mengxi Yuan</i>	
Explaining the exceptive-additive ambiguity in Mandarin and other languages .....	475
<i>Mingming Liu</i>	
Learnability as a window into universal constraints on person systems .....	484
<i>Mora Maldonado and Jennifer Culbertson</i>	
A variable-force variable-flavor attitude verb in Koryak .....	494
<i>Maša Močnik and Rafael Abramovitz</i>	
Are conservative quantifiers easier to learn? Evidence from novel quantifier experiments .....	504
<i>Jennifer Spender and Jill de Villiers</i>	

Quantifiers in natural language optimize the simplicity/informativeness trade-off.....	513
<i>Shane Steinert-Threlkeld</i>	
Equation of degrees or state-kinds, or both.....	523
<i>Yenan Sun</i>	

## **Workshop: Super Linguistics**

Animal linguistics and the puzzle of Titi monkeys alarm sequences .....	533
<i>Mélissa Berthet, Juan Benjumea, Juliette Millet, Cristiane Cäsar, Klaus Zuberbühler and Ewan Dunbar</i>	
Situating rich demonstrations in discourse.....	543
<i>Christian De Leon</i>	
Towards a uniform super-linguistic theory of projection.....	553
<i>Maria Esipova</i>	
Pronominal typology and reference to the external world .....	563
<i>Patrick Grosz</i>	
Gestures as markers of non-canonical questions.....	574
<i>Michela Ippolito</i>	
Picturing words: the semantics of speech balloons .....	584
<i>Emar Maier</i>	
Walk-denoting music: refining music semantics .....	593
<i>Léo Migotti and Léo Zaradzki</i>	
Joker face. Recognizing irony in the visual mode in spoken and signed language .....	603
<i>Francesca Panzeri, Beatrice Giustolisi, Lara Mantovan and Chiara Calderone</i>	
Indexing across media .....	612
<i>Mats Rooth and Dorit Abusch</i>	

# Conditionals in selection semantics <sup>\*</sup>

Fabrizio Cariani<sup>1</sup>

Northwestern University, Evanston, IL, USA  
fabrizio79@gmail.com

## Abstract

I explore the semantic and logical prospects for reconstructing Stalnakerian truth-conditions for conditionals on a theory on which these emerge from the separate contribution of selection modals (*will* and *would*) and a restrictor theory of *if*-clauses.

## 1 Introduction

This paper presents an account of the interaction between conditionals and selection modals. Cariani and Santorio [6, 7] argue that modals such as *will* and *would* are neither necessity nor possibility operators, but instead operate ‘world selection’.<sup>1</sup> I explore how this commitment illuminates some key questions concerning conditionals. The paper complements and overlaps with previous work by some fellow neo-Stalnakerians, specifically Santorio [25] and Mandelkern [22]. The overlap is virtuous: the core ideas of §§4-5 germinated more or less simultaneously, and there are many choice points at which we take different paths.

## 2 Selection modals

Following a strong linguistic tradition, assume that *will* and *would* share a common modal morpheme WOLL; *will* decomposes as PRES + WOLL; *would* decomposes as PAST + WOLL. Semantically, WOLL takes two arguments: a modal base  $\mathbf{f}$  and a prejacent proposition. (I distinguish between the modal base *variable*  $f$  and the set theoretic object  $\mathbf{f}$  it is assigned to.) It is widely believed (but by no means obvious) that the natural modal base for *will* and *would* is historical [4, 7, 13, 15], which I interpret to mean:

in context  $c$ ,  $\mathbf{f}(w) =$  the worlds that are duplicates of  $w$  up to the time of  $c$ .

Properly understood, this appeal to historical modal bases is neutral between several metaphysical options concerning the future, such as branching and divergence.<sup>2</sup>

Onto the semantics. Let  $\sigma$  be a selection function and  $\mathcal{W}$  a background set of worlds. That is,  $\sigma : \mathcal{P}(\mathcal{W}) \times \mathcal{W} \mapsto \mathcal{W}$  satisfying.

success: for  $w \in \mathcal{W}$ ,  $\mathbf{P} \subseteq \mathcal{W}$  with  $\mathbf{P} \neq \emptyset$ ,  $\sigma(w, \mathbf{P}) \in \mathbf{P}$

centering: for  $w \in \mathcal{W}$ ,  $\mathbf{P} \subseteq \mathcal{W}$  with  $w \in \mathbf{P}$ ,  $\sigma(w, \mathbf{P}) = w$

In this note, I adopt a toy treatment of the temporal role of *will* and *would* and tense more generally.<sup>3</sup> Let  $\blacktriangleright$  map a time  $t$  to a time that is no earlier than  $t$ . The basic semantics for WOLL is as follows:

---

<sup>\*</sup>Conversations with David Boylan, Josh Dever, Valentine Hacquard, Matt Mandelkern, Ginger Schultheis, and Paolo Santorio have much improved this paper. Special thanks to Simon Goldstein for feedback on a draft.

<sup>1</sup>See also [3, 4]. Kratzer [18] also independently proposes a semantics along the same lines.

<sup>2</sup>See [20, §4.2] for the metaphysical distinctions, and [3, 7] for the neutrality arguments.

<sup>3</sup>For a fuller, more descriptive, proposal about the future orientation of these modals, see [4] and [3, ch. 7].

$$\llbracket \text{WOLL}_f(\text{P}) \rrbracket^{w,t} = 1 \text{ iff } \llbracket \text{P} \rrbracket^{\sigma(\mathbf{f}(w),w),\blacktriangleright(t)} = 1$$

Present tense is given a redundant interpretation:

$$\llbracket \text{PRES}(\text{P}) \rrbracket^{w,t} = 1 \text{ iff } \llbracket \text{P} \rrbracket^{w,t} = 1$$

We will need the non-selectional modal *might* as well—also with a plain vanilla entry. Following Condoravdi [8], assume that *might* is future-oriented, which in the current toy picture means:

$$\llbracket \text{might}_f(\text{P}) \rrbracket^{w,t} = 1 \text{ iff } \exists v \in \mathbf{f}(w), \llbracket \text{P} \rrbracket^{v,\blacktriangleright(t)} = 1$$

There is evidently much more to be said about the complexities of a integrated semantics for modality, tense and aspect, but none of it will be said here (see among others [4, 8, 12, 32]). I will also not explicitly spell out the consequences of my discussion for *would*-conditionals, though they are very much on the theoretical radar.

### 3 Stalnaker’s conditional.

The selectionist approach to *will* and *would* draws inspiration from Stalnaker’s account of conditionals [28, 29, 30]. According to Stalnaker, *if* P, Q is true at  $w$  iff Q is true at a world  $v$  that is the “closest” P-world to  $w$ . Here is a Stalnaker-inspired entry for a conditional connective ‘>’, with the modification that, with later developments in mind, I throw in a modal base.

$$\llbracket \text{P} >_f \text{Q} \rrbracket^{w,t} = 1 \text{ iff } \llbracket \text{Q} \rrbracket^{\sigma(\mathbf{f}(w) \cap \text{P},w),t} = 1$$

Unlike my selectionist theory, this account associates the selection-behavior with the conditional, as opposed to the modals. Another difference between my account and Stalnaker’s is that Stalnaker imposes stricter conditions on selection.

The lynchpin of Stalnaker’s account is the validity of:

$$\text{CEM. } \vdash (\text{P} >_f \text{Q}) \vee (\text{P} >_f \text{not Q})$$

Here I assume without argument that validating conditional excluded middle is desirable. I defer to the extensive literature to supply the relevant arguments ([5, 6, 22, 25, 29, 33]).

Just as central as CEM is a classic set of invalidities for the Stalnaker conditional.

$$\text{AS. } \text{P} >_f \text{R} \not\vdash (\text{P} \ \& \ \text{Q}) >_f \text{R}$$

$$\text{I/E. } \text{P} >_f (\text{Q} >_f \text{R}) \not\vdash (\text{P} \ \& \ \text{Q}) >_f \text{R}$$

The conventional wisdom surrounding these is that the invalidity of AS is desirable, but the failure of I/E is not, at least for indicatives.<sup>4</sup> Another logical challenge to Stalnaker’s analysis involves the relation between *if* and *might*. These pairs seem contradictory:

- (1) a. If Keith plays, Mick will sing.  
b. If Keith plays, Mick might not sing.
- (2) a. If Keith had played, Mick would have sung.  
b. If Keith had played, Mick might not have sung.

<sup>4</sup>See [23] for treatment of some alleged counterexamples to indicative I/E. Mandelkern also highlights that the subjunctive conditional case is very different when it comes to I/E.

The Stalnaker conditional does not deliver these incompatibilities. This drove Stalnaker to the heroic but not widely accepted idea that the *might* in (1-b) is a wide-scoped epistemic operator.

It is helpful to identify the truth-conditions that are predicted for (1-a). Since the conditional already provides a selection function, a Stalnakerian analysis of the meaning of (1-a) naturally requires a purely temporal analysis of *will*. To keep this purely temporal meaning distinct from the selectionist one, introduce the operator FUT endowed with a purely temporal semantics.

$$\llbracket \text{FUT } P \rrbracket^{w,t} = 1 \text{ iff } \llbracket P \rrbracket^{w, \blacktriangleright(t)} = 1.$$

Letting ‘**keith plays**’ denote the proposition that Keith plays at the relevant time, this semantic package predicts the following truth-conditions:

$$(3) \quad \llbracket (1-a) \rrbracket^{w,t} = \llbracket \text{Keith plays } >_f \text{ Mick sings} \rrbracket^{w,t} = \llbracket \text{Mick sings} \rrbracket^{\sigma(\mathbf{f}(w) \cap \text{keith plays}, w), \blacktriangleright(t)}$$

Any plausible truth-condition for (1-b) should be compatible with this.

## 4 The basic equivalence

The formal fact at the center of this paper is that, within a selectionist framework for *will*, it is possible to reconstruct Stalnakerian truth-conditions for *will*-conditionals by a different route. This route generates a different set of predictions about the validity of inference patterns. (One would want to say a different *logic for the conditional* if the syntactic differences didn’t make a proper logical comparison less than straightforward.) This section takes up these points in turn.

For purposes of integration with the selection semantics for *will*, assume a restrictor analysis for *if* clauses [16, 17]. Since we have syntacticized modal bases, this means adopting a treatment of conditionals antecedent as assignment shifters along lines that are presented by von Stechow [9] and applied to selectionist *will*-conditionals by Cariani and Santorio in [7]. To avoid dealing explicitly with assignment functions, I give a somewhat coarser presentation of the idea than either of these sources. Let  $f + \mathbf{P}$  be the result of operating on the assignment function so that  $f$  is assigned to  $\lambda w. \mathbf{f}(w) \cap \mathbf{P}$ . Whatever the exact entry for *if* (there are choices we don’t need to consider here), it needs to deliver this constraint (where  $\mathbf{P}$  denotes the set of P-worlds):

$$\llbracket (if P)((\text{TENSE+})\text{MODAL}_f Q) \rrbracket^{w,t} = 1 \text{ iff } \llbracket (\text{TENSE+})\text{MODAL}_{f+\mathbf{P}} Q \rrbracket^{w,t} = 1$$

The key claim I intend to establish is that the restriction + selection analysis of *will*-conditionals delivers the same truth-conditions as Stalnaker’s analysis of *if* combined with a temporal analysis of *will* (as specified in the semantics of FUT). This is easily ascertained:

$$\begin{aligned} & \llbracket (if \text{Keith plays})(\text{PRES}(\text{WOLL}_f(\text{Mick sings}))) \rrbracket^{w,t} = \\ & \llbracket \text{PRES}(\text{WOLL}_{f+\text{keith plays}}(\text{Mick sings})) \rrbracket^{w,t} = \llbracket \text{Mick sings} \rrbracket^{\sigma(\mathbf{f}(w) \cap \text{keith plays}, w), \blacktriangleright(t)} \end{aligned}$$

These are precisely the truth-conditions we predicted in (3) within the Stalnakerian framework for *Keith plays*  $>_f$  FUT(*Mick sings*). This moral is fully general: combining selection semantics for *will* with the restrictor analysis predicts Stalnakerian truth-conditions for *will*-conditionals.

Let us move on to the predictions about logic. A form of CEM is valid for *will*-conditionals:

$$\text{CEM-w.} \quad \vdash (if P)(will_f Q) \vee (if P)(will_f \text{ not } Q)$$

Interestingly, in addition to this, the semantics validates an analogue of (I/E), appropriately restricted to *will*-conditionals.



$$\text{I/E-w.} \quad (\text{if } P)(\text{if } Q)(\text{will}_f R) \dashv\vdash (\text{if } P \ \& \ Q)(\text{will}_f R)$$

I'm glossing over some needed syntactic work here by assuming it's possible for two *if*-clauses to independently restrict a single modal.

We still have no account of the duality between *will*-conditionals and their *might*-counterparts. In other words, we predict:

$$\text{WM-w.} \quad (\text{if } P)(\text{will}_f Q), (\text{if } P)(\text{might}_f \text{ not } Q) \not\vdash \perp$$

This inference is not classically valid:  $(\text{if } P)(\text{will}_f Q)$  and  $(\text{if } P)(\text{might}_f \text{ not } Q)$  can both be true. As Santorio [25] notes, the dialectic surrounding WM closely mirrors the dialectic surrounding epistemic contradictions—sentences of the form  $P \ \& \ \text{might not } P$ —on the sort of analysis favored by Yalcin [34]. One paradigmatic approach to these is to invoke a more permissive notion of consequence than the classical concept of preservation of truth at an index. Following the dynamic tradition, and more locally [31, 34], introduce a concept of *acceptance*, and define a concept of entailment as preservation of acceptance, so that every truth-preserving argument is also acceptance-preserving.<sup>5</sup> This move might appear conceptually unmotivated in a straightforwardly contextualist semantic framework. There are two possible paths to address this worry: one is to take the semantic framework in more ‘information’-friendly direction; the other is to suggest that the informational notion of consequence can serve as a pragmatic overlay on top of a contextualist semantics. I believe both paths are viable. Here, however, I illustrate a way of walking down the first path, deferring more extensive discussion to [3].

Let  $s$  be an information state—modeled, as usual a set of worlds. Lift semantic values so that indices also have an information state coordinate. At this point, the standard path would be to define a concept of acceptance and use it to define informational entailment. However, the present framework needs to coordinate the relationship between information states and modal bases. Say that information state  $s$  *accepts*  $P$  (in context  $c$ ) iff for all  $v \in s$ ,  $\llbracket P \rrbracket^{s,v,t_c} = 1$ . Suppose that, for each time  $t$ , the background model fixes the historical modal base at  $t$  ( $\mathbf{h}_t$ ). Say that an information state  $s$  is *historically eligible* in context  $c$  if and only for every world  $w \in s$ ,  $\mathbf{h}_{t_c}(w) \in s$ . Informally, we zero-in on those information states  $s$  such that the relevant historical modal base function in  $c$  does not map from inside to outside of  $s$ . Because historical modal bases represent equivalence relations (specifically: duplication up to a time), they induce partitions. Under these assumptions, the eligible information states (in  $c$ ) can be equivalently characterized as those states that do not cut through the cells of the historical partition determined by  $\mathcal{M}$  at  $t_c$ . Finally, say that an argument with premises  $P_1, \dots, P_x$  and conclusion  $R$  is a *coordinated informational entailment* (written  $P_1, \dots, P_x \vdash_{c\text{-info}} R$ ) iff for any context  $c$ , no information state  $s$  that is *historically eligible* in  $c$  accepts each of the  $P_i$ 's (in  $c$ ) but fails to accept  $R$  (in  $c$ ).

Anything that is valid in the truth-preservation sense (and specifically I/E) is informationally valid—a standard fact about informational consequence that also holds for its coordinated variant. Furthermore, though not classically valid, WM-w is informationally valid.

$$\text{WM-w.} \quad (\text{if } P)(\text{will}_f Q), (\text{if } P)(\text{might}_f \text{ not } Q) \vdash_{c\text{-info}} \perp$$

This completes the presentation of the basic framework.<sup>6</sup>

<sup>5</sup>This is the general approach Santorio takes in his path semantics [25], though Santorio goes for a more radical revision of the fundamental objects of the semantics than what I am exploring here.

<sup>6</sup>If we are willing to move to a more standard dynamic framework and to assume that *will* selects out of an information state, we can avoid some of these novelties. Set times aside and evaluate relative to information state/world pairs:

$$\llbracket P > Q \rrbracket^{s,w} = 1 \text{ iff } \llbracket Q \rrbracket^{s \cap P,w} = 1$$

## 5 Stalnaker's conditional factorized.

So far, I focused on conditionals with *will* in their consequents, elaborating a picture that was already outlined in [7]. We can try something more general and assign *all* conditional sentences broadly Stalnakerian truth-conditions while also having an logic that classically vindicates generalizations of CEM and I/E, as well as informationally vindicating WM.

The standard idea associated with Kratzer's restrictor approach is that, when *if* lacks an overt modal to restrict, a covert necessity modal is posited in the LF. In particular:

(4) If Keith played, Mick sang.

gets reconstructed as:

(5) (*if* Keith played)(MUST<sub>f</sub> Mick sang )

It is natural to wonder what would happen if instead of positing a covert necessity modals, we posited covert *selection* modals (this analysis is independently entertained but not endorsed in [22]). Suppose for instance that (4) gets assigned the logical form in (6) (from now on, I omit the redundant present tense):

(6) (*if* Keith played)(WOLL<sub>f</sub> Mick sang )

Evidently, (5) and (6) differ in truth-conditions.

(7) a.  $\llbracket (5) \rrbracket^{w,t} = 1$  iff  $\forall v \in \mathbf{f}(w) \cap \mathbf{keith\ played}, \llbracket \text{PAST Mick sing} \rrbracket^{v,t} = 1$   
 b.  $\llbracket (6) \rrbracket^{w,t} = 1$  iff  $\llbracket \text{PAST Mick sing} \rrbracket^{\sigma(\mathbf{f}(w) \cap \mathbf{keith\ plays}, w), \blacktriangleright(t)} = 1$

Remarkably, the logical form in (6) delivers fully Stalnakerian truth-conditions for (4).

This analysis accounts for the plausibility of those surface forms that appear to instantiate conditional excluded middle like:

(8) Either Mick sang if Keith played or Mick didn't sing if Keith played.

More precisely, the following holds:

CEM-c.  $\vdash (if P)(WOLL_f Q) \vee (if P)(WOLL_f \text{ not } Q)$

As before in addition to CEM-c the semantics classically (and thus informationally) validates a pattern that accounts for instances of import/export involving bare conditionals.

I/E-c.  $(if P)(if Q)(WOLL_f R) \dashv\vdash (if P \ \& \ Q)(WOLL_f R)$

Interestingly, the system makes available a second logical form for I/E, which is however invalid.

I/E-c2.  $(if P)(WOLL_f (if Q)(WOLL_f R)) \not\equiv (if P \ \& \ Q)(WOLL_f R)$

There are some interesting questions to press about I/E-c2. On the one hand, it incorporates the fascinating suggestion that there might be a way of defining a binary conditional connective with Stalnakerian truth-conditions *and* logic inside a restrictor theory: just say that  $P \rightarrow Q = (if P)(WOLL_f Q)$ . On the other, it raises question about how responsible the theory ought to be

---

$\llbracket will P \rrbracket^{s,w} = 1$  iff  $\llbracket P \rrbracket^{s,\sigma(s,w)} = 1$

This will deliver CEM-w, I/E-w and WM-w on a standard dynamic notion of entailment. Thanks to Simon Goldstein for making the point. The original inspiration for this account is Mandelkern's [21] reconstruction of McGee's semantics from [24].

for the predictions of these additional structures. Does the validity of I/E-c properly account for the felt intuitive validity of import/export, given that there is a *very nearby* LF that is invalid? I will say a bit more about this towards the end.

A generalization of the approach in the previous section will also deliver the informational validity of an analogue of WM.

$$\text{WM-c. } (if P)(\text{WOLL}_f Q), (if P)(\text{might}_f \text{ not } Q) \vdash_{c\text{-info}} \perp$$

This, then, is the core of the account of conditionals that emerges if we start with selection modals and work our way out to a general approach to conditional meaning.

The news value of this tale is this: conventional wisdom has it that Stalnakerians are forced into a tradeoff between CEM on the one hand, and I/E and WM on the other. Stalnaker's semantics accounts for the evidence for CEM at the cost of invalidating I/E and WM, which has invited some uncomfortable resistance strategies. Associating selection behavior with some modals shakes up the I/E dialectic, while moving towards coordinated informational consequence recovers WM.<sup>7</sup>

## 6 Counterhistorical restriction.

The proposal of §5 posits a covert WOLL in the logical structure of bare conditionals. What modal base should it have? If this covert operator is to parallel overt *will*, we should expect it (in light of our present assumptions) to have a historical modal base. As Simon Goldstein and John Hawthorne (p.c.) note, this hypothesis falters on counterhistorical antecedents—those antecedents that are incompatible with the settled history up to the time of the context. Suppose that in  $w$ , Ann takes a test on Tuesday. On Wednesday, I say:

- (9) If Ann took her test on Monday, it was graded on Tuesday.

If the restriction of WOLL is historical,  $\mathbf{f}(w)$  is the set of worlds that duplicate  $w$  up to the time of my utterance. That means that, at every world in  $\mathbf{f}(w)$ , Ann took the test on Tuesday, and so the restriction with the antecedent of (11) is vacuous.<sup>8</sup>

A natural alternative—and a solution to this problem—is to assign that covert WOLL an epistemic modal base. It is convenient, but philosophically quite substantial, to assume that at any given time, the epistemic modal base is a coarsening of the historical one. Let  $\mathbf{h}_t$  be the historical modal base at  $t$  and  $\mathbf{e}_t$  be the epistemic modal base at  $t$ .

$$\forall w, v, \forall t : \text{if } v \in \mathbf{h}_t(w), v \in \mathbf{e}_t(w)$$

The informal meaning of this constraint is that worlds can only be distinguished by the epistemic modal base (at  $t$ ) if they there is some qualitative difference between them (at  $t$ ).

The assumption that the covert WOLL has an epistemic modal base correctly handles (11). The spirit of the objection is not yet defeated, however: *will*-conditionals still rely on historical modal bases. Given that, the problem seems to reappear for counterhistorical *will*-conditionals, like:

- (10) If Ann took her test on Monday, it will be graded on Thursday.

<sup>7</sup>The most prominent alternative for validating all of these principles goes *homogeneity*-based account of conditionals [10, 1, 26, 5]. In this paper, I am not concerned with arguing against homogeneity-based accounts—in part because the theory of homogeneity is itself in flux ([19]) and in part because there are some problems for some of the existing formulation ([7, 25]).

<sup>8</sup>Though I haven't specified how the selection function is to operate on the empty set, this is certainly problematic: all counterhistorical conditionals would depend only on what goes on at a single world.

I propose that examples like (10) motivate importing an idea that proponents of restrictor approaches have been advocating independently. Sometimes, a covert modal is posited even in the presence of an overt modal [11, 14]. As before, I differ from this tradition in positing selection modals all the way down. The result is that the logical form of (10) is something like:

$$(11) \quad (if_1 \text{ test monday})(WOLL_{f1} \text{ will}_{f2} \text{ grade thursday})$$

Suppose that  $f1$ , which is co-indexed with the *if*-clause, is assigned to the epistemic modal base  $\mathbf{e}$ , while  $f2$  is assigned to the historic modal base  $\mathbf{h}$  (at the relevant times). This allows for non-vacuous counterhistorical restriction, because it is the  $\mathbf{e}$ , and not  $\mathbf{h}$ , that gets the restriction:

$$(12) \quad \llbracket (11) \rrbracket^{w,t} = \llbracket \text{grade thursday} \rrbracket^{\sigma(\mathbf{h}(\sigma(\mathbf{e}(w)) \cap \text{test monday}), \blacktriangleright(t)), \blacktriangleright(t))}$$

In support of this move, I note that it is independently needed, not only in the class of cases that standardly motivate double modalization, but even in cases that are extremely close to the present dialectic. Suppose that the language contains *some* historical modals like the cumbersome but intelligible *it is historically necessary that*. Then, consider the conditional,

$$(13) \quad \text{If Ann took her test on Monday, it is historically necessary that it will be graded on Tuesday.}$$

For the reasons explored above, we cannot treat (13) as involving a counterhistorical restriction on the historical necessity modal. The only sensible strategy within a restrictor analysis is to posit a covert modal for the *if*-clause to restrict.

## 7 On the proliferation of covert modals.

Mandelkern [22] sketches and rejects an account of conditionals like the one from §5. His first concern is that the heavy reliance on covert modals is suspect for reasons having to do with learnability of conditional constructions across languages (see also [27]).

How do we learn to insert covert modals in all the needed places? And how do we learn which modal to put in? We can imagine a wide array of options that would seem to be open to children concerning what kind of modal we put in (existential? universal? epistemic, deontic, metaphysical?) as well as when to insert them (always? sometimes? never?). How do children (both within and across languages) converge on the correct combination? [22, p. 312]

It is hard to address this challenge without a major digression, and this is not the place for that. But I believe the challenge is not as pressing as Mandelkern suggests. For one thing, figuring out force and flavor of covert modals is not significantly harder than figuring out force and flavor of overt ones. Setting this point aside, it is not clear exactly what children are supposed to not be able to do. On most standard models for acquisition, children can entertain and test syntactic and semantic hypotheses of the relevant complexity: slotting covert elements into the relevant LFs and fixing the relevant parameter values involves searching and testing a relatively small hypothesis space.

There are also less direct reasons for caution. If there is a problem with covert modals here, then there is a problem in all the other places where semanticists have found it plausible to resort to covert modalities. Furthermore, if there is a problem with covert *modals*, then there likely is a problem with covert *elements* more generally. Semantics without covert elements

would be a respectable research program, but it comes with a large unfinished agenda. In light of these considerations, I submit that we have no evidence that the theory is unlearnable.

## 8 Collapse and the identity principle.

Mandelkern’s second concern concerns the logic of the conditional. Accounts that validate I/E fail to deliver the validity of an impressively intuitively compelling schema—the identity schema *if P, then P*. More specifically, [23] argues that Identity is involved in a striking collapse result in which the only other reasonable culprit is I/E. Let ‘ $\rightarrow$ ’ be a conditional connective. Identity, together with a weak monotonicity constraint, yields the principle:

(EC) if P entails Q,  $\vdash P \rightarrow Q$

Here are two consequences of (EC):

- (14) a.  $\vdash (\neg(P \rightarrow Q) \ \& \ Q) \rightarrow \neg(P \rightarrow Q)$   
 b.  $\vdash ((\neg(P \rightarrow Q) \ \& \ Q) \ \& \ P) \rightarrow Q$

However, applying import/export reasoning to (14-b) yields:

(15)  $\vdash (\neg(P \rightarrow Q) \ \& \ Q) \rightarrow (P \rightarrow Q)$

If so,  $(\neg(P \rightarrow Q) \ \& \ Q)$  (the common antecedent of (14-a) and (15)) is logically false, and collapse follows (details in §3 of [23]). Mandelkern rejects the I/E step and develops an impressive theory which accounts for why I/E—though invalid—seems to not fail in the indicative domain.

My first reaction to this argument would be to reach for a bullet-biting, I/E-supporting response. Restrict Identity to its non-junk instances (this is part of a more general line of responses to collapse results that is briefly entertained in [5]). The strategy would grant that Identity is a simple-looking principle in conditional semantics, but resist the idea that the intuitive support for it extends to cases in which P is instantiated by complex, conditional-embedding (and possibly even contradictory) sentences such as  $((\neg(P \rightarrow Q) \ \& \ Q) \ \& \ P)$ .<sup>9</sup>

However, a more concessive reply is also available, if one were so inclined. It is possible within the present framework to embrace the core tenets of Mandelkern’s [23, §6-7] positive proposal. This involves invalidating I/E but recovering the prediction that there are no outright counterexamples to I/E by capturing it as valid under a more permissive, ‘pragmatic’, notion such as (a modernized version of) Strawson entailment.

Here is a (too) compressed presentation of Mandelkern’s framework. First: have the interpretation function keep track of the presuppositional content of expressions, though not in the standard trivalent way. In particular, given input sentence P, the interpretation function outputs a pair whose first coordinate is 1 if the presuppositions of P are satisfied and 0 otherwise; the second coordinate of the pair is the standard truth-value of P. Second: say that P is *Strawson-valid* iff for all models and indices P’s semantic value is  $\langle 1, 1 \rangle$  or  $\langle 0, * \rangle$  where ‘\*’ is a placeholder for either 1 or 0. By contrast, P is *logically valid* iff it is guaranteed to have semantic value  $\langle *, 1 \rangle$ . Informally: logical validity is guaranteed truth, while Strawson validity is guarantee of a status that is either truth with presuppositions being satisfied or presupposition violation. Third: enrich indices a parameter for keeping track of ‘local contexts’. Fourth: assume that the

<sup>9</sup>Even if this were right, this argument would not end here. Someone who wanted to bite the bullet would have to reply to Mandelkern’s case for Identity in §4.2 of [23].

indicative conditional  $P \rightarrow Q$  presupposes a ‘local’ version of Stalnaker’s indicative constraint.<sup>10</sup> If that was too dense, here is an implementation of the semantics of the indicative conditional:

$$(16) \quad \llbracket P \rightarrow Q \rrbracket^{\kappa, w} = \begin{aligned} &\langle 1, * \rangle \text{ iff } \forall v \in \kappa, \sigma(\llbracket P \rrbracket^{\kappa}, v) \in \kappa \\ &\langle *, 1 \rangle \text{ iff } \llbracket Q \rrbracket^{\kappa \cap \llbracket P \rrbracket^{\kappa}, \sigma(\llbracket P \rrbracket^{\kappa}, w)} = \langle 1, 1 \rangle \end{aligned}$$

This semantics Strawson-validates I/E, even though the inference is not logically valid.

My concluding point is that it is possible to replicate Mandelkern’s story within my framework. Earlier, I noted that we have built a kind of conditional connective with the form  $(if P)(WOLL_f Q)$ . The present idea is to translate everything that Mandelkern says about  $P \rightarrow Q$  in terms of  $(if P)(WOLL_f Q)$ . Nearly every part of the story carries over smoothly. The only element that requires special attention is the implementation of the indicative constraint. Instead of associating this with the conditional, associate it with  $WOLL$ . As a result,  $WOLL_f(Q)$  presupposes in  $w$  that for every world  $v \in \kappa$ ,  $\sigma(\mathbf{f}(w), v) \in \kappa$  (note the reinstatement of modal bases which are not part of Mandelkern’s setup). For  $WOLL$ -conditionals, this analysis works out to:

$$(17) \quad \llbracket (if P)(WOLL_f Q) \rrbracket^{\kappa, w} = \begin{aligned} &\langle 1, * \rangle \text{ iff } \forall v \in \kappa, \sigma(\mathbf{f}(w) \cap \llbracket P \rrbracket^{\kappa}, v) \in \kappa \\ &\langle *, 1 \rangle \text{ iff } \llbracket Q \rrbracket^{\kappa \cap \llbracket P \rrbracket^{\kappa}, \sigma(\mathbf{f}(w) \cap \llbracket P \rrbracket^{\kappa}, w)} = \langle 1, 1 \rangle \end{aligned}$$

For full comparability with the theories in this paper we’d have to reinstate a time coordinate, but that can be done modularly.

In §5, I pointed out that I/E might get one of two forms: either I/E-c or:

$$I/E\text{-c2.} \quad (if P)(WOLL_f(if Q)(WOLL_f R)) \dashv\vdash (if P \ \& \ Q)(WOLL_f R)$$

I also noted then that it is puzzling that we don’t detect a truth-conditional difference corresponding to this distinction. Interestingly, under (17), I/E-c2 goes through as a Strawson equivalence in the same, rather nuanced sense, in which Mandelkern’s conditional Strawson-validates I/E.

The upshot of this round of Mandelkern vs. Mandelkern is this: contrary to a claim in [22], the development in [23] is not an obstacle for the factorized Stalnakerian analysis. It might even illuminate why we don’t detect a difference in readings corresponding to two possible construals of I/E. Whether the factorized analysis is tenable in full generality remains to be seen.

## References

- [1] Maria Bittner. Topical referents for individuals and possibilities. In R. Hastings, B. Jackson, and Z. Zvolenszky, editors, *Semantics and Linguistic Theory*, volume 11, pages 36–55, 2001.
- [2] David Boylan and Ginger Schultheis. The qualitative thesis. Rutgers University and the University of Chicago, ms.
- [3] Fabrizio Cariani. *The Modal Future*. Cambridge University Press, forthcoming.
- [4] Fabrizio Cariani. The modal future hypothesis debugged. Northwestern University, ms.
- [5] Fabrizio Cariani and Simon Goldstein. Conditional heresies. *Philosophy and Phenomenological Research*, forthcoming.
- [6] Fabrizio Cariani and Paolo Santorio. Selection semantics for *Will*. In *Proceedings of the 19<sup>th</sup> Amsterdam Colloquium*. 2017.
- [7] Fabrizio Cariani and Paolo Santorio. *Will* done better: Selection semantics, future credence, and indeterminacy. *Mind*, 127:129–165, 2018.

<sup>10</sup>See also [2] for an alternative, non-presuppositional ‘localization’ of the indicative constraint.

- [8] Cleo Condoravdi. Temporal interpretation of modals: Modals for the present and for the past. In D. Beaver, L. Casillas, B. Clark, and S. Kaufmann, editors, *The Construction of Meaning*, pages 59–88. CSLI Publications, Palo Alto, CA, 2002.
- [9] Kai von Fintel. *Restrictions on Quantifier Domains*. PhD thesis, UMass, Amherst, 1994.
- [10] Kai von Fintel. Bare plurals, bare conditionals, and *Only*. *Journal of Semantics*, 14:1–56, 1997.
- [11] Annette Frank. *Context Dependence in Modal Constructions*. PhD thesis, University of Stuttgart, 1997.
- [12] Atle Grønn and Arnim von Stechow. Tense. In M. Aloni and S. Dekker, editors, *The Cambridge Handbook of Formal Semantics, Part III*, pages 313–341. CUP, 2016.
- [13] Michela Ippolito. *Subjunctive Conditionals*. MIT Press, 2013.
- [14] Magdalena Kaufmann and Stefan Kaufmann. Conditionals and modality. In S. Lappin, editor, *The handbook of contemporary semantic theory*, pages 274–312. Wiley, 2015.
- [15] Stefan Kaufmann. Conditional truth and future reference. *Journal of Semantics*, 22(3):231–280, 2005.
- [16] Angelika Kratzer. Conditionals. In A. von Stechow & D. Wunderlich, editor, *Semantics: An International Handbook of Contemporary Research*. De Gruyter, 1991. from the *Semantics Archive*.
- [17] Angelika Kratzer. *Modals and Conditionals*. Oxford University Press, 2012.
- [18] Angelika Kratzer. Chasing hook. In Lee Walters and John Hawthorne, editors, *Conditionals, Probability, and Paradox*. Oxford University Press, forthcoming.
- [19] Manuel Križ. *Aspects of homogeneity in the semantics of natural language*. PhD thesis, University of Vienna, 2015.
- [20] David Lewis. *On the Plurality of Worlds*. Blackwell, 1986.
- [21] Matthew Mandelkern. Import-export and ‘and’. *Philosophy and Phenomenological Research*, forthcoming.
- [22] Matthew Mandelkern. Talking about worlds. *Philosophical Perspectives*, forthcoming.
- [23] Matthew Mandelkern. If  $p$ , then  $p$ . Oxford University, ms.
- [24] Vann McGee. A counterexample to *Modus Ponens*. *Journal of Philosophy*, 82:462–471, 1985.
- [25] Paolo Santorio. Conditional excluded middle in informational semantics. In *Proceedings of the 21<sup>st</sup> Amsterdam Colloquium*, pages 385–394, 2017.
- [26] Philippe Schlenker. Conditionals as definite descriptions. *Research on language and computation*, 2(3):417–462, 2004.
- [27] Katrin Schulz. Troubles at the semantics/syntax interface: Some thoughts about the modal approach to conditionals. In *Proceedings of sinn und bedeutung*, volume 14, pages 388–404, 2010.
- [28] Robert Stalnaker. A theory of conditionals. *Studies in Logical Theory, American Philosophical Quarterly*, Monograph Series 2:98–112, 1968.
- [29] Robert Stalnaker. A defense of conditional excluded middle. In W.L. Harper, G.A. Pearce, and R. Stalnaker, editors, *Ifs*, pages 87–104. Springer, 1981.
- [30] Robert Stalnaker and Richmond H. Thomason. A semantic analysis of conditional logic. *Theoria*, 36(1):23–42, 1970.
- [31] Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25:221–261, 1996.
- [32] Arnim von Stechow and Sigrid Beck. Events, times and worlds - an LF architecture. In C. Fortmann, A. Lübbe, and I. Rapp, editors, *Situationsargumente im Nominalbereich*, pages 13–47. de Gruyter, 2015.
- [33] Robert J. Williams. Defending conditional excluded middle. *Noûs*, 44(4):650–668, 2010.
- [34] Seth Yalcin. Epistemic modals. *Mind*, 116(4):983–1027, 2007.



# Tagging: Semantics at the Iconic/Symbolic Interface

Gabriel Greenberg

University of California, Los Angeles  
gabriel.greenberg@gmail.com

## Abstract

*Tagging* is the phenomenon in which regions of a picture, map, or diagram are annotated with words or other symbols, to provide descriptive information about a depicted object. The interpretive principles that govern tagged images are not well understood, due in part to the difficulty of integrating pictorial and linguistic semantic rules. Rather than directly combining these rules, I propose to use the framework of *perspectival feature maps* as an intermediary representation of content, in which the outputs of pictorial and linguistic interpretation may be assimilated. The result is a simple and compositional semantics for tagged images.

## 1 Tagging

Human communication is multi-modal. Spoken and signed words are accompanied by pictorial gestures and emotive facial expressions. Written words are enriched with illustrations, diagrams, and emoji. Newspaper articles come with photographs, and photographs come with captions. Maps are annotated with detailed geographic labels. And technical illustrations contain numerals, labels, call-out boxes, and more. In nearly every domain of human transaction, symbolic and iconic signs are integrated to efficiently express rich tapestries of information. Ultimately, the science of semantics must come to terms with this multi-modal outpouring.

We can loosely categorize multi-modal signs into three basic types. In **egalitarian** representations, icons and linguistic signs each express their own modality-specific content, which together contribute to a richer discourse content. Egalitarian representation are exemplified by sentential captions on photographs (Alikhani and Stone, 2019), by illustrated narratives and instructions (Alikhani and Stone, 2018b), and by many cases of coverbal gesture (Lascarides and Stone, 2009a,b).

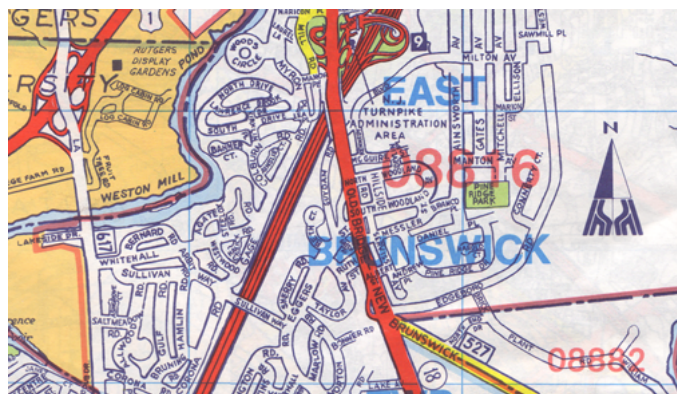
In **language-dominant representation**, icons are used to enrich a linguistic expression; the interpretation of the iconic elements modulate the semantic contribution of the linguistic whole. Language-dominant multi-modality has been widely studied in recent years. Examples include pro-speech, co-speech, and post-speech gesture (Schlenker, 2018a; Tieu et al., 2017), iconic modulation of words (Schlenker, 2018a), and a wide variety of iconic enrichments in sign language (Schlenker, 2018b). See Schlenker (2019) for an overview of the state of the art.

The last category of multi-modal representation has received relatively little attention from semanticists, but is ubiquitous in human affairs. In **icon-dominant representation**, words, phrases, and non-linguistic symbols are used to enrich a picture, map, or diagram. The interpretation of the symbolic elements modulate the semantic contribution of the dominant image. **Tagging** is the phenomena in which symbolic **tags** are associated with specific subregions of the dominant icon.

Maps provide a vivid example of tagging, where names are inscribed throughout the map to indicate the location and identity of landmarks. Other cases include the descriptions, labels, and numerals featured in technical drawings and mass-media imagery, and the variable letters that are used to supplement mathematical and logical diagrams. Speech balloons in comics are



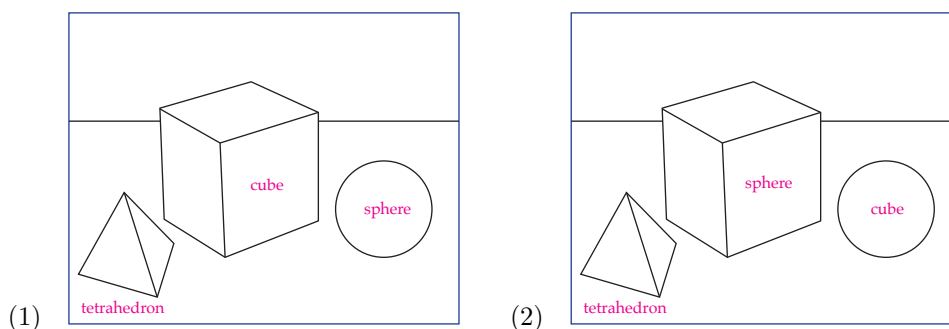
a genre-specific form tagging. And in, sign languages, classifier constructions involve the use of symbolic hand shapes to tag iconically presented paths and locations.



**Figure 1:** Tagging as the annotation of an icon with linguistic symbols. Excerpt from from *Middlesex County Atlas* (2002), pgs. 32-33.

In this paper I’ll focus on the paradigm case of **tagged images**: perspectival pictures enriched with linguistic tags. (I’ll use “picture” and “image” interchangeably; I treat maps as a class of picture.) In tagged images, symbolic tags play the role of contributing identifying and descriptive information about particular *objects* (broadly construed) for which the picture provides pictorial information. An adequate semantics of tagging must address three central explanatory problems.

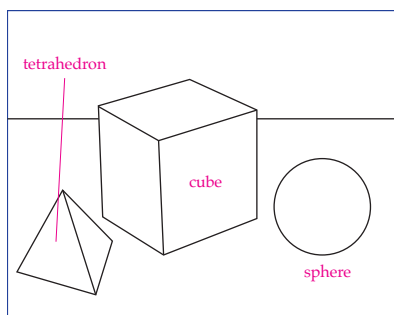
**Problem 1: semantic significance of tag placement.** The placement of tags within a printed images contribute to the accuracy-conditions of the whole tagged image by indicating which depicted objects are associated with the contents of the tags. For example, relative to a realistic scene, (1) is accurate. But swapping the position of “sphere” and “cube” results in (2) which is not accurate. In effect, tagging locates linguistically expressed properties within pictorial space. Understanding how this is done is the central challenge for a semantics of tagging.



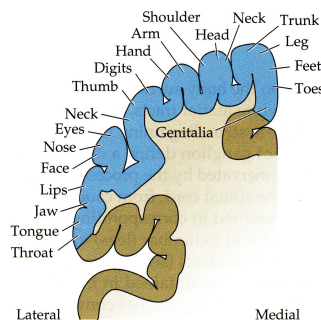
**Figure 2:** Relative to a realistic scene, image (1) is accurate, but (2) is not. They differ only with respect to the placement of “sphere” and “cube.”

**Problem 2: flexibility of tag placement.** The ultimate semantic contribution of tags—to supply descriptive information about depicted objects— can be achieved through a variety of

expressive means. In Figure 3, for example, the tag “sphere” is associated with the image by placement *proximal* the part of the image it tags; “cube” is associated with a part of the image by placement *within* it; and “tetrahedron” is associated with the image through *line linking* (or *indication* in Alikhani and Stone 2018a, pg. 3555).



**Figure 3:** Tagging can be expressed through a variety of visual means: line linking, inclusion, and proximity. (Problem 2)



**Figure 4:** Tagging relations beyond predication in a diagram of the sensory cortex. (Problem 3) From Purves et al. (1997) *Neuroscience*, pg. 22.

**Problem 3: variety of tagging relations.** Nouns and names can both be tags, but because nouns and names denote objects in different semantic categories, the semantic significance of tagging must itself be allowed to vary. A picture of a person tagged with the name “Kiara” expresses the content that the person *is identical with* [[Kiara]]; a picture of person tagged with the noun “professor” expresses the content that the person *has the property* [[professor]]. More extreme variations are common. Consider the tag “08816” on a map; here the relevant relations is *has the zipcode*. In Figure 4, the relation *brain region X processes information from body region Y* is put to work in a standard depiction of the somatosensory cortex. As these examples show, the correct tagging relation cannot be determined simply as a matter of syntactic or semantic type, but must advert to contextual and discourse-sensitive constraints. Instead, I view tagging relations as a species of multi-modal **coherence relation** (Alikhani and Stone, 2018b), a structural link in discourse which functions to bind together independent discursive elements.

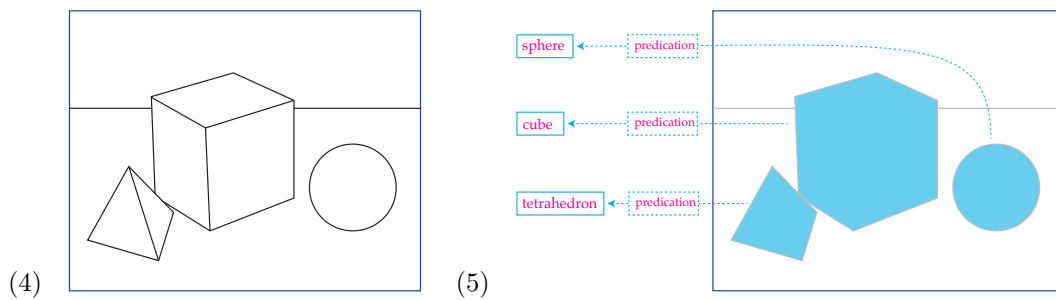
The prospect of a tagging semantics which answers Problems 1-3 presents us with a theoretical puzzle. On one hand, linguistic and pictorial elements demand radically different kinds of semantic analyses (Giardino and Greenberg, 2015; Schlenker, 2019). On the other hand, they cannot be entirely separated: to compute the content of a tagged image, one cannot simply divide it into a picture and a set of tags, compute their respective contents, and put them back together again. There would be no way of tracking which properties went with which depicted objects when they were recombined. In this paper, I’ll recruit the theoretical apparatus of *feature maps* to serve as the nexus point where linguistic and pictorial information streams may come together. This approach will ultimately make it possible to formulate a simple and compositional semantics for tagged images.

## 2 Syntax

The underlying syntactic structure of a tagged image can be divided into a **pure image** that is free from tags, a set of tags, and a set of **linking** relations that hold between regions of the pure image and pairs of tags and relation-symbols (cf. Alikhani and Stone, 2018b, pg. 2). On

this model, the location of a tag on the printed page is not itself part of syntax, but is a *signal* of a syntactic relation. Tags themselves have no location; they are associated with regions of the picture plane by abstract syntactic links. The rationale for this way of approaching tag placement will emerge shortly. Here the syntax of a tagged image stands to the image on the printed page roughly as a sentence's syntax stands to its phonology. I model this structure formally as follows, illustrated in Figure 5.

- (3) A **tagged image**  $T = \langle I, tag \rangle$ , where:
- (i)  $I$  is a pure image;
  - (ii)  $tag$  is a (partial) function from regions of  $I$  to pairs  $\langle s, r \rangle$  where  $s$  is a tag-symbol, and  $r$  is a relation-symbol.



**Figure 5:** The syntax of a tagged image: (4) a pure image; (5) a set of linking relations between picture regions and pairs of tags and relation-symbols.

The structure of the pure image itself can be understood as a 2D plane segment where regions are associated with colors. (A more complete account might include lines and line-types, textures, and color regions.) Formally:

- (6) A **pure image**  $I = \langle P, d, color \rangle$  where:
- (i)  $P$  is a set of points;
  - (ii)  $d$  is a Euclidean metric over  $P$  which defines a 2D rectilinear space;
  - (iii)  $color$  is a (total) function from points or small regions of  $P$  to colors (or values).

Meanwhile the linguistic constituents of tagged images are sub-clausal phrases that include names, numerals, nouns, adjectives, as well as definite and indefinite descriptions. In this paper, I'll set aside indexical sentences, like “you are here,” which can also play a tagging-like role. Non-linguistic tags, as in a map, include a variety of specialized symbols which may be listed in a legend or conventional for a type of discourse.

Tags themselves are associated with regions of the picture plane by abstract structural links.<sup>1</sup> I'll assume that the regions in question are normally contiguous and correspond to psychologically natural segmentations of the graphical space. The links which connect tags to regions need not be explicitly marked on the printed page, but they are nevertheless signaled through a variety of means. Here a range of defeasible conventions may be invoked to indicate structural links:

<sup>1</sup>Such links are a sub-segmental variant of the text-to-image links posited by [Alikhani and Stone \(2018b\)](#). The semantic function of image sub-regions is anticipated in [Abusch's analysis of visual co-reference \(Abusch 2012, §3-5; Abusch 2015, §4\)](#). I'll assume here that the shapes of the regions in question are perfectly definite, though this is certainly an idealization in many cases.

1. **Inclusion:** a symbol tags the region it is located inside of.
2. **Proximity:** a symbol tags the region whose perimeter it is closest to.
3. **Alignment:** a symbol tags the region with an edge it is spatially aligned with.
4. **Line Linking:** a symbols tags the regions it is linked to by a line or arrow.

Figure 3 features line linking, proximity, and inclusion. Alignment can be seen in Figure 1, in the placement of road and river names.

These conventions often come into conflict. If a tag is *proximal* to one region, then it is necessarily *included* in a different region; which convention applies? The adjudication of these conflicts is sometimes guided by strict selection rules; for example, line-linking always trumps proximity and inclusion. But choosing between proximity and inclusion seems to be more open-ended, informed by spatial cues (e.g. degree of proximity), by semantic match between a picture region and the tag (e.g. “sphere” probably goes with the picture of the sphere), and by known design constraints (e.g. a long word like “tetrahedron” cannot be included in the region it tags). Determining which region is tagged by a symbol is a complex problem that may involve visual cognition, world-knowledge, and general purpose reasoning, in addition to specific conventions.

The proposed solution to Problem 2 is that the variety of expressions of tagging on the page all correspond to a single underlying relation of linking at the syntactic level. Inclusion, proximity, and line-linking are all signals of the underlying syntax, but they are not part of it. This analysis reflects a choice about where to draw the syntax/semantics boundary for the interpretation of tagged images. It divides interpretation into the pre-semantic process of disambiguating tagging links between symbols and regions, on one hand, and the semantic process of computing their meanings, on the other. Part of the rationale for this division of labor is that the two processes demand different kinds of cognitive capacities. The pre-semantic process of disambiguation requires defeasible reasoning and world-knowledge. The semantic process, by contrast, follows a set of narrowly defined interpretive rules, as I’ll show. This bifurcation reflects the traditional view in linguistics and philosophy of language, which allows that general purpose reasoning may be enlisted in syntactic disambiguation, whereas semantics follows monotonic and compositional rules.

The final ingredient in the syntax of tagged images addresses the variety of tagging relations from Problem 3. I propose a set of **relation-symbols** which are explicit in the syntax, but implicit on the printed page. Each link between a symbol and a region is associated with one such symbol. Formally, I treat the *tag* function as a mapping from picture regions to pairs of tag symbols and relation-symbols. I’ll represent *identity* and *predication*, the two most common tagging relations, by the relation-symbols as **id** and **pred** respectively.<sup>2</sup>

### 3 Content

What kind of contents are expressed by tagged images? Because tagging ultimately involves the location of linguistic information within pictorial space, we should model the contents of tagged images after the contents of pure images, rather than those for words or sentences. A popular approach to the contents of pictures understands them as sets of viewpoint-centered

---

<sup>2</sup>Despite variation, there are constraints on how tagging relations may be expressed. Identity and predication appear to be defaults. Other relations are inferred when these defaults are incoherent or otherwise ruled out in context. A further constraint is typographic consistency: tags with the same typographic features are expected to encode the same tagging relation. In Figure 4, for example, tags presented with line linking express the relation *processes information from*, while the two tags “Lateral” and “Medial” in the bottom corners, which use proximity, express predication. Ultimately these factors must be considered within the context of discourse coherence theory more generally.

worlds (Blumson, 2009; Abusch, 2015). But a centered-worlds approach is an awkward fit for the object-oriented semantics of tagging. Tags are associated with *objects*, not entire scenes. Once the set of centered-worlds associated with a pure image is fixed, there is no way of going back into the set and introducing the semantic contribution of the tags, without re-computing the content of the image.

Instead, I propose to use a level of structured content which is intermediate between syntax and accuracy conditions. We can define a simple semantics which separately maps visual and symbolic contents into this intermediate level, which in turn is subject to a general definition of accuracy. The intermediate level is a type of **feature map**, a 2D array whose regions are associated with objects and properties. Feature maps preserve the visual structure of the underlying image while trading the syntactic constituents of the picture for the semantic elements they express. In Greenberg (2019b) I develop a model of pictorial contents as **perspectival feature maps** (PFMs), a type of feature map where each point in the array is associated with a viewpoint-centered direction. Feature maps provides an intuitive interface between the pictorial sign and the background projection semantics, and are straightforwardly extended to incorporate the contents of tagged images.

A perspectival feature map is a two-dimensional array where each point in the array is associated with a viewpoint-centric direction in three-dimensional space; and regions of the array are associated with clusters of objects and properties.<sup>3</sup>

- (7) A **perspectival feature map**  $M = \langle Field, Direction, Cluster \rangle$  where:
- (i) *Field* is a two-dimensional array;
  - (ii) *Direction* is a total function from points in *Field* to 3D directions which satisfy a viewpoint condition.
  - (iii) *Cluster* is a (partial) function from regions of *Field* to feature-clusters.

A **feature cluster** is a sequence  $\langle o, G_1, \dots, G_n \rangle$  where  $o$  is an object and  $G_1, \dots, G_n$  are properties. If a picture express a feature map as content, then the objects of the map's feature clusters correspond to the singular contents of the picture. These are the objects it depicts. The properties of the feature cluster correspond to the attributive contents of the picture. These are the properties it depicts its objects as having (Greenberg, 2018). The structure of PFMs is illustrated in the two figures below.

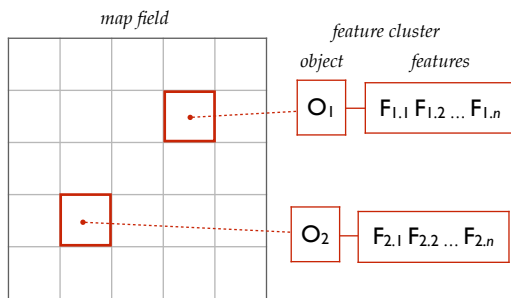


Figure 6: Feature map with feature clusters.

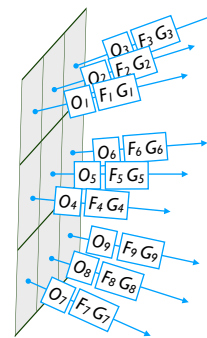


Figure 7: Perspectival feature map with feature clusters and viewpoint-centric directions.

<sup>3</sup>A more complete definition of feature maps would allow for variations in acuity, the representation of relations, the expression of more than one represented object per region (Greenberg, 2019b).

A perspectival feature map can be thought of as a kind of *directional space*— a space whose “dimensions” are directions emanating from a viewpoint, and whose constituents are the objects and properties laid out in that space. It locates each of the objects in its feature clusters in a given direction, and attributes to each its associated properties. A PFM is accurate relative to a world  $w$  and viewpoint  $v$  if and only if the attributions it makes are correct, when it is fixed to the location of  $v$  within  $w$ . In the definition below, given a PFM and its associated *Cluster* function, let *object*( $r$ ) be a function from a region  $r$  to the object of *Cluster*( $r$ ), and *properties*( $r$ ) to the set of properties in *Cluster*( $r$ ).

- (8) A perspectival feature map  $M = \langle Field, Direction, Cluster \rangle$  is **accurate at** a world  $w$  and viewpoint  $v$  iff for every region  $r \in dom(Cluster)$ :
- (i) *object*( $r$ ) is located in *Direction*( $r$ ) from  $v$  in  $w$ ;
  - (ii) *object*( $r$ ) realizes each  $F \in properties(r)$  in  $w$ .

PFMs are a natural choice for the content of tagged images, because they easily accommodate the accumulation of features associated with different regions of the visual field. For pure images, the constituents of the feature clusters are, to a first approximation, entirely visual properties. For tagged images, the content of a tag is simply construed as yet another property which is added to the feature cluster of a depicted object. Thus the word “cube”, used as a tag, expresses the property *cube*. Since the properties expressed by tags enter into feature clusters in the feature map, they automatically inherit the spatial significance of the structure of the map itself. Thus, if the property *cube* is associated with an object  $o$  in the feature map, the final accuracy conditions will simultaneously attribute *cube* to  $o$  and attribute a specific viewpoint-relative direction to  $o$ . The result is that tagged properties are projected out through pictorial space. In this way, the PFM framework allows us to directly co-index tagged properties with visually depicted properties within a single semantic structure. And this is ultimately what is required to capture the distinctive semantic contribution of tagging.

## 4 Semantics

The central challenge for a semantics of tagged images is the problem of integrating the deeply divergent semantic rules for words and pictures. In previous work I’ve advocated a projection semantics for pictures, where a picture is mapped to the set of centered-worlds from which it can be projected (Greenberg, 2019a). But this approach is not easily extended to tagging, since there is no way to locate the properties expressed by tags within the worlds expressed by the pure image, without drawing on the projective rule which defined the pure image content in the first place. Instead, the interpretation of tags would have to be integrated into the definition of projection from the start. The resulting theory could be systematic, but it would not be compositional, since the content of the pure image would not be computed independently of the contents of the tags.<sup>4</sup> On this approach, one can’t simply use one’s semantics for pure images “out of the box.”

To get at the semantics of tagging, we must find a way to integrate the interpretation of the tags with the interpretation of the pure picture. It is here that the feature map framework comes into its own, for PFMs are naturally suited to the task of coordinating diverse streams of object-oriented information within a unified visual frame. My strategy is to exploit the parallelism between tagged image syntax and feature map structure: as each tag is associated

---

<sup>4</sup>Compare the theory of indexing in Abusch (2012, 2015), which integrates object-coreference with rules of projection.



with a region, it contributes a property to the feature cluster expressed by that region. The resulting account smoothly assimilates the content of the pure image and those of the tags.

By stating the semantic rules as mappings from syntactic elements to elements of a feature map, the interpretation of a complete tagged image can be neatly divided into three sub-problems: (i) the semantics for pure images; and (ii) the semantics for the tags themselves; and (iii) the semantics of the region-to-tag links. The resulting semantics for tagged images is compositional, in the sense that the content of the whole is a function of the content of each part (the pure image, the tags, and the relation symbols), and the way they are put together (the linking relations).

Each sub-problem is governed by distinct interpretive rules. The linguistic expressions in a tagged image are interpreted relative to a language. So too, pictures are governed by **systems of depiction**, the pictorial analogues of languages. The pictures of platonic solids used in this paper belong to a simple system of black and white line drawing. But systems vary in their treatment of line, color, shading, stylization, and more. Tagged images themselves belong to **hybrid systems**: the combination of system of depiction and a language. Where  $L$  is a language and  $D$  is a system of depiction, a tagged image belongs to the hybrid system  $L/D$ .

### (9) Tagging Semantics

Given a hybrid system  $L/D$ , for any tagged image  $T = \langle I, tag \rangle$  in  $L/D$ , and context  $c$ :  $\llbracket T \rrbracket_{L/D,c}$  = the minimal feature map  $M = \langle Field, Direction, Cluster \rangle$  such that:

- (i) **Congruence**:  
there is a unique  $f : P \mapsto Field$  such that  $I$  and  $Field$  are congruent wrt to  $f$ ;
- (ii) **Pictorial Semantics**:  $\llbracket I \rrbracket_{D,c} \sqsubseteq M$ ;
- (iii) **Tagging**:  
 $\forall r \in dom(tag) : \text{where } tag(r) = \langle S, R \rangle : \llbracket R \rrbracket_{L/D}(\llbracket S \rrbracket_L) \in properties(f(r))$ .

Clause (i) serves the same function as before, imposing congruency between regions of the image surface and feature map regions. Clause (ii) introduces the semantic contribution of the pure image, presupposing something like the semantics specified above. The “ $\sqsubseteq$ ” relation is the part-hood relation for feature maps.

Clause (iii) states that, for every tagged region  $r$  in the image, a corresponding property should be added to the feature cluster which is expressed by  $r$  in the feature map. The property in question is not simply  $\llbracket S \rrbracket_L$ , the content of the tagged symbol, but rather  $\llbracket R \rrbracket_{L/D}(\llbracket S \rrbracket_L)$ , the content of the tagged symbol as it is modulated by the content of the relevant tagging relation, as required by Problem 3. The denotations for two most common relation-symbols, “**id**” (*identity*) and “**pred**” (*predication*) are:

$$(10) \quad \llbracket \mathbf{id} \rrbracket_{L/D} = \lambda x \lambda y. x = y$$

$$(11) \quad \llbracket \mathbf{pred} \rrbracket_{L/D} = \lambda F. F$$

This semantics has the desired effect of allowing the content of symbolic tags to enter into the content of the tagged image at specific, object-dependent locations in pictorial space. The resulting analysis provides a satisfactory account of Problem 1, the semantic contribution of tags to accuracy conditions. To see this, recall the accurate image (1) and inaccurate image (2) from Figure 2, which differed only in the placement of the tags “cube” and “sphere”. Suppose that, in (1), “cube” is linked by predication to region  $r_1$ . And assume that  $\llbracket \text{cube} \rrbracket_L =$  the property *cube*. By clause (iii) from tagging semantics, the “cube” tag imposes the following condition on the resulting feature map:

$$(12) \quad \text{a. } \llbracket \mathbf{pred} \rrbracket_{L/D}(\llbracket \text{cube} \rrbracket_L) \in properties(f(r_1)) \Leftrightarrow$$

- b.  $cube \in properties(f(r_1))$

By the definition of accuracy for feature maps, it follows that the tagged image is accurate at a centered-world  $\langle w, v \rangle$  only if:

- (13) a.  $object(f(r_1))$  is located in  $Direction(f(r_1))$  from  $v$  in  $w$ ;
- b.  $object(f(r_1))$  realizes the property  $cube$  in  $w$ .

The inaccurate image (2) is exactly like the accurate (1), except that “cube” is associated with  $r_2$ , rather than  $r_1$ . When  $r_2$  is substituted for  $r_1$  in the accuracy conditions above, the picture locates a cube in a different direction within pictorial space. As a result, (1) and (2) express different accuracy conditions.

I turn next to pictorial semantics, the semantics governing pure images. Pictorial semantics is itself a multi-faceted problem, where vision, convention, and context all play a role in determining meaning (Kulvicki, 2006; Greenberg, 2018). I’ll assume that, given a system of depiction  $D$ , image  $I$ , and context  $c$ ,  $\llbracket I \rrbracket_{D,c}$  is a PFM (Greenberg, 2019a). Context determines the singular content of a PFM by associating regions of  $I$  with objects. Systems of depiction, in turn, determine part of the attributive content of a PFM by associating regions with directions and basic properties. Further stages of visual processing contribute additional visual features like depth, 3D shape, and category. The semantics sketched below shows how these different interpretive vectors can be brought together:

(14) **Pictorial Semantics**

Given a system of depiction  $D$ , for any image  $I = \langle P, d, color \rangle$  in  $D$ , and context  $c$ :

$\llbracket I \rrbracket_{D,c}$  = the minimal feature map  $M = \langle Field, Direction, Cluster \rangle$  such that:

- (i) **Congruence:**  
there is a unique  $f : P \mapsto Field$  such that  $I$  and  $Field$  are congruent wrt  $f$ ;
- (ii) **Reference:**  
 $\forall r \subseteq P$  : if  $r \in dom(ref_c)$ , then  $ref_c(r) = object(f(r))$ ;
- (iii) **System of depiction:**  
there is a viewpoint  $v = vp_c(Field)$  such that  $\forall p \in P$  :
  - (a) **Projection condition:**  
 $Direction(f(p))$  is co-directional with  $projection_D(v, f(p))$ ;
  - (b) **Marking condition:**  
 $marking_D(color(p)) \in properties(f(r))$ ;
- (iv) **System of vision:**  
 $V_f(P) \sqsubseteq M$ ;

Clause (i) imposes a spatial congruency  $f$  between regions of the image surface and feature map regions, which is used to preserved consistency between the other clauses. Where  $r$  is an image region,  $f(r)$  is the corresponding map region. Two 2D fields are **congruent** just in case there is a metric isomorphism between them that preserves up/down and front/back orientations. (I’ve suppressed orientation vectors in the definitions of images and map fields here.) Clause (ii) accounts for the singular content of an image, as determined by context, which in turn reflects artist’s intentions and the causal history of the picture’s production. Contextual reference is modeled as a function  $ref_c$ , part of a context  $c$ , which associates (some) regions of  $I$  with objects.

Clause (iii) specifies the contribution of the system of depiction within a projection semantics (Greenberg, 2013; Abusch, 2015). In Greenberg (2019a), I’ve shown how a projection semantics can be translated into a feature map framework; in short, projection semantics implies that each



feature cluster in a picture’s feature map be association with (a) a viewpoint-relative direction; and (b) a basic feature (such as *surface*, *edge*, or color). These are determined, in turn, by the **projection condition** and **marking condition** which characterize the system of depiction. Formally,  $vp_c$  fixes the position of the viewpoint, relative to the map field, within the 3-space of the feature map;  $projection_D$  is a function from viewpoints and points on the map field to rays in the 3-space of the feature map;  $marking_D$  is a function from colors in the picture to basic features. Clause (iv) is a catch-all for the contribution of visual computation to pictorial content.

Drawing upon the feature map analysis of visual content, I have sketched a theory of the syntax, content, and semantics of tagged images. While the account leaves significant issues unresolved, I hope the general analytical strategy I’ve pursued here is flexible enough to extend to other kinds of tagging in maps, comics, sign language, and beyond.

## References

- Dorit Abusch. Applying discourse semantics and pragmatics to co-reference in picture sequences. *Proceedings of Sinn und Bedeutung 17*, 2012.
- Dorit Abusch. Possible worlds semantics for pictures. In Lisa Mathewson, Cécile Meier, Hotze Pullman, and Thomas Ede Zimmermann, editors, *Blackwell Companion to Semantics*. Wiley, New York, 2015. Forthcoming.
- Malihe Alikhani and Matthew Stone. Arrows are the verbs of diagrams. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3552–3563, 2018a.
- Malihe Alikhani and Matthew Stone. Exploring coherence in visual explanations. In *2018 IEEE Conference on MIPR Processing and Retrieval (MIPR)*, pages 272–277.
- Malihe Alikhani and Matthew Stone. “Caption” as a coherence relation implications. In *Proceedings of the 2nd Workshop on Shortcomings in Vision and Language*, pages 58–67, 2019.
- Ben Blumson. Pictures, perspective and possibility. *Philosophical Studies*, 149(2):2009.
- Valeria Giardino and Gabriel Greenberg. Varieties of iconicity. *Review of Philosophy and Psychology*, 6(1):1–25, 2015.
- Gabriel Greenberg. Beyond resemblance. *Philosophical Review*, 122(2):215–287, 2013.
- Gabriel Greenberg. Content and target in pictorial representation. *Ergo*, 5(23), 2018.
- Gabriel Greenberg. Semantics of pictorial space. Manuscript, 2019a.
- Gabriel Greenberg. The structure of visual content. Manuscript, 2019b.
- John Kulvicki. *On Images: Their Structure and Content*. Clarendon, Oxford, 2006.
- Alex Lascarides and Matthew Stone. Discourse coherence and gesture interpretation. *Gesture*, 9(2):147–180, 2009a.
- Alex Lascarides and Matthew Stone. A formal semantic analysis of gesture. *Journal of Semantics*, 26(4):393–449, 2009b.
- Philippe Schlenker. Iconic pragmatics. *Natural Language & Linguistic Theory*, 36(3):877–936, 2018a.
- Philippe Schlenker. Visible meaning: Sign language and the foundations of semantics. *Theoretical Linguistics*, 44(3-4):123–208, 2018b.
- Philippe Schlenker. What is super semantics? *Philosophical Perspectives*, 2019.
- Lyn Tieu, Robert Pasternak, Philippe Schlenker, and Emmanuel Chemla. Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: a journal of general linguistics*, 2(1), 2017.

# Neg-raising *believe* and *Maximize Presupposition*\*

Sam Alxatib<sup>1</sup> and Ivana Đurović<sup>2</sup>

<sup>1</sup> CUNY Graduate Center

alxatib@alum.mit.edu

<sup>2</sup> CUNY Graduate Center

idurovic@gradcenter.cuny.edu

## Abstract

We point out a puzzle that comes from the interaction of *Maximize Presupposition* (MP) with the Neg-raising verb *believe*. In response, we offer a reformulation of MP that we call *Maximize Update* (MU). MU compares not only the presuppositions of the relevant competing forms, but the conjunctions of the forms' presuppositions with their assertions. The principle favors the stronger of these. We compare MU with the possibility of combining MP with other accounts of neg-raising.

## 1 Introduction

In contexts where it is common ground that Tesla designed the Tesla coil, it is odd to say (B):

(B) Kim believes that Tesla designed the Tesla coil.

(B) is odd because it seems to suggest that Tesla did not in fact design the Tesla coil. This inference, on current thinking, results from competition between (B) and an alternative of it, (K), where the factive verb *know* replaces *believe*:

(K) Kim knows that Tesla designed the Tesla coil.

The context we have in mind here supports the factive presupposition of (K), and because (B) and (K) otherwise assert the same thing, (K) blocks (B). The principle that makes this blocking happen, which finds its roots in [Hawkins 1978](#), [Hawkins 1991](#), and [Heim 1991](#), states that a form  $\phi$  is to be favored over another  $\psi$  iff  $\phi$ 's presupposition asymmetrically entails  $\psi$ 's, and  $\phi$  and  $\psi$  are otherwise contextually equivalent. In current literature the principle is called *Maximize Presupposition* (MP). We borrow the formulation in (1) from [Spector and Sudo 2017](#):

(1) **Maximize Presupposition (MP):**

If  $S'$  is an alternative to  $S$ , and  $\text{Dom}(\llbracket S' \rrbracket^c) \subset \text{Dom}(\llbracket S \rrbracket^c)$ , then favor  $S'$  to  $S$  in any context  $c$  where  $c \cap \llbracket S \rrbracket^c = c \cap \llbracket S' \rrbracket^c$ .<sup>1</sup>

From (1), and assuming the lexical entries in (2) and (3) for *believe* and *know*,

(2)  $\llbracket \textit{believe} \rrbracket^w = [\lambda p_{\langle s,t \rangle} . \lambda x_e . \text{BEL}_{x,w} \subseteq p]$  (to be revisited)

(3)  $\llbracket \textit{know} \rrbracket^w = [\lambda p_{\langle s,t \rangle} : p(w) = 1 . \lambda x_e . \text{BEL}_{x,w} \subseteq p]$

---

\*For helpful discussions, we thank Elena Herburger, Orin Percus, Jacopo Romoli, Florian Schwarz, Philippe Schlenker, Yael Sharvit, and Yasutada Sudo.

<sup>1</sup>We assume that a context  $c$  is a set of possible worlds (the context set), and we write  $\llbracket \rrbracket^c$  to mean  $[\lambda w . \llbracket \rrbracket^w]$ .

it follows that (K) is to be preferred to (B) by speakers who share the opinion that Tesla designed the Tesla coil. Conversely, (1-3) lead to the result that (B) is felicitous only in contexts where Tesla is *not* presumed to have invented the Tesla coil.<sup>2</sup>

With this in mind, consider (nB), the negation of (B):

(nB) Kim does not believe that Tesla designed the Tesla coil.

Unlike (B), (nB) is natural in our context. Yet by the same assumptions reviewed above, (nB) should be odd given the availability of the alternative (nK):

(nK) Kim does not know that Tesla designed the Tesla coil.

The only difference between (nB,nK) and (B,K) is the presence of negation. But at first glance it isn't obvious why negation should stop MP from favoring (nK) to (nB) in the same way that it favors (K) to (B). The presuppositions of (nB,nK), and the logical relation between them, are the same as those of (B,K), and it is reasonable to think that the negations of two assertorically equivalent sentences like (B,K) are also assertorically equivalent. What, then, changes with (nB,nK)?

The answer is that *believe*, unlike *know*, is a “neg-raising” verb: while (nK) says — setting aside its factive presupposition — that it is not the case that Kim believes Tesla to have designed the Tesla coil, (nB) says that Kim believes that he did not. Let us use the term NR to refer to this property of the verb *believe*; that when it is negated, it is understood to say that the (relevant) attitude holder *believes* that the (relevant) attitude is *false*. We summarize this, informally and theory-neutrally, in (4):

(4) **NR:** [*not believe S*] implies [*believe not S*]

It might seem that (4) by itself explains why (nB) and (nK) are acceptable: the two sentences are not equivalent in the given context, hence not subject to MP, because (nB) asserts something stronger than (nK). But this explanation is incomplete, because it is not accompanied by a theory of where NR inference comes from. If it comes from an opinionatedness presupposition, as Bartsch (1973) and Gajewski (2005, 2007) have proposed, then indeed (nB) and (nK) would no longer be in the domain of MP, because neither would have presuppositions that entail the other's. But then, for the same reason, (B) and (K) should also no longer be in the domain of MP, leaving the oddness of (B) unexplained. If NR comes from a scalar implicature, as proposed by Romoli (2012, 2013), then (nB) and (nK) could be explained if we show independent evidence that MP can operate on exhausted sentences, not just on their implicature-less literal meanings. This too is not trivial, and we will explain why it is difficult to find independent support for it. Finally, if NR came from the assertoric meaning of (nB), then we would have to assume either a theory where *believe* is ambiguous, or a theory that derives NR from a syntactic mechanism (Fillmore 1963; Collins and Postal 2014, 2018). The first of these options is stipulative; the second, we will argue, makes incorrect predictions elsewhere. Our own proposal is a replacement of MP with another principle, which we call *Maximize Update* (MU), that compares not only the presuppositions of the given alternatives but the conjunctions of their presuppositions together with their assertions.

<sup>2</sup>Careful readers may have noticed a slight change in our narrative: We opened with the claim that (B) *presumes* that Tesla did *not* design the Tesla coil, but from MP we derived the weaker result that the context of (B) does *not presume* that Tesla designed the Tesla coil. The stronger of these two results – the first – is an example of an *antipresupposition*. Readers are referred to Percus 2006 and Chemla 2008 for discussion of the connection between MP and antipresuppositions.

Before we proceed, we want to note that we are not the first to observe this interaction between MP and the NR property of *believe*. Percus (2006), in fn. 14. credits Gillian Ramchand with an observation like ours, and speculates on a connection to NR but does not develop a detailed account. To our knowledge, the interaction has not been discussed elsewhere.

We will begin our presentation in Section 2 with the “semantic” view of NR, i.e. the view that derives inference from a presupposition of opinionatedness, and after some discussion we propose our own solution. We then discuss the implicature-based view (Section 3), and the syntactic view (Section 4).

## 2 The semantic view of NR and *Maximize Update*

Proponents of the semantic account of NR write an “opinionatedness” presupposition into the semantics of *believe*, as in (5):

$$(5) \quad \llbracket \textit{believe} \rrbracket^w = [\lambda p_{\langle s,t \rangle} \cdot \lambda x_e : \text{OPN}_{x,w}(p) \cdot \text{BEL}_{x,w} \subseteq p] \quad (\text{revised from (2)})$$

$$(6) \quad \text{OPN}_{x,w}(p) \text{ iff } \text{BEL}_{x,w} \subseteq p \vee \text{BEL}_{x,w} \subseteq \bar{p}$$

As the reader can see, the opinionatedness presupposition in (5) does not strengthen the truth conditions in positive belief ascriptions, like (B), because opinionatedness follows logically from the assertion itself. However, in cases like (nB) where the verb *believe* appears under negation, the opinionatedness presupposition leads to NR, as desired:

(nB) Kim does not believe that Tesla designed the Tesla coil

$$\text{a. } \llbracket (\text{nB}) \rrbracket^w \text{ is defined only if } \text{OPN}_{k,w}(\llbracket \textit{Tesla designed the Tesla coil} \rrbracket^c) \quad (\text{OPN}_{k,w}(t))$$

$$\text{b. If defined, } \llbracket (\text{nB}) \rrbracket^w = 1 \text{ iff } \text{BEL}_{k,w} \subseteq \overline{\llbracket \textit{Tesla designed the Tesla coil} \rrbracket^c} \quad (\text{BEL}_{k,w} \subseteq \bar{t})$$

In words, (nB) is only interpretable if Kim has an opinion about whether Tesla designed the Tesla coil, and if it is interpretable, (nB) is true iff it is not the case that Kim believes that Tesla designed the Tesla coil. Since (by the presupposition) Kim has an opinion on the matter, it follows that Kim believes that Tesla did not design the Tesla coil. This is the NR inference.

But adding a presupposition of opinionatedness into the meaning of *believe* presents us with a problem. Being opinionated about  $p$  does not entail  $p$ , and  $p$  does not entail being opinionated about  $p$ . This means that neither of (B,nB) should compete with (K,nK) by MP, because the presuppositions of (K,nK) no longer entail those of (B,nB). In the case of (nB) and (nK) this is a good result, because both examples are acceptable, but in the case of (B) and (K) the result is not welcome.

Note that as long as asymmetric entailment is required to hold between the presuppositions of the two competing forms, any alternative to MP will fail to distinguish between the negated pair of alternatives (nB,nK), which are both acceptable, and the unnegated pair (B,K), which are not both acceptable. For example, Spector and Sudo’s (2017) Presupposed Ignorance Principle (7) favors the (presuppositionally-stronger) alternative that current context supports. But because neither *know* nor the presuppositional *believe* is presuppositionally-stronger than the other, there is no way to predict the oddness of (B).

### (7) Presupposed Ignorance Principle:

If  $S'$  is an alternative to  $S$ , and  $\text{Dom}(\llbracket S' \rrbracket^c) \subset \text{Dom}(\llbracket S \rrbracket^c)$ , then  $S$  is infelicitous in any context  $c$  where  $c \subseteq \text{Dom}(\llbracket S' \rrbracket^c)$ .

A more radical revision of MP, one that does not make direct reference to presuppositions, is proposed by Anvari (2018). The principle requires that context preserve entailment between alternatives: a form  $S$  is infelicitous iff it entails an alternative  $S'$  in context, but does not entail it logically:

(8) **Logical Integrity (LI):**

If  $S'$  is an alternative to  $S$ , and  $\llbracket S \rrbracket^c \not\subseteq \llbracket S' \rrbracket^c$ , then  $S$  is infelicitous in any context  $c$  where  $c \cap \llbracket S \rrbracket^c \subseteq c \cap \llbracket S' \rrbracket^c$

But LI also fails to distinguish the unnegated (B,K) from the negated (nB,nK). The principle correctly blocks “*believe p*” (e.g. (B)) in  $p$ -contexts, because in  $p$ -contexts “*believe p*” entails “*know p*”, but it does not entail it logically. However, the same holds of “*not believe p*” (e.g. (nB)) in  $p$  contexts: “*not believe p*” entails “*not know p*” in a  $p$ -context, but it does not entail it logically.

We propose to replace *Maximize Presupposition* with *Maximize Update*:

(9) **Maximize Update (MU):**

If  $S'$  is an alternative to  $S$ , and  $\llbracket S' \rrbracket^c \subset \llbracket S \rrbracket^c$ , then favor  $S'$  to  $S$  in any context  $c$  where  $c \cap \llbracket S \rrbracket^c = c \cap \llbracket S' \rrbracket^c$ .

MU, like LI, does not make any direct reference to the presuppositions of the forms under comparison. Instead it compares their update potential, as it were, by comparing the sets of worlds in which they are true: if  $S'$  is defined and true in a strictly narrower set of worlds than  $S$ , and if the two forms are otherwise contextually equivalent, then  $S'$  is to be favored.

The effect of MU on (B,K) is as we want it to be: the worlds where (B) is true are worlds where Kim’s belief state supports the proposition that Tesla designed the Tesla coil (10a); the worlds where (K) is true make up a proper subset of these — they are worlds where Tesla designed the Tesla coil *and* where Kim has that belief (10b). Crucially, proper subsethood holds between these two sets of worlds despite the lack of entailment between the presuppositions of *believe* and *know*. When their presuppositions and assertions are considered together, *know* comes out stronger, and is therefore preferred when supported.

- (10) a.  $\llbracket (B) \rrbracket^c = \{w : \text{BEL}_{k,w} \subseteq t\}$   
 b.  $\llbracket (K) \rrbracket^c = \{w : t(w)=1 \ \& \ \text{BEL}_{k,w} \subseteq t\}$

In the case of (nB,nK), however, things are different. Worlds where (nB) is true (and defined) are worlds where Kim is opinionated about whether Tesla designed the coil, and does not have the opinion that he did. These are therefore worlds where Kim believes that Tesla did not design the Tesla coil (11a). On the other hand, worlds where (nK) is true (and defined) are worlds where Tesla designed the Tesla coil, but where it is not the case that Kim has that belief (11b):

- (11) a.  $\llbracket (nB) \rrbracket^c = \{w : \text{BEL}_{k,w} \subseteq \bar{t}\}$   
 b.  $\llbracket (nK) \rrbracket^c = \{w : t(w)=1 \ \& \ \text{BEL}_{k,w} \not\subseteq t\}$

Clearly, neither of (11a,11b) is a subset of the other, so (nB,nK) are correctly predicted not to compete under MU.

Note that, by its definition, MU makes the same prediction as MP in every case where two forms  $S, S'$  are equivalent in their assertions but where  $S'$  has a stronger presupposition than

$S$ . In these cases,  $\llbracket S' \rrbracket^c$  is a proper subset of  $\llbracket S \rrbracket^c$ , because  $\llbracket S' \rrbracket^c$  is effectively a conjunction of two propositions  $\phi \& \psi'$  — the assertion and presupposition of  $S'$  respectively — while  $\llbracket S \rrbracket^c$  is by assumption a weaker conjunction  $\phi \& \psi$ ,  $\psi$  being weaker than  $\psi'$ . MP favors  $S'$  because of the presuppositions alone; MU favors  $S'$  because of the “conjunction” of the presuppositions with their assertions.

Note also that MU, like Anvari’s LI, correctly blocks uses of weaker scalar items like *some* in contexts where they are equivalent to their stronger alternatives, like *all*. Cases of this kind were discussed in Magri 2009, 2011. (12), from Magri 2011, is an example:

- (12) Context: Prof. Smith assigns the same grade (possibly a different one every term) to all of his students.
- a. #This year, Prof. Smith assigned an A to some of his students.
  - b. ✓This year, Prof. Smith assigned an A to all of his students.

Magri took data like these to show that scalar implicature calculation is not based on purely pragmatic grounds; given that the context in (12) does not distinguish between the *some* and the *all* alternatives, there is no reason for the imagined listener in this context to infer, on hearing (12a), that the professor did not assign an A to all of his students. Either of (12a,b) should therefore be acceptable, yet only (12b) is.

Magri’s conclusion that extra-literal inferences, like scalar implicatures, are not governed solely by contextual factors does not change under LI, since LI makes reference both to logical entailment and contextual entailment. MU is similar in this respect. (12a) is dispreferred even though it is contextually equivalent to (12b); it is dispreferred because it is *logically* weaker than (12b). We therefore do not see MU as a pragmatic principle.

We now turn our attention to alternative ways of explaining (nB,nK) without altering the basic formulation of MP. These are (a) the possibility of deriving NR as a scalar implicature, and (b) the possibility that NR results from interpreting negation in an embedded position.

### 3 NR as a scalar implicature

The idea of deriving NR as an implicature comes from Romoli 2012. A central assumption in the proposal is that *believe* has a formal alternative that states opinionatedness. Let us abstract away from the form of this alternative, and call it *opn* for short. *Opn* has the following semantics:

$$(13) \quad \llbracket opn \rrbracket^w = [\lambda p_{\langle s,t \rangle} . \lambda x_e . \text{OPN}_{x,w}(p)],$$

where as before:  $\text{OPN}_{x,w}$  iff  $\text{BEL}_{x,w} \subseteq p \vee \text{BEL}_{x,w} \subseteq \bar{p}$

And *believe* itself has no presupposition of opinionatedness. This is the original entry we saw in (2):

$$(2) \quad \llbracket believe \rrbracket^w = [\lambda p_{\langle s,t \rangle} . \lambda x_e . \text{BEL}_{x,w} \subseteq p]$$

In positive contexts, *opn* does not participate in the implicatures of *believe*, because its truth conditions follow from those of *believe*:  $\text{BEL}_{x,w} \subseteq p$  asymmetrically entails  $\text{OPN}_{x,w}(p)$ .

In negative contexts, however, the alternative where *believe* is replaced with *opn* gives rise to an implicature that leads to NR: “*not believe S*” implicates that the alternative “*not opn S*” is false, and “*not opn S*” is false iff “*opn S*” is true. So “*not believe S*” brings with it the implicature that the given attitude holder is opinionated about  $S$ . And since the sentence

asserts that the attitude holder does not have the belief that  $S$ , it follows that s/he believes that not  $S$ . These steps are summarized in (14):

- (14) Kim does not believe that  $S$
- a.  $\text{ALT}(14) = \{\text{Kim does not } \textit{opn} \text{ that } S, \dots\}$
  - b.  $\text{EXH}(14\text{a})(14) = 1$  iff  $\llbracket \textit{Kim not believe } S \rrbracket^w = 1$  and  $\llbracket \textit{Kim not opn } S \rrbracket^w = 0$   
 $= 1$  iff  $\llbracket \textit{believe} \rrbracket^w(\llbracket S \rrbracket^c)(k) = 0$  and  $\llbracket \textit{opn} \rrbracket^w(\llbracket S \rrbracket^c)(k) = 1$   
 $= 1$  iff  $\text{BEL}_{k,w} \not\subseteq \llbracket S \rrbracket^c$  and  $\text{OPN}_{k,w}(p)$   
 $= 1$  iff  $\text{BEL}_{k,w} \subseteq \llbracket S \rrbracket^c$

Now we can revisit the problem we pointed out in the context of MP. The semantic account of NR assigns independent presuppositions to *believe* and *know*, and by doing so breaks the entailment relation between them that feeds MP. The implicature-based account of NR does not have this problem, however; by its semantics, *believe* has no presupposition of opinionatedness, so its competition with *know* works as initially thought: *know*  $S$  is favored over *believe*  $S$  in contexts where  $S$  is part of the common ground.

What about negated cases? Recall that we want both (nB) and (nK) to be acceptable in our Tesla context. But if we go by literal meaning alone, MP will incorrectly favor (nK) to (nB), because (nK) has a factive presupposition and (nB) presupposes nothing. The question that comes up now is whether MP is indeed sensitive only to literal meanings, or whether it compares the exhausted meanings of the given forms. If the latter, then neither (nB) nor (nK) would be favored over the other, because the exhausted meaning of (nB) includes NR (see (14b)), while the exhausted meaning of (nK) does not.

We do not know how to find independent evidence to support this possibility. The task is difficult because it requires that we compare two forms  $\phi$  and  $\psi$ , where  $\phi$  is presuppositionally stronger than  $\psi$ , but where  $\psi$  with its implicatures assert something beyond what  $\phi$  asserts. Now, suppose that  $\phi$  and  $\psi$  are assertorically equivalent. Then the needed implicatures of  $\psi$  would have to come from alternatives that  $\psi$  has, but that  $\phi$  does not have — indeed, this seems to be the case for *believe* and *know* on Romoli's theory (*opn* is a formal alternative to *believe*, but not to *know*). We cannot at the moment think of other cases that fit this description. Suppose, on the other hand, that  $\phi$  and  $\psi$  are not assertorically equivalent. Then they may not be subject to MP in the first place, because MP targets only alternatives that make equivalent assertions in context. We will leave this matter to future work.

## 4 NR as syntactic neg-raising

For our purposes we will assume a descriptive version of the syntactic account of NR. It is one where (nB) has an LF like (15):

- (15) [Kim believes that [NEG [Tesla designed the Tesla coil]]]

Clearly, the interpretation of (15) generates the NR reading of (nB) as part of its *literal* meaning, and because of this, the form's assertion is no longer equivalent to the assertion of (nK). Both sentences are therefore predicted to be acceptable in our context, and the prediction does not change the result we get for the unnegated (B,K).

For the sake of discussion, assume a view that derives all instances of NR from LFs like (15).<sup>3</sup> While (15) is correctly predicted not to compete with (nK), as we just noted, it is predicted to compete with (Kn):

<sup>3</sup>As will point out shortly, not all proponents of the syntactic view believe this.



(Kn) [Kim knows that [NEG [Tesla designed the Tesla coil]]]

Of course, (Kn) is not defined in the context that we are working with, because its presupposition is false. But suppose we change this. Take (Kn′):

(Kn′) Kim knows that Tesla did not found the rock band Tesla.

In our world, (Kn′) has a true factive presupposition.<sup>4</sup> As expected, (Bn′) below is odd, and is predicted to be odd by MP as well as LI and MU:

(Bn′) Kim believes that Tesla did not found the rock band Tesla.

What we would like to point out is that (Bn′) should resemble in all relevant respects the LF of (nB′), assuming that the source of the NR inference in (nB′) is syntactic:

(nB′) Kim does not believe that Tesla founded the rock band Tesla.  
(*LF on the syntactic view*: [Kim believes [that NEG Tesla founded the rock band Tesla]])

But (nB′) is not very odd — certainly not as odd as (Bn′). As far as we can see, the syntactic view of NR does not predict this difference. Perhaps one could assume that, despite the similarity of their LFs, (Kn′) is not a formal alternative to (nB′). However, we do not know if this assumption can be defended. Note by comparison that, on the semantic view of NR, (nB′) does not have an LF where negation appears in the scope of *believe*, which means that (nB′) will not have (Kn′) as a formal alternative. The account therefore predicts (nB′) and (Bn′) to behave differently, as desired.

There are more questions to raise. Even if the underlying syntax of (nB) turns out to be the right explanation of its acceptability, and even if something else can be said to distinguish (nB′) from (Bn′), the overall story cannot be extended to other cases of NR that seem to us to have the same profile as *believe*. Let us expand on this.

Collins and Postal (2018) (C&P), citing Horn 1978, have argued that the expressions *is of the opinion* and *it is X's opinion* license an NR inference when they are negated, but that unlike *believe*, they do not license strict NPIs, and do not allow negative auxiliary inversion. From this, C&P conclude that the NR inference in these cases cannot have a syntactic source. The examples below show the relevant data. (16a,b) show that NR follows from the negations of *be of the opinion* and *it is X's opinion*. (17a,b) show that these negation do not allow embedded strict NPIs; (17c) shows they they do not allow auxiliary inversion. (18a,b) show that negated *believe* allows both.

- (16) a. Kim is not of the opinion that Mars can be colonized (Based on C&P, ex. 18)  
       ≈ Kim is of the opinion that Mars cannot be colonized  
       b. It is not Kim's opinion that Mars can be colonized  
       ≈ It is Kim's opinion that Mars cannot be colonized
- (17) a. \*Kim is not of the opinion that Chris will leave until Friday.  
       b. \*It is not Kim's opinion that Chris will leave until Friday.  
       c. \*It is not my opinion that at any time did he commit perjury.
- (18) a. Kim does not believe that Chris will leave until Friday.  
       b. I do not believe that at any time did he commit perjury.

<sup>4</sup>Tesla is an American rock band that was formed in the early 1980s. Its members have no relation to Nikola Tesla, who died in 1943.



Let us accept the conclusion that the NR inference of *is of the opinion* and *it is X's opinion* does not result from syntactic neg-raising. Now we ask whether the two predicates interact with MP in a similar way to *believe*, specifically, (i) whether they license an “anti-factive” inference when unnegated, and (ii) whether they do not license the inference when negated. We think that the answer is yes to both of these questions: (19a,b) are odd, like (B), and (20a,b) are acceptable, like (nB):

- (19) a. ??Kim is of the opinion that Tesla designed the Tesla coil.  
 b. ??It is Kim's opinion that Tesla designed the Tesla coil.
- (20) a. ✓Kim is not of the opinion that Tesla designed the Tesla coil.  
 b. ✓It is not Kim's opinion that Tesla designed the Tesla coil.

Given this finding, it follows that the syntactic answer to why (nB) is acceptable is not enough, since it cannot apply to cases like (20).

## 5 Summary

We began with *Maximize Presupposition* (MP) and with the standard account of how it favors (K) over (B) in contexts where (K) is licit. We then showed that both of (nB) and (nK) are licit in these contexts, and speculated on possible reasons behind this. We assumed that the NR property of *believe* plays a key role in the interaction, and with that, we turned the question about (nB,nK) into a question about two things: the different accounts of NR, and the correctness of MP. Our proposal is faithful to the semantic view of NR, but includes a reformulation of MP as *Maximize Update* (MU). MU compares the force of not only the presuppositions of its competitors, but the result of conjoining those presuppositions with the assertions of the given forms. We compared the account to Romoli's (2012) implicature-based theory of NR, and to syntactic views of NR. On the implicature theory, we concluded that MP can explain the acceptability of (nB) *if* the principle is assumed to operate on exhaustified expressions. We were not able to find evidence in favor of or against this possibility, and we explained why finding that evidence is difficult. On the syntactic theory, we argued first that the theory predicts (nB) to behave like (Bn), given that the underlying form in the first case is structurally like the surface form of the second. We claimed, however, that the two forms behave differently. More generally, we pointed out that other NR expressions behave just like *believe* in their interaction with MP, but are not amenable to a syntactic account of the NR inference.

## References

- Anvari, Amir. 2018. Logical Integrity. In *Proceedings of the 28th Semantics and Linguistic Theory Conference (SALT 28)*, ed. Sireemas Maspong, Brynhildur Stefánsdóttir, Katherine Blake, and Forrest Davis, 711–726.
- Bartsch, Renate. 1973. ‘Negative Transportation’ gibt es nicht. *Linguistische Berichte* 27:1–7.
- Chemla, Emmanuel. 2008. An epistemic step for anti-presuppositions. *Journal of Semantics* 25:141–173.
- Collins, Chris, and Paul M. Postal. 2014. *Classical NEG Raising: An Essay on the Syntax of Negation*. Cambridge, MA: MIT Press.

- Collins, Chris, and Paul M. Postal. 2018. Disentangling two distinct notions of NEG raising. *Semantics and Pragmatics* 11.
- Fillmore, Charles J. 1963. The position of embedding transformations in a grammar. *Word* 19:208–231.
- Gajewski, Jon Robert. 2005. Neg-Raising: Polarity and Presupposition. Doctoral Dissertation, MIT.
- Gajewski, Jon Robert. 2007. Neg-raising and polarity. *Natural Language Semantics* 30:289–328.
- Hawkins, John A. 1978. *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croom Helm.
- Hawkins, John A. 1991. On (in)definite articles: implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27:405–442.
- Heim, Irene. 1991. Artikel und Definitheit. In *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung/Semantics: An International Handbook of Contemporary Research*, ed. Arnim von Stechow and Dieter Wunderlich. Berlin: de Gruyter.
- Horn, Laurence R. 1978. Remarks on neg-raising. In *Pragmatics*, ed. Peter Cole, volume 9 of *Syntax and Semantics*, 129–220. New York, NY: Academic Press.
- Magri, Giorgio. 2009. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics* 17:245–297.
- Magri, Giorgio. 2011. Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics* 4:1–51.
- Percus, Orin. 2006. Antipresuppositions. In *Theoretical and Empirical Studies of Reference and Anaphora: Toward the Establishment of Generative Grammar as an Empirical Science*, ed. Ayumi Ueyama, 52–73. Washington: Japan Society for the Promotion of Science.
- Romoli, Jacopo. 2012. Soft but Strong. Neg-raising, Soft triggers, and Exhaustification. Doctoral Dissertation, Harvard University.
- Romoli, Jacopo. 2013. A scalar implicature-based approach to neg-raising. *Linguistics and Philosophy* 36:291–353.
- Spector, Benjamin, and Yasutada Sudo. 2017. Presupposed ignorance and exhaustification: how scalar implicatures and presuppositions interact. *Linguistics and Philosophy* 40:473–517.

# Abilities and Success

David Boylan\*

Rutgers University, New Brunswick, New Brunswick, NJ, USA  
db1082@philosophy.rutgers.edu

Does success entail ability? Call the principle that it does *Success*:

*Success*.  $S \phi$ 's  $\models$  S can  $\phi$

When we focus on successful action, *Success* is compelling: when someone succeeds in something, like sinking a putt or surfing a wave, one is forced to concede they were able to do that. This is what *Success* would lead us to expect. But when success is not yet assured, the lesson seems different. When said before the fact, the claim that I can surf that wave is strong — it says that surfing that wave is within my control. This intuition drives against *Success*. Just doing something does not demonstrate it is within my control: flukes do happen. So, if the control intuition is right, success should *not* demonstrate ability.

First I try to make the above tension precise. I argue that the appeal of *Success* is connected to two plausible and related principles: that *past* success entails *past* ability, which I call *Past Success*; and that *cannot* seems to entail *will not*, which I call *Can't-entails-won't*. But, on the other, I show we can find counterexamples to *Success* in cases of inexact ability discussed by [Kenny \[1976\]](#). To explain these data, I maintain we must connect the truth of ability claims to the facts about what our options *settle* and what they leave *open*, in the sense familiar from the literature on future contingents. I do this within a kind of conditional analysis of ability ascriptions. I first define an operator  $\mathcal{W}$  with features attributed to 'will' in the literature on future contingents. In particular,  $\mathcal{W}\phi$  is indeterminate in truth-value, when  $\phi$  is unsettled. Building on previous joint work in [Mandelkern et al. \[2016\]](#), I state my conditional analysis in terms of  $\mathcal{W}$ -conditionals: on my view,  $\lceil$ S can  $\phi \rceil$  says, roughly, there's some action available to S such that if S does it, then  $\mathcal{W}(S \phi)$  is true. By thus building a connection between unsettledness and indeterminacy into ability claims, my conditional account of abilities reconciles the motivations for *Success* with its counterexamples.

## 1 The Status of *Success*

Two facts are easy to explain, if *Success* is valid and hard to explain otherwise. To appreciate the first, let's focus on relatively mundane cases inspired by [Kenny \[1976\]](#)'s discussion of abilities:

**Fluky Dartboard.** I am a terrible dartplayer. I struggle to even hit the board whenever I take a shot. However, I take my shot and I flukily hit the bullseye.

Once I have taken the shot and hit the bullseye, I can compellingly argue:

- (1) I hit the bullseye on that throw.  
So, I was able to hit the bullseye on that throw.

If you know that I have been successful, you must concede I was able to.

---

\*Thanks to Fabrizio Cariani, Melissa Fusco, Jeff King, Justin Khoo, Matthew Mandelkern, Milo Phillips-Brown, Robert Stalnaker, and, especially, Ginger Schultheis.

This is our first point in favour of *Success*: *past* success feels like it entails past ability. Where  $\rightsquigarrow$  denotes a felt entailment, we have:<sup>1</sup>

*Past Success*.  $PAST (S \phi'd) \rightsquigarrow PAST (S \text{ can } \phi)$

This seems like strong motivation for *Success* itself. How could *Past Success* be valid, if *Success* is not? After all, *I* haven't changed since I hit the bullseye. My abilities are what they were some moments ago.

The second fact is that it sounds incoherent to affirm or leave open the premise of *Success*, but deny the conclusion. Consider:

- (2) # I can't hit the bullseye on this shot, but I will.
- (3) # I can't hit the bullseye on this shot, but I might.

This is our second point in favour of *Success*. In general, it seems that *can't* entails *won't*:

*Can't-entails-won't*.  $\neg(S \text{ can } \phi) \rightsquigarrow \neg(S \phi's)$

And *Can't-entails-won't* is just the contrapositive of *Success*.

That is the case for *Success*, as I see it. But at the end of the day, I think it cannot be valid. Recall the control intuition: if I say something like

- (4) I can hit the bullseye on this throw.

I say something quite strong. (4) is not verified by the small chance of me hitting the dartboard. We can leverage this intuition to find counterexamples to *Success*. Take the following variation on the dartboard case inspired by [Kenny \[1976\]](#):

**Unreliable Dartboard.** I am a fairly bad dartplayer. I regularly hit the bottom half when I aim for the top; and vice versa. But I never miss the board entirely.

I am about to take a shot. I am skilled enough to know I will hit the board; so I know that I will either hit the top half of the board or the bottom half of the board.<sup>2</sup> But it does not seem that I should ascribe myself either of the following abilities here:

- (5) I can hit the top on this throw.
- (6) I can hit the bottom on this throw.

Even the disjunction does not seem true:

- (7) I can hit the top of the board or I can hit the bottom of the board.

So, in advance of the shot, even if I know I will hit the board somewhere, I do not have the ability to hit the top or the ability to hit the bottom.

---

<sup>1</sup>This is something like the converse of the actuality entailments discussed by [Bhatt \[1999\]](#) and [Hacquard \[2006\]](#).

<sup>2</sup>Note that here and throughout I assume the following principle:

*Will Excluded Middle*.  $Will \phi \vee Will \neg\phi$

This principle is widely taken to be extremely plausible. And, even given a view that *denies* Will Excluded Middle, no existing view of abilities will deliver both the validity of *Past Success* and *Can't-entails-won't* and the failure of  $\vee$ -*Success* below.

But I would be predicted to have one ability or the other if *Success* were valid. For I will hit either the top or the bottom of the dartboard. In the first case, *Success* says I'm able to hit the top, in the second that I'm able to hit the bottom. In either case, then, (7) is true, if *Success* is valid.

The consequences of *Success* become more absurd as the disjunctions get longer. Suppose we divide the dartboard into a million tiny, numbered regions. I can see that I will hit (at least) one of these regions because I know I will hit the dartboard, yielding:

(8) I will hit region 1 or 2 or 3 or ... or 1,000,000 on this throw.

If *Success* is valid, then the following incredible claim is true:

(9) I can hit region 1 on this throw or I can hit region 2 on this throw or ... or I can hit region 1,000,000 on this throw.

And, to put an even finer point on it, notice that each disjunct of (9) entails:

(10) There is a certain point that I can hit on this throw.

*Success* predicts that (10) should just be a truism here. In fact, (10) would be an incredible boast for me to make in **Unreliable Dartboard**.

So you can be sure that you will make true some disjunction, while failing to have the ability to make true either disjunct. This gives us the failure of an instance of *Success*:

$\forall$ - *Success*.  $S \phi$ 's  $\vee \psi$ 's  $\not\rightarrow S \text{ can } \phi \vee S \text{ can } \psi$

Predicting this combination of data is a serious challenge. On standard theories, *Past Success* and *Can't-entails-won't* are both equivalent to *Success*. On the modal analysis of ability ascriptions, defended by Hilpinen [1969], Lewis [1976], Kratzer [1977] and Kratzer [1981a], reflexivity of the modal domain is necessary and sufficient for all three conditions. Brown [1988], Horty and Belnap [1995] and Horty [2001] all defend a view equivalent to the following:

**Boxy Analysis.**  $\llbracket S \text{ can } \phi \rrbracket^w = 1$  iff  $\exists p \in \mathcal{A}(w) : \forall w' \in p : \llbracket S \phi \text{'s} \rrbracket^{w'} = 1$ .

But again, all three inference patterns are equivalent: each is characterised by the condition that  $\{w\} \in \mathcal{A}(w)$ . Since we have seen that they are not equivalent, a new semantics is needed. (This also illustrates how the puzzle here goes beyond Kenny [1976]'s puzzle about disjunction: even views predict ability fails to distribute over disjunction fail to solve my puzzle.)

## 2 Future Contingents and $\mathcal{W}$

I will explain these data by connecting ability claims to *future contingents*. Future contingents have been argued to have three special properties. As we will see, each property mirrors a property of ability. Our assessments of 'will'-claims, it has been argued, involve a kind of temporal asymmetry. Before the fact, the future seems open and 'will'-claims seem unsettled; but after the fact, we seem happy to talk as if they were settled all along. 'Will' is also scopeless with respect to negation:  $\lceil \neg \text{Will } \phi \rceil$  seems equivalent to  $\lceil \text{Will } \neg \phi \rceil$ .

Here I introduce a modal operator  $\mathcal{W}$  that captures this behaviour.  $\mathcal{W}$  is a selection modal, in the sense of Cariani and Santorio [2018]:  $\lceil \mathcal{W}\phi \rceil$  says that  $\phi$  is true in the closest world to actuality. But crucially, on my theory, the closest world leaves various facts unsettled. This ensures  $\mathcal{W}$  has the right properties for giving the semantics of ability modals.

## 2.1 Features of Future Contingents

Let's first take the temporal asymmetry in assessing future contingents.

Recall Aristotle's famous case of the sea battle. On Monday, it is not yet settled whether there will be a battle or not on Tuesday: a capricious ruler decides by flipping a fair coin this evening. I make the following prediction:

(11) There will be a sea battle on Tuesday.

There is a long tradition of thinking that because the future is unsettled, sentences like (11) must be *indeterminate* in truth-value.<sup>3</sup> Things could go either way, depending on how the coin lands. If the coin comes up heads, there will be a battle; if not then not. But the outcome of the toss is not settled; and so whether there will be a sea battle tomorrow is unsettled too.

Now suppose the sea battle does take place on Tuesday. When I look back on my earlier prediction, what should I think? It seems I can say either of the following:

(12) There would in fact be a sea battle on Tuesday.

(13) I said there would be a sea battle on Tuesday; and indeed there would be.

'Would' is generally regarded to be the past tense of 'will'.<sup>4</sup> But then it is surprising that we can say either of (12) or (13): if my assertion was indeterminate when I said it, why do I now say that it was true?<sup>5</sup>

These are the first two properties I want  $\mathcal{W}$  to have. I want  $\mathcal{W}$  to obey two inference patterns. Where  $\diamond\phi$  says that  $\phi$  is circumstantially possible and  $\nabla\phi$  says that  $\phi$  is indeterminate:

*Openness.*  $\diamond\phi, \diamond\neg\phi \rightsquigarrow \nabla \mathcal{W}\phi$

*Past Settledness.*  $PAST \phi \rightsquigarrow PAST \mathcal{W}\phi$

This will eventually allow me to validate *Past Success* without validating *Success*.

The third feature is the way that 'will' interacts with negation. 'Will' does not give rise to any scope distinctions with respect to negation. Take a predicate like 'is absent' that includes a negation as part of its meaning. Consider the following example:

(14) I doubt that John will be present.

This says that I think it is *not* the case that John will be present. But it quite clearly entails

(15) I think that John will be absent.

In general, saying that it is not the case that  $\phi$  will happen just is to say that  $\neg\phi$  will happen.<sup>6</sup> That is,  $\mathcal{W}$  should be *scopeless* with respect to negation, as Cariani and Santorio [2018] put it:

*Scopelessness.*  $\neg\mathcal{W}\phi \rightsquigarrow \mathcal{W}\neg\phi$

*Scopelessness* will secure *Can't-entails-won't*.

<sup>3</sup>See, among others, Aristotle, *De Interpretatione*, Lukasiewicz [1920], Lukasiewicz [1951].

<sup>4</sup>See Abusch [1997] and Condoravdi [2002].

<sup>5</sup>This point was first raised in Prior [1976] and later repeated in MacFarlane [2003].

<sup>6</sup>This point has been recognised at least since Thomason [1970].

## 2.2 Semantics for $\mathcal{W}$

Following Cariani and Santorio [2018], I say that  $\lceil \mathcal{W}\phi \rceil$  is true iff  $\phi$  is true in the *closest world* to the actual world. Unlike Cariani and Santorio, however, I say that the closest world can be unsettled in various respects. As shown by Prior [1967], Thomason [1970] and Thomason [1984], a world that is unsettled past a certain time can be represented using a *set* of worlds which agree in all (relevant) matters of fact up to that time, but diverge afterwards; I'll call such a set an *unsettled world*. In the sea battle case, we represent the earlier, indeterminate state of the world with a set of worlds agreeing on all (relevant) matters of fact up until today and then diverging on whether a sea battle occurs tomorrow. A proposition is true at an unsettled world iff it is true at all worlds in that set; it is false iff it is false at all worlds in that set; and indeterminate, otherwise.

To state the lexical entry for  $\mathcal{W}$ , let's first say how unsettled worlds get into the semantics. I add an unsettled world,  $\mathcal{I}$ , to the index of the semantic evaluation function  $\llbracket \cdot \rrbracket$ . I assume that the unsettled world is supplied to the semantics by context: we form the unsettled world of the context  $\mathcal{I}_c$  by taking the set of worlds that are duplicates of the (determinate) context world up until the context time:

*Unsettled World.*  $\mathcal{I}_c = \{w \mid w \text{ is identical to } w_c \text{ up until } t_c\}$

Now let's consider how to model closeness. Following Stalnaker [1968], I use a selection function to supply the closest worlds. A Stalnakerian selection function  $s$  takes a world  $w$  and a proposition  $\mathbf{A}$  and returns the closest world to  $w$  where  $\mathbf{A}$  is true. My selection functions take an unsettled world as input and can also return an unsettled world as output:  $s(\mathcal{I}, \mathbf{A})$  picks out the closest (possibly unsettled) world to  $\mathcal{I}$  which settles that  $\mathbf{A}$  is true.

What if we want the selection function to give us the closest world to  $\mathcal{I}$  simpliciter? We simply let the other argument be the tautology  $\top$ . I say that  $\top$  is supplied by a *modal base*, a function  $f$  from a world and a time to a set of worlds.<sup>7</sup> I assume that  $f$  is supplied by the index and that  $f$  does not include any information by itself:

*Modal Bases.*  $f_c(w, t) = W$

Thus, to find the closest world simpliciter to the unsettled world of the context,  $\mathcal{I}_c$ , we find  $s(\mathcal{I}_c, f_c(w_c, t_c))$ .

I will make a structural assumption about closeness, which I call *Overlap*:<sup>8</sup>

*Overlap.* If  $\mathbf{A} \cap \mathbf{B} \neq \emptyset$ , then  $s(\mathbf{A}, \mathbf{B}) = \mathbf{A} \cap \mathbf{B}$ .

Suppose we want the closest worlds to an unsettled world  $\mathbf{B}$  where  $\mathbf{A}$  happens and that  $\mathbf{B}$  contains some  $\mathbf{A}$ -worlds. *Overlap* tells us that those closest worlds will be the  $\mathbf{A}$ -worlds in  $\mathbf{B}$ .

*Overlap* guarantees that  $\mathcal{I}_c$  is the closest world to itself. Since  $f_c(w_c, t_c)$  is the set of all worlds,  $s(\mathcal{I}_c, f_c(w_c, t_c))$  must be  $\mathcal{I}_c$ . This allows us to ignore the modal base when  $\mathcal{W}$  is unembedded or under past tense. (The modal base will be relevant, however, when we consider conditionals in the next section.)

Now let's state the semantics. To see if  $\lceil \mathcal{W}\phi \rceil$  is true, we find  $s(\mathcal{I}, f(w, t))$ , the closest world to the unsettled world where  $\phi$  is true. If  $\phi$  is true at  $s(\mathcal{I}, f(w, t))$  (i.e. true throughout  $\mathcal{I}$ ),  $\lceil \mathcal{W}\phi \rceil$  is true; if  $\phi$  is false at  $s(\mathcal{I}, f(w, t))$  (i.e. false throughout  $\mathcal{I}$ ),  $\lceil \mathcal{W}\phi \rceil$  is false; but if  $\phi$  is neither true nor false at  $s(\mathcal{I}, f(w, t))$  (i.e. is true at some worlds in  $\mathcal{I}$  but false at others),

<sup>7</sup>Cariani and Santorio also use a modal base, but do not assume it is empty.

<sup>8</sup>This is an analogue of the Strong Centering principle on Stalnaker selection functions, which says that  $w$  is always the closest world to itself.

$\lceil \mathcal{W}\phi \rceil$  is indeterminate. Say that the value of  $\llbracket \phi \rrbracket$  is *determinate* at a point just it is not indeterminate. Then we can make this precise as follows:

- (16) a.  $\llbracket \mathcal{W}\phi \rrbracket^{w,t,f,\mathcal{I}}$  is determinate only if either  
 (i)  $s(\mathcal{I}, f(w, t)) \subseteq \llbracket \phi \rrbracket^{t,f,\mathcal{I}}$  or  
 (ii)  $s(\mathcal{I}, f(w, t)) \subseteq \llbracket \neg\phi \rrbracket^{t,f,\mathcal{I}}$   
 b. If determinate,  $\llbracket \mathcal{W}\phi \rrbracket^{w,t,f,\mathcal{I}} = 1$  iff  $s(\mathcal{I}, f(w, t)) \subseteq \llbracket \phi \rrbracket^{t,f,\mathcal{I}}$

We also need to make explicit some background assumptions. I assume the following standard semantics for negation, the past and the indeterminacy operator.

- (17)  $\llbracket \neg\phi \rrbracket^{w,t,f,\mathcal{I}} = 1$  iff  $\llbracket \phi \rrbracket^{w,t,f,\mathcal{I}} = 0$   
 $\llbracket \neg\phi \rrbracket^{w,t,f,\mathcal{I}} = 0$  iff  $\llbracket \phi \rrbracket^{w,t,f,\mathcal{I}} = 1$ .  
 $\llbracket \neg\phi \rrbracket^{w,t,f,\mathcal{I}} = \#$  iff  $\llbracket \phi \rrbracket^{w,t,f,\mathcal{I}} = \#$ .  
 (18)  $\llbracket PAST \phi \rrbracket^{w,t,f,\mathcal{I}} = 1$  iff  $\exists t' < t : \llbracket \phi \rrbracket^{w,t',f,\mathcal{I}} = 1$   
 (19)  $\llbracket \nabla\phi \rrbracket^{w,t,f,\mathcal{I}} = 1$ , if  $\llbracket \phi \rrbracket^{w,t,f,\mathcal{I}} = \#$ ;  
 $\llbracket \nabla\phi \rrbracket^{w,t,f,\mathcal{I}} = 0$ , otherwise.

I assume a standard semantics for the circumstantial modal and make the standard assumption about its accessibility relation  $C(w, t)$ :

- (20)  $\llbracket \diamond\phi \rrbracket^{w,t,f,\mathcal{I}} = 1$  iff some  $w' \in C(w, t) : \llbracket \phi \rrbracket^{w',t,f,\mathcal{I}} = 1$

*Circumstantial.*  $C(w, t) = \{w' \mid w' \text{ is identical to } w \text{ up until } t\}$

Finally, I assume entailment is preservation of truth at a context:

*Truth at a context.*  $\llbracket \phi \rrbracket^c = \llbracket \phi \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c}$

*Diagonal validity.*  $\phi_1, \dots, \phi_n \mid_c \psi$  iff whenever  $\llbracket \phi_1 \rrbracket^c = \dots = \llbracket \phi_n \rrbracket^c = 1$  then  $\llbracket \psi \rrbracket^c = 1$ .

## 2.3 Delivering Openness, Past Settledness and Asymmetry

This package delivers our three features of future contingents:

**Fact 1.**  $\diamond\phi, \diamond\neg\phi \mid_c \nabla(\mathcal{W}\phi)$

**Proof.** Suppose  $\llbracket \diamond\phi \rrbracket^c = \llbracket \diamond\neg\phi \rrbracket^c = 1$ . Then  $C(w_c, t_c) \not\subseteq \llbracket \phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$  and  $C(w_c, t_c) \not\subseteq \llbracket \neg\phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$ . By *Unsettled World*,  $\mathcal{I}_c = C(w_c, t_c)$ , so  $\mathcal{I}_c \not\subseteq \llbracket \phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$  and  $\mathcal{I}_c \not\subseteq \llbracket \neg\phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$ . By *Overlap* and *Modal Bases*,  $s(\mathcal{I}_c, f_c(w_c, t_c)) = \mathcal{I}_c$ . Since then  $s(\mathcal{I}_c, f_c(w_c, t_c)) \not\subseteq \llbracket \phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$  and  $s(\mathcal{I}_c, f_c(w_c, t_c)) \not\subseteq \llbracket \neg\phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$ ,  $\llbracket \mathcal{W}\phi \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = \#$ ; so  $\llbracket \nabla(\mathcal{W}\phi) \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ .

**Fact 2.** Where  $\phi$  is non-modal,  $PAST \phi \mid_c PAST(\mathcal{W}\phi)$

**Proof.** Suppose  $\llbracket PAST \phi \rrbracket^c = 1$ . Then there's some  $t'$  earlier than  $t_c$  such that  $\llbracket \phi \rrbracket^{w_c, t', f_c, \mathcal{I}_c} = 1$ . By *Circumstantial* and *Unsettled World*, it follows that  $\mathcal{I}_c \subseteq \llbracket \phi \rrbracket^{t', f_c, \mathcal{I}_c}$ . By *Modal Bases* and *Overlap*, we have  $s(\mathcal{I}_c, f(w, t')) = \mathcal{I}_c$ . So there is some  $t' < t_c$ , such that  $s(\mathcal{I}_c, f(w, t')) \subseteq \llbracket \phi \rrbracket^{t', f_c, \mathcal{I}_c}$ . So  $\llbracket PAST(\mathcal{W}\phi) \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ .

**Fact 3.**  $\neg\mathcal{W}\phi \mid_c \mathcal{W}\neg\phi$



**Proof.** Suppose  $\llbracket \neg \mathcal{W}\phi \rrbracket^c = 1$ . Since  $\llbracket \neg \mathcal{W}\phi \rrbracket^c = 1$  is determinate, it must be that either i)  $s(\mathcal{I}_c, f_c(w_c, t_c)) \subseteq \llbracket \phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$  or ii)  $s(\mathcal{I}_c, f_c(w_c, t_c)) \subseteq \llbracket \neg \phi \rrbracket^{t_c, f_c, \mathcal{I}_c}$ . Case i) is inconsistent with our initial assumption, so case ii) holds. That suffices for  $\llbracket \mathcal{W}\neg\phi \rrbracket^c = 1$

### 3 A Conditional Semantics for Abilities

Now that we have introduced  $\mathcal{W}$ , I can give my semantics for abilities. On my theory,  $\lceil S$  can  $\phi \rceil$  is true iff, for some available action  $\alpha$ , if  $S$   $\alpha$ 's then  $\mathcal{W}(S \phi)$ 's.<sup>9</sup> To build up to a precise statement, we will first say more about actions, conditionals and the projection of indeterminacy.

To simplify, propositions about actions stand in for actions. I represent the available actions using a function  $\mathcal{A}$  that takes a world and a time and yields a set of propositions.  $\mathcal{A}(w, t)$  is the set of actions available to the subject at  $w$  and  $t$ .

I make two important formal assumptions about actions. First, I assume the set of available actions is a *partition* of the circumstantially possible worlds:<sup>10</sup>

*Partitionality.*  $\mathcal{A}(w, t)$  is a partition of  $C(w, t)$ , the circumstantially accessible worlds at  $w$  and  $t$ .

Importantly, since the actual world is always circumstantially accessible to itself, the actual world is always a member of an available action. Secondly, I add the assumption that if an action available in the past was performed, then it is settled that it was performed:

*Action Time.* If  $t < t_c$ , then if  $\alpha \in \mathcal{A}(w_c, t)$  and  $w_c \in \alpha$  then  $\mathcal{I}_c \subseteq \alpha$ .

To fully spell out my conditional analysis, we need to be clear on how the conditional works. I give it a *restrictor* semantics, a la Kratzer [1981b, 2012]: a conditional restricts the domain for a modal in the consequent to worlds where the antecedent is true.<sup>11</sup> More precisely, where  $f^{\mathbf{A}}$  is the function such that  $f^{\mathbf{A}}(w) = f(w) \cap \mathbf{A}$ , the conditional has the following truth-conditions.

$$(21) \quad \llbracket \text{if } \phi, \psi \rrbracket^{w, t, f, \mathcal{I}} = 1 \text{ iff } \llbracket \psi \rrbracket^{w, t, f^{\llbracket \phi \rrbracket^{t, f, \mathcal{I}}}, \mathcal{I}} = 1$$

When  $\mathcal{W}$  is embedded in the consequent, a restrictor conditional restricts the modal base argument to the selection function.  $\lceil \text{if } \phi, \mathcal{W}\psi \rceil$  is true iff the closest world to the actual unsettled world *where  $\phi$  is true* is one where  $\psi$  is true:

$$(22) \quad \llbracket \text{if } \phi, \mathcal{W}\psi \rrbracket^{w, t, f, \mathcal{I}} = 1 \text{ iff } \llbracket \mathcal{W}\psi \rrbracket^{w, t, f^{\llbracket \phi \rrbracket^{t, f, \mathcal{I}}}, \mathcal{I}} = 1 \\ \text{iff } s(\mathcal{I}, f^{\llbracket \phi \rrbracket^{t, f, \mathcal{I}}}(w, t)) \subseteq \llbracket \psi \rrbracket^{t, f^{\llbracket \phi \rrbracket^{t, f, \mathcal{I}}}, \mathcal{I}} = 1$$

Finally, we need to say how indeterminacy projects. I assume a Strong Kleene approach. On a Strong Kleene approach to disjunction,  $\lceil \phi \text{ or } \psi \rceil$  is determinate when we have enough information to determine a unique truth-value using the classical truth-table for 'or'. If at least one of  $\phi$  and  $\psi$  is true,  $\lceil \phi \text{ or } \psi \rceil$  is true; if both are false,  $\lceil \phi \text{ or } \psi \rceil$  is false; in all remaining cases it is indeterminate. This idea carries over to existential quantifiers: an existentially quantified sentence is true if it has a true instance; it is false if it has only false instances; and indeterminate otherwise.

<sup>9</sup>As mentioned, this semantics builds on the account of Mandelkern et al. [2016].

<sup>10</sup>This is a standard move in the literature in deontic modals. See for instance Cariani [2013].

<sup>11</sup>Kratzer assumes bare conditionals involve a tacit epistemic 'must' in the consequent.

Now apply this to ‘can’. I say that ‘S can  $\phi$ ’ is true when, for some available  $\alpha$ , ‘If S  $\alpha$ ’s,  $\mathcal{W}(S \phi$ ’s)’ is true; ‘S can  $\phi$ ’ is false when, for every available  $\alpha$ , ‘If S  $\alpha$ ’s,  $\mathcal{W}(S \phi$ ’s)’ is false; and is indeterminate, otherwise. Spelled out precisely:

- (23)  $\llbracket S \text{ can } \phi \rrbracket^{w,t,f,\mathcal{I}}$  is determinate only if either
- there is some  $\alpha \in \mathcal{A}(w,t)$  such that  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I}} = 1;
  - or for all  $\alpha \in \mathcal{A}(w,t)$ ,  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I}} = 0.

Putting this altogether, here is the full statement of the view:

- (24) a.  $\llbracket S \text{ can } \phi \rrbracket^{w,t,f,\mathcal{I}}$  is determinate only if either
- there is some  $\alpha \in \mathcal{A}(w,t)$  such that  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I}} = 1;
  - or for all  $\alpha \in \mathcal{A}(w,t)$ ,  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I}} = 0.
- b. If determinate,  $\llbracket S \text{ can } \phi \rrbracket^{w,t,f,\mathcal{I}} = 1$  iff for some  $\alpha \in \mathcal{A}(w,t)$  :  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I}} = 1
- i.e. iff for some  $\alpha \in \mathcal{A}(w,t)$  :  $s(\mathcal{I}, f(w,t) \cap \alpha) \subseteq \llbracket S \phi$ ’s \rrbracket^{t,f^\alpha,\mathcal{I}}

Given our earlier entry for negation, this yields the following entry for ‘cannot’:

- (25) a.  $\llbracket S \text{ cannot } \phi \rrbracket^{w,t,f,\mathcal{I}}$  is determinate only if either
- there is some  $\alpha \in \mathcal{A}(w,t)$  such that  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I}} = 1;
  - or for all  $\alpha \in \mathcal{A}(w,t)$ ,  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{w,t,f^\alpha,\mathcal{I},\mathcal{A}} = 0.
- b. If determinate,  $\llbracket S \text{ cannot } \phi \rrbracket^{w,t,f,\mathcal{I}} = 1$  iff for all  $\alpha \in \mathcal{A}(w,t)$  :  $\llbracket \mathcal{W}(S \phi$ ’s) \rrbracket^{t,f^\alpha,\mathcal{I}} \neq 1
- i.e. iff for all  $\alpha \in \mathcal{A}(w,t)$  :  $s(\mathcal{I}, f(w,t) \cap \alpha) \not\subseteq \llbracket S \phi$ ’s \rrbracket^{t,f^\alpha,\mathcal{I}}
- i.e., given the determinacy conditions, iff for all  $\alpha \in \mathcal{A}(w,t)$  :  $s(\mathcal{I}, f(w,t) \cap \alpha) \subseteq \llbracket \neg(S \phi$ ’s) \rrbracket^{t,f^\alpha,\mathcal{I}}

## 4 Predictions

I distilled the tension involving *Success* into three data points about ability modals:

Validity of *Past Success*.  $PAST(S \phi$ ’s)  $\rightsquigarrow$   $PAST(S \text{ can } \phi)$

Validity of *Can’t-entails-won’t*.  $\neg(S \text{ can } \phi) \rightsquigarrow \neg(S \phi$ ’s)

Invalidity of  $\vee$ -*Success*.  $S \phi$ ’s  $\vee S \psi$ ’s  $\not\rightsquigarrow S \text{ can } \phi \vee S \text{ can } \psi$

My semantics predicts all of these data. *Past Success* holds because if S actually did  $\alpha$  and  $\phi$  in the past, this suffices for the truth of ‘PAST(if S  $\alpha$ ’s,  $\mathcal{W}(S \phi$ ’s)’. *Can’t-entails-won’t* is valid because ‘S cannot  $\phi$ ’ says that for all available actions  $\alpha$ , the closest world where S  $\alpha$ ’s settles that S does not  $\phi$ . Since some available action must always be performed, this ensures that S  $\phi$ ’s.  $\vee$ -*Success* fails because S may end up  $\phi$ -ing even though no available action *settles* whether S  $\phi$ ’s.

More precisely, we can prove the following facts:

**Fact 4.**  $S \phi$ ’s  $\vee S \psi$ ’s  $\not\equiv$   $S \text{ can } \phi \vee S \text{ can } \psi$ .

**Proof.** Take S  $\phi$ ’s and (S  $\neg\phi$ ’s). Suppose:

- $\llbracket S \phi$ ’s \rrbracket^{w\_c,t\_c,f\_c,\mathcal{I}\_c} = 1
- $\mathcal{A}(w_c,t_c) = \{S \text{ tries to } \phi, S \text{ tries to } \neg\phi\}$
- $S \text{ tries to } \phi$  and  $S \text{ tries to } \neg\phi$  are consistent with both  $\llbracket S \phi$ ’s \rrbracket^{t\_c,f\_c,\mathcal{I}\_c} and  $\llbracket (S \neg\phi$ ’s) \rrbracket^{t\_c,f\_c,\mathcal{I}\_c}

By 1,  $\llbracket S \phi's \vee S \neg\phi's \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ . But the determinacy conditions for  $\lceil S \text{ can } \phi \rceil$  are not met: for no  $\alpha \in \mathcal{A}_c(w_c, t_c)$ :  $\alpha \subseteq \llbracket S \phi's \rrbracket^{t_c, f_c, \mathcal{I}_c}$  or  $\alpha \subseteq \llbracket S \neg\phi's \rrbracket^{t_c, f_c, \mathcal{I}_c}$ . Similarly the determinacy conditions for  $\lceil S \text{ can } \neg\phi \rceil$  are not met. So  $\llbracket S \text{ can } \phi \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = \llbracket S \text{ can } \neg\phi \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = \#$ . So  $S \phi's \vee S \psi's \not\equiv_{\mathcal{C}} S \text{ can } \phi \vee S \text{ can } \psi$ .

**Fact 5.** For non-modal  $\phi$ ,  $PAST(S \phi's) \equiv_{\mathcal{C}} PAST(S \text{ can } \phi)$

**Proof.** Suppose  $\llbracket PAST(S \phi's) \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ . By our entry for *PAST*, there is some  $t' < t$  such that  $\llbracket S \phi's \rrbracket^{w_c, t', f_c, \mathcal{I}_c} = 1$ . By *Circumstantial*, that  $C(w, t) \subseteq \llbracket S \phi's \rrbracket^{t', f_c, \mathcal{I}_c}$ ; and by *Unsettled World*,  $\mathcal{I}_c \subseteq \llbracket S \phi's \rrbracket^{t', f_c, \mathcal{I}_c}$ . Now by *Partition*, we know that there is some  $\alpha \in \mathcal{A}_c(w_c, t')$  such that  $w_c \in \alpha$ . *Action Time* gives us that  $\mathcal{I}_c \subseteq \alpha$ .

Now we can show that  $s(\mathcal{I}_c, f_c(w_c, t') \cap \alpha) \subseteq \llbracket S \phi's \rrbracket^{t', f_c, \mathcal{I}_c}$ . By *Modal Bases*,  $(f_c(w_c, t') \cap \alpha) = \alpha$ . So  $\mathcal{I}_c \subseteq f_c(w_c, t') \cap \alpha$ . So, by *Overlap*, we have that  $s(\mathcal{I}_c, f_c(w_c, t') \cap \alpha) = \mathcal{I}_c$ . But we already know that  $\mathcal{I}_c \subseteq \llbracket S \phi's \rrbracket^{t', f_c, \mathcal{I}_c}$ . So  $\exists \alpha \in \mathcal{A}_c(w_c, t') : s(\mathcal{I}_c, f_c(w_c, t') \cap \alpha) \subseteq \llbracket S \phi's \rrbracket^{t', f_c, \mathcal{I}_c}$ . Since  $\phi$  is non-modal,  $\llbracket S \phi's \rrbracket^{t', f_c, \mathcal{I}_c} = \llbracket S \phi's \rrbracket^{t', f_c^\alpha, \mathcal{I}_c}$ ; and so for some  $t' < t_c$ ,  $\llbracket S \text{ can } \phi \rrbracket^{w_c, t', f_c, \mathcal{I}_c} = 1$ . So  $\llbracket PAST(S \text{ can } \phi) \rrbracket^c = 1$ .

**Fact 6.** When  $\phi$  is non-modal,  $S \text{ cannot } \phi \equiv_{\mathcal{C}} \neg(S \phi's)$

**Proof.** Suppose  $\llbracket S \text{ cannot } \phi \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ . Since it is determinate, either i) there is some  $\alpha \in \mathcal{A}_c(w_c, t_c)$  such that  $\llbracket \mathcal{W}(S \phi's) \rrbracket^{w_c, t_c, f_c^\alpha, \mathcal{I}_c} = 1$ ; or ii) for all  $\alpha \in \mathcal{A}_c(w_c, t_c)$ ,  $\llbracket \mathcal{W}(S \phi's) \rrbracket^{w_c, t_c, f_c^\alpha, \mathcal{I}_c} = 0$ . Since  $\llbracket S \text{ cannot } \phi \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ , i) cannot hold. ii) and the determinacy conditions for  $\mathcal{W}$  gives us that for all  $\alpha \in \mathcal{A}_c(w_c, t_c)$ ,  $\llbracket \mathcal{W}\neg(S \phi's) \rrbracket^{w_c, t_c, f_c + \alpha, \mathcal{I}_c} = 1$ . In other words, for all  $\alpha \in \mathcal{A}_c(w_c, t_c)$ ,  $s(\mathcal{I}_c, f_c(w_c, t_c) \cap \alpha) \subseteq \llbracket \neg(S \phi's) \rrbracket^{t_c, f_c, \mathcal{I}_c}$ .

By *Partitionality*, *Circumstantial* and *Unsettled World*, there's some  $\alpha \in \mathcal{A}_c(w_c, t_c)$  such that  $w_c \in \alpha$ . This, together with *Modal Bases*, ensures  $w_c \in f^\alpha(w_c, t_c)$ . By *Circumstantial* and *Unsettled World*, we know  $w_c \in \mathcal{I}_c$ . By *Overlap*, we then know  $w_c \in s(\mathcal{I}_c, f_c^\alpha(w_c, t_c))$ . But then, since  $\forall \alpha \in \mathcal{A}_c(w_c, t_c)$ ,  $s(\mathcal{I}_c, f_c(w_c, t_c) \cap \alpha) \subseteq \llbracket \neg(S \phi's) \rrbracket^{t_c, f_c^\alpha, \mathcal{I}_c}$ ,  $w_c \in \llbracket \neg(S \phi's) \rrbracket^{t_c, f_c^\alpha, \mathcal{I}_c}$ , i.e.  $\llbracket \neg(S \phi's) \rrbracket^{w_c, t_c, f_c^\alpha, \mathcal{I}_c} = 1$ . When  $\phi$  is non-modal,  $\llbracket \neg(S \phi's) \rrbracket^{w_c, t_c, f_c^\alpha, \mathcal{I}_c} = 1$  iff  $\llbracket \neg(S \phi's) \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ . So  $\llbracket \neg(S \phi's) \rrbracket^{w_c, t_c, f_c, \mathcal{I}_c} = 1$ .

## References

- Dorit Abusch. Sequence of tense and temporal de re. *Linguistics and Philosophy*, 20(1):1–50, 1997.
- Rajesh Bhatt. Ability Modals and Their Actuality Entailments. In K. Shahin, S. Blake, and E.-S. Kim, editors, *WCCFL 17 Proceedings*, volume 17, pages 74–87, 1999.
- Mark A. Brown. On the Logic of Ability. *Journal of Philosophical Logic*, 17:1–26, 1988.
- Fabrizio Cariani. ‘Ought’ and Resolution Semantics. *Nous*, 47(3):534–558, 2013.
- Fabrizio Cariani and Paolo Santorio. Will done better: Selection semantics, future credence, and indeterminacy. *Mind*, 127(505):129–165, 2018.

- Cleo Condoravdi. Temporal interpretations of modals. In David Beaver, Stefan Kaufman, Brady Clark, and Luis Casillas, editors, *The Construction of Meaning*, pages 59–88. CSLI Publications, 2002.
- Valentine Hacquard. *Aspects of Modality*. PhD thesis, MIT, 2006.
- Risto Hilpinen. An Analysis of Relativised Modalities. In J.W. Davis, D.J. Hockney, and W.K. Wilson, editors, *Philosophical Logic*, volume 20. Springer, 1969.
- John F. Horty. *Agency and Deontic Logic*. Oxford University Press, 2001.
- John F. Horty and Nuel Belnap. The deliberative stit: A study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24(6):583–644, 1995.
- Anthony Kenny. Human Abilities and Dynamic Modalities. In J. Manninen and R. Tuomela, editors, *Essays on Explanation and Understanding*, pages 209–232. D. Reidel, 1976.
- Angelika Kratzer. What ‘Must’ and ‘Can’ Must and Can Mean. *Linguistics and Philosophy*, 1 (337-355), 1977.
- Angelika Kratzer. The Notional Category of Modality. In H. Eikmeyer and H. Rieser, editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*. de Gruyter, 1981a.
- Angelika Kratzer. The notional category of modality. In Hans-Jürgen Eikmeyer and Hannes Reiser, editors, *Words, Worlds and Contexts*, pages 38–74. Walter de Gruyter and Co., 1981b.
- Angelika Kratzer. Conditionals. In *Modals and Conditionals*, chapter 4. Oxford University Press, 2012.
- David Lewis. The Paradoxes of Time Travel. *American Philosophical Quarterly*, (145-152), 1976.
- Jan Lukasiewicz. On three-valued logic. *Ruch filozoficzny*, 5(170-171), 1920.
- Jan Lukasiewicz. *Aristotle’s Syllogistic From the Standpoint of Modern Formal Logic*. Garland, 1951.
- John MacFarlane. Future contingents and relative truth. *Philosophical Quarterly*, 53(212): 321–336, 2003. doi: 10.1111/1467-9213.00315.
- Matthew Mandelkern, Ginger Schultheis, and David Boylan. Agentive modals. To appear, *The Philosophical Review*, 2016.
- A.N. Prior. It was to be. In Peter Geach and Anthony Kenny, editors, *Papers in Semantics and Ethics*, pages 97–108. Duckworth, 1976.
- Arthur Prior. *Past, Present and Future*. Clarendon Press, Oxford, 1967.
- Robert C. Stalnaker. A theory of conditionals. *American Philosophical Quarterly*, pages 98–112, 1968.
- Richmond H. Thomason. Indeterminist time and truth-value gaps. *Theoria*, 36(3):264–281, 1970. doi: 10.1111/j.1755-2567.1970.tb00427.x.
- Richmond H Thomason. Combinations of tense and modality. In D Gabbay and F Guenther, editors, *Handbook of Philosophical Logic: Extensions of Classical Logic*. Reidel, 1984.

# Attitudes, Conditionals and Margins for Error

David Boylan and Ginger Schultheis\*

<sup>1</sup> Rutgers University, New Brunswick, New Brunswick, NJ, USA  
db1082@philosophy.rutgers.edu

<sup>2</sup> University of Chicago, Chicago, IL, USA  
vks@uchicago.edu

This paper is about *The Qualitative Thesis*, the thesis that if you are not sure that  $\varphi$  is false, then you are sure of the indicative conditional  $\varphi > \psi$  just in case you are sure of the material conditional  $\varphi \supset \psi$ . Following contextualists about indicative conditionals like Bacon [2015], we will understand this thesis in a *local* way—roughly as saying that if you are not sure that  $\varphi$  is false, then you are sure of the proposition expressed by  $\varphi > \psi$  *in your context* just in case you are sure of the material conditional  $\varphi \supset \psi$ . To state this precisely, let  $S^{c,w}(\llbracket \varphi \rrbracket^c)$  mean that the speakers in  $c$  are sure of  $\llbracket \varphi \rrbracket^c$  in  $w$ . Then:

**The Local Qualitative Thesis.** For any world  $w$  and context  $c$ , if  $\neg S^{c,w}(\llbracket \neg \varphi \rrbracket^c)$ , then:  $S^{c,w}(\llbracket \text{if } \varphi, \text{ then } \psi \rrbracket^c)$  if and only if  $S^{c,w}(\llbracket \varphi \supset \psi \rrbracket^c)$ .

We investigate the epistemological consequences of The Qualitative Thesis. We characterize The Qualitative Thesis in standard formal frameworks for studying the logic of attitudes and conditionals. With these characterization results in hand, we develop a connection first observed by Ben Holguín (p.c.) between The Qualitative Thesis and a plausible margin-for-error requirement on rational sureness. We show that The Qualitative Thesis is inconsistent with the margin-for-error principle. We propose a new *shifty* semantics for indicative conditionals. We say that the meaning of an indicative conditional is partly determined by the conditional’s local informational environment—the conditional’s *local context*—which, in turn, is systematically shifted by attitude operators. Our account validates The Qualitative Thesis, but dispenses with its undesirable epistemological consequences.

## 1 Motivating The Qualitative Thesis

The first argument for the Qualitative Thesis is that it follows from the conjunction of two standard claims about reasoning with conditionals. The first claim is that Modus Ponens is valid. This entails one half of the Qualitative Thesis—if you are sure of the indicative conditional  $\varphi > \psi$ , then you are sure of the corresponding material conditional  $\varphi \supset \psi$  (regardless of whether you are sure of  $\neg \varphi$ ). The second claim is that Stalnaker’s *Direct Argument* is a reasonable inference. This entails the second half of the Qualitative Thesis, namely, that if you are not sure that  $\neg \varphi$  and you’re sure of the material conditional  $\varphi \supset \psi$ , then you are also sure of the indicative conditional  $\varphi > \psi$ .

The Direct Argument is the argument from the disjunction  $\varphi \vee \psi$  to the indicative conditional  $\neg \varphi > \psi$ . The argument is compelling, as the following example shows.

- (1) Matt is either in Los Angeles or London.

---

\*The authors made equal contributions to the paper. Thanks to Justin Bledin, Sam Carter, Lucas Champollion, Cian Dorr, Kevin Dorst, Simon Goldstein, Justin Khoo, Harvey Lederman, Bernhard Salow, and Una Stojnić for helpful feedback. We are especially grateful to Ben Holguín and Matthew Mandelkern for numerous helpful conversations and comments on earlier drafts.

(2) So, if Matt is not in Los Angeles, he is in London.

We should not say that the Direct Argument is a *valid* inference. For (1) is equivalent to the material conditional *Matt's not in Los Angeles*  $\supset$  *Matt's in London*. So to say that (1) entails (2) would be to say that the material conditional entails the indicative conditional, a notoriously unacceptable consequence. Following Stalnaker, we should instead say that the Direct Argument is a *reasonable inference*—roughly, if you are sure of the disjunction  $\varphi \vee \psi$ , and are not sure that  $\varphi$ , then you are sure that  $\neg\varphi > \psi$ . This claim is equivalent to the second half of the Qualitative Thesis: if you are sure of the material conditional  $\varphi \supset \psi$  and you are not sure that  $\neg\varphi$ , then you are sure of the indicative conditional  $\varphi > \psi$ .<sup>1</sup>

The second argument for The Qualitative Thesis is that, given plausible assumptions, it follows from Stalnaker's Thesis, stated informally below.

**Stalnaker's Thesis.** The probability of  $\varphi > \psi$  is equal to the probability of  $\psi$  conditional on  $\varphi$ .

Stalnaker's Thesis is strongly supported both by intuition and experimental data. Take an example. You are holding a standard 52-card deck of cards, and you draw one at random. Ask yourself how confident you are in the following conditional.

(3) The selected card is a jack if it's a red card.

If you are like most, you will judge the probability of (3) to be 1/13. There are 26 red cards, and 2 of them are jacks. So the probability that the selected card is a jack given that it is red is 1/13. That is the probability that you assign to (3), in conformity with Stalnaker's Thesis.<sup>2</sup> It is easy to multiply examples like this. In general, we calculate the probability of a conditional  $\varphi > \psi$  by calculating the probability of  $\psi$  conditional on  $\varphi$ . This is just what we would expect if Stalnaker's Thesis were true. If we assume a plausible probabilistic account of being sure—specifically, that one is sure of some proposition just in case one assigns credence 1 to that proposition—then Stalnaker's Thesis entails The Qualitative Thesis.

## 2 The Qualitative Thesis in the Standard Framework

Here we present a standard formal framework for thinking about The Qualitative Thesis. This framework gives sureness ascriptions a Hintikka semantics. And, following Kratzer [2012] and Stalnaker [1975a], it gives the conditional a variably strict semantics, where  $\lceil$ if  $\phi$ , then  $\psi\rceil$  says, roughly, that  $\psi$  is true in the closest  $\phi$ -worlds.<sup>3</sup> We characterize the Qualitative Thesis in this framework and then use this result to show that the Qualitative Thesis puts a significant constraint on the logic of sureness, entailing a principle we call *No Opposite Materials*.

<sup>1</sup>Note that it doesn't follow from the Qualitative Thesis that *whenever* the you are sure of (1), you are in position to infer (2). You might be sure of (1) without leaving open that Matt is in Los Angeles, and the Qualitative Thesis is silent about that case. But, as Stalnaker points out, it is felicitous to assert (1) only if the context leaves open that Matt is not in Los Angeles, and so whenever (1) is felicitously asserted, the posterior context will entail that Matt is in Los Angeles or London, but leave open that Matt is in Los Angeles. This means that The Qualitative Thesis predicts that the speakers can infer (2) from (1) whenever they have become sure of (1) on the basis of a successful assertion of (1).

<sup>2</sup>See Adams (1975), Jeffrey and Edgington (1991), Stalnaker (1970), and van Fraassen (1976) for semantic theories that are designed to predict Stalnaker's Thesis. See Douven and Verbugge (2013) and Evans and Over (2004) for empirical work supporting Stalnaker's Thesis.

<sup>3</sup>In Boylan and Schultheis [2019], we prove that analogous results hold in a strict conditional framework, defended by Gillies [2004], Gillies [2009], Rothschild [2013], and Willer [2017], where  $\phi > \psi$  says that  $\phi \supset \psi$  holds throughout some fixed set of closest worlds.

## 2.1 A Standard Framework

We begin by constructing a propositional modal language that we can use to describe what a subject is sure of. The set of sentences of the language  $\mathcal{L}$  is the set of sentences generated by the following grammar:

- $\phi ::= p \mid \neg\phi \mid \phi \wedge \psi \mid \phi > \psi \mid S\phi$

The propositional connectives  $\supset$ ,  $\equiv$ , and  $\vee$  are defined as usual;  $>$  is our conditional operator. We read  $S\phi$  as *the subject is sure of  $\phi$* .

Next, the interpretation of the language. We assume that we are in some fixed arbitrary context with some relevant speaker who determines the particular interpretation of the conditional; that is, our semantic evaluation function,  $\llbracket \cdot \rrbracket$ , specifies only the *content* of the sentences in our language in this context.

We interpret the logical connectives in the standard way. To give the truth-conditions of the conditional, we use a *selection function*, which we assume is supplied by the background context. Where  $f(\mathbf{A}, w)$  is the set of selected  $\mathbf{A}$ -worlds at  $w$  and  $\llbracket \phi \rrbracket = \{w : \llbracket \phi \rrbracket^w = 1\}$ , we say:

**Standard Variably Strict Semantics.**  $\llbracket \phi > \psi \rrbracket^w = 1$  iff  $f(w, \llbracket \phi \rrbracket) \subseteq \llbracket \psi \rrbracket$

This clause says that  $\phi > \psi$  is true at a world  $w$  just in case all of the selected  $\phi$ -worlds at  $w$  are  $\psi$ -worlds. We stipulate that the selection function has the following natural properties:

**Success.**  $f(w, \mathbf{A}) \subseteq \mathbf{A}$

**Minimality.** If  $w \in \mathbf{A}$ , then  $w \in f(w, \mathbf{A})$

**Non-Vacuity.** If  $R(w) \cap \mathbf{A} \neq \emptyset$  then  $f(w, \mathbf{A}) \neq \emptyset$

Success and Minimality are standard assumptions.<sup>4</sup> Success says that the selected  $\mathbf{A}$ -worlds at  $w$  must be  $\mathbf{A}$ -worlds; it's needed to validate  $\phi > \phi$ . Minimality says that if  $w$  is an  $\mathbf{A}$ -world, then it must be among the selected  $\mathbf{A}$ -worlds at  $w$ ; it's needed to validate Modus Ponens. Non-Vacuity says that if there are accessible  $\mathbf{A}$ -worlds at  $w$ , then the set of selected  $\mathbf{A}$ -worlds at  $w$  isn't empty. It's needed to validate a form of Conditional Non-Contradiction, specifically:

**Weak Conditional Non-Contradiction.**  $\neg S\neg\phi \supset \neg((\phi > \psi) \wedge (\phi > \neg\psi))$

Weak Conditional Non-Contradiction says that if  $\phi$  is a live possibility, then  $\phi > \psi$  and  $\phi > \neg\psi$  are not consistent. This is a standard—and desirable—principle in conditional logic.<sup>5</sup> In general, there is something very wrong with asserting both  $\phi > \psi$  and  $\phi > \neg\psi$ .

Truth for the sureness operator  $S$  is defined in terms of an accessibility relation  $R$ :  $wRw'$  means that  $w'$  is compatible with what the subject is sure of in  $w$ .<sup>6</sup>

**Standard Hintikka Semantics.**  $\llbracket S\phi \rrbracket^w = 1$  iff  $\forall w' \in R(w) : \llbracket \phi \rrbracket^{w'} = 1$

<sup>4</sup>See, for example, Stalnaker [1968] and Lewis [1973].

<sup>5</sup>Why not a stronger version of Conditional Non-Contradiction that just says  $\phi > \psi$  and  $\phi > \neg\psi$  are not consistent? This stronger principle is inconsistent with Logical Implication, which says that  $\phi > \psi$  is always true when  $\phi$  entails  $\psi$ . Weak Conditional Non-Contradiction, by contrast, is consistent with Logical Implication. See Stalnaker [1968] and Lewis [1973] for theories that validate a version of Conditional Non-Contradiction that is at least as strong as Weak Conditional Non-Contradiction.

<sup>6</sup>We use the term *doxastic accessibility* to mean compatibility with what the subject *is sure of*, not what she believes.

We assume only that  $R$  is serial: at every world the subject has consistent beliefs and so what they know is compatible with *some* world. We assume that the accessibility relation  $R$  is that of the relevant agent in the arbitrary context we interpret our language in.

Given how we understand the interpretation of our language, we can characterize The Qualitative Thesis by characterizing the following object language principle:

$$\mathbf{QT}. \neg S\neg\phi \supset (S(\phi > \psi) \equiv S(\phi \supset \psi))$$

Our interpretation of the language forces us to understand QT *locally*—specifically, as saying that if the speaker of a given context  $c$  leaves open  $\llbracket\varphi\rrbracket$ , then she is sure of the proposition expressed by  $\phi > \psi$  relative to the information in *her* context just in case she is sure of  $\llbracket\varphi \supset \psi\rrbracket$ .

## 2.2 Characterizing the Qualitative Thesis

We will now characterize QT and show that it requires a strong constraint on the logic of sureness. Consider Stalnaker’s *Indicative Constraint*:

**Indicative Constraint.** If  $R(w) \cap \mathbf{A} \neq \emptyset$ , then if  $w' \in R(w)$ , then  $f(w', \mathbf{A}) \subseteq R(w)$ .<sup>7</sup>

This says that if  $\mathbf{A}$  is compatible with what the speaker is sure of in a world  $w$ , then for any world  $w'$  that is compatible with what the speaker is sure of in  $w$ , the selected  $\mathbf{A}$ -worlds at  $w'$  are a subset of the worlds compatible with what the subject is sure of at  $w$ . We prove:

**Fact 1.** QT is valid iff the Indicative Constraint holds.

**Proof.**  $\Leftarrow$ : We split QT into the following two principles and show that both must be valid on  $\mathcal{F}$ , if it meets the Indicative Constraint:

$$\mathbf{QT}_{\Rightarrow} \neg S\neg\phi \supset (S(\phi > \psi) \supset S(\phi \supset \psi))$$

$$\mathbf{QT}_{\Leftarrow} \neg S\neg\phi \supset (S(\phi \supset \psi) \supset S(\phi > \psi))$$

First we show  $\mathbf{QT}_{\Rightarrow}$  cannot fail.. Suppose for contradiction it did. Then, for some  $w$ ,  $\llbracket\neg S\neg\phi\rrbracket^w = \llbracket S(\phi > \psi)\rrbracket^w = 1$  but  $\llbracket S(\phi \supset \psi)\rrbracket^w = 0$ . So, for some  $w' \in R(w)$  :  $\llbracket\phi\rrbracket^{w'} = 1$  but  $\llbracket\psi\rrbracket^{w'} = 0$ . But, by Minimality,  $w' \in f(\llbracket\phi\rrbracket, w')$ . So  $\llbracket\phi > \psi\rrbracket^{w'} = 0$  and  $\llbracket S(\phi > \psi)\rrbracket^w = 0$  after all; contradiction. So  $\mathbf{QT}_{\Rightarrow}$  holds on any normal frame; and in particular it holds on any normal frame that meets the Indicative Constraint.

Now suppose that  $\mathbf{QT}_{\Leftarrow}$  fails. Then, for some  $w$ ,  $\llbracket\neg S\neg\phi\rrbracket^w = \llbracket S(\phi \supset \psi)\rrbracket^w = 1$  but  $\llbracket S(\phi > \psi)\rrbracket^w = 0$ . So, for some  $w' \in R(w)$ ,  $\llbracket\phi > \psi\rrbracket^{w'} = 0$ . This means there is some  $w''$  such that  $w'' \in f(\llbracket\phi\rrbracket, w')$  and  $w'' \notin \llbracket\psi\rrbracket$ . So, by Success,  $w'' \notin \llbracket\phi \supset \psi\rrbracket$ . But, since  $\llbracket S(\phi \supset \psi)\rrbracket^w$  it follows  $R(w) \subseteq \llbracket\phi \supset \psi\rrbracket$ . So  $w'' \in R(w)$ ; the Indicative Constraint fails.

$\Rightarrow$ : Suppose that the Indicative Constraint does not hold. Then for some  $\mathbf{A}$ , there’s some  $w$  and  $w'$  such that  $R(w) \cap \mathbf{A} \neq \emptyset$ ,  $w' \in R(w)$  but  $f(\mathbf{A}, w') \not\subseteq R(w)$ . So there’s some  $w'' \in f(\mathbf{A}, w)$  such that  $w'' \notin R(w)$ . But now we can build a model where QT fails. Let  $V(p) = \mathbf{A}$  and  $V(q) = \{w''\}$ . We can see that for all  $w' \in R(w)$   $\llbracket p \supset \neg q\rrbracket^{w'} = 1$ , as  $w'' \notin R(w)$ . So  $\llbracket S(p \supset q)\rrbracket^w = 1$ . But  $\llbracket p > q\rrbracket^{w'} = 0$ , since  $w'' \in f(\llbracket p\rrbracket, w')$ . But  $w' \in R(w)$ , so  $\llbracket S(p > q)\rrbracket^w = 0$ .  $\square$

<sup>7</sup>Versions of the Indicative Constraint are defended by von Fintel [1998], Bacon [2015], Khoo [2019], Mandelkern and Khoo [2019] and Mandelkern [2019b].



Given Fact 1, we can show that the Qualitative Thesis has important epistemological upshots. Consider the following property on frames:

**No Opposite Materials.** For any two worlds  $w_1, w_2$ , if there's some  $w_3$  such that  $w_1 R w_3$  and  $w_2 R w_3$ , then, for any  $\mathbf{A} \subseteq W$ : if  $R(w_1) \cap \mathbf{A} \neq \emptyset$ ,  $R(w_2) \cap \mathbf{A} \neq \emptyset$  and  $R(w_3) \cap \mathbf{A} \neq \emptyset$ , then there's no  $\mathbf{C} \subseteq W$  such that  $R(w_1) \subseteq \mathbf{A} \supset \mathbf{C}$  and  $R(w_2) \subseteq \mathbf{A} \supset \neg\mathbf{C}$ .

No Opposite Materials says that for certain pairs of worlds, and certain propositions  $\mathbf{A}$ , you can't be sure of a material conditional  $\mathbf{A} \supset \mathbf{C}$  at the first world and sure of the 'opposite' material conditional,  $\mathbf{A} \supset \neg\mathbf{C}$ , at the second. Which pairs of worlds? Any two worlds that see a world in common. And for which propositions? Any proposition that is consistent with what you're sure of at all three worlds.

We prove that No Opposite Materials is a consequence of QT:

**Fact 2.** QT is valid only if No Opposite Materials holds.

**Proof.** Suppose for contradiction that QT holds but No Opposite Materials does not. Then there are  $w_1, w_2, w_3$  and  $\mathbf{A}$  such that (i)  $R(w_1) \cap \mathbf{A} \neq \emptyset$ ,  $R(w_2) \cap \mathbf{A} \neq \emptyset$  and  $R(w_3) \cap \mathbf{A} \neq \emptyset$  but (ii) for some  $\mathbf{C}$ ,  $R(w_1) \subseteq \mathbf{A} \supset \mathbf{C}$  and  $R(w_2) \subseteq \mathbf{A} \supset \neg\mathbf{C}$ . Since QT is valid on  $\mathcal{F}$ ,  $\mathcal{F}$  obeys the Indicative Constraint. This means that  $f(\mathbf{A}, w_3) \subseteq R(w_1)$  and  $f(\mathbf{A}, w_3) \subseteq R(w_2)$ . So  $f(\mathbf{A}, w_3) \subseteq \mathbf{A} \supset \mathbf{C}$  and  $f(\mathbf{A}, w_3) \subseteq \mathbf{A} \supset \neg\mathbf{C}$ . Given Success, this means  $f(\mathbf{A}, w_3) \subseteq \mathbf{C}$  and  $f(\mathbf{A}, w_3) \subseteq \neg\mathbf{C}$ . But this can only happen if  $f(\mathbf{A}, w_3) = \emptyset$ . But this is already ruled out by Non-Vacuity. Contradiction.  $\square$

In the next section we develop a connection noted first by Ben Holguín (p.c.) and show that No Opposite Materials is inconsistent with a plausible *margin for error* requirement on rational sureness.<sup>8</sup> Fact 2 tells us that QT entails No Opposite Materials. It follows that QT is itself inconsistent with the margin for error requirement.

### 3 No Opposite Materials and Margin for Error Principles

To illustrate the margin for error requirement, we begin with a case from Timothy Williamson.<sup>9</sup> Mr. Magoo is staring out the window at a tree some distance off, wondering how tall it is. Assuming his only sources of information are reflection and present perception of the tree, what should he believe? That depends on how tall the tree actually is. If the tree is 100 inches tall, Mr. Magoo's visual information rules out possibilities in which the tree is 200 inches tall, or so we can imagine. So it would be reasonable for Magoo to be sure that the tree is not 200 inches tall. On the other hand, Magoo's visual information does not rule out possibilities in which the tree is 101 inches tall; his eyesight is simply nowhere near that good. It would not be reasonable for Magoo to be sure that the tree is not 101 inches tall. There's a general principle underlying these observations. Mr. Magoo's beliefs about the height of the tree are rational only if they leave a *margin for error*.<sup>10</sup>

<sup>8</sup>Holguín [Forthcoming] draws a very different moral from his argument, concluding that if you accept the margin for error principle you should reject The Qualitative Thesis. We think these can be reconciled.

<sup>9</sup>See Williamson [2000].

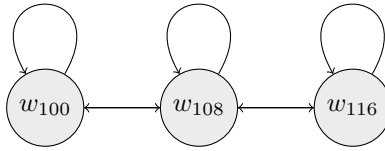
<sup>10</sup>Williamson introduces the margin for error principle as a requirement on knowledge, but as Hawthorne and Magidor [2009], Hawthorne and Magidor [2010] suggest, the principle is equally plausible for other attitudes. Hawthorne and Magidor focus on Stalnaker's attitude of *presupposition*, but similar considerations apply to rational sureness.

To state the margin for error requirement, we introduce a *margin for error frame*  $\langle W, R \rangle$ .  $W$  is a set of worlds representing possible tree heights. Where  $i$  is the height in inches of the tree in  $w$ ,  $W = \{w_i : i \in \mathbb{R} \text{ and } i > 0\}$ .  $R$  is a binary doxastic accessibility relation on  $W$ :  $w_i R w_j$  means that, in a world where the tree is  $i$  inches tall, it is compatible with everything Magoo is rationally sure of that the tree is  $j$  inches tall.  $R$  is defined as follows, relative to an arbitrarily chosen positive constant  $h$ .

**Magoo's Margin.**  $w_i R w_j$  if and only if  $|j - i| < h$ .

$h$  is Magoo's margin for error;  $h$  is positive, for otherwise his discrimination would be perfect.

No Opposite Materials fails on every margin for error frame. To see this, suppose that  $h = 10$ , and consider three worlds in  $W$ :  $w_{100}$ ,  $w_{108}$ , and  $w_{116}$ . Here is a diagram depicting Mr. Magoo's beliefs in these three worlds.



Mr. Magoo's belief worlds at  $w_{116}$  overlap with his belief worlds at  $w_{100}$ :  $w_{108}$  is consistent with what he is sure of in  $w_{116}$  and consistent with what he is sure of in  $w_{100}$ . Moreover, it's consistent with what Magoo is sure of at each world that the tree is either 100 inches tall or 116 inches tall. This means that the antecedent of No Opposite Materials is satisfied. The right and left worlds see a world in common,  $w_{108}$ . And the proposition that the tree is either 100 inches tall or 116 inches tall is consistent with what Magoo is sure of at all three worlds. But the consequent of No Opposite Materials is not satisfied. Since Magoo's margin for error is 10,  $w_{100}$  does not see  $w_{116}$  and  $w_{116}$  does not see  $w_{100}$ . As a result, Mr. Magoo is sure of 'opposite' material conditionals at  $w_{100}$  and  $w_{116}$ . At  $w_{100}$ , Mr. Magoo is sure that (4) is true; at  $w_{116}$ , Mr. Magoo is sure that (5) is true:

$$(4) \quad (116 \vee 100) \supset 100$$

$$(5) \quad (116 \vee 100) \supset 116$$

This shows that No Opposite Materials fails on every margin for error frame when  $h = 10$ . But the choice of 10 inches for  $h$  was arbitrary. It is not hard to see that No Opposite Materials will fail on every margin for error frame, regardless of the value of  $h$ .

In models that violate No Opposite Materials, The Qualitative Thesis places inconsistent demands on the selection function. At  $w_{100}$ , Magoo is sure of (4) and so by the Qualitative Thesis it follows that he is sure of the corresponding indicative conditional. Hence, at  $w_{108}$ , the selected  $(116 \vee 100)$ -worlds are worlds where the tree is 100 inches tall. On the other hand, at  $w_{116}$ , Magoo is sure of (5) and so by the Qualitative Thesis it follows that he is sure of the corresponding indicative conditional. Thus, at  $w_{108}$ , the selected  $(116 \vee 100)$ -worlds are worlds where the tree is 116 inches tall. But the selection function cannot meet both of these demands on pain of violating Non-Vacuity. Putting the problem this way suggests a solution. Instead of just one selection function, which we use to evaluate an indicative relative to just any belief state, we have multiple selection functions, indexed to different belief states. We develop this idea in the next section, showing how it allows us to validate The Qualitative Thesis in models like Williamson's Tree.

## 4 The Local Shifty Account of Conditionals

Starting in the early 1970s, theorists such as Stalnaker [1975b], Karttunen [1974], and Heim [1992] noticed that a sentence’s *local* informational environment can also influence its interpretation. Specifically, how we interpret an expression in a sentence is partly determined by the information contained in the rest of the sentence, its *local context*. This idea has been applied, in both static and dynamic frameworks, to both presupposition projection (Heim [1992] and Schlenker [2009]) and the phenomenon of epistemic contradictions (Veltman [1996], Gillies [2001], Yalcin [2007] and Mandelkern [2019a]).

We develop the idea sketched in the previous section by making the conditional’s contribution sensitive to its local context. Following Schlenker [2009], we add a local context parameter to the index of the semantic evaluation function. This ensures that when a conditional occurs under an attitude verb, the conditional is evaluated relative to the local context introduced by the attitude verb. We then validate The Qualitative Thesis using a version of the Indicative Constraint. But importantly, our account is not subject to the problem of conflicting demands. That is because the selection function for the conditional is indexed to the conditional’s local context. When the local context changes, the selection function does, too.

### 4.1 The Theory

We state our theory in a static, variably strict framework. Where  $\kappa$  is the conditional’s local context, here’s our semantic entry.

**Local Shifty Variably Strict Semantics.**  $\llbracket \text{if } \varphi, \text{ then } \psi \rrbracket^{\kappa, w} = 1$  if and only if:  $\forall w' \in f_{\kappa}(w, \llbracket \phi \rrbracket^{\kappa}) : \llbracket \psi \rrbracket^{\kappa, w'} = 1$

The Local Shifty Variably Strict Semantics is similar to the Standard Variably Strict Semantics. The difference is that there is a new parameter—a local context parameter—and the selection function is indexed to that parameter. Since selection functions are indexed to local contexts, we can impose constraints on selection functions that make reference to local contexts. We propose to replace Stalnaker’s Indicative Constraint with the following *Localized Indicative Constraint*:

**Localized Indicative Constraint.** If  $\mathbf{A} \cap \kappa \neq \emptyset$ , then  $\forall w' \in \kappa : f_{\kappa}(w', \mathbf{A}) \subseteq \kappa$

The Localized Indicative Constraint tells us that the selected antecedent worlds relative to a world  $w$  in the local context for the conditional must be a subset of the local context (so long as the antecedent is compatible with the local context).

With this new parameter, we restate the remaining constraints on the selection function.

**Success.**  $f_{\kappa}(w, \mathbf{A}) \subseteq \mathbf{A}$

**Minimality.** If  $w \in \mathbf{A}$ , then  $w \in f_{\kappa}(w, \mathbf{A})$ .

**Non-Vacuity.** If  $\kappa \cap \mathbf{A} \neq \emptyset$  then  $f_{\kappa}(w, \mathbf{A}) \neq \emptyset$ .

Success says that the selected  $\mathbf{A}$ -worlds are a subset of  $\mathbf{A}$ . Minimality says that if  $w$  is an  $\mathbf{A}$ -world, then  $w$  is one of the selected  $\mathbf{A}$ -worlds at  $w$ . We assume Success and Minimality for the same reasons as the standard framework does. Non-Vacuity says that if there are some  $\mathbf{A}$ -worlds in  $\kappa$ , then the set of selected  $\mathbf{A}$ -worlds at  $w$  is not empty. This constraint guarantees a *local* version of Weak Conditional Non-Contradiction: whenever there are  $\varphi$ -worlds in  $\kappa$ , at most one of  $\phi > \psi$  and  $\phi > \neg\psi$  can be true at a point of evaluation  $\langle \kappa, w \rangle$ .

We said that selection functions are indexed to local contexts and obey the Localized Indicative Constraint. The reason this matters, of course, is that local contexts are *shiftable*. In particular, they can be shifted by attitude predicates, such as *believe*, *want*, and, our focus in this paper, *is sure that*. Following Schlenker [2009], we assume that the local context introduced by an attitude predicate like *is sure that* at a world  $w$  is the set of worlds compatible with what the subject is sure of in  $w$ . Where  $R$  is a doxastic accessibility relation representing what an arbitrary subject is sure of and  $R(w)$  is the set of worlds compatible with what that subject is sure of in  $w$ :

**Shifty Hintikka Semantics.**  $\llbracket S\varphi \rrbracket^{\kappa,w} = 1$  if and only if:  $\forall w' \in R(w) : \llbracket \varphi \rrbracket^{R(w),w'} = 1$

**Shifty Hintikka Semantics** treats ‘is sure that’ as a necessity operator, just as the standard Hintikka semantics does. But now we’ve added a new parameter, a local context parameter, to the index. Shifty Hintikka Semantics says that attitude operators shift this parameter to  $R(w)$ , the set of worlds compatible with what the subject is sure of in  $w$ . This means that when we evaluate an attitude ascription like *Magoo is sure that if  $\varphi$ , then  $\psi$*  at a world  $w$ , we evaluate the embedded conditional relative to Magoo’s belief state *at*  $w$ . As we show in the next section, this is exactly what we need to validate The Qualitative Thesis without falling prey to the problem of conflicting demands.

## 4.2 Local Shifty Indicatives and The Qualitative Thesis

We prove that, on the Local Shifty Variably Strict Semantics, the Localized Indicative constraint is sufficient for QT:<sup>11</sup>

**Fact 3.** If the Local Indicative Constraint holds, then QT is valid.

**Proof.** Suppose the QT fails. Then for some  $\kappa$  and  $w$ , one of two cases obtains: i)  $\llbracket \neg S\neg\phi \rrbracket^{\kappa,w} = 1$ ,  $\llbracket S(\phi > \psi) \rrbracket^{\kappa,w} = 1$  and  $\llbracket S(\phi \supset \psi) \rrbracket^{\kappa,w} = 0$ ; or ii)  $\llbracket \neg S\neg\phi \rrbracket^{\kappa,w} = 1$ ,  $\llbracket S(\phi \supset \psi) \rrbracket^{\kappa,w} = 1$  and  $\llbracket S(\phi > \psi) \rrbracket^{\kappa,w} = 0$ .

Case i) is ruled out by Minimality. For suppose i) obtains. Since  $\llbracket S(\phi > \psi) \rrbracket^{\kappa,w} = 1$ , for all  $w' \in R(w) : f_{R(w)}(w', \llbracket \phi \rrbracket^{R(w)}) \subseteq \llbracket \psi \rrbracket^{R(w)}$ . Since  $\llbracket S(\phi \supset \psi) \rrbracket^{\kappa,w} = 0$ , there is some  $w' \in R(w) : \llbracket \phi \rrbracket^{R(w),w'} = 1$  and  $\llbracket \psi \rrbracket^{R(w),w'} = 0$ . But by Minimality, this  $w' \in f(w', \llbracket \phi \rrbracket^{R(w)})$ . So  $\llbracket \psi \rrbracket^{R(w),w'} = 1$  after all. Contradiction.

In case ii), the Local Indicative Constraint fails. Since  $\llbracket \neg S\neg\phi \rrbracket^{\kappa,w} = 1$ , there is some  $w' \in R(w)$  s.t.  $\llbracket \phi \rrbracket^{R(w),w'} = 1$ ; so the antecedent of the Local Indicative Constraint is satisfied when  $\kappa = R(w)$  and  $\mathbf{A} = \llbracket \phi \rrbracket^{R(w)}$ . Since  $\llbracket S(\phi \supset \psi) \rrbracket^{\kappa,w} = 1$ , for all  $w' \in R(w) : \text{either } \llbracket \phi \rrbracket^{R(w),w'} = 0 \text{ or } \llbracket \psi \rrbracket^{R(w),w'} = 1$ . Since  $\llbracket S(\phi > \psi) \rrbracket^{\kappa,w} = 0$ , there is some  $w' \in R(w)$  such that  $f_{R(w)}(w', \llbracket \phi \rrbracket^{R(w)}) \not\subseteq \llbracket \psi \rrbracket^{R(w)}$ . Since by Success  $f_{R(w)}(w', \llbracket \phi \rrbracket^{R(w)}) \subseteq \llbracket \phi \rrbracket^{R(w)}$ , it cannot be that  $f_{R(w)}(w', \llbracket \phi \rrbracket^{R(w)}) \subseteq R(w)$ . So the Indicative Constraint fails.  $\square$

Now that we’ve shown that the Qualitative Thesis is valid on our theory, the last thing to do is explain why we do not fall prey to the problem of conflicting demands in models where No Opposite Materials fails. Recall that in Williamson’s Tree, Magoo is sure of the material conditional (4) in  $w_{100}$  and he is sure of the material conditional (5) in  $w_{116}$ .

<sup>11</sup>Note that the Localized Indicative Constraint is not *necessary* for validating QT: we only need the instances where  $\kappa = R(w)$  for some  $w$ . But it seems to us that, from a semantic point of view, the more general principle is the more natural one.

(4)  $(100 \vee 116) \supset 100$

(5)  $(100 \vee 116) \supset 116$

In the standard variably strict framework, there is no way to guarantee that The Qualitative Thesis holds at both  $w_{100}$  and  $w_{116}$  without placing conflicting demands on the selection function at the overlap world  $w_{108}$ . To secure The Qualitative Thesis  $w_{100}$ , the selected  $(100 \vee 116)$ -worlds at  $w_{108}$  must be a subset of  $\{w_{100}\}$ ; otherwise  $(100 \vee 116) > 100$  would be false at  $w_{108}$ , and so Magoo would not be sure of it at  $w_{100}$ . To secure The Qualitative Thesis at  $w_{116}$ , the selected  $(100 \vee 116)$ -worlds at  $w_{108}$  must be a subset of  $\{w_{116}\}$ ; otherwise  $(100 \vee 116) > 116$  would be false at  $w_{108}$  Magoo would not be sure of it at  $w_{116}$ . The selection function cannot meet both of these demands on pain of violating Non-Vacuity.

In the local, shifty framework, by contrast, different belief states correspond to different selection functions. When we evaluate an indicative conditional relative to Magoo's belief state at  $w_{116}$ , we use one selection function; when we evaluate a conditional relative to his belief state at  $w_{100}$ , we use a different selection function. Consider (6) and (7):

(6)  $\llbracket \text{Magoo is sure that: } 100 \vee 116 > 100 \rrbracket^{\kappa}$

(7)  $\llbracket \text{Magoo is sure that: } 100 \vee 116 > 116 \rrbracket^{\kappa}$

Where  $R$  is an accessibility relation representing Magoo's beliefs, (6) is true at  $w_{100}$  just in case (8) is true at every world in  $R(w_{100})$ :  $w_{100}$  and  $w_{108}$ . (7) is true at  $w_{116}$  just in case (9) is true at every world in  $R(w_{116})$ :  $w_{108}$  and  $w_{116}$ .

(8)  $\llbracket (100 \vee 116) > 100 \rrbracket^{R(w_{100})}$

(9)  $\llbracket (100 \vee 116) > 116 \rrbracket^{R(w_{116})}$

But (8) and (9) do not place incompatible demands on the selection function at the overlap world  $w_{108}$ . (8) is true at  $w_{108}$  only if the selected  $(100 \vee 116)$ -world at  $w_{108}$ , *relative to Magoo's belief state at  $w_{100}$* , is  $w_{100}$ , whereas (9) is true at  $w_{108}$  only if the selected  $(100 \vee 116)$ -world at  $w_{108}$ , *relative to Magoo's belief state at  $w_{116}$* , is  $w_{116}$ . These are simply different demands on different selection functions, so there is no inconsistency.

## References

- Andrew Bacon. Stalnaker's Thesis in context. *Review of Symbolic Logic*, 8(1):131–163, 2015.
- David Boylan and Ginger Schultheis. The Qualitative Thesis. Ms, 2019.
- Anthony S. Gillies. A new solution to Moore's paradox. *Philosophical Studies*, 105(3):237–250, 2001.
- Anthony S. Gillies. Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616, 2004.
- Anthony S. Gillies. On truth-conditions for if (but not quite only if). *Philosophical Review*, 118(3):325–349, 2009.
- John Hawthorne and Ofra Magidor. Assertion, context, and epistemic accessibility. *Mind*, 118(470):377–397, 2009.

- John Hawthorne and Ofra Magidor. Assertion and epistemic opacity. *Mind*, 119(476):1087–1105, 2010.
- Irene Heim. Presupposition projection and the semantics of attitude verbs. *Journal of semantics*, 9(3):183–221, 1992.
- Ben Holguín. Indicative conditionals without iterative epistemology. *Noûs*, Forthcoming.
- Lauri Karttunen. Presupposition and linguistic context. *Theoretical linguistics*, 1(1-3):181–194, 1974.
- Justin Khoo. The meaning of *if*. Ms, 2019.
- Angelika Kratzer. Conditionals. In *Modals and Conditionals*, chapter 4. Oxford University Press, 2012.
- David Lewis. *Counterfactuals*. Blackwell, 1973.
- Matthew Mandelkern. Bounded modality. *Philosophical Review*, 128(1):1–61, 2019a.
- Matthew Mandelkern. If p, then p! Ms, 2019b.
- Matthew Mandelkern and Justin Khoo. Against preservation. *Analysis*, 79(3):424–436, 2019.
- Daniel Rothschild. Do indicative conditionals express propositions? *Noûs*, 47(1):49–68, 2013.
- Philippe Schlenker. Local contexts. *Semantics and pragmatics*, 2:1–78, 2009.
- Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975a.
- Robert Stalnaker. Presuppositions. In *Contemporary Research in Philosophical Logic and Linguistic Semantics*, pages 31–41. Springer, Dordrecht, 1975b.
- Robert C. Stalnaker. A theory of conditionals. *American Philosophical Quarterly*, pages 98–112, 1968.
- Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.
- Kai von Stechow. The presupposition of subjunctive conditionals. In Uli Sauerland and Orin Percus, editors, *MIT Working Papers in Linguistics 25*, pages 29–44. MITWP, Cambridge, MA, 1998.
- Malte Willer. Lessons from sobel sequences. *Semantics and Pragmatics*, 10, 2017.
- Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.
- Seth Yalcin. Epistemic modals. *Mind*, 116, 2007.

# Cognitive Modeling for Formal Semantics: The Organization of DRSs in Declarative Memory\*

Adrian Brasoveanu<sup>1</sup> and Jakub Dotlačil<sup>2</sup>

<sup>1</sup> Linguistics, UC Santa Cruz, Santa Cruz, CA, USA  
abrsvn@gmail.com, abrsvn@ucsc.edu

<sup>2</sup> University of Utrecht, Utrecht, The Netherlands  
j.dotlacil@gmail.com

## Abstract

One of the key and still outstanding challenges for formal semantics is “how to build formal semantics into real-time processing models – whether psychological or computational – that involve the integration of linguistic and not-specifically linguistic knowledge.” [11] In this paper, we outline the structure of a cognitively realistic semantic processor that we have fully implemented elsewhere [5]. The processor is able to incrementally construct DRT semantic representations in response to linguistic stimuli. We show how this semantic parser can be used to account for the reaction time (RT) data in the fan experiment reported in [4], which investigates how propositional information of the kind encoded by atomic DRSs is (stored and) retrieved from declarative memory. We provide the first (to our knowledge) fully implemented and cognitively realistic model of incremental interpretation and semantic evaluation that is systematically informed by formal semantics and can be used to model RT data.

## 1 Introduction

One of the key and still outstanding challenges for formal semantics is “how to build formal semantics into real-time processing models – whether psychological or computational – that involve the integration of linguistic and not-specifically linguistic knowledge.” [11, 4]) In this paper, we outline the structure of a cognitively realistic semantic processor – that is, a parser that incrementally constructs semantic representations in response to linguistic stimuli of the kind presented in self-paced reading or eye-tracking experiments. The parser has been implemented in [5].

Our choice for a processing-friendly semantics framework is Discourse Representation Theory (DRT) [9, 10]. We chose DRT because atomic DRSs and the compositional construction principles used to build them provide meaning representations and elementary compositional operations that are well understood mathematically, widely used in formal semantics, and can simultaneously function as both meaning representations and models (at least when we restrict ourselves to persistent DRSs [10, 96-97]). Thus, atomic DRSs can be thought of as mental models in the sense of [7, 8].

Due to the incremental nature of the semantic parser and its incremental-interpretation friendly cognitive components, the parser can be used to predict reaction time (RT) data. In particular, we show how it can model time latencies in the fan experiment reported in [4]. This experiment investigates how basic propositional information of the kind encoded by

---

\*We would like to thank the Amsterdam Colloquium 2019 anonymous reviewers and the reviewers of [5], on which this paper builds. Jakub Dotlačil was supported by the NWO grant VC.GW17.112. The usual disclaimers apply.

atomic DRSs is stored and retrieved from memory, which is an essential component of real-time semantic interpretation in at least two respects: (i) incrementally processing semantic representations involves composing and integrating semantic representations introduced by new sentences or new parts of a sentence with semantic representations of the previous discourse; (ii) incremental interpretation also involves evaluating new semantic representations relative to our mental model of the world, and integrating their content into our world knowledge database stored in declarative memory. We provide the first (to our knowledge) cognitively realistic model of incremental interpretation and truth evaluation that is systematically informed by formal semantics and can be used to model RT data.

## 2 The fan effect: three generalizations

The fan effect “refers to the phenomenon that, as participants study more facts about a particular concept, their time to retrieve a particular fact about that concept increases.” ([3, 186]). The original experiment in [4] demonstrated the fan effect in recognition memory. Participants studied 26 facts about people being in various locations, ten of which are exemplified in (1).

- (1)    **a.** A lawyer is in a cave.    **b.** A debutante is in a bank.    **c.** A doctor is in a bank.  
          **d.** A doctor is in a shop.    **e.** A captain is in a church.    **f.** A captain is in a park.  
          **g.** A fireman is in a park.    **h.** A hippie is in a park.    **i.** A hippie is in a church.  
          **j.** A hippie is in a town.

In the training part of the experiment, participants committed 26 items of this kind to memory. In the test part of the experiment, participants were presented with a series of sentences, some of which they had studied in the training part (targets) and some of which were novel (foils). They had to recognize the targets and had to reject the foils, which were novel combinations of the same people and locations.

The ten items in (1) form a minimal network of facts that instantiates the 9 experimental conditions in [4]. Each condition uniquely specifies how often a person or location concept has been used in (1). In case a person concept has been used only once (in one propositional fact), we say it has a fan of 1, if it has been used twice, we say it has a fan of 2 etc. In (1), and in the experiment as a whole, person or location concepts can have a fan of 1, 2 or 3. For example, *lawyer* and *cave* have a fan of 1, *captain* and *bank* have a fan of 2, and *hippie* and *park* have a fan of 3. The mean RTs for target recognition and foil rejection in this fan experiment, measured in seconds (s), are provided in Table 1 (from [3, 187]).

Several generalizations become apparent based on this data and various follow-up experiments, as discussed in [3]. **Generalization (i)**: averaging over targets and foils, the effect of 1-fan (both person and location) was about 1.2 s and increased by about 50 ms for each additional fan. **Generalization (ii)** (the min effect): retrieval latency is a function of the minimum fan associated with a probe, e.g., participants tend to respond more slowly to the 2-2 fan items than to the 1-3 or 3-1 items. **Generalization (iii)**: the fan effects are approximately equal for targets and foils, suggesting that foil rejection is not done by a serial (possibly exhaustive) search of the facts one knows about a concept.

## 3 The account in a nutshell: DRSs in declarative memory

We can reformulate the notion of fan as a relation between the main DRS contributed by a sentence and the sub-DRSs contributed by its three parts: the person indefinite, the location



<b>Target</b>		location fan			<b>Foil</b>		location fan				
		<b>RTs</b>	<b>1</b>	<b>2</b>			<b>3</b>	<b>RTs</b>	<b>1</b>	<b>2</b>	<b>3</b>
person	fan	<b>1</b>	1.11	1.17	1.15	person	fan	<b>1</b>	1.20	1.25	1.26
		<b>2</b>	1.17	1.20	1.23			<b>2</b>	1.22	1.36	1.47
		<b>3</b>	1.22	1.22	1.36			<b>3</b>	1.26	1.29	1.47

Table 1: The mean RTs (in seconds) for target recognition and foil rejection in the fan experiment in [4], as reported in [3].

indefinite, and the relational predicate *in*. Consider the fan example in (1-j). The DRSs (meaning representations) of the three major components of the sentence – the indefinites *a hippie* and *a town* and the binary predicate *in* – are composed / combined together to form the DRS / meaning representation for the full sentence. The content of *A hippie is in a town* ends up represented in memory as the attribute-value matrix (AVM) below: the DREF attribute specifies the new discourse referent (dref) introduced by indefinites, PRED specifies the predicate, and ARG1 and ARG2 specify which drefs fill the first and second argument slots.

$$\text{MAIN-DRS} \left[ \begin{array}{l} \text{SUB-DRS}_1: \left[ \begin{array}{l} \text{DREF: } 1 \\ \text{PRED } \text{hippie} \\ \text{ARG1: } 1 \end{array} \right] \\ \text{SUB-DRS}_2: \left[ \begin{array}{l} \text{DREF: } 2 \\ \text{PRED } \text{town} \\ \text{ARG1: } 2 \end{array} \right] \\ \text{SUB-DRS}_3: \left[ \begin{array}{l} \text{PRED } \text{in} \\ \text{ARG1: } 1 \\ \text{ARG2: } 2 \end{array} \right] \end{array} \right]$$

This partitioning into 3 sub-DRSs matches the basic compositional skeleton generally assumed in the formal semantics literature for this type of sentences, as well as the real-time incremental comprehension process the ACT-R cognitive architecture [1, 2] imposes on us. ACT-R, a widely used architecture that includes an explicit model of declarative memory, among other things, provides the cognitive infrastructure for our incremental semantic processor / parser.

The evaluation of the sentence as true (target) or false (foil) is accomplished by recalling from declarative memory the DRS with the highest activation, and checking whether the DRS incrementally constructed for the current sentence matches the components (sub-DRSs) of the DRS retrieved from memory. In principle, any DRS could be recalled from memory, but sub-DRSs in the currently constructed DRS spread activation to the same sub-DRSs in declarative memory, increasing the activation of the main DRS with matching information.

[3, 188-189] account for these generalizations in ACT-R. **Generalization (iii)** above is captured because memory search in ACT-R is not serial. **Generalizations (i) and (ii)** follow from properties of spreading activation. More specifically, in ACT-R, the total activation  $A_i$  of a fact/main DRS (a.k.a. chunk)  $i$  like the AVM above is:

$$A_i = B_i + \sum_j W_j S_{ji} \quad (1)$$

The base activation  $B_i$  in eq. (1) is determined by the history of usage of fact  $i$ : how many times  $i$  was retrieved and how much time elapsed since each of those retrievals. The spreading

activation component  $\sum_j W_j S_{ji}$  boosts the activation of fact  $i$  based on the concepts  $j$  that fact  $i$  is associated with, e.g., the person and location concepts/sub-DRSs in the AVM above.

For each concept  $j$ ,  $W_j$  is the extra activation flowing from concept  $j$ , and  $S_{ji}$  is the strength of the connection between concept  $j$  and fact  $i$  that modulates how much of the extra activation  $W_j$  actually ends up flowing to fact  $i$ . Source activation  $W$  (for simplicity, let's assume it is the same for all concepts  $j$  and drop the subscript) is inversely proportional to the number of concepts  $j$  in fact  $i$ . For example,  $W = 1/3$  for a fact/main DRS with 3 concepts/sub-DRSs.

The account of the fan effect crucially relies on the strengths of association  $S_{ji}$  between concepts  $j$  and facts  $i$ . The strengths of association are taken to be:

$$S_{ji} = S - \log(\text{fan}_j) \quad (2)$$

In eq. (2),  $S$  is a constant (baseline strength) to be estimated for specific experiments, and  $\text{fan}_j$  is the fan of concept  $j$ . Thus, strength of association, and therefore fact activation, decreases as a logarithmic function of concept fan (see Chapter 8 in [5] for a detailed justification).

The time  $T$  it takes to retrieve fact  $i$  from memory, i.e., its retrieval latency, is a function of fact activation  $A_i$ , as specified in eq. (3), where  $F$ , the latency factor, is a free parameter.

$$T = F e^{-A_i} \quad (3)$$

In an experiment, we usually do not observe the retrieval time directly, rather, we observe the latency in carrying out the task. We can represent this latency as  $T_{total}$ , such that  $T_{total} = I + T$ , where  $I$  is an intercept, i.e., the time needed to carry out cognitive tasks other than retrieval from declarative memory. We can simplify  $T_{total}$  as follows ():

$$T_{total} = I + F' \prod_j \text{fan}_j^W, \quad \text{where } F' = F e^{-B_i - S} \quad (4)$$

This latency equation shows how the ACT-R model captures the fan effect, i.e., **Generalization (i)** that recognition latency increases with fan. The retrieval latency for an AVM like the one above is:  $T_{total} = I + F' (\text{fan}_{person} \cdot \text{fan}_{location} \cdot \text{fan}_{in})^{\frac{1}{3}} = I + F' \sqrt[3]{\text{fan}_{person} \cdot \text{fan}_{location} \cdot \text{fan}_{in}}$ . The predicate *in* is connected to all the sentences/facts, so it has a constant fan and contributes a constant amount of latency across all conditions in the experiment. But the person and location fan values  $\text{fan}_{person}$  and  $\text{fan}_{location}$  are manipulated in the experiment, and the equation correctly predicts that as they increase, the corresponding latency increases.

This equation also provides an account of **Generalization (ii)** (the min effect). The reason is that the product of a set of numbers with a constant sum, specifically the product  $\text{fan}_{person} \cdot \text{fan}_{location}$ , is maximal when the numbers are equal, e.g.,  $2 \times 2 > 3 \times 1$ .

The remainder of the paper discusses in more detail how DRSs like the AVM above are constructed incrementally by means of production rules stored in procedural memory (Section §4) and how semantic (truth) evaluation is modeled as DRS/mental model retrieval from declarative memory (Section §5). In Section §6, we show how the complete cognitive model is fit to the fan-experiment data by embedding it into a Bayesian model, for which it provides the likelihood function; we also compare the posterior predictions of the Bayes+ACT-R model with the RTs observed in the fan experiment and discuss some limitations of the model. Finally, we outline some prospects for cognitive modeling in formal semantics (Section §7).

## 4 Incremental DRS construction and procedural memory

The AVM in Sect. 3 is an ACT-R friendly representation of the DRS  $[x, y | \text{hippie}(x), \text{town}(y), \text{in}(x, y)]$ , and we will assume here that the intuitive correspondence between these alternative represen-

tations is obvious (see Chapter 8 in [5] for a much more detailed discussion).

With these kind of representations in place, we can move to the core component of our cognitive model, which is an eager left-corner parser that parses syntactic trees and DRSs simultaneously and in parallel. As a new word is incrementally read in the usual left-to-right order for English, the model eagerly and predictively builds as much of the syntactic and semantic representation as it can, before deciding to move the eyes to the next word. This process of syntactic and semantic representation building is the process of comprehending the new word and integrating it into the currently available, partial syntactic and semantic structure. The comprehension and integration result in a new partial syntactic and semantic structure, which, in turn, provides the context relative to which the next word is interpreted.

The overall dynamics of parsing is thus very similar to dynamic semantics: a sequence of parsing actions or a sequence of dynamic updates charts a path through the space of information states. Info states serve a dual function: they provide the context relative to which a semantic update is executed, and they encode the result of executing that update. Information states are simpler for dynamic semantics: they are basically variable assignments or similar structures. For our ACT-R incremental parser, information states are complex entities consisting of the partial syntactic and DRS structures built up to that point, as well as the state of the buffers and modules of the ACT-R mind at that point in the comprehension process.

Implicit in the above description of the parsing process is the fact that an ACT-R mind is composed of modules, which include declarative and procedural memory, but also visual and motor modules etc. Modules are not directly accessible: they can only be accessed through associated buffers, e.g., the retrieval buffer is associated with declarative memory. Buffers serve a dual purpose: individually, they provide the input/output interface to specific modules; as a whole, however, buffers represent the current cognitive state of the mind. Crucially, productions fire based on the current cognitive state, i.e., conditioned on the contents of various buffers. The ACT-R architecture constrains cognitive behavior in various ways, two of which are that buffers can hold only one chunk (AVM), and only one production can fire at any given time.

Production rules are **condition-action** pairs stored in procedural memory. These rules trigger a cognitive **action** if the cognitive context (the buffers) satisfy a range of **conditions**. Let us see some of these rules in action by showing how our model incrementally interprets our running example *A hippie is in a town*. Assume that the first word *A* has already been read and recognized as a determiner. At that point, the PROJECT NP rule below is selected and fired:

PROJECT: NP ==> Det N

conditions		=>	actions	
goal>	TASK: parsing STACK1: S DREF_PEG: =peg	=>	goal>	TASK: move_peg STACK1: N STACK2: NP STACK3: S
retrieval>	ISA: word CAT: Det		+imaginal>	ISA: parse_state NODE_CAT: NP DAUGHTER1: Det
			+context>	ISA: drs DREF: =peg ARG1: =peg

The conditions of this rule involve the goal buffer, whose contents drive the cognitive process, and the retrieval buffer, which is the interface to the declarative memory module. The lexical entry for the determiner *a* has to be available in the retrieval buffer, and our current top parsing goal is to parse a sentence (S). Furthermore, the rule assumes that a fresh discourse referent index, a.k.a. DREF\_PEG, is available in the goal buffer. The value of the DREF\_PEG attribute is

assigned to an ACT-R variable =peg (in this context, ‘=’ indicates that we’re dealing with a variable). The variable =peg (which stands for a specific value, 1 in this case) will be used to specify the cognitive actions triggered by this rule. The index =peg is fresh in the sense that no discourse referent with that index was introduced in the semantic representation up to this point (it was never the argument of a predicate, it couldn’t have served as the antecedent for a pronoun etc.) The term ‘peg’ and its specific usage here originates in [12].

Assuming its preconditions are satisfied, the rule triggers several actions. Most importantly, it simultaneously builds a syntactic representation in the imaginal buffer and a semantic representation in the (discourse) context buffer. The ‘+’ sign (+imaginal, +context) indicates that new representations are added to these buffers. In the imaginal buffer (a standard buffer in ACT-R that maintains an internal representation of the current stage of the cognitive process), we build the unary branching node NP dominating the Det node (which in turn dominates the word *A*). In the context buffer, we start a new DRS that will be further specified by the upcoming N (*hippie*). This DRS introduces a new discourse referent with the index =peg and requires this discourse referent to be the first argument (ARG1) of the still-unspecified predicate contributed by the upcoming N.

The context buffer is not a standard buffer in ACT-R. We postulate it in our model for pedagogical purposes, so that we can neatly separate the syntactic and semantic components of our parse states and be able to focus on the semantic aspects of the incremental interpretation process. The DRS contributed by the PROJECT NP rule to the context buffer is provided below. Note that the value of the =peg index is specified as 1, since this is the first discourse referent introduced in the sentence.

AVM format:	DRS format:								
<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">drs</td></tr> <tr><td style="padding: 2px 5px;">DREF:</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">ARG1:</td><td style="padding: 2px 5px;">1</td></tr> </table>	ISA:	drs	DREF:	1	ARG1:	1	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">STILL-UNSPECIFIED-PREDICATE(1)</td></tr> </table>	1	STILL-UNSPECIFIED-PREDICATE(1)
ISA:	drs								
DREF:	1								
ARG1:	1								
1									
STILL-UNSPECIFIED-PREDICATE(1)									

In addition to adding chunks to the imaginal and (discourse) context buffers, the PROJECT NP production rule also specifies a new task in the goal buffer, which is to move.peg. This will trigger a rule that advances the ‘fresh discourse referent’ index to the next number – in our case, it will advance it to 2. This way, we ensure that we have a fresh discourse referent available for any future expression that might introduce a new discourse referent, e.g., another indefinite.

We also have a new stack of expected syntactic categories in the goal buffer: we first expect an N, at which point we will be able to complete the subject NP. Once that is completed, we can return to our overarching goal of parsing an S.

Once the discourse referent peg is advanced, a sequence of rules fires that leads to reading the next word, namely *hippie*, and retrieving its lexical entry from declarative memory. At that point, the PROJECT AND COMPLETE N rule below is fired.

PROJECT AND COMPLETE: N																					
	<b>conditions</b>		<b>actions</b>																		
goal>	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">STACK1:</td><td style="padding: 2px 5px;">N</td></tr> <tr><td style="padding: 2px 5px;">STACK2:</td><td style="padding: 2px 5px;">NP</td></tr> <tr><td style="padding: 2px 5px;">STACK3:</td><td style="padding: 2px 5px;">S</td></tr> </table>	STACK1:	N	STACK2:	NP	STACK3:	S	goal>	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">STACK1:</td><td style="padding: 2px 5px;">NP</td></tr> <tr><td style="padding: 2px 5px;">STACK2:</td><td style="padding: 2px 5px;">S</td></tr> </table>	STACK1:	NP	STACK2:	S								
STACK1:	N																				
STACK2:	NP																				
STACK3:	S																				
STACK1:	NP																				
STACK2:	S																				
retrieval>	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">word</td></tr> <tr><td style="padding: 2px 5px;">CAT:</td><td style="padding: 2px 5px;">N</td></tr> <tr><td style="padding: 2px 5px;">PRED:</td><td style="padding: 2px 5px;">=p</td></tr> </table>	ISA:	word	CAT:	N	PRED:	=p	⇒	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">+imaginal&gt;</td><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">parse_state</td></tr> <tr><td></td><td style="padding: 2px 5px;">NODE_CAT:</td><td style="padding: 2px 5px;">NP</td></tr> <tr><td></td><td style="padding: 2px 5px;">DAUGHTER1:</td><td style="padding: 2px 5px;">Det</td></tr> <tr><td></td><td style="padding: 2px 5px;">DAUGHTER2:</td><td style="padding: 2px 5px;">N</td></tr> </table>	+imaginal>	ISA:	parse_state		NODE_CAT:	NP		DAUGHTER1:	Det		DAUGHTER2:	N
ISA:	word																				
CAT:	N																				
PRED:	=p																				
+imaginal>	ISA:	parse_state																			
	NODE_CAT:	NP																			
	DAUGHTER1:	Det																			
	DAUGHTER2:	N																			
		context>	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">drs</td></tr> <tr><td style="padding: 2px 5px;">PRED:</td><td style="padding: 2px 5px;">=p</td></tr> </table>	ISA:	drs	PRED:	=p														
ISA:	drs																				
PRED:	=p																				

This rule requires the lexical entry for an N (*hippie*, in our case) to be in the retrieval buffer. It also requires the top of the goal stack, that is, the most immediate syntactic expectation, to

be N, followed by NP and S in the subsequent stack positions.

Once these preconditions are satisfied, the rule triggers several actions. First, N is popped off the goal stack. Second, a new part of the tree is added to the imaginal buffer: the binary branching node NP with two daughters, a Det on the left branch and an N on the right. Finally, the DRS in the context buffer is updated with a specification of the PRED attribute: the new value is the predicate =p, contributed by the word *hippie* in the retrieval buffer.

After this rule fires, the DRS in the context buffer becomes SUB-DRS<sub>1</sub> of the MAIN-DRS for the sentence *A hippie is in a town*:

AVM format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">drs</td></tr> <tr><td style="padding: 2px 5px;">DREF:</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">PRED:</td><td style="padding: 2px 5px;">hippie</td></tr> <tr><td style="padding: 2px 5px;">ARG1:</td><td style="padding: 2px 5px;">1</td></tr> </table>	ISA:	drs	DREF:	1	PRED:	hippie	ARG1:	1	DRS format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="text-align: center; padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">hippie(1)</td></tr> </table>	1	hippie(1)
ISA:	drs												
DREF:	1												
PRED:	hippie												
ARG1:	1												
1													
hippie(1)													

At this point, we fire the production rule PROJECT AND COMPLETE S ==> NP VP (not listed here; see Chapter 8 in [5] for details), which eagerly discharges the expectations for an NP and an S from the goal stack and replaces them both with a VP expectation. At the same time, a new part of the syntactic tree is built in the imaginal buffer, namely the top node S and its two daughters NP and VP. This rule also performs an important semantic operation: it transfers the discourse referent 1 introduced by the subject NP to the goal buffer, so that it is available as the first argument of the VP we are about to parse.

We assume, for simplicity, that the copula (*is*) is semantically vacuous. After that, the rule PROJECT AND COMPLETE PP ==> P NP parses the preposition *in*. Once again, we refer the reader to Chapter 8 of [5] for details. The sub-DRS contributed by the subject NP is cleared from the context buffer and the partially constructed sub-DRS below is placed there:

AVM format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">drs</td></tr> <tr><td style="padding: 2px 5px;">PRED:</td><td style="padding: 2px 5px;">in</td></tr> <tr><td style="padding: 2px 5px;">ARG1:</td><td style="padding: 2px 5px;">1</td></tr> </table>	ISA:	drs	PRED:	in	ARG1:	1	DRS format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">in(1, -)</td></tr> </table>	in(1, -)
ISA:	drs									
PRED:	in									
ARG1:	1									
in(1, -)										

The fact that the previous sub-DRS is cleared from the context buffer means that we never have the full view of the main DRS we are constructing. During the incremental interpretation process, we hold only parts of it in our cognitive state (working memory). This is a consequence of essential constraints imposed by the ACT-R architecture, which reflect experimentally established constraints on human cognitive processes.

We can now move on to parsing the location NP *a town*. Once the indefinite determiner *a* is read and lexically retrieved, the PROJECT AND COMPLETE NP ==> Det N rule is fired, which sets the second argument slot of the predicate *in* to a newly introduced discourse referent 2. The result is SUB-DRS<sub>3</sub> of the main DRS we need to construct for our running example:

AVM format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">drs</td></tr> <tr><td style="padding: 2px 5px;">PRED:</td><td style="padding: 2px 5px;">in</td></tr> <tr><td style="padding: 2px 5px;">ARG1:</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">ARG2:</td><td style="padding: 2px 5px;">2</td></tr> </table>	ISA:	drs	PRED:	in	ARG1:	1	ARG2:	2	DRS format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">in(1, 2)</td></tr> </table>	in(1, 2)
ISA:	drs											
PRED:	in											
ARG1:	1											
ARG2:	2											
in(1, 2)												

We then clear this sub-DRS from the context buffer and start constructing a new DRS in there that, once the subsequent word *town* is parsed, will be SUB-DRS<sub>2</sub> of the main DRS for our running example *A hippie is in a town*. This is the final sub-DRS we needed to construct:

AVM format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="padding: 2px 5px;">ISA:</td><td style="padding: 2px 5px;">drs</td></tr> <tr><td style="padding: 2px 5px;">DREF:</td><td style="padding: 2px 5px;">2</td></tr> <tr><td style="padding: 2px 5px;">PRED:</td><td style="padding: 2px 5px;">town</td></tr> <tr><td style="padding: 2px 5px;">ARG1:</td><td style="padding: 2px 5px;">2</td></tr> </table>	ISA:	drs	DREF:	2	PRED:	town	ARG1:	2	DRS format:	<table style="border-collapse: collapse; width: 100%;"> <tr><td style="text-align: center; padding: 2px 5px;">2</td></tr> <tr><td style="padding: 2px 5px;">town(2)</td></tr> </table>	2	town(2)
ISA:	drs												
DREF:	2												
PRED:	town												
ARG1:	2												
2													
town(2)													

## 5 Truth evaluation as retrieval from declarative memory

Our model has completely parsed the sentence *A hippie is in a town*. With the DRS for this sentence in hand, we move on to establishing whether the sentence was studied in the training phase or not. To put it differently, the training phase presented a set of facts, i.e., a *model* in formal semantics terms, and now we have to evaluate whether the test sentence is true relative to the model. That is, we view *semantic evaluation as memory retrieval*.

Thus, the bigger picture behind our DRT-based model of the fan effect is that the process of semantic interpretation proceeds in two stages, similar to the way interpretation proceeds in DRT. In the initial stage, we incrementally construct the semantic representation (DRS, mental discourse model) for the sentence to be evaluated. In DRT [10], this first stage involves a step-by-step transformation of a complete syntactic representation of the sentence into a DRS by means of a series of construction-rule applications. In our processing model, this stage consists of applying an eager, left-corner, syntax-and-semantics parser to the current test sentence.

The DRS/mental discourse model that is the result of the first stage encodes constraints that the set of facts/actual model in declarative memory has to satisfy. In simple cases like the one we are considering, the mental model is homomorphic to a fact in declarative memory.

We can therefore move to the second stage, in which we evaluate the truth of this DRS by connecting it to the actual, ‘real-world’ model, which is our background database of facts stored in declarative memory. In DRT, the second stage involves constructing an embedding function (a partial variable assignment) that verifies the DRS relative to the model. In our processing model, truth/falsity evaluation involves retrieving – or failing to retrieve – a fact from declarative memory that has the same structure as the DRS we have just constructed.

Once again, the full details of the process of semantic evaluation as memory retrieval are provided in Chapter 8 of [5]. The main idea is that the semantic evaluation process sets up a cognitive state that provides the correct configuration for spreading activation from sub-DRSs to main DRSs stored in memory, so that **Generalizations (i), (ii) and (iii)** above are captured.

## 6 Fitting cognitive models to data: Bayes+ACT-R

We fit our cognitive model to the fan-experiment data by embedding the ACT-R-based DRS parser into a Bayesian model. We estimate four model parameters: (i) source activation (weight)  $W$ ; (ii) baseline strength of association  $S$ ; (iii) rule firing  $r$ , which is the time required for a rule to fire; and, finally, (iv) the latency factor  $F$ . The structure of the Bayesian model is provided in the left panel of Fig. 1: low-information priors for the parameters are listed at the top and the entire ACT-R model provides the likelihood function, which outputs semantic evaluation (main DRS retrieval) latencies.

The discussion in Section §3 above provides the motivation for estimating the  $W$ ,  $S$  and  $F$  parameters. We have also decided to estimate the  $r$  parameter, which can be seen as a way to capture the intercept  $I$  in eq. (4).

The posterior estimates for these four parameters are provided and discussed in Chapter 8 of [5]. The plot of the posterior predictions of the model is given in the right panel of Fig. 1. The plot compares the RTs predicted by our model (y-axis) against the observed, experimentally-obtained RTs. The blue dots represent the means, the blue whiskers are 95% credible intervals.

The model captures the data well: the dots are close to the red diagonal line, where predicted and observed RTs are identical. There are slight discrepancies between the predictions of the model and the actual observations, which are due to variance in the data that our fan model ignores from the start. For instance, when we inspect the Target sentence RTs in Table 1, we

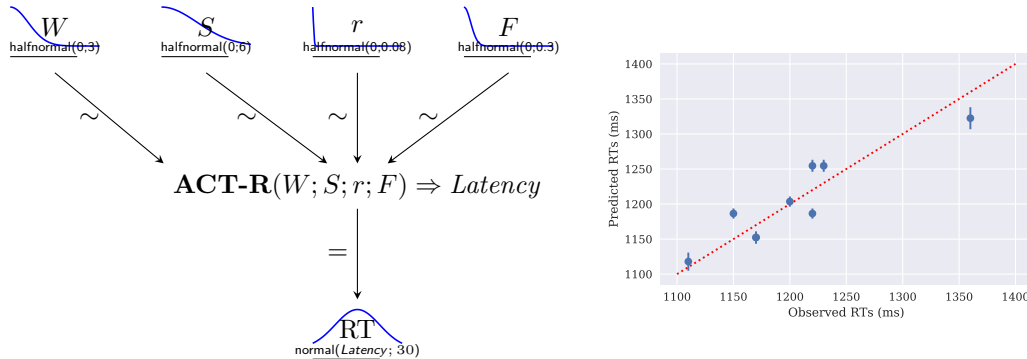


Figure 1: Bayesian model that embeds the ACT-R semantic parser (left-side), and fit of the model to the fan experiment (right-side).

see that the mean RT for targets with a fan of 3 for person and 1 for location is 1.22 s, while the mean RT for targets with a fan of 3 for location and 1 for person is 1.15 s. This difference of 70 ms cannot be captured in our model, which treats the two cases as identical.

## 7 Prospects for cognitive modeling in formal semantics

We modeled the fan experiment in [4], bringing together the ACT-R account of the fan effect in [3] and formal semantics theories of natural language meaning and interpretation. The resulting incremental semantic processor is the first one to integrate, in a computationally explicit way, dynamic semantics in its DRT incarnation on one hand, and mechanistic processing models formulated within an independently motivated cognitive architecture on the other.

In developing this cognitive model, we argued that the fan effect provides fundamental insights into the memory structures and cognitive processes that underlie semantic interpretation and evaluation, which is the process of determining whether a given sentence is true relative to a database of known facts, i.e., relative to a model in the sense of model-theoretic semantics.

Future directions for this line of research include investigating whether a partial match of known facts is considered good enough for language users in comprehension. This could provide an integrated account of a variety of interpretation-related phenomena, e.g., Moses illusions, the interpretation of plural definites like *The reporters asked questions*, where the sentence is true without every single reporter asking a question, and ‘partial presupposition resolution’ cases, where part of the presupposition is resolved and part of it is accommodated.

Another direction for future research is reexamining the decisions we made when developing our incremental DRT parser that were not based on cognitive plausibility, but instead were made for pedagogical reasons – in an effort to ensure that the contributions made by semantic theories were still recognizable in the final syntax-and-semantics parser. These decisions led to posterior estimates of the parameters in the model that differ from the previous literature. For example, rule firing time  $r$ , traditionally set to 50 ms in ACT-R, was estimated to be 15 ms in our parser.

We also oversimplified the model in various ways, e.g., we initiated the semantic evaluation process (the search for a matching fact in declarative memory) only when the sentence was fully parsed syntactically and semantically. This is unrealistic: parsing, disambiguation and semantic evaluation are most probably interspersed processes, and searches for matching facts/main



DRSs in declarative memory are probably launched eagerly after every parsed sub-DRS, if not even more frequently. See [6] for a similar proposal, and for an argument that such an approach, coupled with a judicious use of spreading activation, might explain the preference to provide given (topic) information earlier in the sentence, and new (focused) information later.

Another oversimplification was the decision to add sub-DRSs to the goal buffer to initiate spreading activation only after the entire sentence was parsed. It is likely that some, possibly all, sub-DRSs are added to the goal buffer and start spreading activation to matching main DRSs as soon as they are parsed. It might even be that such sub-DRSs are added to the goal buffer and start spreading activation even before they are completely parsed. This could be the case for NPs with a relative clause, where the partial sub-DRS obtained by processing the nominal part before the relative clause is added to the goal buffer, and then it is revised once the relative clause is processed.

We have barely scratched the surface with the model introduced in this paper. There are many semantic phenomena for which we have detailed formal semantics theories, but no similarly detailed and formalized theories and models on the processing side. The variety of semantic phenomena to be investigated and the variety of semantic frameworks and processing hypotheses that can be formulated and computationally implemented offer a rich and largely unexplored theoretical, empirical and modeling territory.

## References

- [1] John R. Anderson. *How can the human mind occur in the physical universe?* Oxford University Press, 2007.
- [2] John R. Anderson and Christian Lebiere. *The Atomic Components of Thought*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1998.
- [3] John R. Anderson and Lynne M. Reder. The Fan Effect: New Results and New Theories. *Journal of Experimental Psychology: General*, 128(2):186–197, 1999.
- [4] John Robert Anderson. Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6(4):451 – 474, 1974.
- [5] Adrian Brasoveanu and Jakub Dotlačil. *Computational Cognitive Modeling and Linguistic Theory*. Language, Cognition, and Mind (LCAM) Series. Springer, Dordrecht, 2020. In production. The current version of the *pyactr* library (Python3 ACT-R) is available here: <https://github.com/jakdot/pyactr>; the frozen version used in the book, together with the code for all the models, is available here: <https://github.com/abrsvn/pyactr-book>.
- [6] Raluca Budiú and John R. Anderson. Interpretation-based processing: A unified theory of semantic sentence processing. *Cognitive Science*, 28:1–44, 2004.
- [7] Philip N Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press, 1983.
- [8] Philip N. Johnson-Laird. The history of mental models. *Psychology of reasoning: Theoretical and historical perspectives*, pages 179–212, 2004.
- [9] Hans Kamp. A theory of truth and semantic representation. In Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, editors, *Formal Methods in the Study of Language*, pages 277–322. Mathematical Centre Tracts, Amsterdam, 1981.
- [10] Hans Kamp and Uwe Reyle. *From Discourse to Logic. Introduction to Model theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer, Dordrecht, 1993.
- [11] Barbara Partee. The semantics adventure. ms., 2011.
- [12] C.F.M. Vermeulen. Merging without mystery: Variables in dynamic semantics. *Journal of Philosophical Logic*, 24:405–450, 1995.



# Definiteness across Languages: from German to Mandarin

David Bremmers, Jianan Liu, Martijn van der Klis and Bert Le Bruyn<sup>1</sup>

Utrecht University

d.j.e.bremmers@students.uu.nl, j.liu4@uu.nl,  
m.h.vanderklis@uu.nl, b.s.w.lebruyne@uu.nl

## Abstract

We showcase the potential of a data-driven methodology for cross-linguistic research: *Translation Mining*. We introduce the technique and put it to use in the domain of definiteness. We show how Translation Mining confirms existing insights about definiteness in English, German (Schwarz 2009) and Mandarin (Jenks 2018) while at the same time leading to novel insights for Mandarin.

## 1 Introduction

When approaching a phenomenon  $\alpha$  from a language  $\beta$  for the first time, linguists rely on the knowledge about  $\alpha$  that was built up based on other languages. This is the standard way of doing cross-linguistic research: we interpret data against existing accounts and – where necessary – extend these or develop new ones.

In this paper, we showcase the potential of a more data-driven methodology: *Translation Mining* (henceforth *TM*). We present the technique and put it to use in the domain of definiteness. In our presentation of the technique, we compare it to the semantic map method (e.g. Haspelmath 1997). The semantic map method is an example of the standard approach sketched above but also shares some features with TM and provides for an insightful comparison.

The paper is organized as follows. Section 1 introduces TM, Section 2 provides a quick review of recent work on definiteness in the languages under consideration and Section 3 presents our application of TM. Section 4 discusses and concludes.

## 2 From Semantic Maps to Translation Mining

TM can be considered the data-driven variant of semantic maps. We introduce semantic maps (2.1) and then turn to TM (2.2).

### 2.1 Semantic maps

A semantic map is ‘a geometrical representation of functions in "conceptual/semantic space" which are linked by connecting lines’ (Haspelmath 2003). Semantic maps are mainly found in

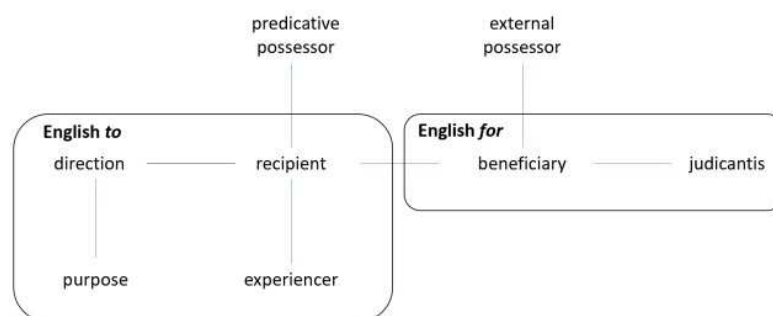
---

<sup>1</sup>David covered the German data collection and analysis, Jianan the Mandarin one. Martijn designed the Translation Mining interface and contributed to the methodological part of this paper. Bert supervised David and Jianan, brought together the different parts and worked them into the current paper. We thank the members of the Time in Translation project for valuable feedback and gratefully acknowledge financial support of NWO (grant 360-80-070).

typological work: ‘the map describes and constrains languages that venture their grammars and/or lexicons into this space, both with respect to diachrony and synchrony.’ (Van der Auwera & Plungian 1998). An example of a semantic map is given in Figure 1. We first look into the research that underlies it and then turn to its interpretation.

Semantic maps are developed for specific domains. The map in Figure 1 was developed by Haspelmath (2003) for datives. Researchers start from existing analyses of the domain and cast the widest possible net to identify as many functions as possible. E.g., in the domain of datives, these would include functions like *recipient*, *beneficiary*, etc.

The next step is to establish whether these functions should indeed be distinguished. The criterion for this is empirical: a function is ‘put on the map if there is at least one pair of languages that differ with respect to this function’ (Haspelmath 2003). An example is the distinction English makes between ‘give something *to* someone’ and ‘buy something *for* someone’: many languages allow their dative to appear in both contexts but the fact that English makes a formal distinction between the two warrants the inclusion in the map of the functions *recipient* and *beneficiary*.



**Figure 1:** A semantic map of typical dative functions / boundaries of English *to* and *for* (Haspelmath 2003)

Turning to the interpretation of a map as in Figure 1, we find functions connected by lines, and boundaries of specific lexical items/constructions:

Connecting lines are added between two functions – as for *direction* and *purpose* – to indicate that there are lexical items/constructions that combine these functions. The absence of a connecting line between two functions – as for *purpose* and *experiencer* – entails that no lexical items/constructions combine these functions without also conveying the functions that connect them – *direction* and *recipient* in this particular case. The exact geometrical lay-out of the map (e.g. the fact that *direction* and *purpose* are arranged on the vertical axis) does not reflect any claim and in general depends on the graphical creativity of the researcher.

For concreteness, we have added the boundaries of English *to* and *for* but in principle we should be able to add any lexical item/construction that conveys one of the functions that is included in the map. The claim the map makes is that these lexical items/constructions convey functions that are connected on the map, both in diachrony and synchrony. Herein lies the predictive power of semantic maps.

## 2.2 Translation Mining

TM can be considered the data-driven variant of semantic maps.<sup>2</sup> We briefly compare the two at the levels of data collection and analysis.

### *Data collection*

As we indicated above, classical work in semantic maps relies on existing literature to establish which functions might be relevant for a given domain. The empirical contribution of a semantic map then lies in establishing which functions should indeed be included and how they should be arranged.

The data collection in TM is different. To illustrate, let us switch from the dative domain to the domain of definiteness. Where classical work in semantic maps would look into existing literature and focus on what has been said about definiteness in different languages, TM would aim for neutrality as to what has been claimed before. The way to achieve this is to start from a marker of definiteness – say English *the* – and to select all contexts that contain this marker in the English version of a parallel corpus (i.e. a corpus containing texts and their translations into different languages). The next step is to establish how *the* in these contexts is rendered in the other languages and to repeat the whole procedure for all the equivalents of *the* that are found in the different languages. The output of this data collection is a set of datapoints consisting of contexts with the lexical items/constructions that are used in the different languages of the corpus. An example is given in (1):

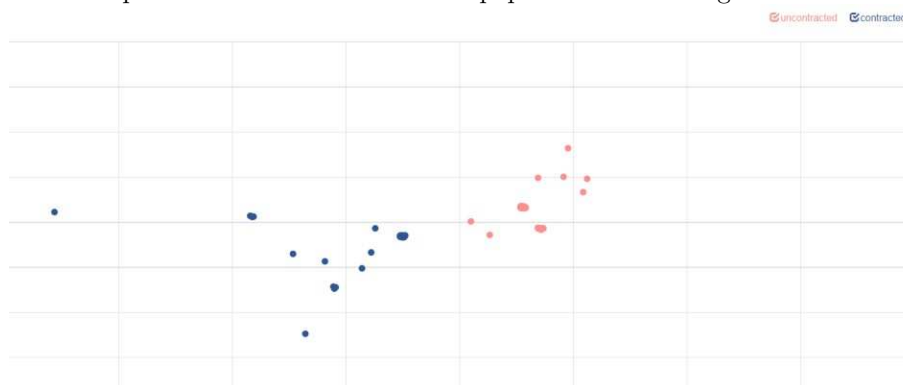
- (1) ‘I’m not having one **in the house**, Petunia!’ English: *the*, German: *contracted definite*, Mandarin: *demonstrative*

This datapoint involves the use of a definite in English (*in the house*), a definite that contracts with the preceding preposition in German (*im Haus* as opposed to *in dem Haus*) and a proximal demonstrative in Mandarin (*zhè CL fángzi*).

### *Analysis*

The analysis in TM starts with a graphical representation of the data, known as a map. We first present TM maps and then compare them to classical semantic maps.

One of the maps that will come back in this paper is the following:



**Figure 2:** A TM map of definiteness / distribution of German lexical items/constructions

<sup>2</sup>Variants of core components of TM can – among others – be found in Wälchli & Cysouw (2012) and Beekhuizen et al. (2017). For a comparison between classical semantic maps and work like that of Wälchli & Cysouw (2012), see also Georgakopoulos & Polis (2018). See van der Klis et al. (2017, 2019) for an application of TM to the semantic domain of the (*have*)-Perfect.

This map is based on a subset of data that were collected along the lines set out above. We get back to the selection criteria in Section 3. For now, we note that we have used data from English, German and Mandarin and comment on what the map represents and how it has been generated.

Every dot on the map stands for a datapoint like (1). The organization of the dots on the map is created through Multi-Dimensional Scaling (MDS). Intuitively speaking, this algorithm groups contexts that use the same form in a given language. By doing this in parallel for the three languages, a single cross-linguistic pattern of groupings obtains.

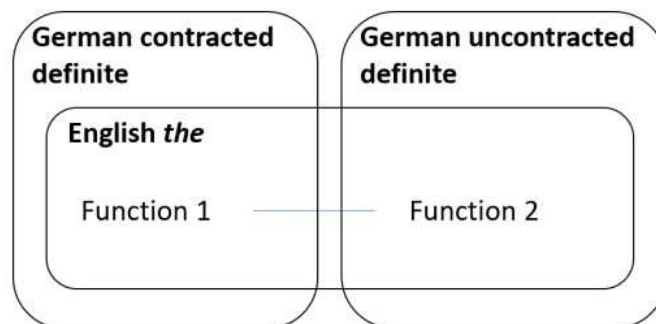
With language specific color schemes, TM visualizes variation across languages. A comparison between the English map in Figure 3 and the German one in Figure 2 allows us to establish that the formal distinction between contracted and uncontracted definites has no formal equivalent in English. This is not immediately clear from the static version of the maps: the purple dots in the English map make up very tight clusters and this makes it hard to establish that they outnumber the other forms.



**Figure 3:** A TM map of definiteness / distribution of English lexical items/constructions

The maps created in TM are however interactive and allow us to zoom in and establish that out of the 96 datapoints on the map, 80 involve the definite article, making it by far the most frequent equivalent of both contracted (30 out of 40) and uncontracted definites (50 out of 56).

Now that we have established how TM maps are generated and what they represent, we can compare them to classical semantic maps. The best way to do so is to think of groupings of dots in TM maps as functions. The classical semantic map fragment that can be derived from Figures 2 and 3 would then look as in Figure 4:



**Figure 4:** A classical semantic map fragment based on the TM maps of definiteness for German and English

Figure 4 distinguishes two functions. German makes a formal distinction between the two using contracted and uncontracted definites whereas English *the* covers both functions. The crucial point to be made is that no interpretive labels are given to the functions, unlike what we saw for the classical semantic map in Figure 1. Indeed, TM maps are different from classical semantic maps in that they merely record formal distinctions between groups of contexts. Under the assumption that formal distinctions reflect functional ones, TM maps then invite the researcher to investigate differences between groups of contexts and establish which functional distinctions underlie the formal ones. In the classical map method, these two steps are not strictly separated: formal distinctions are used to confirm the relevance of functional ones, not to discover them. In this sense, TM maps and the TM method count as the data-driven variant of classical semantic maps and the classical semantic map method.

### 3 Definiteness across languages

Putting TM to work for the analysis of a full semantic domain like definiteness across a typologically balanced sample of languages is – at the moment – utopian. We restrict our enterprise in two ways. The first is to limit ourselves to three languages: German, Mandarin and English. The second is to only look at a subset of contexts, *viz.* definites that occur in PPs in German. To justify these restrictions, we first need to give a brief sketch of the literature on definiteness.

The literature typically distinguishes between two types of definiteness, one based on uniqueness, the other on anaphoricity (Russell 1905; Strawson 1950; Kamp 1981; Heim 1982). English is believed to combine the two types in a single lexical item (*the*). German, on the other hand, has been claimed to formally distinguish between the two in the prepositional domain: Schwarz (2009) claims that uniqueness definites can contract with a subset of prepositions whereas anaphoric definites resist contraction. Examples from Schwarz are given in (2) and (3):

- (2) Der Empfang wurde **vom/#von dem Bürgermeister** eröffnet.  
‘The reception was opened **by the mayor**.’
- (3) In der New Yorker Bibliothek gibt es ein Buch über Topinambur. Neulich war ich dort und habe **#im / in dem Buch** nach einer Antwort auf die Frage gesucht, ob man Topinambur grillen kann.  
‘In the New York public library, there is a book about topinambur. Recently, I was there and searched **in the book** for an answer to the question of whether one can grill topinambur.’

The mayor in (2) has not been introduced before but is the unique mayor of the contextually salient town. This is a uniqueness context and the definite contracts with the preposition. In (3), a book is introduced in the first sentence and referred back to in the second. The second sentence constitutes a familiarity context and contraction is not allowed.

The contraction facts from German led Schwarz to coin the terms *weak* and *strong* definiteness. The first applies to contracted (uniqueness) definites, the latter to uncontracted (anaphoric) ones. We adopt this terminology here as it has become standard in studies looking at definiteness distinctions in typologically diverse languages (see Aguilar-Guevara et al. 2019). One such language is Mandarin.

Jenks (2018) argues that Mandarin makes a formal distinction between weak and strong definiteness in its use of bare nouns and demonstratives (examples from Jenks 2018):<sup>3</sup>

---

<sup>3</sup>This distinction disappears in subject position. See Jenks (2018) for discussion.

- (2) a. (#Na/Zhe ge) Taiwan (de) zongtong hen shengqi.  
 DEM/DEM CL Taiwan DE president very angry  
 ‘The president of Taiwan is very angry.’
- b. Jiaoshi li zuo-zhe yi ge nansheng he yi ge nüsheng. Wo zuotian yudao  
 classroom in sit-ASP one CL boy and one CL girl I yesterday meet  
 #(na ge) nansheng.  
 DEM CL boy  
 ‘There are a boy and a girl sitting in the classroom. I met the boy yesterday.’

(2) shows demonstratives are proscribed in uniqueness contexts (2a) but are mandatory in anaphoric contexts (2b). Bare nouns come out as the mirror image of demonstratives: they are proscribed in anaphoric contexts and mandatory in uniqueness contexts.

With the preliminary data and intuitions about definiteness in German, Mandarin and English in place, we can return to the restrictions we introduced above. To showcase the potential of TM, we need a semantic domain with a solid theoretical basis and a literature that has used this basis to explore cross-linguistic variation. Definiteness qualifies with the literature on English definiteness going back over a century. The recent literature on German and Mandarin further neatly illustrates how new data are approached from existing insights: the distinction between uniqueness and anaphoric definiteness has made its way into the cross-linguistic literature under the label of the weak/strong distinction. The restriction to definites occurring in PPs in German is inspired by Schwarz, who suggests that PPs are the only place where we can find form variation in German.<sup>4</sup> This restriction allows for a more focused data collection.

## 4 Definiteness in Translation Mining

In this section, we put TM to work for an analysis of the semantic domain of definiteness with the two restrictions we introduced in Section 3: we focus on German, Mandarin and English and restrict our attention to contexts that are rendered with a definite in German PPs. We introduce our corpus, briefly comment on the data collection and then present our results.

### 4.1 The corpus

The corpus we selected is the first volume of the Harry Potter series and its translations to German and Mandarin. An important asset of this corpus is that the source text is English. Given that English does not make a formal distinction between different types of definiteness, the formal distinctions we find in the German and Mandarin translations are independent from each other and translation biases are kept to a minimum. Other assets of this corpus are its recency and the availability of many other languages.

---

<sup>4</sup>We independently checked this suggestion through the study of referential expressions in German, English, Dutch and French. We used TM as our methodology and the first three chapters of the novel *L'Étranger* and its translations as our parallel corpus. We found that there is indeed little variation in the expression of definiteness in these languages. Based on Löbner (2011) we had expected some cross-linguistic variation between the definite article and demonstratives but we failed to establish this variation.

## 4.2 The data collection

Our first step was to extract PPs in German and specifically those in which the preposition contracted with the definite following it or in which the preposition could have contracted. For the uncontracted forms we selected all PPs in the novel. For the contracted forms we limited ourselves to forms in the first three chapters except for those contracted PPs that had an uncontracted counterpart with the same noun. In the latter case we extracted all occurrences in the novel. The goal of our selection procedure was to end up with a dataset with a more or less even distribution of contracted and uncontracted PPs, while at the same time maximizing the likelihood of including minimal pairs.

Once the set of German PPs established, we aligned them with the English original and the Mandarin translation. Alignment was done by two of the authors, one a native speaker of German and the other a native speaker of Mandarin. The number of contexts we ended up with for German-English-Mandarin amounts to 96.

## 4.3 Results

We discuss the results on the basis of TM maps. We start with German, then move to English and end with Mandarin.

### *German*

The TM map with the color scheme for German was introduced above as Figure 2. Unsurprisingly, we find clear-cut groups of contracted and uncontracted forms. There are 40 contracted cases and 56 uncontracted ones.

Given that TM maps merely record formal distinctions, the map in Figure 2 does not allow us to confirm or refute Schwarz's claim that the groups oppose weak/uniqueness and strong/anaphoric definiteness. With the interactive TM interface we can however zoom in on the contexts themselves and confirm the predictions an analysis along the lines of the weak/strong distinction makes.

- (3) **E:** 'I suppose we could take him to the zoo,' said Aunt Petunia slowly, '... and leave him **in the car** ...'  
**G:** 'Ich denke, wir könnten ihn in den Zoo mitnehmen', sagte Tante Petunia langsam, '... und ihn **im Wagen** lassen ...'
- (4) [Context: 'As the owls flooded into the Great Hall as usual, everyone's attention was caught at once by a long thin package carried by six large screech owls. Harry was just as interested as everyone else to see what was in this large parcel and was amazed when the owls soared down and dropped it right in front of him, knocking his bacon to the floor.']  
**E:** They had hardly fluttered out of the way when another owl dropped a letter **on top of the parcel**.  
**G:** Sie waren kaum aus dem Weg geflattert, als eine andere Eule einen Brief **auf das Paket** warf.

The car in (3) does not refer back to a previously introduced car but to the unique family car. It consequently counts as a weak definite. The parcel in (4) refers back to the package that was introduced before and counts as a strong definite. As predicted by Schwarz, German relies on a contracted definite in (3) and an uncontracted definite in (4).

In the case of German, TM does not lead to the discovery of new semantic insights in the typology of definiteness but does allow us to check the basic predictions existing analyses

make. It furthermore allows us to establish the opposition between contracted and uncontracted forms in German as a baseline for studying weak/uniqueness and strong/anaphoric contexts in our corpus.<sup>5</sup>

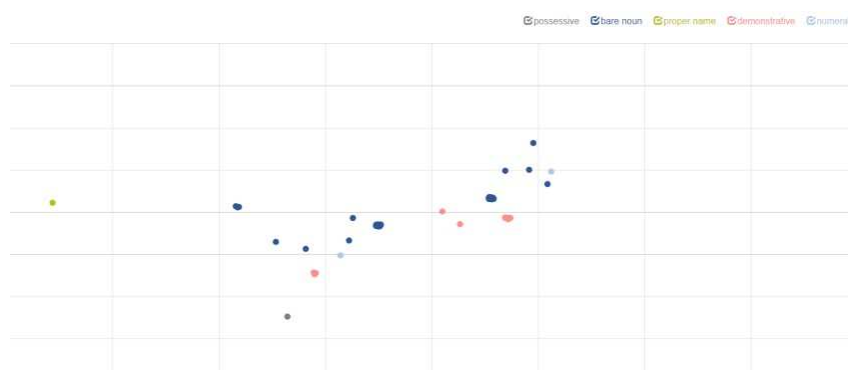
### *English*

The TM map with the color scheme for English was introduced above as Figure 3. The main equivalents of the German definites are the definite article (N=80), the bare singular (N=5) and the demonstrative (N=4). In line with the literature, the comparison with the German TM map shows that the definite article can be found both in weak and strong contexts. The distribution of bare singulars is clearly on the side of the weak/uniqueness definites. This is also in line with the literature that has often hinted at the complementary distribution of bare nouns and weak definites (e.g. Carlson & Sussman 2005; Aguilar-Guevara & Zwarts 2010). The distribution of demonstratives deserves closer scrutiny but the fact that they allow for a deictic and an anaphoric use might explain their appearance in weak and strong contexts.

For English, the contribution of TM is similar to that in German. We do not find new semantic insights. The technique however does allow us to confirm claims in the literature. This reinforces both these claims and the value of the technique.

### *Mandarin*

The TM map with the color scheme for Mandarin is given below as Figure 5. The main equivalents of the German definites are bare nouns (N=79) and demonstratives (N=13).



**Figure 5:** A TM map of definiteness / definiteness / distribution of Mandarin lexical items/constructions

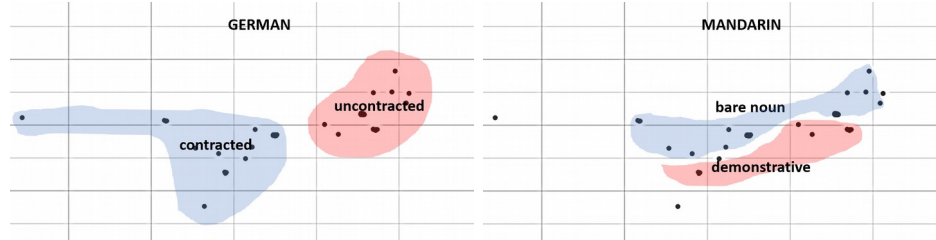
The comparison with the German TM map reveals that the distribution of bare nouns and demonstratives in Mandarin is unexpected on Jenks' analysis. Jenks predicts (i) Mandarin bare nouns to appear in the same contexts as German contracted forms, and (ii) Mandarin demonstratives to appear in the same contexts as German uncontracted forms.

In Figure 6 we compare the TM maps of German and Mandarin side by side. The colors are a visualization of the predictions Jenks makes. Blue covers contexts with contracted

<sup>5</sup>We refer to Bremmers (2019) for an in-depth discussion of the German facts.



forms in German and bare nouns in Mandarin. Red covers contexts with uncontracted forms in German and demonstratives in Mandarin. If Jenks' predictions were borne out, we would have expected the colors to cover the same contexts in the two languages. This is not what we find. Rather, the contracted/uncontracted distinction and the bare noun/demonstrative distinction seem orthogonal to one another.



**Figure 6:** Enhanced TM maps for German and Mandarin

Closer study of the data reveals that the use of demonstratives in contracted contexts (N=3) probably relies on the deictic interpretation of demonstratives and not on their anaphoric use. This is illustrated in (5):

- (5) **M:** Pèinī, wǒ juébù ràng tāmen rènhrén jìn zhè dòng fángzi.  
 Petunia I not have them anyone enter this CL house

(5) is the Mandarin version of example (1) and is uttered by a house owner who assures his wife that certain people will never be welcome in their house. The proximal demonstrative that is used refers to the house the two are in at the moment of speech.

With examples like (5) out of the way, the comparison between German and Mandarin becomes simpler. Demonstratives come out – as predicted by Jenks – as equivalents of German uncontracted forms (N=10). The problem that remains is that bare nouns not only occur as the equivalent of German contracted forms (N=34) but also of uncontracted forms (N=45). An example of a bare noun occurring as the equivalent of a German contracted form is given in (6). This is the Mandarin version of example (4) (see example (4) for the broader context):

- (6) **M:** Tāmen pūshan-zhe chìbǎng gānggāng fēi zǒu, yòu  
 They flutter-ASP wings right fly away, and  
 yǒu yī zhǐ māotóuyīng xié lái yī fēng xìn,  
 have one CL owl bring come one CL letter  
 rēng zài bāoguǒ shàngmiàn.  
 throw to parcel on.

Unlike what we found for German and English, the Mandarin facts do not follow from current wisdom on how differences in form relate to differences in function. What they suggest is that weak/uniqueness definites are indeed conveyed by bare nouns but that strong/familiarity definites are conveyed both by bare nouns and demonstratives. These facts invite linguists to have a closer look at what distinguishes between bare nouns and demonstratives in strong/familiarity contexts.

The conclusion that imposes itself is that TM points to there being more functions in the definiteness domain than previously anticipated. In particular, Mandarin seems to come with

two types of strong/familiarity definiteness, one conveyed by bare nouns, the other conveyed by demonstratives.

## 5 Discussion and conclusion

In this paper, we have introduced Translation Mining as a data-driven way of doing cross-linguistic research. We have compared it to the classical semantic maps method and shown how it can be put to use in the study of definiteness. Our results show that Translation Mining not only allows us to confirm existing intuitions (see our discussion of German and English) but is also able to identify new areas of research (see our discussion of Mandarin).

The division of labor between Mandarin bare nouns and demonstratives in strong/anaphoric definite contexts suggests that there is a dimension of definiteness that has hitherto gone unnoticed. In Bremmers et al. (*ms.*) we argue that this new dimension is a stable one and not an accident of our data collection. We furthermore develop the intuition that bare nouns in anaphoric contexts need to be part of the same narrative sequence as the one their antecedent was introduced in (example (6)).<sup>6</sup> Outside these sequences, a demonstrative is required for anaphoric reference (example (2b)).

We would like to end by noting that it is clear to us that a standard approach to cross-linguistic research would sooner or later also establish a more fine-grained account of the Mandarin data. The value of TM lies in that it allows us to anticipate this result, even with a very restricted application like the one in this paper.

## References

- Aguilar-Guevara, A. et al. 2019. *Definiteness across languages*. Language Science Press.
- Aguilar-Guevara, A. & J. Zwarts. 2010. Weak definites and reference to kinds. In N. Li and D. Lutz (eds.), *Semantics and Linguistic Theory (SALT) 20*, 179-196.
- Beekhuizen, B., J. Watson & S. Stevenson. 2017. Semantic Typology and Parallel Corpora: Something about Indefinite Pronouns. In *CogSci*.
- Bremmers, D. 2019. La définitude en français, anglais, allemand et mandarin. Un published BA thesis, Utrecht University. Available at <https://dspace.library.uu.nl/handle/1874/384039>.
- Bremmers, D., J. Liu, M. van der Klis & B. Le Bruyn. 2019. *Translation Mining: definiteness across languages. A reply to Jenks (2018)*. Manuscript, Utrecht University. Available at <https://time-in-translation.hum.uu.nl/>.
- Carlson, G. & R. Sussman. 2005. Seemingly indefinite definites. *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 85, 71-85.
- Georgakopoulos, T. & S. Polis. 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass*, 12(2).
- Haspelmath, M. 1997. Explaining article-possessor complementarity: economic motivation in noun phrase syntax. *Language*, 227-243.
- Haspelmath, M. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In *The new psychology of language*, 277-322. Psychology Press.
- Heim, I. 1982. *The semantics of definite and indefinite noun phrases*. Doctoral dissertation, University of Massachusetts, Amherst.
- Jenks, P. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49(3), 501- 536.

---

<sup>6</sup>The term narrative sequence goes back to the term *narration* in SDRT (e.g. Lascarides & Asher 1993) and involves sequences of (chronologically ordered) events.

- Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof (eds.) *Formal Methods in the Study of Language*, 277-322. Mathematical Center Tract 135, Amsterdam.
- Klis, M. van der, B. Le Bruyn & H. de Swart. 2017. Mapping the perfect via translation mining. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 497-502.
- Klis, M. van der, B. Le Bruyn & H. de Swart. 2019. De la sémantique des temps verbaux à la traductologie: une comparaison multilingue de *L'Étranger* de Camus. In E. Corre, D.-T. Do-Hurinville, and H.-L. Dao (eds.) *Linguistic approaches to Tense, Aspect, Modality, Evidentiality, based on the Novel L'Étranger ("The Stranger") by Albert Camus, and its Translations*.
- Lascarides, A. & N. Asher. 1993. Temporal interpretation, discourse relations, and common sense Entailment. *Linguistics and Philosophy* 16, 437-493.
- Löbner, S. 2011. Concept types and determination. *Journal of semantics*, 28(3), 279-333.
- Russell, B. 1905. On denoting. *Mind* 14(56), 479-493.
- Schwarz, F. 2009. *Two Types of Definites in Natural Language*. Doctoral dissertation, University of Massachusetts, Amherst.
- Strawson, P. 1950. On referring. *Mind* 59(235), 320-344.
- Van der Auwera, J. & V. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2(1), 79-124.
- Wälchli, Bernhard & Michael Cysouw. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics* 50(3), 671-710.

# Keeping *Dou* as a Simple Distributor\*

Zhuo Chen

The Graduate Center, CUNY, New York, NY, USA  
zchen3@gradcenter.cuny.edu

## Abstract

Against the *even* analysis of *dou*, this paper argues that a simple distributor semantics of *dou* can be maintained with two assumptions: (i) Mandarin sentences allow an optionally overt topic set [12]; (ii) scalar *dou*-sentences have an optionally overt preposition *lian* [1, 15].

## 1 Introduction

The multi-functional particle *dou* has drawn immense attention from Chinese linguists. The two main uses exemplified in (1) and (2), with *dou* as a distributor and a focus operator *even*, respectively, raise an imminent question — what is the semantic core of *dou* that conditions its distribution?

- (1) Tamen dou mai le yi liang che<sup>1</sup>      (2) Zhangsan<sub>F</sub> dou xihuan Lisi  
they DOU buy Asp one CL car      Zhangsan DOU like Lisi  
They each bought a car.      Even Zhangsan likes Lisi.

Linguists take two different routes to answer this question. The first route is to treat the distributive use as the core function of *dou* that subsumes the *even* reading as a special case [12]. To the contrary of the first, the second route is to analyze *dou* as *even* while attributing the distributive reading to a covert *dist* operator [11]. In this paper, I argue that analyses along the second route are on the wrong track, evidenced by a set of incorrect predictions made by the *even* analysis of *dou*. I show that a simple distributor semantics for *dou* can be maintained with a satisfactory account of its distribution in (1) and (2), if we make two assumptions. First, as a topic prominent language, Mandarin allows an optionally overt topic set in a sentence. Second, the *even*-sentences with *dou* have an optionally present preposition *lian* [1, 15] that makes non-trivial contribution to the scalar reading.

## 2 Flaws in the *even* analyses of *dou*

Mandarin Chinese *wh+dou* resembles the *indefinite+even* sequences found cross-linguistically. With a clause-mate negation, it gives rise to an NPI reading; with a generic predicate, it has a universal reading. The parallel is illustrated here with a comparison between Mandarin and Hindi<sup>2</sup>, as in (3) - (4).

---

\*I thank Sam Alxatib, Anna Szabolcsi and the three reviewers of Amsterdam Colloquium for their helpful comments. For native English judgement, I thank Sarah Cocozzello D'Arienzo, Andre Eliatamby, Paul Feitzinger, Yulia Gorokhovskiy, and Cameron Wilson. The remaining errors are mine.

<sup>1</sup>As [10] points out, *dou* does not require to distribute down to atoms. I use *each* here for simplicity.

<sup>2</sup>All Hindi examples are from [7], glosses his. The same behavior of *indefinite+even* sequence is also found in Japanese, Korean, Bangla, Malayalam and many more languages. Due to space constraint, I only present Hindi examples in this paper for illustration.

- (3) a. shei dou mei lai                      b. shei dou keyi taiqi zhe zhang zhuozi  
       who DOU NEG come                    who DOU can lift this CL table  
       No one came.                         Anyone can lift this table.                (Chinese)
- (4) a. koi bhii nahiiN aayaa                b. tum to kuch bhii kah dete ho  
       someone even not come                you PRT something even say  
       No one came.                         You say anything.                         (Hindi)
- (5) a. yi ge ren dou mei lai                b. yi ge ren dou keyi taiqi zhe zhang zhuozi  
       one CL person DOU NEG come        one CL p DOU can lift this CL table  
       No one came.                         Even one person can lift this table.     (Chinese)
- (6) a. ek bhii aadmii nahiiN aayaa        b. ek bhii cingaarii ghar-ko jalaa detii hai  
       one even not come                    one even spark house burns  
       No one came.                         Even one spark burns the house.        (Hindi)

The same alternation between an NPI and a universal reading is also exhibited by minimizers *one NP+dou* and *one NP+even* in Mandarin and Hindi, respectively, in (5)-(6). Given the similarities between Hindi *indefinite+even* and *one NP+even*, Lahiri [7] proposes to reduce the former to the latter. Specifically, an indefinite like ‘someone’ introduces a variable satisfying the cardinality predicate **one** and the NP restriction, i.e.  $one(x) \wedge person(x)$ . In (4a), an existential closure closes off the the variable introduced by *koi*. The negation turns the existential statement into a negative one, serving as the prejacent of the focus operator *bhii*. *Bhii* activates the alternatives of its associate *koi*, each with **one** replaced with a different cardinality predicate *two*, *three*, etc. Moreover, *bhii* imposes two implicatures. The additive implicature says there is at least one true alternative distinct from the prejacent. The least likelihood implicature says the prejacent is the least likely one among its alternatives. The formal representation of the prejacent, the alternative set and the two implicatures are given in (7). Since the prejacent logically entails all other alternatives, the two implicatures are satisfied as long as the prejacent is true.

- (7) a. Assertion:  $a = \neg \exists x[\mathbf{one}(x) \wedge came(x)]$   
       b. ALT =  $\{\neg \exists x[\mathbf{P}(x) \wedge came(x)] \mid \mathbf{P} \in \{two, three, four...\}\}$   
       c. Additive implicature:  $\exists q \in ALT[q \neq a \wedge q]$   
       d. Least likelihood implicature:  $\forall q \in ALT[q \neq a \rightarrow a <_{likelihood} q]$

In a generic sentence (4b), Lahiri gives the same kind of semantics to *kuch bhii*. The generic operator binds the variable introduced by *kuch*, which locates in the restrictor of the universal generic quantifier, a downward-entailing environment. The prejacent, again, logically entails all the other alternatives. The two implicatures are thus satisfied as long as the prejacent is true.

Lee and Horn [8] have very similar ideas as Lahiri, who argue that English NPI *any NP* is underlyingly *one NP+even* in negative sentences. In generic sentences, however, they argue that *any* is a superlative plus *even*, motivated by an observation made by Fauconnier [3] that (8) and (9) mean the same.

- (8) Alfred will eat any food.  
 (9) Alfred will eat the most awful food.

Given the resemblance between Chinese *wh+dou* and Hindi *indefinite+even*, it is tempting to carry over the aforementioned two theories to the Chinese data. Despite the temptation, however, I will show that their theories make incorrect predictions on the Chinese data.

Another recent *even* theory of *dou* is proposed by Liu [11], who assumes the same semantics for *dou* as that given by Karttunen and Peters [5] to English *even*, (10). The *even*-sentences with *dou* are thus captured without ado.

$$(10) \llbracket dou \rrbracket = \lambda p : \forall q [q \in \text{ALT}(p) \rightarrow q <_{\text{likely}} p].p$$

As for a distributive sentence like (1), Liu assumes a covert *dist* operator [14] is present. The alternatives are formed by replacing the plurality associated with *dou* with its subparts. The prejacent logically entails all other alternatives, as shown in Figure 1, making the least likelihood automatically satisfied, wherefore the *even* flavor of *dou* is not sensed. I will show that this theory also makes incorrect predictions.

## 2.1 Problems with Lahiri's theory

Reducing *wh+dou* to *one NP+dou* in Chinese cannot explain their different grammatical status in (11) or (12) as an NPI. When Lisi forgot to count the cardinality of the students, (11b) is a grammatical complaint but (11a) is not. When Lisi held his debut in his hometown instead of any big city, (12a) is a grammatical description but (12b) is not.

- (11) a. \*Lisi shei dou mei shu                      b. Lisi yi ge ren dou mei shu  
       Lisi who DOU NEG count                      Lisi one CL person DOU NEG count  
       Lisi didn't count anyone.                      Lisi didn't count even one person.
- (12) a. Lisi shouyan na ge dachengshi dou mei ban. Ta laojia ban de  
       Lisi debut which CL big city DOU NEG hold. He hometown hold SFP  
       Lisi didn't hold his debut in any big city. He held it in his hometown.
- b. \*Lisi shouyan yi ge dachengshi dou mei ban...  
       Lisi debut one CL big city DOU NEG hold  
       Lisi didn't hold his debut in even one big city...

Moreover, *wh+dou* and *one NP+dou* behave differently in their universal uses. When discussing who can be the President, (13a) is a grammatical statement but (13b) is not. When evaluating the weight of a table, (14b) is a grammatical report but (14a) is not.

- (13) a. shei dou keyi dang zongtong                      b. \*Yi ge ren dou keyi dang zongtong  
       who DOU can be president                      one CL person DOU can be president  
       Anyone can be the President.                      Even one person can be the President.
- (14) a. \*shei dou zugou taiqi zhe zhang zhuozi  
       who DOU sufficient lift this CL table  
       Anyone is sufficient to lift this table.
- b. yi ge ren dou zugou taiqi zhe zhang zhuozi  
       one CL person DOU sufficient lift this CL table  
       Even one person is sufficient to lift this table.

Lahiri's theory that treats *wh+dou* and *one NP+dou* as semantically equivalent fails to explain their different distribution in the above examples.

## 2.2 Problems with Lee and Horn's theory

Lee and Horn [8] account for the NPI use of English *any* in the same spirit as Lahiri [7], analyzing *any* as underlyingly *indefinite+even* argued to be semantically equivalent to the minimizer *one NP+even*. The same problem plaguing Lahiri's theory on the NPI use thus carries over. As for the generic use, Lee and Horn take *any* as a superlative plus *even*, a theory that has its own problem.

- (15) a. Any set is a subset of itself.      b. \*Even the biggest set is a subset of itself.  
 (16) a. shei dou you shengmu  
       who DOU have biological mom  
       Anyone has a biological mother.  
       b. \*zuilaode ren dou you shengmu  
       oldest pers. DOU have biological mom  
       Even the oldest one has a biological mother.

(15a) states an axiom in Mathematics. The satisfaction of the predicate only has to do with the property of being a set, without activating any kind of scalar relationship between the sets. Actually, an attempt to impose such a scale will fail, (15b). Likewise, the predicate ‘has a biological mother’ only concerns the NP restriction *person* of *who* in (16) without reference to any scale. Hence the contrast between (16a) and (16b).

### 2.3 Problem with Liu’s theory

Liu [11] ascribes the absence of the *even* flavor in distributive *dou*-sentences to the automatic satisfaction of the least likelihood presupposition imposed by *dou*. This analysis predicts the *even* flavor of *dou* to be subdued whenever the prejacent logically entails all other alternatives. While it seems to be true in the distributive (1), the prediction is not borne out in a collective sentence (17), even if the latter exhibits the same logical entailment relations between the prejacent and the alternatives as the former, see Figure 2<sup>3</sup>.

- (17) Yuehan, Mali he Bi'er yiqi dou keyi jin zhe ge hezi, geng bu yong  
 John Mary and Bill together DOU can squeeze into this CL box, more NEG need  
 shuo Yuehan he Mali liang ge ren le  
 say John and Mary two CL people SFP  
 EVEN J, M and B together can squeeze into the box. Let alone the two of J and M.

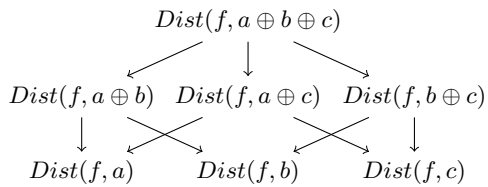


Figure 1: Distributive (1)

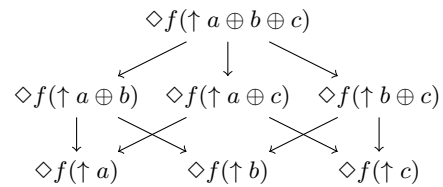


Figure 2: Collective (17)

### 2.4 Interim conclusion

Analyzing *dou* as *even* obligates a scalar relationship between the alternatives triggered by the focus associate of *dou*. This requirement, however, clashes with the unordered set denoted by a singular *wh*-indefinite in *wh*+*dou* constructions. In the NPI use, this clash is obscured when the predicate is distributive (3a), but is foregrounded when the predicate is collective relative to the indefinite (11a). In the universal use, this clash is also veiled by a distributive predicate (3b), but is spotlighted when the predicate comes with an obligatory scalar requirement on the argument indefinite (14b). Reducing *wh*-indefinite+*dou* to a scalar phrase+*dou*, be it a minimizer or a superlative, therefore, cannot explain the drastic differences in their compatibility/incompatibility with the same predicate.

<sup>3</sup> *f* in Figure 1 and Figure 2 refers to the predicate in the respective sentences.

The distinction between Chinese *wh+dou* and *one NP/superlative+dou* is also corroborated by an old observation in English by Heim [4], who classified NPIs into ones with an inherent *even* and ones without. The former type includes *any/ever* and the latter includes minimizers *one NP+even*. Heim shows that minimizers are only grammatical in sentences that facilitate the satisfaction of their scalar implicatures, but *any* does not have such requirement. Hence the contrast between (18a) and (18b).

- (18) a. ??Every restaurant that charges so much as a dime for iceberg lettuce happens to have four stars in the handbook.  
 b. Every restaurant that advertises in any of these papers happens to have four stars in the handbook.

Since *wh+dou* behaves in line with English *any* in its NPI and universal use, the conclusion drawn by Heim [4] casts further doubt on the *even* analyses of *dou*. In light of all the incorrect predictions made by the current *even* theories of *dou* and the corroboration from English data, I argue that the *even* analyses of *dou* are on the wrong track.

### 3 *Dou* as a simple distributor

I propose that *dou* can be treated as a simple distributor once we make two assumptions, both having been argued for elsewhere by other linguists. First, Mandarin Chinese allows a covert topic set present in sentences [12]. Second, the *even*-sentences with *dou* always have a semantically active preposition *lian*, overt or covert [15]. The semantic entry for *dou* in this proposal is as in (19).

$$(19) \llbracket dou \rrbracket^g = \lambda P_{\tau t} \lambda Q_{\tau t} . \forall x [Q(x) \rightarrow P(x)]^4$$

#### 3.1 The distributive sentences with *dou*

Chinese subjects are moved to the topic position by default [9], shown by the felicitous insertion of topic markers between the subject and the predicate (20). Moreover, with contextual support, it's free to drop the topic (21), wherefore the semantic equivalence of the following two sentences in the given context.

*Context: You ask John whether the students have come for class. John answers:*

- |  |   |
|--|---|
| <p>(20) Xueshengmen a/na dou lai le<br/>         student TOP DOU come SFP<br/>         Students have all come.</p> | <p>(21) dou lai le<br/>         DOU come SFP<br/>         Students have all come.</p> |
|--|---|

In a simple distributive sentence like (1), *dou* distributes over the subject topic *they*. Suppose *they* denotes the set of John and Mary. We get the truth condition in (22) accordingly. When the topic set is covert whose value is determined by the context like (21), *dou* does the same job. Suppose the covert topic has an index *j*, we get the truth condition in (23) accordingly.

$$(22) \llbracket (1) \rrbracket^g = 1 \text{ iff } \forall x [x \in \{j, m\} \rightarrow \exists y [car(y) \wedge bgt(y)(x)]]$$

$$(23) \llbracket (21) \rrbracket^g = 1 \text{ iff } \forall x [x \in g(j) \rightarrow come(x)]$$

<sup>4</sup>As pointed out before, *dou* does not require to distribute down to atoms [10]. Here, I give this entry to simplify the illustration. A more rigorous entry is as in (i), where *C* is a cover on the plurality associates with *dou* in the sense of Schwarzschild [14].

i  $\llbracket dou \rrbracket^g = \lambda P_{\tau t} \lambda Q_{\tau t} . \forall x [Q(x) \wedge C(x) \rightarrow P(x)]$



The composition of distributive sentences with *dou* is quite straightforward. We now turn to the scalar *even* sentences with *dou*.

### 3.2 The *even* sentences with *dou*

In an *even* sentence with *dou*, the topic set is again optionally overt. Moreover, a preposition *lian* literally meaning ‘including’ that precedes the focus associate is optionally present as well. As a result, the following four sentences all mean the same in the given context.

*Context: John thinks Mary is the most popular student in their class because Bill, the aloofest student, likes her.*

- |  |  |
|--|--|
| <p>(24) tamenban lian Bi'er<sub>F</sub> dou xihuan Mali<br/>they class LIAN Bill DOU like Mary<br/>(In their class), even Bill likes Mary.</p> | <p>(25) tamenban Bi'er<sub>F</sub> dou xihuan Mali<br/>they class Bill DOU like Mary<br/>(In their class), even Bill likes Mary.</p> |
| <p>(26) lian Bi'er<sub>F</sub> dou xihuan Mali<br/>LIAN Bill DOU like Mary<br/>(In their class), even Bill likes Mary.</p>                     | <p>(27) Bi'er<sub>F</sub> dou xihuan Mali<br/>Bill DOU like Mary<br/>(In their class), even Bill likes Mary.</p>                     |

Since *dou* is not *even*, as we have concluded, the semantic contribution of *even* is naturally shifted to *lian*<sup>5</sup>. I propose a semantic entry for *lian* as in (28). Specifically, its semantics is broken down to three ingredients. I name them likelihood presupposition, membership condition and property condition respectively.

$$(28) \llbracket \textit{lian} \rrbracket^w = \lambda x \lambda Y \lambda P : \underbrace{\forall z [(z \in (\text{F-ALT}(x) \cap Y) \wedge z \neq x) \rightarrow \lambda w'. P_{w'}(Y \setminus \{x\}) >_{\textit{likely}} \lambda w'. P_{w'}(Y \setminus \{z\})]}_{\text{likelihood presupposition}} \\ \underbrace{x \in Y}_{\text{membership condition}} \wedge \underbrace{P_w(Y)}_{\text{property condition}}$$

Take (24), whose logical form is given in Figure 3 for example. Suppose ‘their class’, the topic set, denotes the set of students  $\{\textit{john}, \textit{mary}, \textit{bill}, \textit{sue}\}$ . The focus alternatives  $\text{F-ALT}(\textit{John})$  is the set of all elements of the same type as John, i.e.  $D_e$  [13]. The topic set confines the quantificational domain of *lian* to the members in it only.

The composition of (24) is shown in (29).

$$(29) \text{ a. } \llbracket \textcircled{3} \rrbracket^{w,g} = \lambda Q. \forall x [x \in Q \rightarrow \textit{like}(w)(m)(x)] \\ \text{ b. } \llbracket \textcircled{2} \rrbracket^{w,g} = \lambda P : \underbrace{\forall z [(z \in \{j, m, s\}) \rightarrow \lambda w'. P_{w'}(\{j, m, s\}) >_{\textit{likely}} \lambda w'. P_{w'}(\{j, m, b, s\} \setminus \{z\})]}_{\text{likelihood presupposition}} \\ \underbrace{b \in \{j, m, b, s\}}_{\text{membership condition}} \wedge \underbrace{P_w(\{j, m, b, s\})}_{\text{property condition}}$$

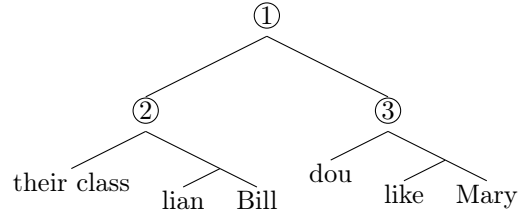


Figure 3: LF of (24)

<sup>5</sup>A reviewer has brought to my attention a recent thesis by Zhao [16], who also argues to shift the *even* contribution to *lian*. A comparison between the current theory and his analysis will have to be left for another occasion.

$$\begin{aligned}
c. \quad & \llbracket \textcircled{1} \rrbracket^{w,g} = 1 \text{ iff} \\
& \underbrace{\forall z[(z \in \{j, m, s\}) \rightarrow [\forall y \in \{j, m, s\}[\textit{like}(y)]] >_{\textit{likely}} [\forall y \in \{j, m, b, s\} \setminus \{z\}[\textit{like}(m)(y)]]]}_{\text{likelihood presupposition}} \\
& \quad \underbrace{b \in \{j, m, b, s\}}_{\text{membership condition}} \wedge \underbrace{\forall y \in \{j, m, b, s\}[\textit{like}_w(m)(y)]}_{\text{property condition}}
\end{aligned}$$

The membership condition says Bill is a member of the topic set of students. The property condition says *dou*-VP holds of the topic set, i.e. every student in the topic set likes Mary. We deduce from the two conditions that Bill likes Mary. The likelihood presupposition says for every other student  $z$  distinct from Bill in the topic set, the likelihood that *dou*-VP holds of the topic set with Bill subtracted from the set is higher than the likelihood that *dou*-VP holds of the topic set with  $z$  subtracted from it. If  $z$  is Sue, the likelihood presupposition is reduced to (30), i.e. the likelihood that Sue likes Mary is greater than the likelihood that Bill likes Mary. Iterating through all the other students will lead to the conclusion that Bill is the least likely one among the students to like Mary.

$$\begin{aligned}
(30) \quad & [\forall y \in \{j, m, b, s\} \setminus \{b\}[\textit{like}(m)(y)]] >_{\textit{likely}} [\forall y \in \{j, m, b, s\} \setminus \{s\}[\textit{like}(m)(y)]] \\
& = [\textit{like}_{w'}(m)(j) \wedge \textit{like}_{w'}(m)(m) \wedge \textit{like}_{w'}(m)(s)] >_{\textit{likely}} \\
& \quad [\textit{like}_{w'}(m)(j) \wedge \textit{like}_{w'}(m)(m) \wedge \textit{like}_{w'}(m)(b)] \\
& = p_j \times p_m \times p_s > p_j \times p_m \times p_b \\
& = p_s > p_b \qquad (p_x \text{ stands for the likelihood that } x \text{ likes Mary})
\end{aligned}$$

The proposed semantics for *lian* can easily explain the *even* sentences with *dou* when the focus associate is a scalar term. In (31), the topic set is the set of time instants that have arrived, a set with the utterance time  $u$  as the ending element  $\{\dots u - \epsilon, \dots u\}$ . A covert *lian* is assumed to be present, giving the logical form as in (32). The truth condition derived for this sentence is in (33).

$$\begin{aligned}
(31) \quad & \text{wudian}_F \text{ dou dao le, zenme hai mei hao} \\
& 5 \text{ o'clock DOU arrive Asp, how come still NEG good} \\
& \text{Even 5 o'clock has arrived. How come you are still not done?}
\end{aligned}$$

$$(32) \quad \text{LF: } [_{\text{IP}}[_{\text{DP}}\{\dots u - \epsilon, \dots u\}] [_{\text{PP}}\textit{lian } 5]] [\textit{dou } [_{\text{VP}} \textit{has arrived}]]$$

$$\begin{aligned}
(33) \quad & \llbracket \textcircled{31} \rrbracket^{w,g} = 1 \text{ iff} \\
& \underbrace{\forall z[(z \in \{\dots, u\} \setminus \{5\}) \rightarrow [\forall y \in \{\dots, u\} \setminus \{5\}[\hat{\textit{arr}}(y)]] >_{\textit{likely}} [\forall y \in \{\dots, u\} \setminus \{z\}[\hat{\textit{arr}}(y)]]]}_{\text{likelihood presupposition}} \\
& \quad \underbrace{5 \in \{\dots, u\}}_{\text{membership condition}} \wedge \underbrace{\forall y \in \{\dots, u\}[\hat{\textit{arr}}_w(y)]}_{\text{property condition}}
\end{aligned}$$

The membership condition says 5 o'clock is at least as early as the utterance time. The property condition says the utterance time and all the instants preceding it have arrived. From these two conditions, we know that 5 o'clock has arrived. The likelihood presupposition requires the likelihood of the instants in the topic set with 5 o'clock subtracted having all arrived be higher than the likelihood of the instants in the topic set with any non-5 o'clock instant subtracted having all arrived. This presupposition can only be satisfied when 5 o'clock is itself the utterance time. If 5 is not the utterance time, the likelihood that the instants in the topic set with 5 subtracted have all arrived is the same as the likelihood that the utterance time has arrived (due to the fact that for an instant to have arrived means for this instant and every instant preceding it to have arrived). As a result, for any instant  $t$  in the topic set that is distinct from 5, the likelihood presupposition is reduced to (34). If  $t$  is not the utterance time,

the likelihood on either side will be the same, both being the likelihood of the utterance time having arrived. If  $t$  is the utterance time, the likelihood on the left will be smaller, not greater than the likelihood on the right, as the utterance time is later than any instant in the topic set with itself subtracted. In other words, when 5 is not the utterance time, the likelihood presupposition cannot be satisfied for any instant  $t$  distinct from 5, be it the utterance time or not.

$$(34) \quad \lambda w'.arr_{w'}(u) >_{likely} [\forall y \in \{\dots, u\} \setminus \{t\} [\lambda w'.arr_{w'}(y)]]$$

If 5 is the utterance time, however,  $t$  has to be an instant preceding 5. The likelihood presupposition for any  $t$  is accordingly reduced to (35). A later instant having arrived entails any earlier instant having arrived, thus is less likely. Since the utterance time 5 o'clock is later than any instant in the topic set with 5 subtracted, the presupposition is satisfied in this case.

$$(35) \quad [\forall y \in \{\dots, 5\} \setminus \{5\} [\lambda w'.arr_{w'}(y)]] >_{likely} [\lambda w'.arr_{w'}(5)]$$

I have shown that we can keep *dou* as a simple distributor with the semantics in (19), as long as two assumptions are adopted. First, Mandarin Chinese allows a covert topic set. Second, in *even* sentences with *dou*, a focus operator *lian* is always present, covertly or overtly. These two assumptions have been argued for in previous literature and are not concocted specifically for the current proposal on *dou*.

## 4 Discussion

Taking *dou* as a distributor in (19) necessarily implements a universal quantification over the topic set. Recall that in (24), the property condition imposed by *lian* requires that every student in the class like Mary. Two reviewers wonder whether the universal quantification is overly strong, considering especially the fact that linguists [2, 5, 6] have encoded the quantification in the presupposition rather than the assertion of English *even*. Moreover, a sentence like (36) seems to be felicitous in a context where the speaker expresses her surprise that other people who are more punctual than John are late.

(36) Even John has arrived. How come the other people are still not here?

I do not have a satisfactory answer to this question yet and leave it for future research<sup>6</sup>.

Another reviewer raises the question of how to account for a free choice sentence with *dou* (37) using the current proposal. While a complete discussion of the free choice *wh+dou* construction in Chinese is beyond this paper, I give the main idea of how the proposed theory will explain (37) here.

(37) Yuehan huozhe Mali dou keyi dang zongtong  
 John or Mary DOU can be president  
 John can be the president and Mary can be the president.

The disjunctive marker *or* forms a set out of its flanking arguments. The subject topic in (37), therefore, is a set of John and Mary  $\{j, m\}$ . The VP that *dou* takes as its first input is a modal one — ‘can be the president’, which is interpreted as a set of individuals, with each individual being the president in an epistemically accessible world, formally represented in (38). *Dou* then claims that each member in the subject topic set is in the set denoted by the VP, giving a truth condition as in (39).

<sup>6</sup>Rooth [13] treats the quantification of the alternatives as part of the conventional implicature of *even*, similar to my stance in this paper. However, he argues for an existential quantification rather than a universal one.

(38)  $\lambda x. \exists w'[w' \text{ is epistemically accessible to } w_{\text{@}} \wedge \textit{President}(x) \text{ in } w']$

(39)  $\llbracket (37) \rrbracket = 1$  iff  
 $\forall x[x \in \{j, m\} \rightarrow \exists w'[w' \text{ is epistemically accessible to } w_{\text{@}} \wedge \textit{President}(x) \text{ in } w']]$

Still another question from the reviewers is why a negation can be inserted before *dou* in a simple distributive sentence (40), but not in an *even* sentence (41). The answer has two parts. First of all, the negation is a focal negation locating syntactically higher in the structure with an optional focal marker *shi* after it. Second, the negation follows the topic set. In (40), the topic set is overt but in (41) it is covert. An analogous *even* sentence to be compared to (40) should be the grammatical (42) rather than the ungrammatical (41).

(40) Tamenban bu (shi) dou xihuan Mali (41) \*Lian Yuehan bu dou xihuan Mali  
 they class NEG FOC DOU like Mary LIAN John NEG DOU like Mary  
 Their class don't all like Mary. It's not that even John likes Mary.

(42) (tamenban) bu (shi) lian Yuehan dou xihuan Mali  
 their class NEG FOC LIAN John DOU like Mary  
 Among their class, it is not the case that even John likes Mary.

## 5 Conclusion

This paper discusses the multi-functional particle *dou* in Mandarin Chinese. Specifically, I argue to keep *dou* as a simple distributor. The recent *even* theories of *dou* have been shown to make incorrect predictions on the (un)grammaticality of the NPI and universal use of *wh+dou* constructions with certain predicates as well as on the presence/absence of the *even* flavor of a *dou*-sentence. Adopting two assumptions that have been argued for in Mandarin — Mandarin Chinese allows a covert topic set and *even*-sentences in Mandarin allow an optionally overt focus operator *lian* — I propose a simple distributor semantics of *dou* and shift the labor of *even* to the focus operator *lian*. The proposal is able to capture the distributive and *even* sentences with *dou*. The question of whether it is too strong to make a universal assertion in the *even* sentences with *dou* that the predicate holds of all the members in the topic set is left for future research.

## References

- [1] L. Badan. The *even*-construction in Mandarin Chinese. In *Chinese Linguistics in Leipzig*. EHESS-CRLAO, 2008.
- [2] J. Bennett. Even if. *Linguistics and Philosophy*, 5(3):403–418, 1982.
- [3] G. Fauconnier. Pragmatic scales and logical structure. *Linguistic Inquiry*, 6(3):353–375, 1975.
- [4] I. Heim. A note on negative polarity and downward entailingness. In *Proceedings of NELS*, volume 14, pages 98–107, 1984.
- [5] L. Karttunen and S. Peters. Conventional implicature. *Syntax and Semantics*, 11:1–56, 1979.
- [6] P. Kay. Even. *Linguistics and Philosophy*, 13(1):59–111, 1990.

- [7] U. Lahiri. Focus and negative polarity in Hindi. *Natural Language Semantics*, 6(1):57–123, 1998.
- [8] Y.-S. Lee and L. Horn. Any as indefinite plus even. *Manuscript, Yale University*, 1995.
- [9] C. Li and S. Thompson. *A Functional Reference Grammar of Mandarin Chinese*. Berkeley, CA: University of California Press., 1981.
- [10] J.-W. Lin. Distributivity in Chinese and its implications. *Natural Language Semantics*, 6(2):201–243, 1998b.
- [11] M. Liu. Varieties of alternatives: Mandarin focus particles. *Linguistics and Philosophy*, 40(1):61–95, 2017.
- [12] P. Portner. Topicality and (non-) specificity in Mandarin. *Journal of Semantics*, 19(3): 275–287, 2002.
- [13] M. Rooth. *Association with Focus*. PhD thesis, University of Massachusetts, Amherst, MA, 1985.
- [14] R. Schwarzschild. *Pluralities*, volume 61. Springer Science & Business Media, 1996.
- [15] M. Xiang. Plurality, maximality and scalar inferences: A case study of Mandarin *dou*. *Journal of East Asian Linguistics*, 17(3):227–245, 2008.
- [16] Z. Zhao. Varieties of Distributivity: From Mandarin *dou* to Plurality, Free Choice and Scalarity. Master’s thesis, University of Amsterdam, 2019.

# Attitude Semantics\*

Ivano Ciardelli

Munich Center for Mathematical Philosophy, LMU Munich, Germany  
ivano.ciardelli@lrz.uni-muenchen.de

## Abstract

Many recent theories treat indicative conditionals as restricted necessity modals. I discuss two problems for this view. First, indicative conditionals do not behave like necessity modals in embedded contexts, e.g., under ‘might’ and ‘probably’: in these contexts, conditionals do not contribute a universal quantification over epistemic possibilities. Second, when we assess the probability of a conditional, we do not assess how likely it is that the consequent is *necessary* given the antecedent, but how likely it is that it is *true* given the antecedent. I propose an account which predicts the embedding behavior of conditionals under modals and the way we assign probabilities to conditionals. The account is based on the idea that the semantics of conditionals involves only a restriction of the relevant epistemic state, and no quantification over epistemic possibilities. The relevant quantification is contributed by an attitude parameter in the semantics, which is shifted by epistemic modals. If the conditional is asserted, the designated attitude is acceptance, which contributes a universal quantifier, producing the overall effect of a restricted necessity modal.

## 1 Introduction

**The Box View.** Many recent theories of indicative conditionals analyze them as restricted epistemic necessity modals. This family includes dynamic semantics accounts (Gillies, 2004, 2009; Willer, 2014, 2018; Starr, 2014) and other semi-dynamic information-based accounts (Yalcin, 2007; Gillies, 2009, 2010; Bledin, 2014).<sup>1</sup> The core idea is the following:<sup>2</sup>

*Box View.* The semantics of an indicative conditional  $p \Rightarrow q$  involves two components:

1. Restriction: restrict the set of epistemic possibilities to the  $p$ -worlds;
2. Quantification: check that  $q$  is true at every world in the restricted set.

The conditional is true/accepted iff the check in the second step is successful.<sup>3</sup>

This view is motivated by an intuitively compelling analysis of how indicative conditionals are assessed, known as the *Ramsey test view* (after Ramsey (1929)). The idea is that, in order to assess  $p \Rightarrow q$  in an information state  $s$ , we proceed in two steps: first, we add  $p$  to  $s$ , resulting in a hypothetical state  $s[p]$ ; second, we check whether we accept the consequent  $q$  in this hypothetical state; if so (and only in this case) we accept the conditional.

---

\*Thanks to Maria Aloni, Paul Egré, Peter Hawke, Luca Incurvati, Hannes Leitgeb, Matt Mandelkern, Floris Roelofsen, Katrin Schulz, Shane Steinert-Threlkeld and Frank Veltman for discussion of these ideas.

<sup>1</sup>The restrictor account of Kratzer (1986) also fits broadly within this line, although due to its specific assumptions about the syntax of conditionals it needs to be discussed separately. We will do so in Section 4.

<sup>2</sup>In this paper I focus on indicative conditionals, although the puzzles that I will raise concern subjunctive conditionals as well. Proposal extends straightforwardly to the subjunctive case. However, making subjunctive assumptions involves a different process than making indicative assumptions; spelling out the details of this process is orthogonal to our present concerns; thus, I leave subjunctive conditionals out of consideration here.

<sup>3</sup>In some of the theories mentioned above, the view is implemented as giving truth-conditions for conditionals relative to a world with an associated set of epistemic possibilities; in other theories, conditionals lack truth-conditions, and the semantics delivers acceptance conditions relative to an information state.

In possible world semantics, an information state  $s$  is typically modeled as a set of worlds—those worlds which are possible according to the available information. Adding  $p$  to  $s$  corresponds to restricting to  $s$  to the  $p$ -worlds, obtaining the state  $s[p] = s \cap |p|$ . Checking whether we accept  $q$  in the resulting hypothetical state amounts to checking whether  $q$  is true in all the worlds in  $s[p]$ . Thus, the two steps of the Ramsey test procedure correspond exactly to the two components of the semantics of conditionals as postulated by the Box View.

The Box View of conditionals also makes several welcome empirical predictions. For instance, it explains why the discourse in (1) sounds contradictory:

- (1) If Alice left, she went to London; #but it might be that she left and went to Paris.

In a discourse of the form  $p \Rightarrow q; \diamond(p \wedge \neg q)$ , the conditional requires all  $p$ -possibilities to be  $q$ -possibilities, while the might-continuation requires the existence of  $p \wedge \neg q$ -possibilities. In spite of these attractions, however, the Box View also faces some important problems.

**Problem 1: Embedding.** Consider a conditional in the scope of an epistemic ‘might’:

- (2) It might be that if Alice left she went to London.  $\diamond(p \Rightarrow q)$

Intuitively, (2) is a conditional possibility claim: it says that, among the epistemic possibilities where Alice left, there are some where she went to London. This is not what the Box View would lead us to expect. According to this view, (2) should be a second-order epistemic statement: it is possible that, relative to the worlds where Alice left, it is necessary that she went to London. This does not seem right. Moreover, (2) seems to mean just the same as (3).

- (3) If Alice left, it might be that she went to London.  $p \Rightarrow \diamond q$

This, too, is puzzling from the point of view of the Box View: if conditionals contribute a universal quantification, then  $\diamond(p \Rightarrow q)$  should correspond to a  $\exists\forall$  statement, while  $p \Rightarrow \diamond q$  should correspond to a  $\forall\exists$  statement, so it is hard to see how the two could be equivalent. Indeed, in all theories cited above, the equivalence  $\diamond(p \Rightarrow q) \equiv p \Rightarrow \diamond q$  is not predicted.<sup>4</sup>

This phenomenon is not restricted to epistemic ‘might’, but concerns epistemic modals quite generally. In particular, consider the case of ‘probably’.

- (4) It is probable that if Alice left she went to London.  $\text{Pr}(p \Rightarrow q)$

Again, what (4) expresses is not that it is probable that Alice going to London is epistemically necessary on the supposition that she left; rather, what it expresses is that Alice going to London is probable on the supposition that she left. This reading does not involve any epistemic necessity. Moreover, in this case too, (4) sounds fully equivalent to (5):

- (5) If Alice left, it is probable that she went to London.  $p \Rightarrow \text{Pr } q$

It is hard to see how this commutation can hold if  $\Rightarrow$  introduces a universal quantifier. Again, to my knowledge, no implementation of the Box View predicts this commutation.

Summing up, then, when embedded under epistemic modals, conditionals do not seem to contribute a universal quantification over epistemic alternatives. Moreover, conditionals seem to commute with epistemic modals. We would like an account that predicts these observations. On the Box View, it is far from clear how this can be achieved.

<sup>4</sup>For recent discussion of this puzzle, see Gillies (2018). Building on data semantics (Veltman, 1981) Gillies proposes a new analysis for ‘might’ and claims that ‘might’ is ambiguous between the standard and the revised analyses. However, the fact that exactly the same problem arises with other operators, like ‘probably’, suggests that it is not the semantics of ‘might’ which needs to be amended, but rather the semantics of conditionals.

**Problem 2: Probability.** Suppose I make the following claim:

(6) If the coin was tossed, it landed heads.

If the coin is fair and no other hint is available, it is natural to assign a probability of 50% to my claim. But this is not the probability that it is *epistemically necessary* that the coin landed heads if it was tossed. That probability is zero, since we are sure that it is *not* epistemically necessary that the coin landed heads if it was tossed. This is odd: if a conditional is a restricted epistemic necessity claim, its probability should just be the probability that this claim obtains.<sup>5</sup>

In general, it seems that the way in which we attribute probabilities to conditionals conforms to the thesis of Adams (1975), according to which the probability of a conditional  $p \Rightarrow q$  equals the conditional probability of  $q$  given  $p$ .<sup>6</sup> That is, when we estimate the probability of  $p \Rightarrow q$ , we do not estimate, as the Box View would have it, how likely it is that  $q$  is *necessary* given  $p$ ; instead, we estimate how likely it is that  $q$  is *true* given  $p$ .

We would like to have an account of conditionals that explains why we behave in this way. That is, we would like an account that predicts that, given what a conditional means, and given a natural way to construe the notion of probability which is the subject of our intuitive judgments, the probability of  $p \Rightarrow q$  just *is* the conditional probability of  $q$  given  $p$ .

**Aim and structure of the paper.** In this paper, I propose an account of conditionals, modals, and probabilities that achieves the following desiderata. First, it accounts for the data that motivate the Box View. In particular, it predicts that a conditional  $p \Rightarrow q$  is acceptable if and only if  $q$  is acceptable on the supposition of  $p$ , vindicating the Ramsey test idea. And it explains the infelicity of (1). Second, it accounts for the peculiar way in which conditionals embed under modal operators. This is achieved without *ad-hoc* syntactic stipulations, in particular, without denying that conditionals can take scope with respect to modals. Finally, it predicts that the probability of  $p \Rightarrow q$  is the conditional probability of  $q$  given  $p$ . This result is obtained without *ad-hoc* stipulations about probabilities of conditionals. Rather, it is derived from the semantics of conditionals and a general definition of probability.

The paper is structured as follows: in Section 2 I present my theory, called Attitude Semantics (AS, for short, in the following); in Section 3 I show that AS provides a solution to the two problems discussed above; finally, in Section 4 I discuss similarities and differences with the restrictor theory of Kratzer (1986).

## 2 Attitude Semantics

**Language and models.** To present the theory explicitly, I will work with a formal language. The base layer of the language is a set  $\mathcal{L}_0$  of *factual sentences*, whose semantics can be given in terms of truth-conditions relative to possible worlds. For our purposes, we may take  $\mathcal{L}_0$  to be the language of propositional logic based on a finite set  $\mathcal{P} = \{p, q, \dots\}$  of atomic sentences.

**Definition 1** (Factual language,  $\mathcal{L}_0$ ).  $\alpha ::= p \mid \neg\alpha \mid \alpha \wedge \alpha \mid \alpha \vee \alpha$  where  $p \in \mathcal{P}$

The full language  $\mathcal{L}$  that I will work with is obtained by enriching  $\mathcal{L}_0$  with operators designed to capture epistemic vocabulary: a binary operator  $\Rightarrow$  for indicative conditionals; unary operators  $\Box$  and  $\Diamond$  for ‘it must/might be that’; and a unary operator  $\text{Pr}$  for ‘it is probable that’. To avoid

<sup>5</sup>See also Mandelkern (2018). On the more general point that our probabilistic assessment of conditionals contrasts with the hypothesis that they are universal modal claims, see DeRose (1994); Edgington (1995); Schulz (2014). For an analogous point about future discourse, see Belnap *et al.* (2001); Cariani and Santorio (2017).

<sup>6</sup>For an overview of the classical experimental literature, see Evans and Over (2004), §8.



further complications which are not essential to our concerns, we restrict to factual antecedents, and we do not consider Boolean compounds of epistemic sentences.<sup>7</sup>

**Definition 2** (Epistemic language,  $\mathcal{L}$ ).  $\phi ::= \alpha \mid \alpha \Rightarrow \phi \mid \Box\phi \mid \Diamond\phi \mid \Pr\phi$  where  $\alpha \in \mathcal{L}_0$

Semantically, I will work with a set  $W$  of possible worlds. We can model a possible world as a valuation function  $w : \mathcal{P} \rightarrow \{0, 1\}$ . Notice that, since  $\mathcal{P}$  is finite, the set  $W$  of possible worlds is also finite.<sup>8</sup> The truth-value of a factual formula  $\alpha$  at a world  $w$ , denoted  $w(\alpha)$ , is defined as usual. The proposition expressed by  $\alpha$ , denoted by  $|\alpha|$ , is the set of worlds where  $\alpha$  is true.

**Information states and suppositions.** In order to spell out our semantics, we will need a formal notion of information states. Since we are concerned not just with qualitative notions, but also with probabilistic ones, I will take an information state to be a probability distribution on  $W$ . Since  $W$  is finite, we can represent such a distribution simply as a map which assigns a probability to each possible world.

**Definition 3** (Information states).

An information state is a function  $s : W \rightarrow [0, 1]$  with the property that  $\sum_{w \in W} s(w) = 1$ .

A world  $w$  is *ruled out* by an information state  $s$  if it is assigned probability 0. Worlds which are not ruled out are referred to as the *live possibilities* in  $s$ .

**Definition 4** (Live possibilities).

If  $s$  is an information state, its set of live possibilities is  $L(s) := \{w \in W \mid s(w) \neq 0\}$ .

The probability of a proposition  $X \subseteq W$  is just the probability that  $X$  is true.

**Definition 5** (Probability of a proposition). If  $X \subseteq W$ ,  $s(X) := \sum_{w \in X} s(w)$ .

Next, we need a modeling of the process of making an indicative supposition. When we suppose  $\alpha$  in a state  $s$ , we enter a new state  $s[\alpha]$  in which  $\alpha$  is treated as certain. Thus, all worlds in which  $\alpha$  is false should be assigned probability 0. The relative probabilities of the  $\alpha$ -worlds are unaffected by the supposition, and should just be rescaled by a factor  $1/s(|\alpha|)$  so that they sum up to 1 again. In other words, supposing can be modeled by the operation of *conditionalization*. If  $\alpha$  is ruled out in  $s$ , i.e., if  $s$  rules out all  $\alpha$ -worlds, then  $\alpha$  is not consistently supposable in  $s$ .<sup>9</sup>

**Definition 6** (Supposing). If  $s$  is an information state and  $\alpha$  a factual sentence with  $s(|\alpha|) \neq 0$ , then  $s[\alpha]$  is the information state defined as follows:

$$s[\alpha](w) = \begin{cases} \frac{s(w)}{s(|\alpha|)} & \text{if } w \in |\alpha| \\ 0 & \text{if } w \notin |\alpha| \end{cases}$$

Notice that the set of live possibilities after the supposition is  $L(s[\alpha]) = L(s) \cap |\alpha|$ .

<sup>7</sup>The reason to restrict to factual antecedents is that it is just not clear how the process of supposing epistemic sentences works (though see Kolodny and MacFarlane, 2010, for a proposal that we may take on board). The reason not to look at compounds of epistemic sentences is that it is unclear, intuitively, how one should assign probabilities to sentences like  $p \wedge (q \Rightarrow r)$  (Egré and Cozic, 2011). In both cases, the complications do not stem from the particular assumptions of our semantics.

<sup>8</sup>This assumption is not essential, although it simplifies the exposition.

<sup>9</sup>At least, not as an indicative assumption. One can suppose  $\alpha$  as a subjunctive assumption, triggering a different kind of change. This matters for the interpretation of subjunctive conditionals, which we set aside here.

**Semantics.** Traditionally, semantics is about specifying truth-conditions for sentences in contexts. With much recent work (e.g. Dekker, 1993; Veltman, 1996; Yalcin, 2007; Willer, 2013; Hawke and Steinert-Threlkeld, 2016), we will assume that epistemic sentences, including indicative conditionals, differ essentially from factual sentences, which have truth-conditions relative to states of affairs. Rather, we will take such sentences to be devices for negotiating the epistemic attitude to be taken towards certain truth-conditional contents.<sup>10</sup> Accordingly, our semantics for the epistemic language  $\mathcal{L}$  specifies what information it takes to bear a certain attitude to a sentence  $\phi \in \mathcal{L}$ . More specifically, the semantics will take the form of a ternary relation between an information state  $s$ , an attitude  $A$ , and a sentence  $\phi$ :

$$s \models_A \phi$$

The possible values for the attitude parameter  $A$  include full acceptance (denoted  $\forall$ ), compatibility (denoted  $\exists$ ), and partial acceptance to degree  $x \in [0, 1]$  (denoted  $\pi_x$ ). When the word ‘acceptance’ is used without further qualification, what I mean is full acceptance.

**Definition 7.** The set of attitude parameters is  $\text{At} = \{\forall, \exists\} \cup \{\pi_x \mid x \in [0, 1]\}$

We now have all ingredients to recursively specify the semantics. For factual sentences  $\alpha \in \mathcal{L}_0$  we have the following natural clauses:

- $s \models_{\forall} \alpha \iff s(|\alpha|) = 1$
- $s \models_{\exists} \alpha \iff s(|\alpha|) \neq 0$
- $s \models_{\pi_x} \alpha \iff s(|\alpha|) \geq x$

Notice that the clauses for full acceptance and compatibility can be rewritten in terms of the set of live possibilities associated with  $s$ :

- $s \models_{\forall} \alpha \iff L(s) \subseteq |\alpha|$
- $s \models_{\exists} \alpha \iff L(s) \cap |\alpha| \neq \emptyset$

The semantic role of an epistemic modal is that of indicating the attitude expressed towards the prejacent: acceptance for ‘must’, compatibility for ‘might’, and partial acceptance to a high degree for ‘probably’. Formally, these operators work by shifting the attitude parameter:

- $s \models_A \Box \phi \iff s \models_{\forall} \phi$
- $s \models_A \Diamond \phi \iff s \models_{\exists} \phi$
- $s \models_A \text{Pr} \phi \iff s \models_{\pi_{x_0}} \phi$  where  $x_0 \in [0, 1]$  is a fixed threshold value

Finally, conditionals are interpreted by a Ramsey test clause: to bear an attitude to the conditional is to bear the attitude to the consequent, on the supposition of the antecedent. What if the antecedent is not supposable in the state? Then intuitively, the conditional cannot be assessed—we have a presupposition failure.<sup>11</sup> Due to space constraints, here I will not introduce presuppositions; I will just assume that, in order to bear any attitude to a conditional, the antecedent must be compatible with the evaluation state. Thus, the official clause is:

- $s \models_A \alpha \Rightarrow \phi \iff s \models_{\exists} \alpha$  and  $s[\alpha] \models_A \phi$

Provided the antecedent is supposable, this reduces to the following simpler clause:

- $s \models_A \alpha \Rightarrow \phi \iff s[\alpha] \models_A \phi$

<sup>10</sup>And for describing the properties of different bodies of information, as in “Alice believes it might rain.” The present proposal can be extended to cover such occurrences, but I will set them aside here.

<sup>11</sup>It is standard in the literature to treat indicative conditionals as presupposing the epistemic possibility of their antecedent. See, a.o., Von Stechow (1998); Gillies (2009, 2010); Starr (2014); Willer (2014, 2018).

**Semantic equivalence.** If the semantics treats two sentences  $\phi$  and  $\psi$  in exactly the same way, we will say that  $\phi$  and  $\psi$  are semantically equivalent and write  $\phi \doteq \psi$ .<sup>12</sup>

**Definition 8** (Semantic equivalence).  $\phi \doteq \psi \iff \forall s \forall A : (s \models_A \phi \iff s \models_A \psi)$

**Assertion.** In the truth-conditional setting an assertion of  $\phi$  is normally construed, along the lines of [Stalnaker \(1978\)](#), as a proposal to the conversational participants to accept the proposition expressed by  $\phi$ . In our setting, many sentences do not express propositions. However, the semantics gives us a notion of what it is to accept these formulas. Therefore, we do not need the detour through the proposition expressed. We can simply say that an assertion of  $\phi$  is a proposal to the conversational participants to coordinate on a state which accepts  $\phi$ .

### 3 Predictions

**Recovering the predictions of the Box View.** Let us first show that some of the key predictions of the Box View are preserved on our account. First, consider the acceptance conditions of a conditional in a state where the antecedent is supposable. We have:

$$s \models_{\forall} p \Rightarrow q \iff s[p] \models_{\forall} q \iff L(s[p]) \subseteq |q| \iff L(s) \cap |p| \subseteq |q|$$

Thus, AS yields the same results as the Box View about the acceptance of factual conditionals: provided the antecedent is supposable,  $p \Rightarrow q$  is accepted if and only if all live  $p$ -possibilities are  $q$ -possibilities. Since an assertion is a proposal to adopt a state in which the sentence is accepted, the effect of asserting  $p \Rightarrow q$  is also in accordance with the Box View.

Notice also that the proposal fully vindicates the Ramsey test idea: accepting  $p \Rightarrow q$  in a state  $s$  amounts to accepting  $p$  in the hypothetical state  $s[p]$  resulting from the supposition of  $p$ .

We also predict the inconsistency of  $p \Rightarrow q$  with  $\diamond(p \wedge \neg q)$ , since we have  $s \models_{\forall} \diamond(p \wedge \neg q) \iff s \models_{\exists} p \wedge \neg q \iff L(s) \cap |p \wedge \neg q| \neq \emptyset \iff L(s) \cap |p| \not\subseteq |q|$ . Thus, it is impossible for a state to accept simultaneously  $p \Rightarrow q$  and  $\diamond(p \wedge \neg q)$ .

**Solving the embedding problem.** Both attitude semantics and the Box View predict  $p \Rightarrow q$  to be acceptable just when all the epistemically possible  $p$ -worlds are  $q$ -worlds. But there is a crucial difference: in AS, the semantics of the conditional operator does not involve a universal quantification over epistemic possibilities. If we look back at the derivation of the acceptance conditions for  $p \Rightarrow q$ , we can see that the source of the universal quantification is the acceptance attitude, not the conditional operator. This makes a crucial difference: when interpreting a conditional embedded under an epistemic modal, the relevant attitude may be shifted away from acceptance; as a result, no universal quantifier will show up in the semantics. For instance, consider the acceptance conditions for a conditional embedded under ‘might’. Again, let us focus on the interesting case in which the antecedent is supposable.

$$s \models_{\forall} \diamond(p \Rightarrow q) \iff s \models_{\exists} p \Rightarrow q \iff s[p] \models_{\exists} q \iff \exists w \in L(s[p]) : w(q) = 1$$

Thus, we predict that  $\diamond(p \Rightarrow q)$  is not a second-order epistemic claim, but a simple claim of conditional possibility:  $q$  is possible on the supposition that  $p$ . Moreover,  $\diamond$  commutes with  $\Rightarrow$ .

**Proposition 1.**  $\diamond(p \Rightarrow q) \doteq p \Rightarrow \diamond q$

<sup>12</sup>Weaker notions, such as *acceptance equivalence* (having the same acceptance conditions) can also be defined. Although they are quite interesting, they are not relevant to the discussion below.

To see that this is the case, take any state  $s$  and attitude  $A$ . We have:

$$\begin{aligned}
s \models_A \diamond(p \Rightarrow q) &\iff s \models_{\exists} p \Rightarrow q \\
&\iff s \models_{\exists} p \text{ and } s[p] \models_{\exists} q \\
&\iff s \models_{\exists} p \text{ and } s[p] \models_A \diamond q \\
&\iff s \models_A p \Rightarrow \diamond q
\end{aligned}$$

The predictions about conditionals embedded under ‘probably’ are analogous:  $\text{Pr}$  shifts the attitude parameter to  $\pi_{x_0}$ , and thus the acceptance conditions for  $\text{Pr}(p \Rightarrow q)$  involve no universal quantification over epistemic possibilities. Instead, we get:

$$s \models_{\forall} \text{Pr}(p \Rightarrow q) \iff s \models_{\pi_{x_0}} p \Rightarrow q \iff s[p] \models_{\pi_{x_0}} q \iff s[p](|q|) \geq x_0$$

Thus,  $\text{Pr}(p \Rightarrow q)$  is acceptable in case, in the state resulting from the supposition of  $p$ , the probability of  $q$  is high. As we will see, this means that the conditional probability  $s(q|p)$  is high. This is the intuitively correct prediction. Moreover, with a proof analogous to the one we saw for  $\diamond$ , we can show that  $\text{Pr}$  and  $\Rightarrow$  commute.

**Proposition 2.**  $\text{Pr}(p \Rightarrow q) \doteq p \Rightarrow \text{Pr} q$

Thus, we have a solution to the embedding problem: we can explain why conditionals embedded under  $\diamond$  and  $\text{Pr}$  do not contribute a universal quantification over epistemic possibilities, although the acceptance conditions for unembedded conditionals involve such a quantification; and we can predict the commutation of epistemic modals with conditionals.

**Solving the probability problem.** How to characterize the probability  $\mathbb{P}_s(\phi)$  of a sentence in our language relative to a state  $s$ ? For factual sentences, this is clear: the probability of  $\alpha \in \mathcal{L}_0$  is the probability that  $\alpha$  is true:

$$\mathbb{P}_s(\alpha) = s(|\alpha|)$$

Conditionals—we are assuming—don’t express propositions, so they cannot be assigned probabilities in this way. Still, as we discussed, it seems that we can meaningfully attribute probabilities to conditionals. Of course, we could simply follow Adams (1975) and stipulate that, for a factual conditional  $\alpha \Rightarrow \beta$ , its probability is just the conditional probability of the consequent given the antecedent. But this would not explain *why* we assign probabilities to conditionals in this way. Is there a sense in which, when we are estimating the probability of  $p \Rightarrow q$ , we are estimating *the same thing* as when we estimate the probability of  $q$ ?

Attitude semantics allows us to define such a general notion: take the probability of a sentence  $\phi$  in a state  $s$  to be the highest degree to which  $\phi$  is accepted in  $s$  (if  $\phi$  is not accepted in  $s$  to any degree, let the probability be zero). The idea is that the probability of a sentence in a state is a measure of *how acceptable* the sentence is based on the available information.

**Definition 9** (Probability of a sentence).  $\mathbb{P}_s(\phi) := \sup_{[0,1]} \{x \mid s \models_{\pi_x} \phi\}$

Let us look at the predictions that this proposal makes. First, we can prove that the probability of a factual sentence is just the probability that the corresponding proposition is true.

**Proposition 3** (Probabilities of factual sentences). If  $\alpha \in \mathcal{L}_0$ ,  $\mathbb{P}_s(\alpha) = s(|\alpha|)$

The proof is simple, since we have  $\mathbb{P}_s(\alpha) = \sup\{x \mid s \models_{\pi_x} \alpha\} = \sup\{x \mid s(|\alpha|) \geq x\} = s(|\alpha|)$ .

Next, let us look at a conditional  $p \Rightarrow q$ . Based on the semantics of conditionals and the definition of probability, we can now *prove* Adams thesis: in a state in which the antecedent is supposable, the probability of  $p \Rightarrow q$  is the conditional probability of  $q$  given  $p$ .

**Proposition 4** (Probabilities of conditionals). Let  $s$  be a state with  $s(|p|) \neq 0$ . Then:

$$\mathbb{P}_s(p \Rightarrow q) = s[p](|q|) = \frac{s(|p \wedge q|)}{s(|p|)}$$

*Proof.*  $\mathbb{P}_s(p \Rightarrow q)$  is defined as the maximum  $x$  for which  $s \models_{\pi_x} p \Rightarrow q$ . Since  $s \models_{\pi_x} p \Rightarrow q \iff s[p] \models_{\pi_x} q \iff s[p](|q|) \geq x$ , the maximum value of  $x$  for which this holds is  $s[p](|q|)$ . This gives the first identity. As for the second identity, we have:

$$s[p](|q|) = \sum_{w \in |q|} P_{s[p]}(w) = \sum_{w \in |p| \cap |q|} \frac{s(w)}{s(|p|)} = \sum_{w \in |p \wedge q|} \frac{s(w)}{s(|p|)} = \frac{s(|p \wedge q|)}{s(|p|)}$$

Thus, now we can explain *why* the probability of a conditional is the conditional probability of the antecedent given the consequent in terms of our semantics of conditionals and our construal of probability. The reason is that all a conditional does is to restrict the evaluation state—and *not* introduce any quantification of its own: to accept  $p \Rightarrow q$  to degree  $x$  in state  $s$  is just to accept  $q$  to degree  $x$  in the restricted state  $s[p]$ . Therefore, the probability of  $p \Rightarrow q$  in  $s$  is equal to the probability of  $q$  in  $s[p]$ , and this is just the conditional probability of  $q$  given  $p$  in  $s$ .<sup>13</sup>

## 4 Comparison with the restrictor theory

According to the restrictor theory of conditionals (Kratzer, 1986) the embedding problem discussed above stems from a fundamental mistake about the syntax of conditionals: there is no operator  $\Rightarrow$  corresponding to the ‘if ... then’ construction in natural language. Rather, *if*-clauses spell out the restrictor of a modal operator. On this view, sentences (2) and (3) have exactly the same logical form, namely,  $\diamond_p q$ . The restricted modal  $\diamond_p$  works like the original operator  $\diamond$ , except that its domain is restricted to the  $p$ -worlds. In this way, the right interpretation for (2) is derived (unlike in AS, this is achieved by denying that (2) involves a modal embedding a conditional). The problem of accounting for the commutation of modals and conditionals also vanishes, since there is no conditional operator with respect to which the modal can take scope. Thus, the embedding problem is dissolved, or rather, turned into a problem of syntax-semantics interface. However, the probability problem still remains. Consider a plain conditional like (7):

(7) If the coin was tossed, it landed heads.  $\square_p q$

In this sentence, no modal occurs. In order to analyze it, the restrictor theory postulates that it contains a silent epistemic ‘must’, so its logical form is  $\square_p q$ . Thus, (7) makes a certain epistemic claim: that the consequent is epistemically necessary given the antecedent. Then, its probability should be the probability that this claim is true. In the case of (7), in a prototypical

<sup>13</sup>On the idea that Adams’ thesis can be explained in terms of the restricting role of if-clauses, see also Egré and Cozic (2011). The present work improves on that proposal in two ways. First, we give a general definition of probability of a sentence from which both the case of factual sentences and the case of conditionals follow as particular cases. Second, Egré and Cozic’s approach deals with conditionals embedded under probabilistic modals (e.g., “There’s a 50% chance that”) but not with the probabilities of simple, unmodalized conditionals. Since Egré and Cozic build on Kratzer’s restrictor view, they face the problem described in the next section.

scenario this will be zero: it is certainly not the case that the outcome heads is epistemically necessary given the toss. That is not the right prediction.<sup>14</sup> Another way to put the problem is to observe that, on the restrictor theory, (7) has exactly the same logical form as (8).

(8) If the coin was tossed, it must have landed heads.  $\Box_p q$

Thus, (7) and (8) should be judged identically in terms of probabilities. This seems wrong: intuitively, (8) can be rejected with certainty; if we have to assign a probability to it, it would be 0.

Comparing AS to the restrictor theory, we find one key similarity and one key difference. The similarity is that, in both theories, conditional constructions do not contribute a quantifier. The source of quantification lies elsewhere. This is what allows both theories to predict, e.g., that (2) involves only an existential quantifier over possible worlds, and no universal quantifier.

The key difference lies in where the two theories locate the source of quantification: in the restrictor theory, the source of quantification is a modal operator; therefore, we need to assume that a modal operator is always present in a sentence involving conditionals, whence the need to postulate silent necessity modals in the logical form of bare conditionals. Besides lacking independent motivation, this stipulation is empirically problematic, since probability judgments show that bare conditionals are not assessed in the same way as conditionals containing an overt ‘must’. In AS, by contrast, the source of quantification is the attitude parameter; while modal operators can shift this parameter, we need not assume that every sentence contains a modal operator. Thus, AS obviates the need for covert modals. If a sentence does *not* contain an epistemic modal that fixes the attitude parameter, the relevant quantification is not determined once and for all from within the sentence, but it is determined from the outside, by the attitude under consideration. If the sentence is asserted, the relevant attitude is full acceptance, which is responsible for introducing a universal quantification, producing the same effect as if the sentence had contained a ‘must’. But if the sentence is assessed for probability, the process will not involve any universal quantification, and the result will differ from that of the corresponding ‘must’ sentence. This allows AS to provide a solution to the probability problem.

## References

- Adams, E. (1975). *The logic of conditionals: An application of probability to deductive logic*, volume 86. Springer Science & Business Media.
- Belnap, N., Perloff, M., and Xu, M. (2001). *Facing the future: agents and choices in our indeterminist world*. Oxford University Press.
- Bledin, J. (2014). Logic informed. *Mind*, **123**(490), 277–316.
- Cariani, F. and Santorio, P. (2017). Will done Better: Selection Semantics, Future Credence, and Indeterminacy. *Mind*, **127**(505), 129–165.
- Dekker, P. (1993). *Transsentential Meditations. Ups and Downs in Dynamic Semantics*. Ph.D. thesis, ILLC, University of Amsterdam.
- DeRose, K. (1994). Lewis on ‘might’ and ‘would’ counterfactual conditionals. *Canadian Journal of Philosophy*, **24**(3), 413–418.

<sup>14</sup>One might respond that, when asked to judge how probable (7) is, what we do is to make claims of the form  $(P^\alpha)_p q$ , where  $P^\alpha$  is the modal “it is  $\alpha$ -probable that”. But this essentially a restatement of the problem. Why would we do that, if such claims bear no relation to the probability of the statement we are asked to judge? For discussion of this point, see also Mandelkern (2018).

- Edgington, D. (1995). On conditionals. *Mind*, **104**, 235–329.
- Egré, P. and Cozic, M. (2011). If-clauses and probability operators. *Topoi*, **30**.
- Evans, J. and Over, D. (2004). *If*. Oxford University Press.
- Gillies, A. (2004). Epistemic conditionals and conditional epistemics. *Noûs*, **38**(4), 585–616.
- Gillies, A. (2009). On truth-conditions for if (but not quite only if). *Philosophical Review*, **118**(3), 325–349.
- Gillies, A. S. (2010). Iffiness. *Semantics and Pragmatics*, **3**, 4–1.
- Gillies, A. S. (2018). Updating Data Semantics. *Mind*.
- Hawke, P. and Steinert-Threlkeld, S. (2016). Informational dynamics of epistemic possibility modals. *Synthese*, pages 1–34.
- Kolodny, N. and MacFarlane, J. (2010). Ifs and oughts. *The Journal of philosophy*, **107**(3), 115–143.
- Kratzer, A. (1986). Conditionals. *Chicago Linguistics Society*, **22**(2), 1–15.
- Mandelkern, M. (2018). Talking about worlds. *Philosophical Perspectives*, **32**(1), 298–325.
- Ramsey, F. P. (1929). General propositions and causality. In H. Mellor, editor, *F. Ramsey, Philosophical Papers*. Cambridge University Press.
- Schulz, M. (2014). Counterfactuals and Arbitrariness. *Mind*, **123**(492), 1021–1055.
- Stalnaker, R. (1978). Assertion. *Syntax and Semantics*, **9**, 315–332.
- Starr, W. (2014). A uniform theory of conditionals. *Journal of Philosophical Logic*, **43**(6), 1019–1064.
- Veltman, F. (1981). Data semantics. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*. Mathematical Centre.
- Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic*, **25**(3), 221–261.
- Von Fintel, K. (1998). The presupposition of subjunctive conditionals. *The interpretive tract*, **25**, 29–44.
- Willer, M. (2013). Dynamics of epistemic modality. *Philosophical Review*, **122**(1), 45–92.
- Willer, M. (2014). Dynamic thoughts on ifs and oughts. *Philosophers' Imprint*, **14**(28), 1–30.
- Willer, M. (2018). Simplifying with free choice. *Topoi*, **37**(3), 379–392.
- Yalcin, S. (2007). Epistemic modals. *Mind*, **116**(464), 983–1026.

# Testing formal pragmatics of questions through their ignorance inferences\*

Zhuoye Zhao<sup>1</sup> and Alexandre Cremers

<sup>1</sup> NYU Department of Linguistics

## Abstract

Questions are well-studied in semantics, including from a psycholinguistics perspective. They also play a key role in pragmatics through *questions under discussions*, which are known to affect a wide range of phenomena (e.g., focus, implicatures). The pragmatics of questions themselves however is largely understudied, with very few theoretical proposals and only a handful of experimental studies. Pragmatics studies how speakers choose an utterance over possible competitors, and how listeners can draw complex inferences by reconstructing the speaker’s reasoning leading to this choice. While we have a good understanding of the factors at play when a speaker utters a declarative sentence, much less is known about questions. The goal of this paper is to test two proposals extending Grice’s Maxims to questions: van Rooij’s entropy as a measure of question utility, and Groenendijk and Roelofsen’s inquisitive pragmatics. The two theories make opposite predictions regarding the choice between polar and wh-questions in ignorance and partial knowledge situations, and therefore the implicatures they give rise to regarding the questioner’s knowledge. The results of two experiments corroborate the predictions of van Rooij’s proposal. In passing, we establish that the cornering effect of negative alternative questions (Biezma 2009) is independent from their ignorance requirements.

## 1 Introduction

### 1.1 Questions in pragmatics, and pragmatics of questions

Since the work of Roberts (1996), questions have played a crucial role in pragmatics via *questions under discussion*—a theoretical construct describing how any sentence should be understood as addressing an (often implicit) question. This approach has been extremely fruitful and has been applied to a wide variety of topics in formal pragmatics (see Benz and Jasinskaja, 2017 for a review). However, much less has been said on the pragmatics of questions themselves. While we have a good understanding of the principles guiding speakers’ choice of a declarative utterance (dating back at least to Grice, 1967), much less is known regarding questions: How does a speaker’s intentions affect their choice of a question? How does the interpretation of a given question vary with context?

Relevant work includes a few studies on contextual dependency of question interpretation (Ginzburg, 1996; Aloni, 2005; Potts, 2012), a very rich literature on so-called *biased questions*, i.e. specific forms of questions which tend to convey a speaker’s bias towards certain answers (e.g., negative questions; Romero and Han, 2004, a.o., negative alternative questions; Biezma, 2009; Biezma and Rawlins, 2012; Beltrama et al., 2018, discourse particles in questions), and—what will interest us here—a few proposals for a formal theory of how speakers choose particular questions to achieve specific goals.

---

\*We would like to thank Manuel Križ, Edgar Onea, Floris Roelofsen and the inquisitive semantics group, as well as audiences at the InqBnB3 workshop and Graz Department of Linguistics for feedback. We are grateful to three anonymous reviewers for very helpful comments. This work was supported by the European Research Council (ERC, grant number 680220) and the Netherlands Organisation for Scientific Research (NWO).



## 1.2 Existing proposals and their predictions

One of the earliest proposals is [van Rooij \(2003\)](#), who defines a measure of utility for questions given a speaker’s goal, formalizing ideas from previous work (in particular [Ginzburg, 1995](#) and [Krifka, 1995](#)). Under the assumption that the speaker is only trying to learn the truth, this measure boils down to the expected informativity of answers to the question, i.e. the question’s entropy ([van Rooij, 2004](#)), as defined in (1) for a question yielding the partition  $Q$ .<sup>1</sup>

$$(1) \quad U(Q) = \sum_{p \in Q} \Pr(p) \times (-\log_2(\Pr(p))) = E(Q)$$

The speaker then selects the question which maximizes entropy, and this turns out to be the most general question available, formalizing an intuition of [Krifka \(1995\)](#).

An alternative proposal has been sketched in [Groenendijk and Roelofsen \(2009\)](#), henceforth: G&R), building on inquisitive semantics. Without going too far into details, their idea is that questioners should favor questions which have a higher chance of being answered, all else being equal. This is essentially the opposite of [van Rooij’s](#) proposal, as it gives higher values to questions with less informative answers (i.e. answers the addressee is more likely to know).

A direct way to tease apart these two accounts is to look at the competition between polar and wh-questions. Taking the example of the 2018 soccer world cup final brackets, we can consider two possible situation: a (partial) knowledge situation, where the speaker knows which two countries made it to the final, but not which of them won, and a (full) ignorance situation where the speaker has no idea which of the 8 countries that reached quarter finals could have won (and considers them equally likely to win). Now let’s assume that the goal of the speaker is to find out who won the world cup.

Let’s first look at [van Rooij’s](#) theory. In the knowledge situation, there are only two options (France and Croatia). The two questions “Did France win?” and “Who won?” have the same entropy, since the relevant probability measure is the questioner’s internal beliefs (the questioner is assumed to be knowledgeable), and for the speaker the two questions are semantically equivalent. All else being equal, the two questions should therefore be equally good (at this point, other factors such as utterance length or complexity should come into play). For the ignorant speaker, the situation is different. The polar question has much lower entropy than the wh-question (about 0.5 versus 3), so the wh-question should be preferred. In short, [van Rooij](#) predicts the wh-question to be acceptable in both ignorance and knowledge situations, while the polar question should only be good in the knowledge situation (unless we assume a different speaker goal). From the addressee’s point of view, hearing a polar question would suggest that the questioner is already quite knowledgeable, whereas hearing a wh-question does not indicate anything about the questioner’s knowledge.

Turning to G&R’s inquisitive pragmatics, we might need to fill in some gaps first. They propose that the least inquisitive question should be preferred, but it’s unclear against what information state this should be measured. Considering the whole logical space, the wh-question asymmetrically entails the polar question, so it should never be used. Considering the speaker’s information, the two questions become equivalent in the knowledge situation, so the polar question would only be preferred in the ignorance situation. However, none of these options make much sense given the motivation for this constraint: the goal is to make sure the addressee can answer the question. What should matter then, is what the speaker knows of her addressee’s information state. If the participants in the conversation know nothing about each other besides what has been said so far, this is just the common ground. If the common ground does not

<sup>1</sup>It is not essential for the analysis to assume that questions denote partitions ([van Rooij, 2004](#) for details).

contain the speaker’s knowledge, the speaker should contribute this knowledge before asking a question. We can therefore assume that in the knowledge situation, the common ground already entails that France or Croatia won, and so the two questions are equally inquisitive. Conversely, the common ground cannot be more informative than the speaker’s knowledge, so in the ignorance situation, the wh-question is strictly more inquisitive than the polar question. In short, assuming that participants have been cooperative so far, measuring inquisitiveness against the common ground amounts to measuring it against the questioner’s information state, so it predicts that polar questions should be preferred in the ignorance situation only. Finally, if the speaker knows that her addressee is fully informed, none of the questions will be inquisitive in the addressee’s information state, therefore no preference is predicted (in this case, one would need to complement the proposal with an independent maxim). In short, inquisitive pragmatics predicts either no preference between the two questions, or a preference for polar questions in ignorance situations. From the hearer’s perspective, wh-questions would suggest that the speaker is quite knowledgeable already, while polar questions do not indicate anything about the questioner’s knowledge as they are always optimal.

## 2 Experiment 1

### 2.1 Goal and Design

The goal of this first experiment was to test whether questions do indeed give rise to ignorance inferences, and to compare the predictions of the different theories. We tested this straightforwardly with an inference task where participants had to evaluate how much a speaker knows about a given situation given the question they asked, as illustrated in Figure 1.

*Dave wants to get an exotic pet. He went to an exotic pet store and there were only snakes and iguanas. Afterwards he called his friend Barbara, and Barbara asked him...*

**Barbara:** “What animal did you get?”

Would you conclude that Barbara knows that there were only snakes and iguanas in the exotic pet store?

Definitely not  1  2  3  4  5  Definitely yes

Figure 1: Example of an item from Experiment 1 with a wh-question.

More specifically, each item introduced a context which established a disjunctive restriction on some predicate (in this case, Dave either got a snake or an iguana), and a character asked a question. Participants then had to judge whether the questioner was aware of the restriction on a scale from “definitely not” to “definitely yes”. We tested the three constructions of interest in (2a-c): wh-questions, plain polar questions, and negative alternative questions (NAQ). The latter has been shown to have specific pragmatic effects (Bolinger, 1978), for which different explanations have been put forward, predicting different patterns of interaction with ignorance effects (we come back to this in the discussion).

There are three possible effects for a given question. It can either convey speaker knowledge, convey speaker ignorance, or not convey anything regarding the speaker’s knowledge state. We therefore used three baselines: the alternative question in (2d) presupposes knowledge of the

disjunction so it offers a knowledge baseline, the polar question in (2e) is fully resolved by the disjunction, so it can only be used if the questioner is ignorant, and the unrelated question in (2f) does not say anything about the questioner’s knowledge so it allows us to probe the prior knowledge participants would attribute in a given scenario.

- (2) a. WH: “Who won the finals?”  
 b. POLAR: “Did France win the finals?”  
 c. NAQ: “Did France win the finals or not?”  
 d. ALTERNATIVE: “Did France or Croatia win the finals?” (knowledge baseline)  
 e. RESOLVED polar: “Did Belgium win the finals?” (ignorance baseline)  
 f. UNRELATED: “Was any player wounded during the finals?” (prior baseline)

## 2.2 Methods and Materials

The survey was built using a latin square design with question type as a within-subject and within-item factor. We used 3 repetitions per level, hence created 18 contexts.<sup>2</sup> We also included two training items at the beginning of the survey and four fillers.

71 participants were recruited on Amazon Mechanical Turk (age range: 23–62). The survey took about 10min and was paid \$1.40. One non-native speaker was removed from the analysis, as well as 7 participants whose error rate on fillers was more than one standard deviation above the mean (threshold: 53%).

## 2.3 Results

The results for each construction are presented in Figure 2. The control items behave as expected: the ALTERNATIVE question clearly conveys knowledge while the RESOLVED polar question clearly conveys ignorance. The UNRELATED question indicates that participants tend to assume that the speaker is ignorant by default.

The results were analysed with a proportional-odds mixed-effects model (using package `ordinal` in R) with maximal by-item and by-subject random-effects structure (Barr et al., 2013). Question type was treatment-coded with the UNRELATED question as baseline. We calculated the 95% confidence interval for each parameter, as indicated in Table 1. The analysis confirms what can be seen on the graph: the WH questions are virtually indistinguishable from the UNRELATED control, the polar questions and NAQ give rise to significantly stronger knowledge inferences, albeit less than the alternative questions. Furthermore, the results confirm that there is no difference whatsoever between polar questions and NAQ with respects to knowledge/ignorance inferences.<sup>3</sup>

## 2.4 Discussion

The results indicate that WH-questions do not in fact convey ignorance, while the polar questions and NAQ do suggest some degree of knowledge, albeit not as much as the alternative questions which presuppose such knowledge is common ground. At first glance, this is support for van

<sup>2</sup>All materials, including a list of all contexts, are available at <https://semanticsarchive.net/Archive/GM1NGI2Z/Zhao-Cremers-QuestPrag.html>, together with anonymized data and analysis scripts.

<sup>3</sup>On a suggestion of Manuel Križ, we tested possible differences between wh-questions with *which*-phrases and with plain wh-word as a posthoc analysis. Nevertheless, a proportional-odds mixed model on WH-questions with type of wh-phrase as a fixed and by-subject random effect revealed no significant difference ( $\chi^2(1) = 1.6, p = .21$ ).

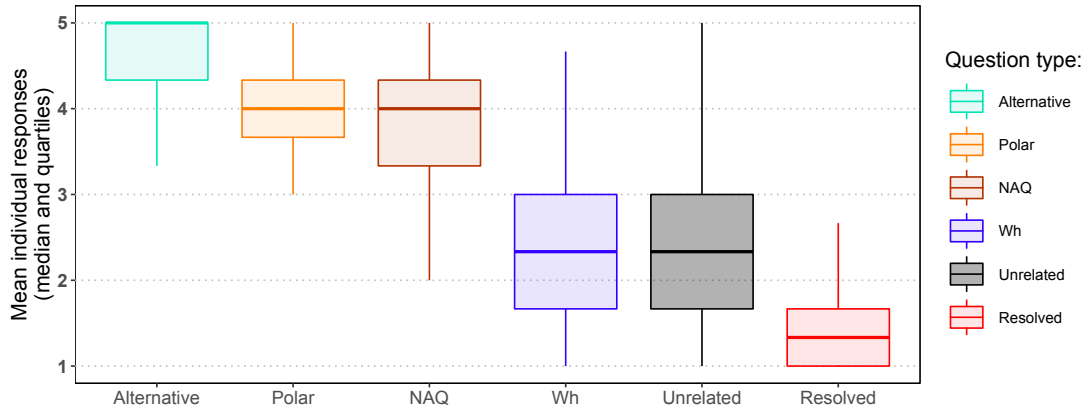


Figure 2: Results from Experiment 1. The ALTERNATIVE and RESOLVED polar questions provide the baseline for knowledge and ignorance respectively. The UNRELATED question serves to probe participants’ prior on knowledgeableability of the speaker when the question offers no indication.

Parameter	$\beta$	95% CI
RESOLVED	-2.1	[-3.1, -1.2]
WH	-0.1	[-0.8, 0.6]
POLAR	2.8	[2.0, 3.5]
NAQ	2.8	[2.0, 3.6]
ALTERNATIVE	5.9	[4.5, 7.3]

Table 1: Estimated parameters and their 95% confidence interval for Experiment 1.

Rooy contra G&R, but before going any further into our discussion of theoretical implications, we want to rule out low-level explanations for the results.

An anonymous reviewer points out that implicit restriction on the domain of the wh-phrases may have affected the results. After all, with explicit restrictors, wh-questions can either convey ignorance (“Which of the 32 competing countries won the world cup?”) or knowledge (“Which of France and Croatia won the finals?”), so leaving the restriction implicit could lead to uncontrolled biases. Note however that in van Rooij’s theory, the probability measure used for entropy is the speaker’s subjective probability, so it isn’t be affected by the explicit restrictor as long as it contains all the answers that the speaker considers possible (answers known to be false do not contribute any entropy). In particular, it is not possible to explain the absence of ignorance inference from wh-questions as covert domain restriction to the two disjuncts: if the speaker is knowledgeable, assuming a smaller domain does not increase the entropy of the wh-question. If the question with an unrestricted domain was infelicitous, the overt “which of A and B” would also be.<sup>4</sup>

<sup>4</sup>The opposite effects of the two restrictors are presumably independent from entropy, although one can be captured by van Rooij’s account. Making overt a smaller domain would reduce entropy, on top of presumably leading to a presupposition failure (assuming the question has an existential presupposition). The entropy account therefore predicts that the questioner considers that the true answer must be in the restrictor (hence the knowledge inference from “which of A and B”), but this is arguably a presupposition to being with. The account does not explain why using a larger domain would lead to an ignorance inference however. The explication seems to be an implicature that each member of the restrictor may be the true answer. Cremers et al. (2019b) discuss

One could imagine that the WH-question fails to give rise to any implicatures simply because participants fail to take into account its polar alternative. Indeed, Cremers et al. (2019a) observed that ignorance inferences from superlative modified numerals were missing in an inference task directly probing them, but were properly detected in an acceptability judgment task. They attribute this effect to the difficulty of the inference task, which led participants to ignore alternatives with a comparative modifier. One reason to suspect that such an effect could be at play here is the asymmetry between the knowledge inference of the polar question and NAQ on the one hand, and the lack of ignorance inference from the wh-questions. Indeed, assuming that participants were fully rational, even if the wh-question does not convey ignorance on its own, the fact that the speaker did not use a polar question should increase the posterior probability for ignorance.<sup>5</sup> We therefore conclude that participants were not fully rational in their inference patterns.

To address this potential limitation of Experiment 1, we ran a second experiment aiming to replicate the results using an acceptability judgment task.

## 3 Experiment 2

### 3.1 Goal and Design

The goal of this experiment was to replicate the results of Experiment 1 using a different design, namely an acceptability judgment task. The idea is to now explicitly include information about the questioner’s knowledge and probe the acceptability of the various questions given this information. The difference between situations in which the questioner is knowledgeable and situations in which they are not indicates whether the question imposes any constraint on the questioner’s internal state. More specifically, if a question requires speaker ignorance, it should be less natural in a situation where the questioner has partial knowledge. Conversely, a question which requires partial knowledge would be less natural when the speaker is fully ignorant.

### 3.2 Methods and Materials

We created two versions of each context in Experiment 1: one in which the questioner was described as aware of the disjunctive state of affairs, and one in which the information corresponding to the disjunctions was omitted entirely and the questioner was described as fully ignorant. The participants’ task was to evaluate whether the question was natural given what they know of the situation and the questioner’s knowledge state. Figure 3 presents an example of ignorance item.

We tested the same questions from Experiment 1 minus alternative questions. Indeed, in an ignorance situation, an alternative question would lead to a presupposition failure. Since

---

the case of questions embedded under *wonder*, and show that this ignorance implicature is sensitive to the form of the restrictor (it is stronger when the restrictor is an explicit conjunction, and possibly when it contains a numeral).

<sup>5</sup>Specifically, assuming that participants apply Bayesian reasoning and that polar questions increase the posterior probability for knowledge (i.e.  $P(\text{knowledge}|\text{polar}) > P(\text{knowledge})$ ), it follows that:

$$P(\text{knowledge}|\neg\text{polar}) = \frac{P(\text{knowledge}) - P(\text{knowledge}|\text{polar})P(\text{polar})}{1 - P(\text{polar})} < P(\text{knowledge})$$

Note that to get close to equality, we would need to assume either that the knowledge inference from polar questions is very small ( $P(\text{knowledge}) \approx P(\text{knowledge}|\text{polar})$ ), in clear contradiction with our results, or that polar questions are unexpected ( $P(\text{polar}) \ll 1$ ), which seems unreasonable, especially when these questions appeared as frequently as wh-questions in our design.

One day, Tony is visiting his friend Betty, who just got a fancy coffee machine.  
Betty doesn't know what coffee Tony likes. Before making coffee, she asks...

**Betty:** "Would you like an espresso?"

How natural is Betty's question in this context?

Odd  1  2  3  4  5  Natural

Figure 3: Sample item from Experiment 2 with a wh-question and a fully ignorant questioner.

this infelicity is presumably stronger than infelicities due to violations of pragmatic constraints, we considered that it risked pushing participants to accept all other questions even when they should have been pragmatically odd (or at least reduce the resolution of our experimental paradigm). We therefore decided to do without an upper bound on knowledge requirements, but kept felicitous alternative questions as true fillers in the experiment (to ensure that they remain salient as alternatives to other questions).

We now had two experimental factors: question type (5 levels)  $\times$  questioner knowledge (2 levels). We reduced the number of repetitions per item to two, and created two new contexts, for a total of 20. The design was again a latin square with both factors within-item and within subject. Each participant only saw a given context under one knowledge condition to avoid potential confusion about what the questioner actually knows.

77 participants were recruited on Amazon Mechanical Turk (age range: 20–65). The survey took about 10min and was paid \$1.40. One non-native speaker was removed from the analysis, as well as 7 participants whose error rate on fillers was more than one standard deviation above the mean (threshold: 53%).

### 3.3 Results

The results for each construction are presented in Figure 4. Control items behave as expected: the UNRELATED question is as acceptable under ignorance as it is under partial knowledge, while the RESOLVED question is clearly degraded under partial knowledge.

The results were analysed with a proportional-odds mixed-effects model (using package `ordinal` in R) with by-item and by-subject random intercepts and slopes for the two main effects, but not for the interaction. Both factors were treatment-coded with UNRELATED and IGNORANCE as baselines. We calculated the 95% confidence interval for each parameter, as indicated in Table 2.

The results overall confirm the first conclusion from Experiment 1: the polar questions and NAQ are much less natural when the questioner is fully ignorant (i.e. they require partial knowledge). In contrast with Experiment 1, WH-questions now seem to be slightly better with ignorance (the effect doesn't reach significance though:  $p = .07$ ).<sup>6</sup>

A posthoc analysis also revealed that NAQs are overall less natural than polar questions ( $\chi^2(1) = 21, p < .001$ )—possibly reflecting the cornering effect—, but that this effect is independent of the questioner knowledge requirement imposed by both questions (no interaction with KNOWLEDGE;  $\chi^2(1) = .05, p = .82$ ).

<sup>6</sup>This time again, there was no difference between which-phrases and plain wh-words ( $\chi^2(1) = 1.1, p = .29$ ).

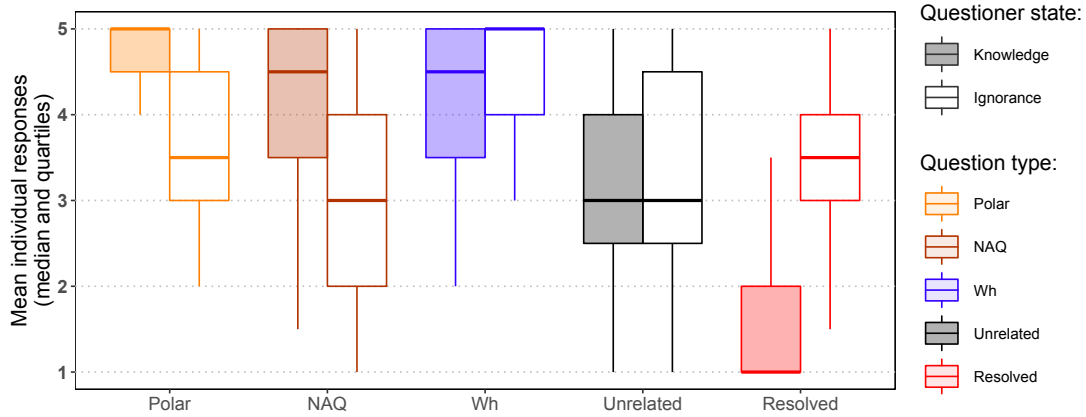


Figure 4: Results from Experiment 2. The UNRELATED question offers a baseline for a question the acceptability of which does not depend on the questioner’s knowledge. The RESOLVED polar questions indicates a maximal contrast, as it’s completely resolved by the disjunctive knowledge.

Parameter	$\beta$	95% CI
RESOLVED	-1.4	[-2.2, -0.5]
WH	2.2	[1.4, 3.1]
POLAR	1.6	[0.7, 2.5]
NAQ	0.5	[-0.3, 1.3]
KNOWLEDGE	-0.1	[-0.7, 0.4]
RESOLVED $\times$ KNOWLEDGE	-3.5	[-4.2, -2.7]
WH $\times$ KNOWLEDGE	-0.7	[-1.4, 0.06]
POLAR $\times$ KNOWLEDGE	2.1	[1.4, 2.8]
NAQ $\times$ KNOWLEDGE	2.0	[1.3, 2.6]

Table 2: Estimated parameters and their 95% confidence interval for Experiment 2.

### 3.4 Discussion

The second experiment closely replicates the results of Experiment 1. Unlike Cremers et al. (2019a), we did not observe a significant task effect on ignorance inferences. We did however observe that the task was easier for participants, as the mean error rate after exclusion was 7%, down from 19%, and a trend for a small ignorance inference with wh-questions. This is still compatible with van Rooij’s predictions that wh-questions are compatible with both ignorance and knowledge, but compete with equivalent polar questions in knowledge situations, slightly reducing their acceptability.

## 4 General discussion

Our results reveal that polar questions are more acceptable when the speaker has partial knowledge, and participants infer knowledge from them, while wh-questions do not convey knowledge and do not seem to convey ignorance either (although there is a small non-significant trend in this direction). Overall these results conform with the predictions of van Rooy (2003).



While some predictions of G&R seem clearly wrong (e.g., that wh-questions should convey partial knowledge), some ideas make intuitive sense and could easily be implemented in van Rooij’s framework. The main intuition is that questioners should take into account the possibility that their addressee doesn’t necessarily have enough information to answer their questions, and that is why they should not always aim for the question with the most informative answers. Concretely, we would replace the probability that an answer is true with the probability that it is true *and* known to the addressee. However, as appealing as it may be, there may be a good reason why we don’t see evidence for this idea in the data: Once we take partial answers into consideration, stronger (more inquisitive) questions may in fact always provide more information than less inquisitive ones.<sup>7</sup>

Finally, an independent contribution of our study is to further demonstrate that the cornering effect of NAQs is orthogonal to their knowledge implicatures. This is new support for Beltrama et al. (2018), who argue against Biezma and Rawlins’s (2017) bundling explanation. Biezma and Rawlins assume that the NAQ bundles together alternatives that the plain polar question leaves open. Framed in van Rooy’s (2003) terms, this would mean that it has a lower entropy (since it does not distinguish between multiple alternatives) and should therefore be more sensitive to ignorance than the polar question as it has even lower utility compared to the wh-question. This would translate into stronger knowledge inferences in Experiment 1 and a greater difference between ignorance and knowledge in Experiment 2, two predictions which are not borne out.<sup>8</sup> By contrast, Beltrama et al.’s (2018) explanation in terms of a ban on discourse-initial focus on polarity is perfectly compatible with our results.

## References

- Aloni, M. (2005). A formal treatment of the pragmatics of questions and attitudes. *Linguistics and Philosophy*, 28(5):505–539.
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3):255–278.
- Beltrama, A., Meertens, E., and Romero, M. (2018). Decomposing cornering effects: an experimental study. In *Proceedings of Sinn und Bedeutung*, volume 22.
- Benz, A. and Jasinskaja, K. (2017). Questions under discussion: From sentence to discourse. *Discourse Processes*, 54(3):177–186.
- Biezma, M. (2009). Alternative vs polar questions: the cornering effect. In *Semantics and Linguistic Theory*, volume 19, pages 37–54.

<sup>7</sup>Take the two questions  $Q_1$  = “Who won?” and  $Q_2$  = “Did Croatia win?”. In inquisitive semantics,  $Q_1 \rightarrow Q_2$ , which means that anyone who can answer  $Q_1$  can also answer  $Q_2$ . Conversely, the chance that the addressee cannot give a complete answer is higher with  $Q_1$  than with  $Q_2$ . In some occasions, she will be able to resolve  $Q_2$  but not  $Q_1$  (e.g., if she only knows that the winner was France or Belgium), hence G&R’s idea that  $Q_2$  is a less risky conversation move than  $Q_1$ . Nevertheless, knowing that the winner was France or Belgium, while not a complete answer to  $Q_1$ , is still relevant as a partial answer. Because the complete answers to  $Q_1$  are at least partial answers to  $Q_2$ , under the assumption that cooperative speakers do give any relevant information (even if it only partially resolves the question),  $Q_1$  always result in at least as much information gain as  $Q_2$ .

<sup>8</sup>Beltrama et al. (2018) do not discuss a closely related proposal by van Rooy and Šafářová (2003), who assume that NAQs require equal utility for the two alternatives, while plain polar question either impose no requirement or favor higher utility for the positive alternative. This would also predict that NAQ are more degraded in ignorance situations than polar questions, and better than polar questions in knowledge situations.



- Biezma, M. and Rawlins, K. (2012). Responding to alternative and polar questions. *Linguistics and Philosophy*, 35(5):361–406.
- Biezma, M. and Rawlins, K. (2017). Or what? *Semantics & Pragmatics*, 10.
- Bolinger, D. (1978). Yes–no questions are not alternative questions. In Hiz, H., editor, *Questions*, pages 87–105. Springer, Dordrecht.
- Christensen, R. H. B. (2015). ordinal—regression models for ordinal data. R package version 2015.6-28. <http://www.cran.r-project.org/package=ordinal/>.
- Cremers, A., Coppock, E., Dotlačil, J., and Roelofsen, F. (2019a). Ignorance implicatures of modified numerals. (under revision for *Linguistics&Philosophy*).
- Cremers, A., Roelofsen, F., and Uegaki, W. (2019b). Distributive ignorance inferences. *Semantics & Pragmatics*. (in press).
- Ginzburg, J. (1995). Resolving questions i. *Linguistics and Philosophy*, 18(5):459–527.
- Ginzburg, J. (1996). Interrogatives: Questions, facts and dialogue. *The handbook of contemporary semantic theory*. Blackwell, Oxford.
- Grice, H. (1967). Logic and conversation. William James Lectures, Harvard University.
- Groenendijk, J. and Roelofsen, F. (2009). Inquisitive semantics and pragmatics. In Larrazabal, J. M. and Zubeldia, L., editors, *Meaning, Content, and Argument: Proceedings of the ILCLI International Workshop on Semantics, Pragmatics, and Rhetoric*. University of the Basque Country Press. Available via [www.illc.uva.nl/inquisitivesemantics](http://www.illc.uva.nl/inquisitivesemantics).
- Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic Analysis*, 25:1–49.
- Potts, C. (2012). Goal-driven answers in the cards dialogue corpus. In *Proceedings of the 30th west coast conference on formal linguistics*, pages 1–20. Cascadilla Proceedings Project.
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Roberts, C. (1996). Information structure in discourse: toward a unified theory of formal pragmatics. *Ohio State University Working Papers in Linguistics*, 49:91–136.
- Romero, M. and Han, C.-h. (2004). On negative yes/no questions. *Linguistics and philosophy*, 27(5):609–658.
- van Rooy, R. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy*, 26(6):727–763.
- van Rooy, R. (2004). Utility, informativity and protocols. *Journal of Philosophical Logic*, 33(4):389–419.
- van Rooy, R. and Šafářová, M. (2003). On polar questions. In Young, R. and Zhou, Y., editors, *Proceedings of SALT 13*, pages 292–309, Ithaca, NY. Cornell University.

# Donkey Anaphora in Non-Monotonic Environments\*

Milica Denić<sup>1</sup> and Yasutada Sudo<sup>2</sup>

<sup>1</sup> Universiteit van Amsterdam, Amsterdam, Netherlands

<sup>2</sup> University College London, London, UK

## 1 Introduction

Donkey anaphora is a type of pronominal anaphora involving a pronoun that is semantically bound by a non-c-commanding indefinite in a quantificational context as in (1).

- (1) Every farmer who owns a donkey loves it.

We will call a pronoun involved in donkey anaphora a *donkey pronoun*.

Donkey pronouns are known to have two possible readings, the  $\exists$ -reading and the  $\forall$ -reading [Bra07, CBH18, Chi95, Fop08, Geu02, Kan94, Kri96, SP89, Yoo94, Yoo96].<sup>1</sup> For instance, without a context, (1) naturally receives a  $\forall$ -reading, (2-a), but as [Chi95] points out, in certain contexts, a weaker  $\exists$ -reading, (2-b), becomes available.

- (2) a. Every donkey-owning farmer loves *all* of their donkeys.  
b. Every donkey-owning farmer loves *at least one* of their donkeys.

The  $\exists$ -reading is observed with examples that make the  $\forall$ -reading implausible to be true due to world knowledge, as in (3) [Chi95, Kan94, SP89].

- (3) a. Every man who had a quarter put it in the parking meter.  
b. Every man who had a credit card paid his bill with it.

In addition to context, the quantifier of a donkey sentence is known to affect the perceived reading [Chi95, Kan94, Yoo94, Yoo96, CBH18]. Generally, donkey pronouns in the scope of universal quantifiers tend to receive  $\forall$ -readings, while those in the scope of *no* and existential quantifiers like *some* preferentially receive  $\exists$ -readings.

- (4) No farmer who owns a donkey hates it.  
 $\approx$  No donkey-owning farmer hates *any* of their donkeys.  
(5) Some farmers who own a donkey love it.  
 $\approx$  Some donkey-owning farmers love *at least one* of their donkeys.

Donkey anaphora involving monotonic quantifiers like the examples we've seen so far is relatively well discussed. In particular, [Kan94] proposes the following generalization (see [Yoo94, Yoo96, Geu02, Fop08] for experimental support).

---

\*We wish to thank Emmanuel Chemla for many important discussions and suggestions. We also thank Keny Chatain, Gennaro Chierchia, Benjamin Spector and the audiences of LINGUAE seminar at ENS and XPRAG 2019 in Edinburgh for useful feedback and discussion. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313610 and n. STG 716230 CoSaQ, and was supported by ANR-17-EURE-0017.

<sup>1</sup>For some speakers (1) is associated with a uniqueness presupposition that every farmer who owns a donkey owns only one, and with this presupposition the two readings collapse to one reading. Our experimental results indicate that this is not obligatory for our participants. See [Hei82, Roo87, Kad87, Hei90, Kad90, Kri96, Chi95, CBH18] for related discussion.

- (6) a. Default readings for  $\uparrow$  MON  $\uparrow$  quantifiers (e.g. *a*, *some*) and  $\downarrow$  MON  $\downarrow$  quantifiers (e.g. *no*) are  $\exists$ -readings.  
 b. Default readings for  $\uparrow$  MON  $\downarrow$  quantifiers (e.g. *all*) and  $\downarrow$  MON  $\uparrow$  quantifiers (e.g. *not all*) are  $\forall$ -readings.

Compared to donkey anaphora involving monotonic quantifiers, the behavior of donkey pronouns in the scope of non-monotonic quantifiers is less well understood. For example, *most* and other proportional quantifiers are non-monotonic with respect to the NP argument, and [Kan94, 132] notes that donkey pronouns under their scope have no noticeable preferences between the two readings (see also [Hei82, §2.1.2] for a similar remark). On the other hand, [Kri96, 151] seems to think otherwise and remarks that proportional quantifiers do often give rise to  $\forall$ -readings.

[Kan94] also points out that ‘existential non-monotonic quantifiers’ like *exactly three*, which are non-monotonic with respect to both arguments, seem to preferentially receive  $\exists$ -readings.

- (7) Exactly three farmers who own a donkey love it.  
 $\approx$  Three donkey-owning farmers love at least one of the donkeys they own and no other donkey-owning farmers love any of the donkeys they own.

More recently [CBH18] put forward an alternative theory to [Kan94] that (modulo contextual factors) predicts a ‘conjunctive reading’ for sentences like (7), which is essentially the conjunction of the existential and  $\forall$ -readings. Thus, under this view, (7) is purported to be true if and only if there are exactly three donkey-owning farmers who love their donkeys, and they all love all of their donkeys.

To our knowledge, no controlled empirical study so far has investigated the readings of donkey pronouns in non-monotonic environments, but as the above short review of the literature reveals, they are of particular theoretical interest, given that conflicting judgments have been reported and at the same time, different theories make different predictions.

The central aim of the present paper is to fill in this empirical gap by reporting on an experimental study that compares donkey anaphora involving two non-monotonic quantifiers, *exactly three* and *all but one*, using truth-value judgment tasks. The results of our experiments suggest that donkey anaphora involving *all but one* receives  $\forall$ -readings more prominently than donkey anaphora involving *exactly three*, for which we obtained no evidence that a  $\forall$ -reading can be accessed. In addition, we observe that the  $\exists$ -reading is easily accessible with both non-monotonic quantifiers, but do not find conclusive evidence that the aforementioned ‘conjunctive’ reading described by [CBH18] is accessed for either of the two quantifiers.

## 2 Experiment 1: *Exactly Three*

### 2.1 Design and Items

The meaning of a non-monotonic quantifier can be decomposed into an upward monotonic and a downward monotonic component. For instance, *exactly three* is semantically equivalent to the conjunction of the upward entailing quantifier *at least 3* and the downward entailing quantifier *at most 3*. Given this, there are in principle four logically possible readings for a donkey pronoun in the scope of *exactly three*: the pronoun could get (i) an  $\exists$ -reading in both components, (ii) a  $\forall$ -reading in both components, and (iii) a  $\forall$ -reading in the upward component combined with an  $\exists$ -reading in the downward component, and (iv) an  $\exists$ -reading in the upward component combined with the  $\forall$ -reading in the downward component. We will label these readings as

(i)  $\frac{\exists}{\exists}$ , (ii)  $\frac{\forall}{\forall}$ , (iii)  $\frac{\forall}{\exists}$ , (iv)  $\frac{\exists}{\forall}$ , respectively. The mnemonic is that what's above the line is the reading of the upward component of the meaning and what's below the line is the reading of the downward component of the meaning. For example, the four logically possible readings for (8) are paraphrased below.

- (8) Exactly three squares that are above a heart are connected to it.
- $\frac{\exists}{\exists}$  At least three squares that are above a heart(s) are connected to **some** of those hearts and at most three squares that are above a heart(s) are connected to **some** of those hearts.
  - $\frac{\forall}{\forall}$  At least three squares that are above a heart(s) are connected to **all** of those hearts and at most three squares that are above a heart(s) are connected to **all** of those hearts.
  - $\frac{\forall}{\exists}$  At least three squares that are above a heart(s) are connected to **all** of those hearts and at most three squares that are above a heart(s) are connected to **some** of those hearts.
  - $\frac{\exists}{\forall}$  At least three squares that are above a heart(s) are connected to **some** of those hearts and at most three squares that are above a heart(s) are connected to **all** of those hearts.

These four possible readings stand in the entailment relation depicted in Figure 1.

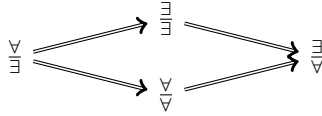


Figure 1: Entailments among the four logically possible readings

Test sentences in Experiment 1 were always of the following form:

- (9) Exactly three of the ⟨squares, triangles⟩ that are above a ⟨star, heart⟩ are connected to it.

Given the entailment relation among the four readings, there are four kinds of situations where at least one of the readings is true, which constitute the target conditions of Experiment 1.

- DEWEAK-UEWEAK: Only the weakest reading  $\frac{\exists}{\forall}$  is true, e.g. Figure 2a.
- DEWEAK-UESTRONG: Only  $\frac{\forall}{\forall}$  and  $\frac{\exists}{\forall}$  are true, e.g. Figure 2b.
- DESTRONG-UEWEAK: Only  $\frac{\exists}{\exists}$  and  $\frac{\exists}{\forall}$  are true, e.g. Figure 2c.
- DESTRONG-UESTRONG: All four readings are true, e.g. Figure 2d.

In addition there were two control conditions in the experiment, where none of the four readings was true. We refer to them as DEFALSE-UESTRONG and DESTRONG-UEFALSE: DEFALSE-UESTRONG is a type of context that makes the upward entailing part of the test sentence true under the  $\forall$ -reading, but falsifies the downward entailing part of the quantifier under both readings (e.g. Figure 2e). DESTRONG-UEFALSE makes the downward-entailing part of the test sentence true under the  $\exists$ -reading, but falsifies the upward entailing part of the quantifier under both readings (e.g. Figure 2f).

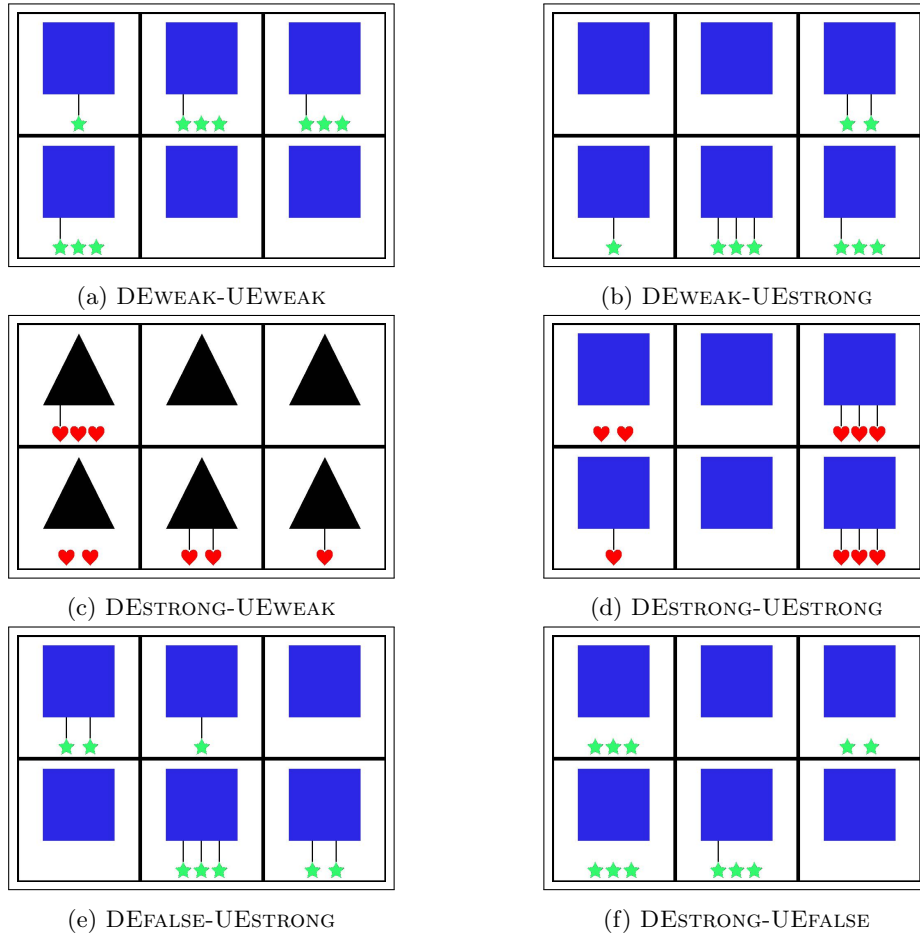


Figure 2: Examples of experimental items.

This amounts to a total of six conditions (four target and two control conditions). Each of the six conditions had six items, meaning there were 36 test items.

Each image consisted of six vignettes as in Figures 2. Each of the vignettes contained a large shape of the same kind (either triangles or squares). In four out of six vignettes the square/triangle was above one, two, or three instances of smaller shapes of the same kind (either stars or a hearts). There were thus two vignettes in which the square/triangle was not above any hearts/stars, which ensured the felicity of the relative clause (cf. test sentence in (9)). In at least one of the vignettes the square or the triangle would appear above exactly one heart or star: this was to ensure the felicity of the singular morphology on the indefinite noun. For each item, a combination of shapes was chosen randomly (i.e. squares+stars, squares+hearts, triangles+stars, triangles+hearts), and the positions of the two vignettes with squares/triangles with no stars/hearts below them were chosen randomly as well. Likewise, the exact number of stars/hearts (one, two, or three) that appeared below the four squares/triangles in an item was chosen randomly for each of the four squares/triangles for each item, granting however that at least one of the squares/triangles would be above exactly one star/heart for felicity reasons mentioned above. We opted for having four squares/triangles that are above a star/heart in

all of the test conditions because this permitted us to use the exact same visual stimuli for this experiment as for Experiment 2 that tested *all but one*.

## 2.2 Procedure and Participants

Participants were directed to a web-based truth-value judgment task, hosted on Alex Drummond’s Ibx platform for psycholinguistic experiments. They were told that they would see sentences paired with images and that their task was to decide whether the sentence was true with respect to the image with which it was paired. The responses were given on a bounded continuous scale, whose ends were labeled as ‘Completely false’ and ‘Completely true’.

The Participants first saw two practice trials, one involving a true sentence and one involving a false sentence, accompanied by suggested responses. The purpose of these examples was to familiarize the participants with the task. They then began the test phase of the experiment, the first two items of which were identical to the two practice trials. These were then followed by the 36 test items, presented in a randomized order for each participant.

65 participants (21 females) were recruited on Amazon Mechanical Turk. One participant was excluded for not being a native speaker of English. We furthermore excluded those participants whose average judgment in the four test conditions combined was lower than their average judgment in the two control conditions combined. The logic behind this exclusion criterion is the following. If they were able to access at least one of the four aforementioned logically possible readings, this should suffice for them to judge the test conditions on average better than the control conditions. If they did not do so, they might have not understood the experimental task, or they were possibly only able to access the uniqueness reading which was not verified in any of the six conditions and hence was not relevant for our purposes (cf. footnote 1). This led to the exclusion of two additional participants. The remaining 62 participants were thus kept for the analyses.

## 2.3 Results and Analysis

The results obtained are summarized in Figure 3a and Table 2. Recall that the target conditions render different logically possible readings true, as summarized in Table 1. Based on this, we will now discuss which readings the results give evidence for.

	$\forall \exists$	$\exists \exists$	$\forall \forall$	$\exists \forall$
DESTRONG-UESTRONG	T	T	T	T
DESTRONG-UEWEAK	F	T	F	T
DEWEAK-UESTRONG	F	F	T	T
DEWEAK-UEWEAK	F	F	F	T
DEFALSE-UESTRONG	F	F	F	F
DESTRONG-UEFALSE	F	F	F	F

Table 1: Target conditions and the readings that they make true.

**No evidence for  $\exists \forall$**  If the weakest reading,  $\exists \forall$ , has been accessed, DEWEAK-UEWEAK, which validates  $\exists \forall$ , should receive higher rating than the control conditions, which validate none of

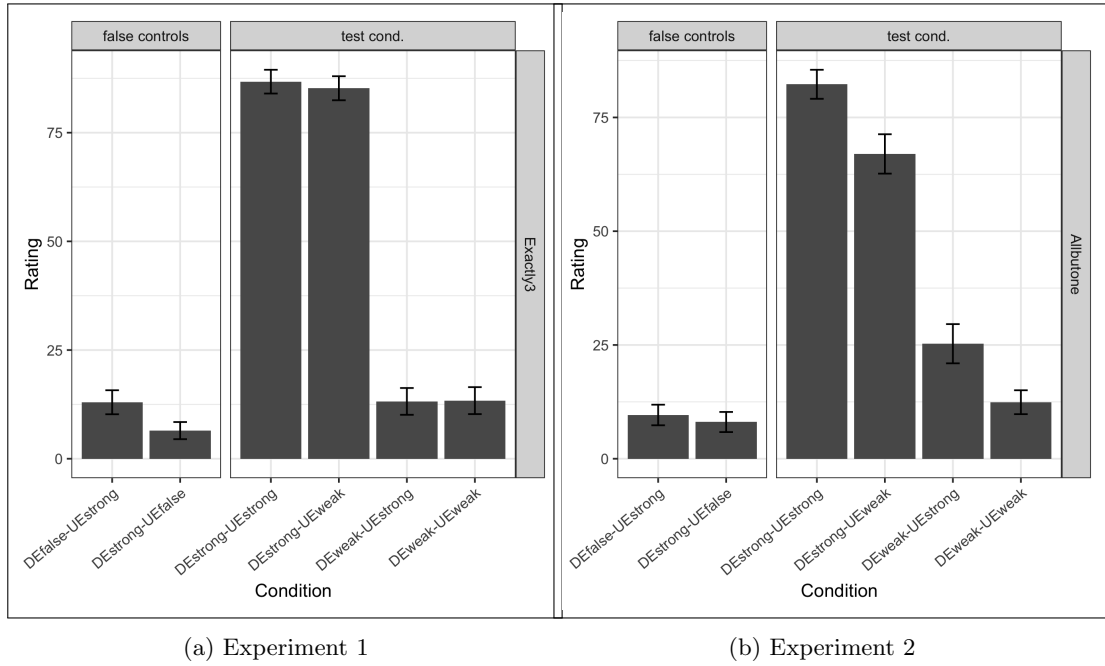


Figure 3: Results of the two experiments per CONDITION. Error bars represent standard errors.

Condition	Mean rating (SD)	
	<i>exactly 3</i>	<i>all but one</i>
DEWEAK-UEWEAK	13.4 (3.1)	12.4 (2.6)
DEWEAK-UESTRONG	13.2 (3.1)	25.3 (4.3)
DESTRONG-UEWEAK	85.2 (2.8)	67 (4.3)
DESTRONG-UESTRONG	86.7 (2.7)	82.3 (3.2)
DEFALSE-UESTRONG	13 (2.8)	9.6 (2.3)
DESTRONG-UEFALSE	6.5 (2)	8.1 (2.2)

Table 2: Experiments 1 and 2: Mean participants' rating and standard deviation per condition.

the readings. The data was subsetting to the items in DEFALSE-UESTRONG and DEWEAK-UEWEAK conditions<sup>2</sup>. A linear mixed model was fitted on this data set with CONDITION as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without CONDITION as a fixed effect revealed no significant effect of CONDITION ( $\chi(1) = 0.06, p = .8$ ). There is thus no evidence for the existence of  $\exists \forall$  reading with *exactly 3*.

**No evidence for  $\forall \forall$**  If  $\forall \forall$  has been accessed, DEWEAK-UESTRONG, which validates both  $\forall \forall$  and  $\exists \forall$ , should receive higher rating than DEWEAK-UEWEAK, which only validates  $\exists \forall$ . The data was subsetting to the items in DEWEAK-UESTRONG and DEWEAK-UEWEAK conditions.

<sup>2</sup>DEFALSE-UESTRONG was chosen rather than DESTRONG-UEFALSE because the mean rating of DEFALSE-UESTRONG was higher than that of DESTRONG-UEFALSE, and thus provides a stricter requirement for the detection of  $\exists \forall$ .

A linear mixed model was fitted on this data set with `CONDITION` as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without `CONDITION` as a fixed effect revealed no significant effect of `CONDITION` ( $\chi(1) = 0.02, p = .9$ ). There is thus no evidence for the existence of  $\forall \exists$  reading with *exactly 3*.

**Evidence for  $\exists \exists$**  If  $\exists \exists$  has been accessed, `DESTRONG-UEWEAK`, which validates both  $\exists \exists$  and  $\forall \exists$ , should receive higher rating than `DEWEAK-UEWEAK`, which only validates  $\exists \exists$ . The data was subsetting to items in `DESTRONG-UEWEAK` and `DEWEAK-UEWEAK` conditions. A linear mixed model was fitted on this data set with `CONDITION` as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without `CONDITION` as a fixed effect revealed a significant effect of `CONDITION` ( $\chi(1) = 100, p < .001$ ). Our results thus provide evidence for the existence of  $\exists \exists$  reading with *exactly 3*.

**No evidence for  $\forall \exists$**  In order to see whether the  $\forall \exists$  reading is available or not, we cannot simply compare `DESTRONG-UESTRONG` condition, which is the only condition validating  $\forall \exists$ , to some other condition. The reason is the following. Suppose the reading  $\forall \exists$  is never accessed, while the other three readings (i.e.  $\forall \forall$ ,  $\exists \exists$ , and  $\forall \forall$ ) are accessed at least to some extent. This would mean that `DESTRONG-UESTRONG` validates all of the three available readings, while all the other conditions validate at most a proper subset thereof. This on its own might suffice to make participants rate items in `DESTRONG-UESTRONG` condition higher than in any of the remaining conditions. Therefore, a significant difference between `DESTRONG-UESTRONG` and any of the other conditions in itself would not constitute strong evidence for the existence of  $\forall \exists$ .

To circumvent this issue, we selected participants with the following property: their mean rating in at least one of `DESTRONG-UEWEAK` and `DEWEAK-UESTRONG` is equal or lower than in `DEWEAK-UEWEAK`. The idea is that these participants accessed at most one of  $\exists \exists$  and  $\forall \forall$ . In other words, they (at most) accessed either (i)  $\forall \forall$ ,  $\forall \exists$ , and  $\exists \exists$  or (ii)  $\exists \exists$ ,  $\forall \exists$ , and  $\forall \forall$ . This further means that, for the participants who did not access  $\forall \forall$ , the only reading which is true in `DESTRONG-UESTRONG` but not in `DESTRONG-UEWEAK` is  $\forall \exists$ . Likewise, for participants who did not access  $\exists \exists$ , the only reading which is true in `DESTRONG-UESTRONG` but not in `DEWEAK-UESTRONG` is  $\forall \exists$ . Thus, if these participants would rate `DESTRONG-UESTRONG` even better than the one they rated better between `DESTRONG-UEWEAK` and `DEWEAK-UESTRONG`, this could be taken as evidence that these participants accessed  $\forall \exists$ . 42 participants fell into this category, and the following analysis was conducted on their responses.

The data was subsetting to items in `DESTRONG-UESTRONG` and the better rated condition between `DESTRONG-UEWEAK` and `DEWEAK-UESTRONG` (as determined for each participant separately). A linear mixed model was fitted on this data set with `CONDITION` (`DESTRONG-UESTRONG` vs. `OTHER`) as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without `CONDITION` as a fixed effect revealed no significant effect of `CONDITION` ( $\chi(1) = 0.01, p = .9$ ). For reference, the mean rating of the better rated conditions between `DESTRONG-UEWEAK` and `DEWEAK-UESTRONG` (as determined for each participant separately) was 88.6 ( $SD = 2.8$ ), while their mean rating in `DESTRONG-UESTRONG` was 88.7 ( $SD = 3.4$ ). There is thus no evidence for the existence of  $\forall \exists$  with *exactly 3*.



### 3 Experiment 2: *All but One*

#### 3.1 Design and Items

Experiment 2 had the exact same design as Experiment 1 except that the test sentences used *all but one*, in place of *exactly three* as in (10).

- (10) All but one of the ⟨squares, triangles⟩ that are above a ⟨star, heart⟩ are connected to it.

As mentioned above, we constructed the pictures for Experiment 1 in such a way that they can be used in Experiment 2 as well. Thus, any differences between the results of the two experiments have to be due to the linguistic, rather than visual, stimuli.

#### 3.2 Procedure and Participants

The procedure was identical to Experiment 1. A new set of 65 participants (25 females) were recruited on Amazon Mechanical Turk, none of whom participated in Experiment 1. One participant was excluded for failing to complete the experiment, two participants were excluded for not being native speakers of English, and six participants were excluded for their average judgment in target conditions not being higher than their average judgment in control conditions (which is the same exclusion criterion as in Experiment 1). 56 participants were thus kept for the analysis.

#### 3.3 Results and Analysis

The results obtained are summarized in Figure 3b and Table 2. The logic of the data analysis is identical to Experiment 1, and we've conducted statistical analyses parallel to Experiment 1 as follows.

**No evidence for  $\exists\forall$**  Statistical analyses on data from DEWEAK-UEWEAK and DEFALSE-UESTRONG revealed no significant effect of CONDITION ( $\chi(1) = 2.35, p = .12$ ). There is thus no evidence for the existence of  $\exists\forall$  with *all but one*.

**Evidence for  $\forall\forall$**  Statistical analyses on data from DEWEAK-UEWEAK and DEWEAK-UESTRONG showed that unlike in Experiment 1, DEWEAK-UESTRONG was judged significantly better than DEWEAK-UEWEAK ( $\chi(1) = 10.5, p < .01$ ). This result provides evidence for the existence of  $\forall\forall$  with *all but one*.

**Evidence for  $\exists\exists$**  Statistical analyses on data from DEWEAK-UEWEAK and DESTRONG-UEWEAK indicate that as in Experiment 1, DESTRONG-UEWEAK is judged significantly better than DEWEAK-UEWEAK ( $\chi(1) = 64.8, p < .001$ ). This result provides evidence for the existence of  $\exists\exists$  with *all but one*.

**Weak evidence for  $\forall\exists$**  As in Experiment 1, in order to determine whether participants have accessed  $\forall\exists$ , we selected participants whose mean rating in at least one of DESTRONG-UEWEAK and DEWEAK-UESTRONG is equal or lower than in DEWEAK-UEWEAK. 33 participants fell into this category in Experiment 2. Analyses parallel to Experiment 1 were conducted on their responses in DESTRONG-UESTRONG and the better rated condition between

DESTRONG-UEWEAK and DEWEAK-UESTRONG (as determined for each participant separately). They revealed a borderline effect of CONDITION (DESTRONG-UESTRONG vs. OTHER) ( $\chi(1) = 3.62, p = .057$ ). For reference, the mean rating of the better rated conditions between DESTRONG-UEWEAK and DEWEAK-UESTRONG (as determined for each participant separately) was 73.2 ( $SD = 5.6$ ), while the mean rating in DESTRONG-UESTRONG was 82.3 ( $SD = 4.6$ ). This suggests that  $\forall_{\exists}$  might be available with *all but one*, but further research is needed to establish this conclusively.

## 4 Concluding Remarks

To summarize the main empirical findings, the results of the two truth-value judgment experiments showed that the  $\exists_{\exists}$  reading of donkey anaphora is available with both *exactly 3* and *all but one*, and we obtained no evidence that  $\exists_{\forall}$  reading is available with either. Furthermore, we find some suggestive evidence that the  $\forall_{\exists}$  reading might be available with *all but one*, but further data should be assessed before a firm conclusion could be drawn. Most interestingly, we find differences between *all but one* and *exactly 3* with respect to the availability of the  $\forall_{\forall}$  reading.

As the logical monotonicity profiles of the two quantifiers are the same, something else must be behind the observed difference. A direction we explored in a follow-up experiment is that subjective, rather than logical, monotonicity might be what matters for the interpretation of donkey anaphora, as [CHR11] claim for NPI licensing. Focusing on quantifier *all*, we tested whether participants' perceived monotonicity of this quantifier explains the extent to which different readings of donkey anaphora are available with it. Due to the limited space available here, we have to omit details of this follow-up experiment, but we found no evidence that subjective monotonicity plays a role in donkey anaphora interpretation with quantifier *all*. It is thus unlikely that subjective perception of monotonicity properties of *all but one* and *exactly 3* are to explain the differences between the two quantifiers in terms of donkey anaphora interpretation.

In light of this, we see two other promising directions to be explored in future research. The first is the possibility that the symmetry profile of the quantifier affects the default reading, as suggested by [Kan94].<sup>3</sup> That is, *exactly 3* is symmetric, but *all but one* is not. The second possibility connects to the well-known fact that context and question under discussion can influence the preferred reading of donkey pronouns (cf. [Chi95, CBH18]). The idea is that *all but one* and *exactly 3*, despite having the same monotonicity properties, might be typically used to answer different questions under discussion, which results in the differences in their default readings.

## References

- [Bra07] Adrian Brasoveanu. *Structured Nominal and Modal Reference*. Ph.D. dissertation, Rutgers University, 2007.
- [CBH18] Lucas Champollion, Dylan Bumford, and Robert Henderson. Donkeys under discussion. To appear in *Semantics and Pragmatics*, 2018.
- [Chi95] Gennaro Chierchia. *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar*. University of Chicago Press, Chicago, 1995.

<sup>3</sup>[Kan94] also mentions Left Continuity as a potentially relevant property here, but since both *exactly 3* and *all but one* are left continuous, it wouldn't explain the difference we observed.

- [CHR11] Emmanuel Chemla, Vincent Homer, and Daniel Rothschild. Modularity and intuitions in formal semantics: the case of polarity items. *Linguistics and Philosophy*, 34(6):537–570, 2011.
- [Fop08] Francesca Foppolo. The puzzle of donkey anaphora resolution. In *NELS*, 2008.
- [Geu02] Bart Geurts. Donkey business. *Linguistics and Philosophy*, 25(2):129–156, 2002.
- [Hei82] Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. dissertation, University of Massachusetts, Amherst, 1982.
- [Hei90] Irene Heim. E-type pronouns and donkey anaphora. *Linguistics and Philosophy*, 13(2):137–177, 1990.
- [Kad87] Nirit Kadmon. *On unique and non-unique reference and asymmetric quantification*. Ph.D. dissertation, University of Massachusetts, Amherst, 1987.
- [Kad90] Nirit Kadmon. Uniqueness. *Linguistics and Philosophy*, 13(3):273–324, 1990.
- [Kan94] Makoto Kanazawa. Weak vs. strong readings of donkey sentences and monotonicity inferences in a dynamic setting. *Linguistics and Philosophy*, 17(2):109–158, 1994.
- [Kri96] Manfred Krifka. Pragmatic strengthening in plural predications and donkey sentences. In *Proceedings of SALT 6*, pages 136–153, 1996.
- [Roo87] Mats Rooth. Noun Phrase Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics. In Peter Gärdenfors, editor, *Generalized Quantifiers: Linguistic and Logical Approaches*, pages 237–268. Reidel, Dordrecht, 1987.
- [SP89] Lenhart K. Schubert and Francis Jeffery Pelletier. Generically speaking, or, using discourse representation theory to interpret generics. In Gennaro Chierchia, Barbara H. Partee, and Raymond Turner, editors, *Properties, Types and Meaning II: Semantic Issues*, pages 193–268. Kluwer Academic Publishers, 1989.
- [Yoo94] Youngeun Yoon. *Weak and Strong Interpretations of Quantifiers and Definite NPs in English and Korean*. PhD thesis, University of Texas at Austin, 1994.
- [Yoo96] Youngeun Yoon. Total and partial predicates and the weak and strong interpretations. *Natural Language Semantics*, 4(3):217–236, 1996.

# Learning modal force: evidence from children’s production and input\*

Anouk Dieuleveut<sup>1</sup>, Annemarie van Dooren<sup>1</sup>,  
Ailís Cournane<sup>2</sup> and Valentine Hacquard<sup>1</sup>

<sup>1</sup>University of Maryland, USA

<sup>2</sup>New York University, USA

adieulev@umd.edu

## Abstract

This paper investigates when and how children figure out that possibility modals express possibility, and necessity modals, necessity. Given that *necessary p* entails *possible p*, what prevents children from hypothesizing possibility meanings for necessity modals? We argue that this *entailment problem* is not a psychological one. On the basis of a corpus study of the modal productions of 2-year-old English children and their mothers and two *Human Simulation Paradigm* experiments (Gillette et al. 1999), we show that children can use cues from the conversational context in which modals are used to learn force, and do not need to rely on negative environments, nor on a bias for necessity meanings.

## 1 Introduction

This paper investigates how children figure out the force of the modals in their language: that *can* or *might* in (1a) express possibility, whereas *must* or *have to* in (1b) express necessity. English modals lexically encode force: they express either possibility or necessity. Standard semantic accounts analyze them as quantifiers over possible worlds: possibility modals introduce existential quantification, whereas necessity modals introduce universal quantification (Lyons 1977; Kratzer 1977). Note that the same modal can be used to express different types, or *flavors* of modality: (1b) can be used to express an epistemic necessity (a likelihood), or various types of ‘root’ (*i.e.* non epistemic) necessities: deontic (obligation), bouletic (desire), or teleological (need).

- (1) a. Jo *can/may/might*... draw.                      possibility  
      b. Jo *must/should/have to*/... draw.            necessity

Experimental literature on modal comprehension suggests that children struggle with modal force until at least age 4: they tend to accept possibility modals in necessity situations, and necessity modals in possibility situations (Noveck 2001; Ozturk and Papafragou 2015). Typically, these studies attribute errors to reasoning difficulties: children over-accept possibility modals in necessity situations because of difficulties reasoning about when a stronger modal would be more appropriate, and necessity modals in possibility situations because of difficulties reasoning about open possibilities. However, they take for granted that children already know modals’ underlying force. In this paper, we address when and how children figure out force by investigating modal talk to and by young children: quantitatively, through corpus analyses, and qualitatively, using two Human Simulation

---

\* We thank our research assistants J. Lee, D. Whitcomb and M. Pizzo, the members of UMD/NYU Modality group S-Lab and Acquisition Lab, especially A. Williams, J. Lidz and Y. Yang. Our project is supported by NSF grant #BCS-1551628.

Paradigm experiments testing how well adults can guess the force of child and adult modal uses from the conversational context.

How do children distinguish necessity from possibility modals? Syntactic information might help narrow candidate meanings to modal meanings (in the spirit of Landau and Gleitman’s 1985 *syntactic bootstrapping hypothesis*), but it cannot help distinguish force. Cues from the physical context are also bound to be limited, since modals express non-actual concepts, with few physical correlates. It thus seems that to figure out modal force, children must rely on cues from the conversational context. But how informative is context about modal force? How easily identifiable are possibility meanings when possibility modals are used, and necessity meanings when necessity modals are used? A second issue might be that necessity entails possibility (the *entailment* or *set/subset problem*).<sup>1</sup> If a possibility meaning is always true when a necessity modal is used, what prevents children from hypothesizing a possibility meaning for a necessity modal?

One solution to the *entailment problem* would be to use evidence from negative (or more generally downward-entailing) environments (Gualmini and Schwarz 2009),<sup>2</sup> as they reverse the direction of logical entailments. A potential issue with this solution is that some necessity modals (*e.g. must*) scope over negation (Iatridou and Zeijlstra 2013). Such cases may be particularly misleading, and reinforce a possibility meaning, as *must not* (*necessary not*) is truth-conditionally equivalent to *cannot* (*not possible*). Another solution would be to equip learners with a bias towards strong meanings, in the spirit of Berwick’s (1985) *strongest meaning hypothesis*. We argue that neither are necessary to solve the problem, as the conversational context in which modals are used provide enough evidence as to their force. Such evidence involves situational cues (*e.g.* who the interlocutors are), pragmatic cues (what the speaker is trying to achieve, in particular, giving orders or permissions), and cues from world knowledge. We show that modal force can be inferred on the basis of these cues alone.

In **section 2**, we provide a quantitative and qualitative assessment of the modals children hear in their *maternal input*. We use the Manchester Corpus of UK English (Theakston et al., 2001) on CHILDES (MacWhinney 2000), and a *Human Simulation Paradigm* (HSP) study (Gillette et al. 1999). This HSP allows us to give a general measure of the informativity of natural conversational contexts about force. In **section 3**, we provide an assessment of *children’s* modals productions, using the same methods: first a corpus study, then an HSP study which assesses the extent to which children’s modal use is adult-like, a novel way of approaching whether children have adult-like grammars. **Section 4** discusses further implications of our findings.

## 2 Children’s modal input

### 2.1 Input corpus study

The Manchester Corpus (Theakston et al., 2001) consists of 12 child-mother pairs (6 females; age range: 1;09-3;00) recorded in unstructured play sessions. We chose this corpus for its relative density, uniformity of sampling, and early age range, and focused on the period between 2;00 and 3;00 y.o. All utterances containing modals<sup>3</sup> (26,598 of 564,625 total

<sup>1</sup>We borrow the term from related discussions involving nouns or quantifiers (see *e.g.* Xu and Tenenbaum 2007).

<sup>2</sup>Gualmini and Schwarz propose this not for modals, but for any entailment problem. See also Musolino et al. (2019).

<sup>3</sup>Modal: possibility: *can, could, might, may; able to*; necessity: *must, should, need, have to, got to, supposed to, need to*. We do not differentiate amongst subtypes of root flavors, and exclude future modals, whose force is a

utterances; adult: 20,755; child: 5,842; excluding repetitions (6.6%): adult: 19,986; child: 4,844) were coded for force (possibility vs. necessity), presence of negative element, and flavor (epistemic vs. root).

Overall, possibility modals are more frequent than necessity modals in adult speech (72.5% of adults’ modal utterances) (**Table 1**). Epistemic uses of modals are overall very rare (**Table 2**). Possibility modals co-occur with negation more frequently than necessity modals (possibility: 20.9% vs. necessity: 10.1%). Most cases of necessity modals occurring with negation involve a modal overscoping negation (*{must, should, ought to}*: 19.4% vs. *{have to, got to, need to, supposed to}*: 7.4%). Negation is significantly more frequent on root than on epistemic modals (epistemic: 4.6% negated, vs. root: 19.1%). Modals rarely occur in other negative environments. Whether children can make use of the relatively infrequent negated necessity cases depends on whatever expectations they have about how modals scope relative to negation, and how clearly meanings of non-necessity or impossibility get conveyed with a negated modal, which we test in the HSP.

	ADULT (n=19,986)		ADULT (n=18,853)			
	all		no negation		negation	
possibility	14,491	72.5%	10,672	79.1%	2,828	20.9%
necessity	5,495	27.5%	4,814	89.9%	539	10.1%

**Table 1:** Counts and percentages of modal uses by force with and without negation, for adults (repetitions excluded: 3.7% of the data)

	all		no negation		negation	
	root	17,190	91.2%	13,896	80.9%	3,293
possibility	12,175	64.6%	9,414	77.3%	2,761	22.6%
necessity	5,015	26.6%	4,482	89.4%	533	10.5%
epistemic	1,662	8.8%	1,590	95.4%	73	4.6%
possibility	1,324	7.0%	1,257	94.9%	67	5.0%
necessity	341	1.8%	332	97.3%	6	2.6%

**Table 2:** Counts and percentages of modal uses, by force, flavor and negation, for adults (n=18,853)

## 2.2 Human Simulation Paradigm Study

To assess the general informativity of natural conversational contexts about force, we use a variant of the *Human Simulation Paradigm* (Gillette et al. 1999), using dialogue contexts extracted from the Manchester corpus. We investigate whether participants can guess the force of a modal based on excerpts of conversations in which it appears, and whether the context is equally informative for necessity and possibility modals, for epistemic and root modals, and for negative vs. positive contexts.

**Procedure** – The experiment was run online on Alex Drummond’s IBEX Farm. Participants recruited via Amazon Mechanical Turk were asked to guess a redacted modal in a dialogue between a child and her mother by choosing between two options, corresponding either to a possibility or a necessity modal, as illustrated in **Figure 1**. All dialogue contexts consisted of the modal sentence with a blank and the 7 preceding utterances. There was first a short training where participants had to choose between the definite vs. indefinite article (*the* vs. *a*) (3 examples with feedback), and then the test phase, without feedback. Overall, each participant had to judge 40 different dialogues (20 trials: 10 possibility, 10 necessity; 20 controls using tense: 10 past, 10 future). The 20 trials were selected randomly from a list of 40 contexts originally extracted from the corpus; the 20 controls were always the same.

---

matter of debate.

**Conditions** – We tested force (possibility vs. necessity) within participants, and flavor (root vs. epistemic) and negation (present vs. absent) between participants. Negation was tested only for root flavor, because negated epistemics were too rare in the corpus. **Table 3** presents the experiment design.

Test condition (between participants)		Modal lemma (within participants)	
		possibility	necessity
EPI-AFF (epistemic affirmative)		<i>might</i>	<i>must</i>
ROOT-AFF (root affirmative) <sup>4</sup>	ROOT-AFF-1	<i>can</i>	<i>must</i>
	ROOT-AFF-2	<i>can/able</i>	<i>have to</i>
ROOT-NEG (root negative)		<i>can't/not able</i>	<i>not have to</i>

**Table 3:** Summary of experimental conditions (Human Simulation Paradigm)

**Material and participants** – 160 contexts (2\*20 per condition) were extracted from the corpus for the different modals (*can*, *able*, *might*, *must*, *have to*). **Exclusion criteria.** We excluded contexts where the target modal was used in preceding utterances. Contexts were not excluded when the adult (or the child) used another non-target modal. **Controls.** Participants had to choose between future and past (*e.g.* [*saw*] vs. [*will see*]). Importantly, the correct answer was not always guessable based on the target sentence only: this required participants to read the entire dialogue. Extraction procedure was the same as for targets. 289 participants were recruited on Amazon Mechanical Turk (4 groups (between participants): ROOT-AFF-1: 73, ROOT-AFF-2: 72; ROOT-NEG: 73; EPI-AFF: 71; language: US English; 156 females, mean age = 40.6 y.o.). We removed from analysis 8 participants (2.8%) who were less than 75% accurate on controls. We thus present results for 281 participants (ROOT-AFF-1: 71, ROOT-AFF-2: 69; ROOT-NEG: 70; EPI-AFF: 71).

CHILD: no.  
 CHILD: it doesn't go there.  
 MOTHER: it doesn't go there.  
 CHILD: oh.  
 MOTHER: does it go there?  
 CHILD: no.  
 MOTHER: no.  
 MOTHER: so it \_\_\_\_\_ go here somewhere.



**Figure 1** Human Simulation Paradigm stimuli: example trial (EPI-AFF, *must*)

**Results** – Data analyses were conducted using R (R Core Team, 2013), using the package lme4 (Bates *et al.* 2014a, 2014b). Overall, participants were highly accurate at guessing modal force (see **Table 4**: general mean accuracy 79.9%). We first ran binomial tests to see whether they differ from chance for each condition. Participants' accuracy significantly differs from chance in each condition. Their lowest performance is found for ROOT-NEG necessity modals (*e.g.* *not have to*) (61.3%).<sup>5</sup> **Force** (possibility vs. necessity). To test whether there was an effect of Force, we used binomial linear mixed effects models, built

<sup>4</sup>We implemented two versions of the **ROOT-AFF** condition: **ROOT-AFF-1** (*can* vs. *must*) allowed us to keep syntactic category of both options identical, **ROOT-AFF-2** (*can/able to* vs. *have to*) to avoid concerns related to the formality of *must* for US English speakers. In cases where *have to* was tensed, we used *able to* as the alternative to avoid losing tense information.

<sup>5</sup>Accuracy for controls was very high (94.6%). There was no difference in accuracy between groups.

with a maximal random effect structure, testing Accuracy with Subject and Item as random factors (following Barr *et al.*, 2013),<sup>6</sup> first overall and then for each condition. We find a general effect of Force, in the direction of a higher accuracy for possibility contexts ( $\chi^2(1)=20.49$ ,  $p=5.9e-6^{***}$ ). Restricting to each comparison group, we find a significant effect in ROOT-AFF-1 ( $\chi^2(1)=61.1$ ,  $p=5.5e-15^{***}$ ) and ROOT-NEG ( $\chi^2(1)=15.6$ ,  $p=7.8e-05^{***}$ ), again in the direction of a higher accuracy for possibility contexts, but not for ROOT-AFF-2 ( $\chi^2(1)=6e-04$ ,  $p=0.98$  (NS)) and EPI-AFF ( $\chi^2(1)=3.73$ ,  $p=.053$  (NS)). **Negation:** we compared ROOT-AFF-2 and ROOT-NEG (these conditions include the same lemmas). We find a significant effect of negation on necessity modals, which leads to lower accuracy (*have to* vs. *not-have to*:  $\chi^2(1)=6.45$ ,  $p=0.011^*$ ). On possibility modals, negation leads to higher accuracy, but the effect is not significant (*can* vs. *can't*:  $\chi^2(1)=2.29$ ,  $p=0.13$  (NS)). We find a strong interaction effect between Force and Negation (**Interaction Force\*Neg**:  $\chi^2(1)=7.9$ ,  $p=0.0047^{**}$ ). **Flavor** (epistemic vs. root): no general effect of flavor ( $\chi^2(1)=0.11$ ,  $p=0.74$  (NS)).

	Mean accuracy <sup>7</sup> (se)		Exact binomial tests (two-sided)	
	possibility	necessity	possibility	necessity
<b>ROOT-AFF-1</b>	91.7% (0.027)	71.7% (0.054)	p <.001*** 95% CI [0.90, 0.94]	p <.001*** 95% CI [0.68, 0.75]
<b>ROOT-AFF-2</b>	81.5% (0.053)	82.0% (0.052)	p <.001*** 95% CI [0.79, 0.85]	p <.001*** 95% CI [0.79, 0.84]
<b>ROOT-NEG</b>	89.5% (0.031)	61.3% (0.065)	p <.001*** 95% CI [0.88, 0.92]	p = 8.95e-08 95% CI [0.56, 0.64]
<b>EPI-AFF</b>	87.2% (0.028)	74.3% (0.049)	p <.001*** 95% CI [0.84, 0.90]	p <.001*** 95% CI [0.71, 0.77]
all	87.5% (0.018)	72.3% (0.028)		
ALL	79.9% (0.018)			

**Table 4:** Accuracy by condition (adult, n = 281\*10)

**Analysis by contexts (post-hoc)** – To get a sense of the kinds of contextual cues that were particularly helpful, we explored the contexts that led to lowest and highest accuracy, for possibility and necessity modals. This informal analysis pointed out two factors, depending on flavor. For root modals, cases where the proposition expressed by the prejacent was effortful or undesirable seem to lead to high accuracy for necessity modals. For epistemic modals, we found high accuracy for necessity modals in contexts that made salient strong evidence for the prejacent.

**Discussion** – Our results show that the conversational context is overall informative about force: adult participants can guess the force of the modal accurately, for both flavors (mean accuracy: 79.9%). This suggests that there are useful cues in the conversational context: if children are sensitive to these cues, they may not need to rely on negation, nor a bias towards necessity meanings. We find a general effect of **force**, with higher accuracy for possibility modals (general accuracy: possibility: 87.5%; necessity: 72.3%). This could be taken as showing that possibility contexts are more informative than necessity contexts, but the effect is found in only 2 sub-conditions, ROOT-AFF-1 and ROOT-NEG (it is near-significant in EPI-AFF:  $\chi^2(1) = 3.73$ ,  $p = .053$ ), but not significant in ROOT-AFF-2). This higher accuracy in possibility contexts might reflect a tendency to answer with possibility modals by

<sup>6</sup>We sometimes had to step back to random-intercepts-only models when the model failed to converge with the full random-effects specification.

<sup>7</sup>Accuracy corresponds to the mean accuracy across the 20 contexts initially extracted for each condition. On average, each context was seen by 34.7 participants (ranging between 24 and 47).



default, maybe because of the relative frequency of possibility and necessity modals,<sup>8</sup> which itself might be due to alternative ways speakers can express necessity (e.g., imperatives for deontic necessity). As for **negation**, we find an opposite effect of negation on possibility and necessity modals: while negation leads to slightly higher accuracy for possibility modals (*can't*: 89.5% vs. *can*: 81.5%), it leads to significantly lower accuracy for necessity modals (*don't have to*: 61.3% vs. *have to*: 82.0%) (interaction effect Force\*Negation). This suggests that negation may not be useful for necessity modals. Thus, contexts containing negated root necessity modals are not only infrequent, they are also not very informative. That said, negation may be useful for possibility modals: negation is frequent on root possibility modals (20.9%), and impossibility contexts are particularly informative. This could help at least for root possibility modals, if children assume that negation scopes over modals.

To summarize, possibility modals are more frequent than necessity modals in the input. Negation is rare on necessity modals, and most cases involve modals that scope over negation, which could be misleading if children assume that negation scopes over modals. Results from our HSP study however show that the conversational context in which modals are used is highly informative about both forces in affirmative contexts. Our posthoc examination of the contexts that led to highest and lowest accuracy in the HSP suggests that the cues adults use may vary with flavor: the desirability of the prejacent seems to matter for root modals, and the salience of evidential support for the prejacent for epistemics.

### 3 Children's productions study

#### 3.1 Corpus study

Like adults, children produce more possibility modals than necessity modals, but the difference is even greater (79.3% of children's modal productions, vs. adults: 72.5%) (**Table 5**). Negated necessity modals are particularly rare in child productions (only 5.1% of necessity modals are negated), but negated possibility modals are extremely frequent: 51.0%. Epistemic modals are overall very rare in child productions (only 2.4% of their modal utterances).

	CHILD (n=4,844)		CHILD (n=4,800)			
	all		no negation		negation	
possibility	3841	79.3%	1861	49.0%	1937	51.0%
necessity	1003	20.7%	950	94.8%	52	5.2%

**Table 5:** Counts and percentages of modal uses by force, ordered by lemma frequency, with and without negation, for children (repetitions excluded: 17.0% of the data)

We find that children use (root) possibility modals frequently, both with and without negation, which we can take as initial evidence of productivity (Stromswold 1990). Children produce fewer necessity modals, and rarely with negation. This difference might be explained in part by their input, as children might grasp more frequent words first, and perhaps by differences in what adults and children like to talk about. Quantitative data about children's productions can only provide a partial picture of whether children use and understand

<sup>8</sup>To control for the effect of frequency, we compared accuracy for *can* and *able-to* (used in ROOT-POS-2 and ROOT-NEG), which are both root possibility modals but strongly differ in frequency (3 *able* for 100 *can* in the Manchester corpus). The general accuracy on *able* was slightly but not significantly lower than on *can* (overall: *able*: 80.8% vs. *can*: 89.8%; vs. *have to*: 71.7%)

modals correctly. To assess these productions in a more qualitative way, we ran an HSP study on children’s modals.

### 3.2 Human Simulation Paradigm Study on child modal usage

The goal of this second HSP study was to investigate children’s early modal productions, and see whether they use modals in an adult-like way by comparing their usage to adult usage. Can adults guess the force of a modal intended by children, given the conversational context in which they use it?<sup>9</sup>

The experiment was identical to the HSP on adult productions, except that we used children’s utterances and made small changes in the instructions. An example of the display is given in **Figure 3**. We implemented the exact same conditions: ROOT-AFF-1; ROOT-AFF-2; ROOT-NEG; EPI-AFF. Controls were also based on tense. **Extraction procedure** – Given the low frequency of negated necessity modals and epistemic necessity modals in child productions, we could test only 10 different contexts for ROOT-NEG necessity and 12 contexts for EPI-AFF necessity. This did not make a difference for the participants, who always had 10 contexts to judge per condition. In all the other conditions, the 10 contexts were still selected randomly out of a list of 20 contexts, randomly extracted from the Manchester corpus. **Exclusion criteria** – We didn’t remove cases where a modal already appeared in the preceding dialogue. **Participants** – 289 participants were recruited on Amazon Mechanical Turk (EPI-AFF: 74, ROOT-AFF-1: 72, ROOT-AFF-2: 73; ROOT-NEG: 72; language: US English; 173 females, mean age = 40.2 y.o.). We removed 18 participants (6.2%) who were less than 75% accurate on controls.<sup>10</sup> We thus present results for 273 participants (EPI-AFF: 68; ROOT-AFF-1: 70; ROOT-AFF-2: 70; ROOT-NEG: 65).

CHILD: train going round minute.  
 MOTHER: in a minute?  
 MOTHER: what... where's the train?  
 CHILD: train... train must be there.  
 MOTHER: it must be there?  
 MOTHER: well.  
 MOTHER: I can't see it.  
 CHILD: \_\_\_\_\_ be at home.



**Figure 2:** Human Simulation Paradigm stimuli: example trial (*must*), child productions

**Analysis – Table 6** reports mean accuracy in each condition. We first ran the same binomial tests as for the adult version. Participants performed better than chance in all conditions involving possibility, but not necessity modals: for ROOT-AFF-2 (*have to*) (mean accuracy: 42.6%) and ROOT-NEG necessity (*not have to*) (mean accuracy: 32.3%), they performed lower than chance. We again used binomial linear mixed effects models (built with

<sup>9</sup>We make the assumption that adults rely on their own competence to judge usage patterns.

<sup>10</sup>For the adult version, errors on controls was very low (5.4%). For the child version, the initial proportion of errors on controls was high (21.6%): post-hoc analysis revealed that this came from 5 control contexts for which accuracy was very low. We removed these 5 controls from our exclusion criteria, as they were particularly difficult, and probably did not indicate that subjects were not doing the task correctly. After restricting to the 15 remaining controls, mean accuracy on controls was 90.0%.

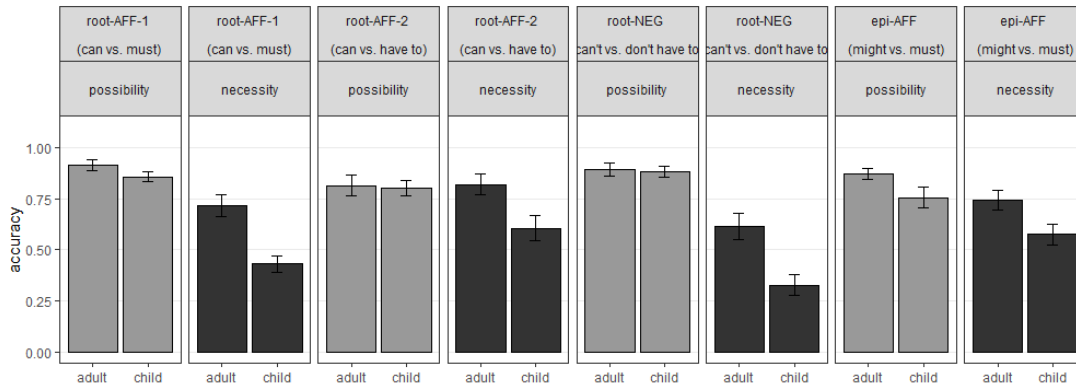
a maximal random effect structure testing Accuracy with Subject and Item as random factors), and find an effect of **Force** in all conditions, with higher accuracy for possibility modals (all:  $\chi^2(1)=20.49$ ,  $p=5.9e-6^{***}$ ; ROOT-AFF-1 (ftc with full spec):  $\chi^2(1)=60.4$   $p=7.7e-15^{***}$ ; ROOT-AFF-2:  $\chi^2(1)=7.37$   $p=0.0066^{**}$ ; ROOT-NEG:  $\chi^2(1)=38.1$ ,  $p=6.6e-10^{***}$ ; EPI-AFF:  $\chi^2(1)=7.93$   $p=0.0048^{**}$ ). We find an effect of Negation, significant for possibility ( $\chi^2(1)=3.65$ ,  $p=0.056^*$ ) and necessity conditions ( $\chi^2(1)=6.74$ ,  $p=0.0093^{**}$ ) (Interaction Force\*Neg:  $\chi^2(1)=9.2374$ ,  $p=0.0024^{**}$ ). No effect of **Flavor** ( $\chi^2(1)=0.14$ ,  $p=0.71$ ). **Age** (adult vs. child productions): we find a general effect of Age, with lower accuracy for child usage ( $\chi^2(1)=260.52$ ,  $p<.001^{***}$ ). Among possibility conditions, only ROOT-AFF-1 is significant; among necessity conditions, all comparisons were significant, except EPI-AFF (**Figure 3**). We find a strong interaction Force\*Age: the difference in accuracy between possibility and necessity modals for child productions is larger than for adult productions ( $\chi^2(1)=32.1$ ,  $p=1.45e-08^{***}$ ).

	Mean accuracy (se)		Exact binomial tests (two-sided)	
	possibility	necessity	possibility	necessity
<b>ROOT-AFF-1</b>	85.1% (0.026)	42.6% (0.039)	p <.001*** 95% CI [0.82, 0.88]	p = 5.085e-05 95% CI [0.39, 0.46]
<b>ROOT-AFF-2</b>	79.6% (0.041)	60.2% (0.060)	p <.001*** 95% CI [0.77, 0.83]	p = 2.05e-07 95% CI [0.56, 0.63]
<b>ROOT-NEG</b>	88.2% (0.027)	32.3% (0.050)	p <.001*** 95% CI [0.86, 0.91]	p <.001*** 95% CI [0.29, 0.36]
<b>EPI-AFF</b>	75.6% (0.050)	56.8% (0.047)	p <.001*** 95% CI [0.73, 0.80]	p = 0.000194 95% CI [0.53, 0.61]
all	82.1% (0.019)	50.1% (0.028)		
ALL	67.4% (0.021)			

**Table 6:** Accuracy rates and significance tests by condition (child, n = 273\*10 observations per cell)

	possibility	necessity
<b>ROOT-AFF-1</b>	$\chi^2(1) = 3.12$ , $p = 0.078$ (NS)	$\chi^2(1) = 35.8$ , $p = 2.1e-09$ ***
<b>ROOT-AFF-2</b>	$\chi^2(1) = 5.80$ , $p = 0.016$ *	$\chi^2(1) = 51.8$ , $p = 6.3e-13$ ***
<b>ROOT-NEG</b>	$\chi^2(1) = 2.78$ , $p = 0.096$ (NS)	$\chi^2(1) = 21.1$ , $p = 4.37e-06$ ***
<b>EPI-AFF</b>	$\chi^2(1) = 3.76$ , $p = 0.053$ (NS)	$\chi^2(1) = 0.22$ , $p = 0.64$ (NS)
all	$\chi^2(1) = 15.9$ , $p = 6.7e-05$ ***	$\chi^2(1) = 175.7$ , $p <.001^{***}$
ALL	$\chi^2(1) = 231.4$ , $p < 2.2e-16$ ***	

**Table 7:** Results of the model testing effect of Age (adult usage vs. child usage)



**Figure 3:** Accuracy by condition: comparison between adult usage and child usage

We find that participants are good at identifying *possibility* modals used by children for both flavors, even if they are less accurate than with adult modals. However, accuracy was much lower for children’s *necessity* modals: while mean accuracy on child possibility modals is 82.1%, it is only 50.1% for necessity modals. The unclarity of children’s necessity modals uses, illustrated in (1) and (2) below, suggests that children may lack adult competency.

- (1) [...] CHILD: what shall I put first?  
 CHILD: that.  
 CHILD: what’s that?  
 MOTHER: pardon?  
 CHILD: I **have to** see a cat. (Becky, 2;08.16) (HSP mean accuracy: 2.9%)
- (2) [...] MOTHER: I thought we had all of these eggs.  
 CHILD: they not.  
 CHILD: they go in the bag.  
 CHILD: they going in there.  
 CHILD: they go in there.  
 MOTHER: oh you’re putting them back in there now, are you?  
 CHILD: you **don’t have to** eat them. (Carl, 2;08.07) (HSP mean accuracy: 20.0%)

## 4 General discussion

When do children figure out the force of modals? Our results suggest that children master possibility modals early: by age 2, they produce them frequently and productively, with and without negation, in an apparently adult-like way given the high accuracy found in the child HSP study. It is less clear, however, whether they master necessity modals at this early age: the few necessity modals they produce do not seem adult-like: accuracy on necessity modals in our child HSP study was below chance, suggesting that children use them in situations where adults would rather use possibility modals. While production data may not fully reflect comprehension, this early asymmetry in children’s mastery of modals might help explain earlier results from comprehension studies with older children, if it persists into the preschool years: if children are uncertain about the underlying force of necessity modals, they will accept them in possibility contexts, but they will also over-accept possibility modals in necessity contexts as they lack a reliable stronger alternative.

How do children eventually figure out modal force? Our results suggest that the conversational context in which modals are used might provide sufficient cues to help them figure out their force. Children don’t need to—and probably can’t—rely on negative contexts, nor do they need a bias toward necessity meanings. Looking at the actual input to children, our results show that negative contexts are not particularly helpful for necessity modals, and in fact it might even be responsible for some of the difficulties children have with these modals. First, negated necessity modals are rare in the input— perhaps for functional reasons, as speakers can express non necessity meanings from the mere use of a possibility modal, via *scalar implicatures* (Horn 1972). Second, negation doesn’t behave uniformly with all modals: some necessity modals like *must* outscope negation (Iatridou and Zeijlstra 2013). If children were to rely on negation to figure out force, they might be misled into thinking that *must* expresses possibility, if they assume that negation scopes over modals. The problem is further complicated by the fact that epistemic and root modals interact differently with negation. Thus, whether children can use negation to figure the force

a modal then depends on what they expect about its scope.<sup>11</sup> Finally, we find in our input HSP that intended force is the least clear for negated necessity modals: adults sometimes use them in situations corresponding to impossibility meanings (e.g. ‘you *don’t have to* break those things’, used as a prohibition). Negation might be more helpful for possibility modals: they cooccur frequently in the input, and negated possibility contexts are particularly informative.

Our input HSP results show that a necessity bias is not necessary for children to solve the *entailment problem*, as the conversational context provides ample cues to figure out force. Our child HSP results further suggest that such a bias is just not at play: if it were, why would children master possibility modals earlier, and struggle with necessity modals? Should our results, in turn, be taken to show that children have a bias towards *possibility* meanings? We do not believe that this has to be the case. Children’s difficulties with necessity modals can be explained by various aspects of the input: first, learners will have fewer opportunities to hear necessity than possibility modals, maybe for functional reasons: speakers have other ways to give orders (using imperatives), or to express certainty (using bare assertion of the prejacent). Second, possibility modals occur in a diverse set of environments (with and without negation, in declarative and interrogative sentences), while necessity modals occur mostly in declarative sentences. Given that the context is highly informative for both possibility and necessity modals in our adult HSP study, we believe that children’s early mastery of possibility compared to necessity modals may be more of a matter of quantity rather than quality of the input.

We have shown that the conversational context is highly informative about force. But what exactly gives away force? One aspect of the context that might be particularly helpful for deontic modality is the kind of (indirect) speech acts that modals are used for: listeners might easily discern orders from permissions, by relying, in part, on the perceived undesirability or effortfulness of the prejacent (e.g. ‘you *#can/have to* eat your broccoli’ vs. ‘you *can/#have to* take a cookie’). For epistemic modals, our posthoc analysis suggests that salient evidence in favor of the prejacent biases towards necessity. We plan to explore this further in future work.

## References

- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in psychology*, 4, 328.
- Berwick, R. C. (1985). *The acquisition of syntactic knowledge* (Vol. 16). MIT press.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73(2), 135-176.
- Gualmini, A., & Schwarz, B. (2009). Solving learnability problems in the acquisition of semantics. *Journal of Semantics*, 26(2), 185-215.
- Horn, L. R. (1972). On the semantic properties of logical operators. Doctoral dissertation, *UCLA*.
- Iatridou, S., & Zeijlstra, H. (2013). Negation, polarity, and deontic modals. *Linguistic inquiry*, 44(4), 529-568.
- Koring, L., Meroni, L., & Moscati, V. (2018). Strong and Weak Readings in the Domain of Worlds: A Negative Polar Modal and Children’s Scope Assignment. *Journal of psycholinguistic research*, 47(6), 1193-1217.

---

<sup>11</sup>A number of studies assess children’s comprehension of sentences containing modals and negation (e.g. Moscati and Crain 2014; Koring et al. 2018), but more research still needs to be done to really understand how and when children learn the right scope interpretations between modals and negation, across force and flavors.

- Kratzer, A. (1977). What ‘must’ and ‘can’ must and can mean. *Linguistics and philosophy*, 1(3), 337-355.
- Landau, B., & Gleitman, L. R. (1985). Cognitive science series, 8. *Language and experience: Evidence from the blind child*. Cambridge, MA, US: Harvard University Press.
- Lyons, J. (1977). *Semantics* (Vol. 2). Cambridge University Press.
- MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.
- Moscatti, V., & Crain, S. (2014). When negation and epistemic modality combine: The role of information strength in child language. *Language Learning and Development*, 10(4), 345-380.
- Musolino, J., Laity d’Agostino, K., & Piantadosi, S. (2019). Why We Should Abandon the Semantic Subset Principle. *Language Learning and Development*, 15(1), 32-46.
- Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development*, 11(3), 191-214.
- Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2013). Modeling the acquisition of quantifier semantics: a case study in function word learnability. *Under review*.
- Portner, P. (2009). *Modality* (Vol. 1). Oxford University Press.
- Stromswold, K. J. (1990). *Learnability and the acquisition of auxiliaries*, Doctoral dissertation, MIT).
- Team, R. C. (2013). R: A language and environment for statistical computing.
- Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of child language*, 28(1), 127-152.
- White, A. S., Hacquard, V., Resnik, P., & Lidz, J. (2017). The contextual modulation of semantic information.
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological review*, 114(2), 245.

# String Iconicity and Granularity\*

Tim Fernando

Trinity College Dublin, Ireland

## Abstract

Given any finite set  $A$  of temporal predicates, expressions iconically representing strings of subsets of  $A$  are defined that depend on whether the predicates in  $A$  are transitional or stative (i.e., homogeneous and to varying degrees, inertial, following Dowty). Subsets  $\Sigma$  of  $A$  represent granularities at which form-meaning resemblance is assessed (through transducers that approximate strings up to  $\Sigma$ ). Complications to iconicity of order are investigated in a first-order fragment of MSO over strings, where experiential factors that trump time are expressible.

## 1 Introduction

A string  $s$  is an *iconic* regular expression inasmuch as the regular language  $\{s\}$  it denotes consists of exactly one entity, an entity  $s$  that undeniably resembles the expression  $s$ . Not all regular expressions represent their denotations quite so faithfully. But pattern matching regular expressions is surely one reason regular expressions are far more popular than formulas of Monadic Second Order logic over strings (MSO; e.g., [Lib04]) for specifying languages accepted by finite automata. That said, conceptualising strings as models of predicate logic suggests useful notions of granularity for form-meaning likeness, supported by a logical apparatus for analysing such likeness. Our specific interest here is *iconicity of order*, linked in [New92] with the principle that “the order of elements in language parallels that in physical experience or the order of knowledge” [Gre63],<sup>1</sup> and understood more broadly below to cover pictorial narratives (e.g., [Abu14, AR17, MB19]). Granularity enters two ways (following [GB92] where it takes the form of *signatures*), syntactically as

- (i) a tag  $\Sigma$  on a string  $s$  (qua expression) for a pair  $(\Sigma, s)$

and semantically as

- (ii) a subscript  $A$  on the interpretation of  $(\Sigma, s)$ , for  $\Sigma \subseteq A$ , as the set

$$\llbracket (\Sigma, s) \rrbracket_A := \{s' \in \mathcal{L}_A \mid f_\Sigma(s') = s\}$$

of strings  $s'$  from a certain set  $\mathcal{L}_A$  whose  $\Sigma$ -approximation  $f_\Sigma(s')$  equals  $s$

with the special case  $\Sigma = A$  leading to the aforementioned instance of iconicity as identity

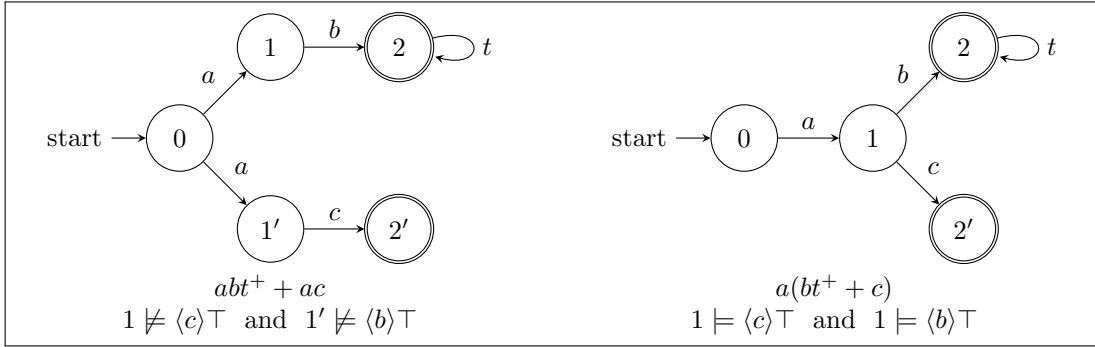
$$f_A(s') = s', \text{ implying } \llbracket (A, s) \rrbracket_A = \{s\} \text{ for } s \in \mathcal{L}_A.$$

The precise definitions of  $\mathcal{L}_A$  and  $f_\Sigma$  (from [Fer19b]) are given below, after pausing to consider if strings are fit for purpose.

---

\*Thanks to three anonymous referees for constructive criticism and forbearance. I am grateful also for feedback from Robin Cooper, Simon Dobnik, Stergios Chatzikyriakidis and others at an initial airing in Gothenburg.

<sup>1</sup>The term in [PV14] is “iconicity of sequence” according to which “the sequence of forms conforms to the sequence of experience, as in the famous collocation *veni, vidi, vici*”

Figure 1: Two automata accepting  $abt^+ + ac$ 

Do strings have sufficient structure to encode the order-related notions of interest? What, for example, are we to do about the continuum  $\mathbb{R}$  (employed in [KR93] and elsewhere as a model of the timeline)? Assuming granularity can be bounded to a finite subset  $A$  of  $\mathbb{R}$ , we can cast  $A$  as a string

$$a_1 a_2 \cdots a_n \quad \text{where } A = \{a_1, a_2, \dots, a_n\} \text{ and } a_1 < a_2 < \cdots < a_n$$

(and, if we wish, reconstruct  $\mathbb{R}$  from a suitable inverse limit of such strings). But, of course, it is not points in  $\mathbb{R}$  that are of linguistic concern, but rather events (and states) spanning intervals. These might — following Russell and Wiener (e.g., [KR93]) and [Ham71] — be deemed primitive, complicating matters in at least two ways. First, the 3 relations  $<, =, >$  between points multiply into 13 (or more) interval relations (e.g., [All83]). And second, a predicate  $P$  that holds at an interval  $I$  immediately raises the question: must  $P$  also hold at a subinterval of  $I$ ? The answer commonly given by linguists appeals to a typology of predicates, including *stative* predicates  $P$ , on which the aspectual calculus of [Dow79] is based, and for which

$$P \text{ holds at an interval } I \text{ iff } P \text{ holds at every subinterval of } I.$$

The present work builds on the Priorean perspective [Pri67], applied in [Fer16, Fer15b] to [Dow79] and other linguistic works. Beyond any particular linear order, there is the additional issue of branching or non-determinism, concerning which it has been argued a process is more than the set of its runs (recorded as strings). An instructive example, adapted from process algebra ([Fok07], pages 10,11), is the pair of automata in Figure 1 accepting the strings

$$abt^+ + ac = a(bt^+ + c). \quad (1)$$

An instance of the distributive law (true of languages, and a semiring axiom), the equality (1) fails to discriminate the left automaton from the right, even though a state *incapable* of a  $c$ -transition is  $a$ -accessible from state 0 in the left but not the right. Suppose, however, we pick out states  $q$  that can make an  $l$ -transition with the formula  $\langle l \rangle\top$  from Dynamic Logic ([HKT00])

$$q \models \langle l \rangle\top \iff (\exists q') q \xrightarrow{l} q' \quad (\text{i.e., } q \in \text{domain}(\xrightarrow{l}))$$

and turn say, the string  $ac$  of arc labels in  $0 \xrightarrow{a} 1' \xrightarrow{c} 2'$  into the string  $\boxed{\langle a \rangle\top} \boxed{\langle c \rangle\top}$  of boxes describing the states 0, 1', 2', respectively. Then equation (1) becomes the *inequation*

$$\boxed{\langle a \rangle\top} \boxed{\langle b \rangle\top} \boxed{\langle t \rangle\top}^+ + \boxed{\langle a \rangle\top} \boxed{\langle c \rangle\top} \neq \boxed{\langle a \rangle\top} \boxed{\langle b \rangle\top, \langle c \rangle\top} (\boxed{\langle t \rangle\top}^+ + \boxed{\phantom{\langle t \rangle\top}}) \quad (2)$$



with distributivity making (2)'s lefthand side equal to

$$\boxed{\langle a \rangle \top} (\boxed{\langle b \rangle \top} \boxed{\langle t \rangle \top} + \boxed{\langle c \rangle \top}).$$

The second box in (2)'s righthand side consists of two formulas,  $\langle b \rangle \top$  and  $\langle c \rangle \top$ , representing state 1 of the right automaton in Figure 1, in accordance with the view that “disjunctions are conjunctive lists of epistemic possibilities” ([Zim00], page 255), perhaps rewriting  $\langle l \rangle \top$  to  $\text{may}(l)$ . Evidently, reducing automata to sets of strings need not conflate the pair in Figure 1, provided we are careful to encode the appropriate information into the strings. Exactly what information is appropriate determines the choice  $\Sigma$  of granularity such as

$$\Sigma = \{ \langle l \rangle \top \mid l \in \{a, b, c, t\} \}$$

for Figure 1. The contrast between the loop label  $t$  in Figure 1 and the other transition labels  $a, b, c$  between distinct states is noteworthy. The label  $t$  can be derived in Dynamic Logic from a formula  $\varphi$  as a test that a state  $q$  satisfies  $\varphi$ , leaving  $q$  fixed if it does and aborting otherwise. More precisely, if the states satisfying  $\varphi$  are collected in the set  $\llbracket \varphi \rrbracket$ , then the input/output interpretation of the *test*  $\varphi?$  is the binary relation

$$q \xrightarrow{\varphi?} q' \iff q \in \llbracket \varphi \rrbracket \text{ and } q = q'$$

testing  $\varphi$  on a state  $q$  without side-effects. Like all programs in Dynamic Logic, the test  $\varphi?$  is interpreted as an input/output relation  $R$ , abstracting away all intermediate states, as in the *relational composition*  $R; R'$  of  $R$  with  $R'$

$$R; R' := \{ (q, q') \mid (\exists \hat{q}) qR\hat{q} \text{ and } \hat{q}R'q' \}$$

with intermediate states  $\hat{q}$  connecting  $R$  to  $R'$  quantified out, and tests  $\varphi?$  idempotent relative to ;

$$\xrightarrow{\varphi?}; \xrightarrow{\varphi?} = \xrightarrow{\varphi?}.$$

The strings  $\alpha_1 \cdots \alpha_n$  of boxes  $\alpha_i$  above keep not only the input  $\alpha_1$  and the output  $\alpha_n$ , but also the intermediate boxes  $\alpha_i$ ,  $1 < i < n$ . Unlike strings  $abt, abtt, \dots$  of transition labels, a string  $s = \alpha_1 \cdots \alpha_n$  of sets  $\alpha_i$  can be intersected componentwise with any set  $\Sigma$  to form its  $\Sigma$ -*reduct*

$$\rho_\Sigma(\alpha_1 \cdots \alpha_n) := (\alpha_1 \cap \Sigma) \cdots (\alpha_n \cap \Sigma)$$

approximating  $s$  up to granularity  $\Sigma$ . [Fer19b] extracts a notion of time dependent on  $\Sigma$ , linked in [FWV19] to *guarded* strings (built with states given by subsets of a finite set  $B$  of Booleans) for *Kleene algebra with tests* [KS96]. Returning to the  $A$ -interpretation  $\llbracket (\Sigma, s) \rrbracket_A$  mentioned above, the idea is that  $f_\Sigma(s)$  is a compressed  $\Sigma$ -reduct of  $s$

$$f_\Sigma(s) = \begin{cases} bc(\rho_\Sigma(s)) & \text{for stative } \Sigma \\ d_\square(\rho_\Sigma(s)) & \text{for transitional } \Sigma \end{cases}$$

where *block compression*  $bc$  compresses substrings  $\alpha\alpha$  to  $\alpha$ , as in

$$bc(\boxed{\langle a \rangle \top} \boxed{\langle b \rangle \top} \boxed{\langle t \rangle \top}^n) = \boxed{\langle a \rangle \top} \boxed{\langle b \rangle \top} \boxed{\langle t \rangle \top} \quad \text{for } n \geq 1$$

while *depadding*  $d_\square$  eliminates all occurrences of  $\boxed{\phantom{\langle a \rangle \top}}$ , as in

$$d_\square(\boxed{\langle a \rangle \top} \boxed{\langle c \rangle \top}) = \boxed{\langle a \rangle \top} \boxed{\langle c \rangle \top}$$

and the line between stative and transitional  $\Sigma$  rests on whether or not  $\Sigma$  consists of tests. Full definitions are given in the next section, which is followed by a delineation of star-free languages for string iconicity and some applications to temporal interpretation.

As the material below can be off-puttingly dry and technical, a few words about its linguistic motivation are in order. The basic challenge is to determine how events and states described by a sequence of sentences or pictures are temporally related. A starting point (suggested by order iconicity) is the order in which expressions are described, refined so that events move the time of a narrative forward, but states do not (e.g., [Dow86]). This refinement coincides with the refinement of concatenation to *coalesced product*  $\diamond$  in *Kleene algebra with tests* ([KS96]), provided statives are, as in [FWV19], understood as tests. The coalesced product  $s \diamond s'$  of strings  $s$  and  $s'$  attaches the right end of  $s$  to the left end of  $s'$ , assuming these match (in which case they are fused). A more flexible alternative associates sets  $\Sigma$  and  $\Sigma'$  with  $s$  and  $s'$ , respectively, and interprets the combination  $\{(\Sigma, s), (\Sigma', s')\}$  at  $A$  as the intersection

$$\llbracket \{(\Sigma, s), (\Sigma', s')\} \rrbracket_A = \llbracket (\Sigma, s) \rrbracket_A \cap \llbracket (\Sigma', s') \rrbracket_A$$

of the  $A$ -interpretations of  $(\Sigma, s)$  and  $(\Sigma', s')$ . The remainder of this paper investigates string expressions built and interpreted along these lines, related to iconicity of order by granularities  $\Sigma, \Sigma'$  based on contextual factors other than time that shape “physical experience or the order of knowledge.” Space is an obvious factor, as are an agent’s cognitive limitations, leading to distinct (if not disjoint) perspectives  $\Sigma$  and  $\Sigma'$ . For a concrete illustration, take two football matches kicking off at the same time in different cities, for distinct views  $\Sigma$  and  $\Sigma'$ , and narratives  $s$  and  $s'$  (respectively) that may temporally overlap here and there.

## 2 Stative and transitional projections

Just as automata transitions  $q \xrightarrow{a} q'$  are formed from states  $q, q'$  and actions  $a$ ,

- (i) events are analysed in [GJW18] in terms of results and actions
- (ii) non-stative verbs are classified in [LRH13] as result or manner, exemplified in [Fil70] by *break* and *hit*
- (iii) telic and iterable transitions are expressed in [Dow79] through operators BEC(ome) and DO, leading in [Fer15a] to strings  $\boxed{\neg\varphi} \boxed{\varphi}$  and  $\boxed{\text{ap}(a)} \boxed{\text{ef}(a)}$  respectively, with result described by a stative  $\varphi$  and expressions  $\text{ef}(a)$  and  $\text{ap}(a)$  of the effect and application of  $a$ .

$$q \xrightarrow{a} q' \quad \left\{ \begin{array}{lll} \text{result } q' & \textit{break}/\textit{result} & \text{BEC} \\ \text{action } a & \textit{hit}/\textit{manner} & \text{DO} \end{array} \right. \quad \begin{array}{|c|c|} \hline \neg\varphi & \varphi \\ \hline \text{ap}(a) & \text{ef}(a) \\ \hline \end{array}$$

Gärdenfors      Fillmore/LevinRH      Dowty

[LRH13] distinguishes between a result being specified lexically,  $\varphi$ , and supplied contextually,  $\text{ef}(a)$ , which we can understand as a context of use supplying MSO-formulas that constrain  $\text{ef}(a)$  and  $\text{ap}(a)$ . An MSO-formula is built from a binary (successor) relation  $S$ , and unary relations  $P_\sigma$  labeled by elements  $\sigma$  of some finite set  $\Sigma$  (including, for example,  $\varphi, \neg\varphi, \text{ap}(a)$  and  $\text{ef}(a)$ ). A string  $s = \alpha_1 \cdots \alpha_n$  of subsets  $\alpha_i \subseteq \Sigma$  is identified with the MSO $_\Sigma$ -model

$$\text{Mod}_\Sigma(s) := \langle [n], S_n, \{U_\sigma\}_{\sigma \in \Sigma} \rangle$$

with string positions forming its universe/domain

$$[n] := \{1, 2, \dots, n\},$$

to which the  $+1$  function is restricted to interpret  $S$  as

$$S_n := \{(i, i+1) \mid i \in [n-1]\} = \{(1, 2), (2, 3), \dots, (n-1, n)\}$$

and each  $P_\sigma$  interpreted as the subset

$$U_\sigma := \{i \in [n] \mid \sigma \in \alpha_i\} \quad (\sigma \in \Sigma)$$

of string positions where  $\sigma$  occurs. Each box  $\alpha_i$  of  $s$  can be recovered by the equation

$$\alpha_i = \{\sigma \in \Sigma \mid i \in U_\sigma\} \quad (i \in [n])$$

establishing an isomorphism between  $(2^\Sigma)^+$  and a family of  $\text{MSO}_\Sigma$ -models. That family of  $\text{MSO}_\Sigma$ -models is large enough to accommodate for any finite set  $A \supseteq \Sigma$ , the  $\Sigma$ -reduct  $M \upharpoonright_\Sigma$  of an  $\text{MSO}_A$ -model  $M = \langle [n], S_n, \{U_a\}_{a \in A} \rangle$  approximating  $M$  by restricting  $\{U_a\}_{a \in A}$  to  $\{U_\sigma\}_{\sigma \in \Sigma}$

$$M \upharpoonright_\Sigma := \langle [n], S_n, \{U_\sigma\}_{\sigma \in \Sigma} \rangle.$$

Proceeding from a string  $s$  of subsets of  $A$ , the  $\Sigma$ -reduct of  $s$  picks out the  $\Sigma$ -reduct of  $\text{Mod}_A(s)$

$$\text{Mod}_A(s) \upharpoonright_\Sigma = \text{Mod}_\Sigma(\rho_\Sigma(s)).$$

For  $f_\Sigma(s)$ , we compress the  $\Sigma$ -reduct of  $s$ , removing stutters  $\alpha\alpha$  through  $\mathit{bc}$

$$\mathit{bc}(s) := s \quad \text{if } \text{length}(s) < 2 \quad \text{and} \quad \mathit{bc}(\alpha\alpha's) := \begin{cases} \mathit{bc}(\alpha's) & \text{if } \alpha = \alpha' \\ \alpha \mathit{bc}(\alpha's) & \text{otherwise} \end{cases}$$

or occurrences of the empty box  $\square$  through  $\mathit{d}_\square$

$$\mathit{d}_\square(\epsilon) := \epsilon \quad \text{and} \quad \mathit{d}_\square(\alpha s) := \begin{cases} \mathit{d}_\square(s) & \text{if } \alpha = \square \\ \alpha \mathit{d}_\square(s) & \text{otherwise.} \end{cases}$$

The functions  $\mathit{bc}$  and  $\mathit{d}_\square$  are related by a certain border translation ([Fer19b]) that illuminates the difference between stative and transitional  $\Sigma$ , determining whether  $f_\Sigma = \rho_\Sigma; \mathit{bc}$  or  $f_\Sigma = \rho_\Sigma; \mathit{d}_\square$ . More precisely, let  $l$  and  $r$  be 1-1 functions with domain  $A$ , whose images form  $A_\bullet$

$$A_\bullet := \{l(a) \mid a \in A\} \cup \{r(a) \mid a \in A\}$$

on the assumption that  $A$ ,  $\{l(a) \mid a \in A\}$  and  $\{r(a) \mid a \in A\}$  are pairwise disjoint. The border translation  $\mathbf{b} : (2^A)^* \rightarrow (2^{A_\bullet})^*$  marks  $a$ 's left borders  $l(a)$  and  $a$ 's right borders  $r(a)$ , mapping  $\alpha_1 \cdots \alpha_n \in (2^A)^*$  to  $\beta_1 \cdots \beta_n$  where

$$\begin{aligned} \beta_i &:= \{l(a) \mid a \in \alpha_{i+1} - \alpha_i\} \cup \{r(a) \mid a \in \alpha_i - \alpha_{i+1}\} \quad \text{for } i < n \\ \beta_n &:= \{r(a) \mid a \in \alpha_n\}. \end{aligned}$$

For example,

$$\mathbf{b}(\boxed{a' \mid a \mid a}) = \boxed{l(a), r(a') \mid \mid r(a)} \quad \text{and} \quad \mathit{d}_\square(\mathbf{b}(\boxed{a' \mid a \mid a})) = \mathbf{b}(\mathit{bc}(\boxed{a' \mid a \mid a}))$$

and in general,

$$\mathbf{b}(\mathit{bc}(s)) = \mathit{d}_\square(\mathbf{b}(s)) \quad \text{for any } s \in (2^A)^+ \text{ not ending in } \square.$$

The border symbols  $l(a)$  and  $r(a)$  specify their effects lexically through the  $\text{MSO}_{\{a\}}$ -formulas

$$\begin{aligned} \chi_{l(a)}(x) &:= \exists y(xSy \wedge P_a(y)) \wedge \neg P_a(x) && \text{“}a \text{ occurs at } x\text{'s successor but not at } x\text{”} \\ \chi_{r(a)}(x) &:= P_a(x) \wedge \neg \exists y(xSy \wedge P_a(y)) && \text{“}a \text{ occurs at } x \text{ but not at any successor of } x\text{”} \end{aligned}$$

characterising  $\mathbf{b}$  in that for any  $s \in (2^{A \cup A_\bullet})^+$ ,

$$\mathbf{b}(\rho_A(s)) = \rho_{A_\bullet}(s) \iff \text{for all } b \in A_\bullet, s \models \forall x(P_b(x) \equiv \chi_b(x)).$$

	stative (result)	transition (action)
compression	$bc : sa\alpha s' \rightsquigarrow sas'$	$d_{\square} : s\square s' \rightsquigarrow ss'$
[All83] $a$ starts $a'$	$\boxed{a, a'   a'}$ interior	$\boxed{l(a), l(a')   r(a)   r(a')}$ border
[vLH05]	fluent ( <i>HoldsAt</i> )	event type ( <i>Happens</i> )
MSO <sub>A</sub>	$\forall x \forall y (xSy \supset \neg \bigwedge_{a \in A} (P_a(x) \equiv P_a(y)))$	$\forall x \bigvee_{a \in A} P_a(x)$

Table 1: Statives vs transitions

More concretely, Table 1 represents an interval relation in [All83] two ways ([Fer19b]): statively in terms of an interval's interior, and transitionally in terms of an interval's borders ([DS08], page 3288). Following section 3, chapter 7 of [vLH05], Table 1 also aligns statives with fluents (comparable to Dynamic logic formulas  $\varphi$  that induce tests  $\varphi?$  which are idempotent relative to relational composition  $;$ ), and transitions with event types (which may, like the program  $x := x + 1$  incrementing  $x$ , change with iteration). Idempotence relative to  $;$  is tied in [FVV19] to the homogeneity of statives assumed in [Dow79].

Under the present approach, the  $bc/d_{\square}$  divide between statives and transitions translates to a contrast between MSO<sub>A</sub>-formulas  $\chi_A^{bc}$  and  $\chi_A^{d_{\square}}$  prescribing “no steps<sub>S</sub> without change<sub>A</sub>”

$$\chi_A^{bc} := \forall x \forall y (xSy \supset \Delta_A(x, y)) \quad \text{where } \Delta_A(x, y) := \neg \bigwedge_{a \in A} (P_a(x) \equiv P_a(y))$$

and “no time without change<sub>A</sub>”

$$\chi_A^{d_{\square}} := \forall x \bigvee_{a \in A} P_a(x) \quad (\text{each } a \in A \text{ specifying a transition})$$

(respectively) in that for any  $s \in (2^A)^*$ ,

$$s = bc(s) \iff \text{Mod}_A(s) \models \chi_A^{bc} \quad \text{and} \quad s = d_{\square}(s) \iff \text{Mod}_A(s) \models \chi_A^{d_{\square}}$$

[Fer19b]. Coming back to the  $A$ -interpretation  $\llbracket (\Sigma, s) \rrbracket_A \subseteq \mathcal{L}_A$  above, for stative  $\Sigma$ , we let  $\mathcal{L}_A$  be the image of  $(2^A)^*$  under  $bc$  for the stutterless strings in  $(2^A)^*$

$$\mathcal{L}_A := \{bc(s) \mid s \in (2^A)^*\} = \{s \in (2^A)^* \mid s = bc(s)\} = \{s \in (2^A)^* \mid \text{Mod}_{\Sigma}(s) \models \chi_A^{bc}\}$$

and similarly for transitional  $\Sigma$ , let  $\mathcal{L}_A$  be the image of  $(2^A)^*$  under  $d_{\square}$  for the  $S$ -words of [DS08]

$$\mathcal{L}_A := \{d_{\square}(s) \mid s \in (2^A)^*\} = (2^A - \{\square\})^* = \{s \in (2^A)^* \mid \text{Mod}_{\Sigma}(s) \models \chi_A^{d_{\square}}\}.$$

In either case,  $f_A$  is the identity function on  $\mathcal{L}_A$

$$f_A(s) = s \quad \text{for any } s \in \mathcal{L}_A$$

and the family  $\{f_{\Sigma}\}_{\Sigma \subseteq A}$  of functions indexed by subsets of  $A$  forms a projective system with composition amounting to intersection

$$f_{\Sigma'}; f_{\Sigma} = f_{\Sigma \cap \Sigma'}. \quad (3)$$

As we will see presently, (3) allows us to simplify sets  $r \subseteq 2^A \times \mathcal{L}_A$  of pairs  $(\Sigma, s)$  interpreted at  $A$  by intersection

$$\llbracket r \rrbracket_A := \bigcap_{(\Sigma, s) \in r} \llbracket (\Sigma, s) \rrbracket_A \quad \text{where} \quad \llbracket (\Sigma, s) \rrbracket_A := \{s' \in \mathcal{L}_A \mid f_{\Sigma}(s') = s\}.$$

### 3 First-order fragments for string iconicity

Throughout this section, a choice between stative and transitional  $\Sigma$  is assumed that fixes  $f_\Sigma$  and  $\mathcal{L}_A$  accordingly (using  $\mathcal{b}$  for the stative case, and  $\mathcal{d}_\square$  for the transitional). That choice is made explicit only as needed, to simplify the notation (suppressing the relativization of the notions of  $L$ -traits and  $A$ -profiles defined below to  $\mathcal{b}$  or  $\mathcal{d}_\square$ ). Given a non-empty language  $L \subseteq (2^A)^*$ , an  $L$ -trait is a pair  $(\Sigma, s)$  of a subset  $\Sigma$  of  $A$  and a string  $s \in \mathcal{L}_\Sigma$  such that

$$f_\Sigma(s') = s \text{ for every } s' \in L.$$

The *trait record* of  $L$ ,  $\text{tr}_A(L)$ , is the set of  $L$ -traits. This subset of  $2^A \times \mathcal{L}_A$  is related to any  $r \subseteq 2^A \times \mathcal{L}_A$  by the (anti-tone) Galois connection

$$r \subseteq \text{tr}_A(L) \iff L \subseteq \llbracket r \rrbracket_A$$

recalling from the previous section that  $\llbracket r \rrbracket_A$  is defined as  $\bigcap_{(\Sigma, s) \in r} \{s' \in \mathcal{L}_A \mid f_\Sigma(s') = s\}$ . As the indexed family  $\{f_\Sigma\}_{\Sigma \subseteq A}$  forms a projective system, we can reduce any  $r \subseteq 2^A \times \mathcal{L}_A$  such that  $\llbracket r \rrbracket_A \neq \emptyset$  as follows, without changing its interpretation  $\llbracket r \rrbracket_A$ . Let us call  $r$  an  $A$ -profile if

- (i)  $r$  is functional (deterministic): for all  $(\Sigma, s), (\Sigma, s') \in r$ ,  $s = s'$
- (ii)  $\text{domain}(r)$  is an anti-chain:

$$(\forall \Sigma, \Sigma' \in \text{domain}(r)) \Sigma \subseteq \Sigma' \text{ implies } \Sigma = \Sigma'$$

- (iii) for all  $(\Sigma, s) \in r$ ,  $s \in \mathcal{L}_\Sigma$ .

For example, we can reduce the trait record  $\text{tr}_A(L)$  to two  $A$ -profiles by restricting its domain to subsets of  $A$  with exactly two elements

$$\text{tr}_{2A}(L) := \{(\Sigma, s) \in \text{tr}_A(L) \mid \text{cardinality}(\Sigma) = 2\}$$

or to  $\subseteq$ -maximal subsets

$$\text{mtr}_A(L) := \{(\Sigma, s) \in \text{tr}_A(L) \mid (\forall \Sigma', s') \in \text{tr}_A(L) \Sigma \subseteq \Sigma' \text{ implies } \Sigma' = \Sigma\}$$

without straying from the  $A$ -interpretation of  $\text{tr}_A(L)$

$$\llbracket \text{tr}_{2A}(L) \rrbracket_A = \llbracket \text{tr}_A(L) \rrbracket_A = \llbracket \text{mtr}_A(L) \rrbracket_A.$$

Restricting  $\text{tr}_A(L)$  to singletons may alter its interpretation, as  $L = \{\boxed{a} \boxed{a'}\}$  shows

$$\llbracket \text{tr}_A(L) \rrbracket_A \neq \llbracket \{(\Sigma, s) \in \text{tr}_A(L) \mid \text{cardinality}(\Sigma) = 1\} \rrbracket_A \ni \boxed{a} \boxed{a, a'} \boxed{a'} \text{ for stative } a, a'.$$

For slightly more complicated languages  $L$ , we can delete a pair from  $\text{tr}_{2A}(L)$  without changing its interpretation, as is clear from the implication

$$\mathcal{b}_{\{a, a'\}}(s) = \boxed{\boxed{a} \boxed{a'}} \text{ and } \mathcal{b}_{\{a, a''\}}(s) = \boxed{\boxed{a'} \boxed{a''}} \text{ imply } \mathcal{b}_{\{a, a''\}}(s) = \boxed{\boxed{a} \boxed{a''}}$$

behind the meets-meets entry for the transitivity table of [All83] (Figure 4, page 836) Redundancies in  $\text{tr}_{2A}(L)$  are an argument for preferring the  $A$ -profile  $\text{mtr}_A(L)$ , although there is no denying the popularity of binary relations on intervals labeling interval networks.

A simple notational variant of  $A$ -profiles is provided by formulas  $\varphi$  generated by

$$\varphi ::= s \mid \langle \Sigma \rangle \varphi \mid \varphi \wedge \varphi' \quad (\Sigma \subseteq A) \quad (4)$$

driving the semantic interpretations

$$\llbracket s \rrbracket_A := \{s\} \cap \mathcal{L}_A \quad \llbracket \langle \Sigma \rangle \varphi \rrbracket_A := \{s \in \mathcal{L}_A \mid f_\Sigma(s) \in \llbracket \varphi \rrbracket_A\} \quad \llbracket \varphi \wedge \varphi' \rrbracket_A := \llbracket \varphi \rrbracket_A \cap \llbracket \varphi' \rrbracket_A$$

matching  $A$ -profiles

$$\{\llbracket \varphi \rrbracket_A \mid \varphi \text{ from (4)}\} = \{\llbracket r \rrbracket_A \mid r \text{ is an } A\text{-profile}\} \cup \{\emptyset\}.$$

A larger family of languages is expressed by *first-order* formulas  $\varphi$  generated by

$$\varphi ::= P_a(x) \mid x < y \mid x = y \mid \neg \varphi \mid \varphi \wedge \varphi' \mid \exists x \varphi \quad (a \in A) \quad (5)$$

with  $<$  in place of  $S$ , to match *star-free* expressions (such as strings  $s$ ), obtained from regular expressions by substituting complementation for Kleene star  $\cdot^*$  (e.g., [Lib04], §7.5). The clause for  $\langle \Sigma \rangle \varphi$  in (4) is mirrored by relativizing a formula  $\varphi$  from (5) to  $\varphi_\Sigma$  by induction

$$\begin{aligned} \varphi_\Sigma &:= \varphi \text{ for atomic } \varphi & (\neg \varphi)_\Sigma &:= \neg(\varphi_\Sigma) & (\varphi \wedge \psi)_\Sigma &:= \varphi_\Sigma \wedge \psi_\Sigma \\ (\exists x \varphi)_\Sigma &:= \exists x(V_\Sigma(x) \wedge \varphi_\Sigma) \text{ where } V_\Sigma(x) &:= \begin{cases} \forall y(xSy \supset \Delta_\Sigma(x, y)) & \text{for stative } \Sigma \\ \bigvee_{a \in \Sigma} P_a(x) & \text{for transitional } \Sigma \end{cases} \end{aligned}$$

supporting the induction hypothesis (on  $\varphi$ ) that

$$\text{Mod}(s), g \models \varphi_\Sigma \iff \text{Mod}(f_\Sigma(s), \hat{g} \models \varphi$$

for all  $s \in (2^A)^*$  and assignments  $g : \text{Var} \rightarrow [\text{length}(s)]$  of variables  $x$  to  $g(x) \in \llbracket V_\Sigma \rrbracket_s$

$$\text{Mod}(s), g \models V_\Sigma(x) \text{ for each } x \in \text{Var}$$

(where  $\hat{g}$  is the obvious  $f_\Sigma$ -adjustment of  $g$ ). Missing from (4) above is a clause for negation, yielding disjunction (by conjunction and De Morgan). It is questionable whether negation or disjunction preserves form-meaning resemblance, which negation denies and spurious disjuncts spoil. (Hence, *no*  $\neg$  or  $\vee$  in (4).) While any non-empty subset of Allen interval relations may label an arc in an interval network, a move away from arbitrary disjunctions towards neighborhoods of similar configurations is advocated in [Fre92], §2.2. Forms  $(\Sigma, s)$  of strings with  $\Sigma$ -projection  $s$  represent such neighborhoods. Beyond (4), we can generalize  $f_\Sigma$  to mereological relations  $R$  (e.g., subsumption  $\triangleright$ ; [Fer15b]) to interpret accessibility in iconic forms  $\diamond_{R\varphi}$ , producing non-determinism via underspecification (as opposed to overspecified disjunction).

## 4 Conclusion

The significance of iconicity, resemblance of form to meaning, is an issue open to dispute even in pictorial representation, positions about which differ based on how the issue is framed (e.g., [Gre13]). Broadly construed, iconicity points to transparent representations that facilitate the communication and inferential processing of information up to a limited but useful extent. The present work applies that broad construal to the experience of change, proposing representations around string iconicity. For any non-empty set  $L$  of strings of sets, an iconic star-free

approximation,  $\text{mtr}_A(L)$ , of  $L$  is formed from maximal  $L$ -traits, eschewing arbitrary disjunctions (inimical to iconicity and underspecification). I close by drawing attention to two tensions running through the work above.

The first is between iconicity and abstraction.  $A$ -profiles are iconic expressions carving out a small fragment of first-order formulas in MSO, as the grammars (4) and (5) bring out. Exactly what the  $\Sigma$ -projections  $f_\Sigma$  (behind  $A$ -profiles) come to, for stative and transitional  $\Sigma$ , are described above using first-order formulas  $V_\Sigma(x)$  that fall outside the fragment generated by (4). That is, decidedly non-iconic MSO formulas are employed to explain iconic expressions. Similarly, pictures (exemplifying iconicity) can be described, when paired with viewpoints, by predicates that are (in some cases) stative (in accordance with [Abu14]’s argument that “all pictures have stative informational content”) and (in other cases) non-stative, accounting for temporal progression under an inertial constraint of change requiring force/action ([Fer19a]). The basic claim of [Fer19a] is that for all its abstractness (non-iconicity), MSO is a helpful tool in breaking down and piecing together the meaning of pictorial narratives.

The second tension is between the temporal ontology expressed by the successor predicate  $S$  and the vocabulary  $\Sigma$  for articulating a particular perspective.  $S$  and  $\Sigma$  are linked by  $V_\Sigma(x)$ , transparently in the stative case (“no  $\text{step}_S$  without  $\text{change}_\Sigma$ ”) and less overtly in the transitional case (“no time without  $\text{change}_\Sigma$ ”),  $S$  being buried in what it means for an element of  $\Sigma$  to be transitional. With respect to iconicity of order,  $\Sigma$  may represent contextual dimensions other than time, constituting, for example, viewpoints in viewpoint-centered propositions for pictorial narratives. It is not unreasonable to assume that “the order of elements in language” (natural or formal) is significant, even if that is complicated by a wealth of factors behind “physical experience or the order of knowledge.” These factors presumably shape discourse structure, used in [AL03] to explain *push-fall* temporal reversals and other challenges to order iconicity, as well as types recording the interaction between speakers and listeners around questions under discussion ([CG15]). [Ste05]’s proposal that there is more to temporality in natural language than temporal order suggests an understanding of iconicity of order that subordinates time given by  $S$  to notions of causality and contingency expressed through  $\Sigma$ .

## References

- [Abu14] D. Abusch. Temporal succession and aspectual type in visual narrative. In L. Crnić and U. Sauerland, editors, *The Art and Craft of Semantics: A Festschrift for Irene Heim*, volume 1, pages 9–29. MIT Working Papers in Linguistics 70, 2014.
- [AR17] D. Abusch and M. Rooth. The formal semantics of free perception in pictorial narratives. In *Proc. 21st Amsterdam Colloquium*, pages 85–95. University of Amsterdam, 2017.
- [All83] J. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.
- [AL03] N. Asher and A. Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.
- [CG15] R. Cooper and J. Ginzburg. TTR for natural language semantics. In S. Lappin and C. Fox, editors, *Handbook of Contemporary Semantic Theory*, pages 375–407. Wiley-Blackwell, second edition, 2015.
- [Dow79] D.R. Dowty. *Word Meaning and Montague Grammar*. Reidel, Dordrecht, 1979.
- [Dow86] D.R. Dowty. The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy*, 9(1):37–61, 1986.
- [DS08] I.A. Durand and S.R. Schwer. A tool for reasoning about qualitative temporal information: the theory of S-languages with a Lisp implementation. *Journal of Universal Computer Science*, 14(20):3282–3306, 2008.

- [Fer15a] T. Fernando. Negation and events as truthmakers. In *Proc. 20th Amsterdam Colloquium*, pages 109–118. University of Amsterdam, 2015.
- [Fer15b] T. Fernando. The semantics of tense and aspect: A finite-state perspective. In S. Lappin and C. Fox, editors, *Handbook of Contemporary Semantic Theory*, pages 203–236. Wiley-Blackwell, second edition, 2015.
- [Fer16] T. Fernando. Prior and temporal sequences for natural language. *Synthese*, 193(11):3625–3637, 2016.
- [Fer19a] T. Fernando. Pictorial narratives and temporal refinement. *Proc. 29th Semantics and Linguistic Theory*, 2019. Draft available at [www.scss.tcd.ie/Tim.Fernando/salt29.pdf](http://www.scss.tcd.ie/Tim.Fernando/salt29.pdf).
- [Fer19b] T. Fernando. Projecting temporal properties, events and actions. In *Proc. 13th International Conference on Computational Semantics*, pages 1–12. [www.aclweb.org/anthology/W19-0401](http://www.aclweb.org/anthology/W19-0401), 2019.
- [FWV19] T. Fernando, D. Woods, and C. Vogel. MSO with tests and reducts. *Proc. 14th International Conference on Finite-State Methods and Natural Language Processing*, pages 27–36, 2019. [www.aclweb.org/anthology/W19-3106/](http://www.aclweb.org/anthology/W19-3106/).
- [Fil70] C.J. Fillmore. The grammar of hitting and breaking. In R.A. Jacobs and P.S. Rosenbaum, editors, *Readings in English Transformational Grammar*, pages 120–133. Ginn, 1970.
- [Fok07] W. Fokink. *Introduction to Process Algebra*. Springer, 2nd edition, 2007. [www.cs.vu.nl/~wanf/BOOKS/procalg.pdf](http://www.cs.vu.nl/~wanf/BOOKS/procalg.pdf).
- [Fre92] C. Freksa. Temporal reasoning based on semi-intervals. *Artificial Intelligence*, 54:199–227, 1992.
- [GJW18] P. Gärdenfors, J. Jost, and M. Warglien. From actions to effects: Three constraints on event mappings. *Frontiers in Psychology*, 9:1391, 2018.
- [GB92] J.A. Goguen and R.M. Burstall. Institutions: abstract model theory for specification and programming. *Journal of the ACM*, 39(1):95–146, 1992.
- [Gre13] G. Greenberg. Beyond resemblance. *Philosophical Review* 122(2), 215–287, 2013.
- [Gre63] J. Greenberg. *Universals of Language*. MIT Press, 1963.
- [Ham71] C.L. Hamblin. Instants and intervals. *Studium generale*, 24:127–134, 1971.
- [HKT00] D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic*. MIT Press, 2000.
- [KR93] H. Kamp and U. Reyle. *From Discourse to Logic*. Kluwer, 1993.
- [KS96] D. Kozen and F. Smith. Kleene algebra with tests: Completeness and decidability. In *Proc. 10th Int. Workshop Computer Science Logic*, LNCS 1258, pages 244–259. Springer, 1996.
- [LRH13] B. Levin and M. Rappaport Hovav. Lexicalized meaning and manner/result complementarity. In B. Arsenijević, B. Gehrke, and R. Marín, editors, *Subatomic Semantics of Event Predicates*, pages 49–70. Springer, 2013.
- [Lib04] L. Libkin. *Elements of Finite Model Theory*. Springer, 2004.
- [MB19] E. Maier and S. Bimpikou. Shifting perspectives in pictorial narratives. In *Proc. Sinn und Bedeutung 23(2)*, pages 91–106. Barcelona, 2019.
- [New92] F.J. Newmeyer. Iconicity and generative grammar. *Language*, 68(4):756–96, 1992.
- [PV14] P. Perniss and G. Vigliocco. The bridge of iconicity: from a world of experience to the experience of language. *Phil. Trans. R. Soc. B*, 369: 20130300, 2014.
- [Pri67] A.N. Prior. *Past, Present and Future*. Clarendon Press, Oxford, 1967.
- [Ste05] M. Steedman. The Productions of Time, 2005. Draft tutorial notes about temporal semantics, <http://homepages.inf.ed.ac.uk/steedman/papers.html>.
- [vLH05] M. van Lambalgen and F. Hamm. *The Proper Treatment of Events*. Blackwell, 2005.
- [Zim00] T.E. Zimmermann. Free choice disjunction and epistemic possibility. *Natural Language Semantics*, 8:255–290, 2000.



# Time matters: the role of temporal boundaries in NPI licensing\*

Julie Goncharov<sup>1</sup> and Lavi Wolf<sup>2</sup>

<sup>1</sup> The Arctic University of Northern Norway, Tromsø, Norway  
julie.goncharov@uit.no

<sup>2</sup> Ben-Gurion University of the Negev, Beer-Sheva, Israel  
wolfla@post.bgu.ac.il

## Abstract

This paper discusses double Negative Polarity Item (NPI) constructions, such as *\*I don't think at all John will leave until 10pm*. These constructions have been claimed to provide a strong argument for a syntactic approach to Neg-raising (NR) (Lakoff, 1969; Prince, 1976; Crowley, 2019). We show that the empirical landscape is more intricate than what has been reported in the literature. As a result, far from supporting the syntactic approach to NR (Fillmore, 1963; Collins and Postal, 2014, a.o.), double NPI constructions provide strong evidence against it and in favour of a semantic/pragmatic approach (Bartsch, 1973; Gajewski, 2005; Romoli, 2012, a.o.). We propose an account for a subset of double NPI constructions capitalizing on three ideas: (i) punctual *until* has a non-cancellable (modal) inference that the action occurs after the time specified by *until* (Karttunen, 1974; Condoravdi, 2008; Iatridou and Zeijlstra, to appear), (ii) epistemic accessibility relations include a time index as well as a world one, and (iii) NPI-like predicate modifiers, such as *at all*, can be re-conceptualized from a domain-widener to an inhibitor of a contextual domain restriction.

## 1 Introduction

Several linguists have pointed out the paradigm in (1), e.g. Lakoff 1969; Prince 1976; Crowley 2019. In (1a), the strict/strong Negative Polarity Item (NPI) punctual *until* is licensed under the negated predicate *think*.<sup>1</sup> In (1b), *ever* and *at all*, two other NPIs, are licensed in the same configuration in the matrix clause. However, when both the low and the high NPI are present, (1c), the sentence becomes ungrammatical. We dub the constructions, such as (1c), *double NPI constructions*. They will be the focus of our investigation here.

- (1) a. I didn't think [John would arrive until 10pm]  
b. I didn't ever/at all think John would arrive  
c. \*I didn't ever/at all think [John would arrive until 10pm]

---

\*We would like to thank Paul Egré, Jon Gajewski, Sabine Iatridou, Gillian Ramchand, the audiences at OASIS 2 and NELS 50, as well as the anonymous reviewers for AC, for their stimulating questions, helpful remarks and comments. All errors are our own.

<sup>1</sup>*Until* was traditionally described as being ambiguous between a durative and a punctual reading. Recently, Iatridou and Zeijlstra (to appear) have provided arguments for a unified analysis of *until*. As only the punctual *until* shows NPI properties, we will keep its specification as punctual, for clarity. As an NPI, punctual *until* has the distribution of the so-called strict/strong NPIs, which is more restricted compared to weak NPIs like *any* or *ever*, see Gajewski 2011; Chierchia 2013; Collins and Postal 2014 for recent accounts.

Double NPI constructions have been claimed to provide a strong argument in favour of a syntactic approach to Neg-raising (NR). NR is a phenomenon in which with certain predicates, like *think*, *believe*, *expect*, and others, matrix negation can be interpreted in the embedded clause, (2). According to the syntactic approach to NR (Fillmore, 1963; Collins and Postal, 2014, a.o.), NEGation originates in the embedded clause and then undergoes syntactic movement to the matrix clause, (3). The NR-reading arises when the low copy of NEG is interpreted.

- (2) John didn't think/believe/expect it would snow  
 $\leadsto$  John thought/believed/expected it wouldn't snow (NR-reading)
- (3) [ John did NEG think [ it would NEG snow ] ]

Under the syntactic approach, the punctual *until* in (1a) is locally (clause-mate) licensed by NEG in the embedded clause, whereas *ever/at all* in (1b) is licensed by the moved NEG in the matrix clause. However, (1c) is unacceptable because, depending on a particular implementation, one of the NPIs remains unlicensed.<sup>2</sup>

An alternative to the syntactic approach to NR is a semantic/pragmatic approach, which derives the embedded reading of negation from the Excluded Middle (EM) inference (Bartsch, 1973; Gajewski, 2005; Romoli, 2012, a.o.). For instance, the assertion in (4a) together with the EM in (4b) entails the NR-reading in (4c).

- (4) a. John didn't think it would snow  
 b. EM: John thought it would snow or John thought it wouldn't snow  
 c.  $\leadsto$  John thought it wouldn't snow (NR-reading)

This approach has had little-to-nothing to say about the deviance of (1c) with the exception of Gajewski 2005. Gajewski (2005, 71-2) suggests that double NPI constructions are unacceptable because with *ever/at all* in the matrix clause, the EM projects existentially rather than universally, which does not guarantee the entailment in (4c) necessary for licensing *until*.

Both the syntactic and the semantic/pragmatic approach to NR explain the ungrammaticality of double NPI constructions by appealing to the way NR is derived. This reasoning predicts that the acceptability of double NPI constructions does not depend on a particular NPI in the embedded and/or matrix clause. For the syntactic approach, what is important is that the low NPI is a strict/strong NPI, which requires a clause-mate negation. The punctual *until* is one such NPI. Other strict/strong NPIs and minimizers, like *in weeks*, *a single soul*, and *sleep a wink*, are predicted to be equally unacceptable in double NPI constructions. For the semantic/pragmatic approach, in addition to the above prediction, the high NPI should be an existential quantifier. In other words, the semantic/pragmatic approach predicts, for example, that substituting *not..ever* in (1c) by *never* should make (1c) fully acceptable for the EM under *never* projects universally and nothing bars the entailment necessary for licensing a strong NPI in the embedded clause.

In the next section, we see that the distribution of double NPI constructions is more complex than what has been reported in the literature and the predictions of neither the syntactic nor the semantic/pragmatic approach to NR are borne out. More specifically, we present the results of a survey study showing that the acceptability of double NPI constructions does depend on the choice of an NPI in the embedded as well as the matrix clause. No amendments can save the syntactic approach to NR since there is no difference between the syntactic licensing of different strict/strong NPIs, e.g. *until* vs. *in weeks*. Therefore, far from being a strong argument in favour

<sup>2</sup>Arguably, this reasoning is supported by examples like *John didn't ever arrive until (after) 10pm*, where the two NPIs are licensed when occurring in the same clause.

of the syntactic approach to NR, double NPI constructions, in fact, provide strong grounds for eliminating this approach (for other arguments against the syntactic approach to NR see e.g. Zeijlstra 2017). In section 3, we propose an account for a subset of double NPI constructions, which is compatible with the semantic/pragmatic approach to NR. Section 4 concludes the paper.

## 2 The intricate empirical landscape

### 2.1 Double NPI constructions in the literature

To the best of our knowledge, double NPI constructions have not been studied as a phenomenon in itself. They have been mentioned as a side (or footnote) phenomenon in the literature on NR and are mostly restricted to the paradigm in (1). Below, we briefly review known to us studies which deal with double NPI constructions indicating the judgments provided by the authors. The upshot of this review is that the empirical grounds are not settled. First, the data sample is not representative as it usually involves the same low and high NPIs (*until* and *ever/at all*). Second, the grammaticality judgments are often controversial or even contradictory.

Lakoff (1969, 142) mentions the contrast between (5a) and (5b-d) and suggests that it casts doubt on the rule of Negative Transportation (Fillmore, 1963).

- (5) a. I didn't think John would leave until tomorrow  
 b. \*I didn't ever think that John would leave until tomorrow  
 c. \*I never thought that John would leave until tomorrow  
 d. \*At no time did I think that John would leave until tomorrow

However, Lakoff (1969) also provides the examples in (6), which have a similar structure, but are grammatical, according to her judgements. She does not propose an explanation for the contrast between (5a) and (5b-d), nor between (5b-d) and (6).

- (6) a. No one thought that John would leave until tomorrow  
 b. It wasn't thought by anyone that John would leave until tomorrow

Seuren (1974, reprinted as Seuren 2001) repeats Lakoff's examples in (5b-d) and adds the examples in (7):

- (7) a. \*I didn't think yet that Fred would get here until midnight  
 b. \*Often I don't think he has got it yet  
 c. \*He usually doesn't think there is as much as a shred of evidence

Seuren (1974) attributes the ungrammaticality of (5b-d) and (7) to the presence of the adverb in the matrix clause that blocks NR. He also points out that it is plausible that there is a deeper semantic reason for the blocking effect, but leaves this reason unexplored.

Prince (1976, fn. 7) independently points out the examples in (8) attributing them to Richard Smaby. For her, the ungrammaticality of (8c) is predictable in terms of the syntactic (raising) analysis of NR, the requirement on *until* to be clause-mate licensed, and *at all* to be under the scope of negation.

- (8) a. I don't at all think that John will leave  
 b. I don't think that John will leave until next week

- c. \*I don't at all think that John will leave until next week

The subsequent literature (Gajewski, 2005; Crowley, 2019) uses Prince's paradigm in (8) for their main (although opposite) claims. Moreover, Gajewski (2005, fn.2) challenges Lakoff's judgements (as well as her conclusions) reporting that according to his informants, (5c) is grammatical whereas (6b) is ungrammatical:

- (9) Lakoff 1969  
 a. It wasn't thought by anyone that John would leave until tomorrow  
 b. \*I never thought that John would leave until tomorrow
- (10) Gajewski 2005  
 a. \*It wasn't thought by anyone that John would leave until tomorrow  
 b. I never thought that John would leave until tomorrow

Overall, the literature does not provide a clear empirical picture. Therefore, we conducted a pilot study to investigate the distribution of double-NPI constructions.

## 2.2 Pilot study

Participants were recruited via Amazon Mechanical Turk (MTurk). Overall 28 participants took part in the study. We excluded subjects who did not complete the survey, failed an attention task or took less than 2 minutes to complete the survey. Nine participants in total were excluded.

In each trial, a participant was presented with a sentence containing target items and then asked to rank its level of acceptability on a 5 point Likert scale where 1 is the least acceptable and 5 is the most acceptable. There were 3 practice trials, 32 experimental trials, out of which 11 were fillers not containing any target items, and one trial checking the participant's attention. Target items and fillers were pseudo-randomized.

Four items were designed to check baseline judgements. They contained the NPIs *until* and *in weeks/months* embedded under *think*, in either a negative environment with a wide scope negation or a non-negative environment.<sup>3</sup> The baseline items are shown in (11) with the mean judgements in square brackets, see also Figure 1. The standard prediction that strict/strong NPIs are licensed under negated Neg-raising predicates was borne out.

- (11) a. I don't think that Mary will arrive until Thursday [4.63/5]  
 b. I don't think that John has been here in weeks [4.31/5]  
 c. I think that Sue will leave until next week [1.89/5]  
 d. I think that Mary has visited us in months [1.73/5]

The other trials in the survey dealt with double NPI constructions and tested for the combination of high NPIs *ever*, *any*, *never*, *yet* with low NPIs *until*, *in days/weeks/years*. The items with corresponding mean judgements are shown in (12) and Figure 2.

- (12) a. I didn't ever think that Harry would arrive until Monday [3.42/5]

<sup>3</sup>The survey also originally contained sentences with *sleep a wink*. The data on this NPI manifested a pattern which did not fit any of the theory-related predictions on NPIs, including high acceptance values in environments that do not license NPIs and low acceptance values in environments that do license NPIs. Since accounting for this puzzling behavior is beyond the scope of this paper, the *sleep a wink* items were not taken into consideration.

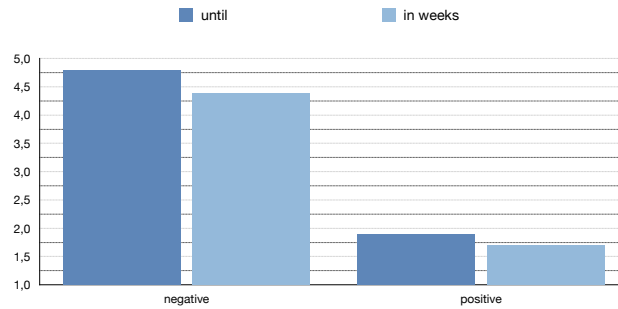


Figure 1: Baseline

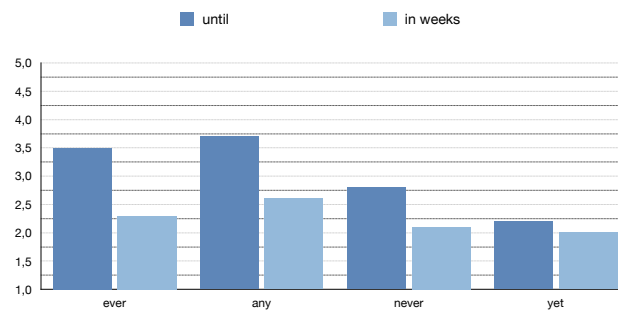


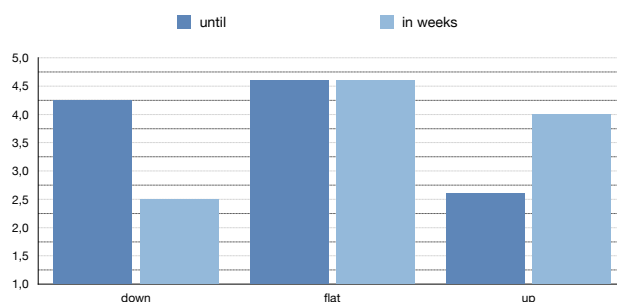
Figure 2: Different combinations of low and high NPIs

- b. I didn't ever think that Kate had been here in years [2.26/5]  
 c. I don't believe to any degree that Miriam will leave until tomorrow [3.57/5]  
 d. I don't believe to any degree that Kyle has come in days [2.52/5]  
 e. I never believed that Dan arrived until Friday [2.73/5]  
 f. I never believed that Charlie had visited us in weeks [2.05/5]  
 g. I didn't think yet that Jack would get here until midnight [2.1/5]  
 h. I didn't think yet that Laura had been here in weeks [2/5]

The relatively high acceptability of *until* with *ever/to any degree*, compared to *in weeks*, goes against the predictions of the syntactic approach to NR. In addition, the low acceptability of sentence with *never* compromises the semantic/pragmatic approach.

An interesting picture emerges when *at all* is used in the matrix clause of double NPI constructions. This picture, as the findings above, cannot be explained by the syntactic approach to NR. The configuration with *at all* shows three distinct patterns depending on whether *until* or *in weeks* is used in the embedded clause. We label these patterns 'Down' (*until* is (significantly) more acceptable than *in weeks*), 'Flat' (both *until* and *in weeks* are acceptable), and 'Up' (*in weeks* is more acceptable than *until*). The items with their means are given in (13) and illustrated in Figure 3. In the next section, we will concern ourselves with this observation.

- (13) a. I don't believe at all that Eric will leave until Wednesday [4.25] [4.6] [2.6]  
 b. I don't believe at all that Fred has talked to Mary in weeks [2.5] [4.6] [4]  
           'Down' 'Flat' 'Up'

Figure 3: Three groups for *not...at all...until/in weeks*

## 2.3 Summary and outlook

After inspecting the data on double NPI constructions, the picture which arises is of a landscape more intricate than what the literature has had. Different NPIs do not behave alike in double NPI constructions. This is contrary to the predictions of the syntactic approach to NR. Moreover, the syntactic approach has no way of accounting for the diverse behaviour of *until* and *in week* that we saw above. Therefore, double NPI constructions pose a serious problem to the syntactic approach to NR. The semantic/pragmatic approach is also undermined as the judgments it is based on do not prove to be robust.

## 3 Proposal for *not...at all...until/in weeks*

In this section, we propose an account for a subset of double NPI constructions. More precisely, we look at the configuration where *at all* is used in the matrix clause and *until* or *in weeks* is used in the embedded clause. This configuration shows three patterns as illustrated in Figure 3. We begin by explaining the pattern labelled ‘Up’. The ‘Up’ pattern is exemplified in (14).

- (14) a. ??I don’t think at all that John arrived until Friday [2.6/5]  
 b. I don’t t hink at all that John has been here in weeks [4/5]

Our proposal capitalizes on three ideas: (i) punctual *until* has a non-cancellable (modal) inference that the action occurs after the time specified by *until* (Karttunen, 1974; Condoravdi, 2008; Iatridou and Zeijlstra, to appear), (ii) epistemic accessibility relations include a time index as well as a world one, and (iii) NPI-like predicate modifiers, such as *at all*, can be re-conceptualized from a domain-widener to an inhibitor of a contextual domain restriction. We elaborate on these ideas in the next three sections (sections 3.1-3.3). In section 3.4, we attend to the remaining two patterns ‘Down’ and ‘Flat’ in Figure 3 and suggest a tentative explanation for them.

### 3.1 Change of State Inference of *until* (CoSI)

The observation that punctual *until* has a Change of State Inference (CoSI) goes back at least to Karttunen 1974. He points out that sentences like (15a) have the presupposition in (15b), which explains the infelicity of (15c). The status of the inference in (15b) (presupposition vs. non-cancellable implicature) is controversial (see Iatridou and Zeijlstra to appear for the discussion). simply as an inference.

- (15) a. John didn't arrive until Friday  
 b.  $\rightsquigarrow$  John arrived after Friday (CoSI)  
 c. # John didn't arrive until Friday and maybe he didn't arrive at all

Following Iatridou and Zeijlstra (to appear), we assume that the non-cancellable CoSI of *until* is modal, i.e. it can be satisfied by a non-actual world, when the proposition is embedded under a modal. This assumption resolves the tension between the infelicitous (15c) and the felicitous (16), where John's leaving after Friday can be true in some non-actual future world.

- (16) John won't leave until Friday, if at all

### 3.2 *Think* and the satisfaction of CoSI

The key point of our proposal is that for the modal CoSI of *until* to be satisfiable, the set of worlds over which an attitude predicate like *think* ranges must be restricted along the time index. To model our account, we take the universal quantifier associated with epistemic attitudes like *think* to range over  $\langle w, t \rangle$ -pairs as shown in (18), where the time span of the doxastic accessibility relation  $R_x$  is contextually restricted by the selection function  $S_{[m,n]}$ .<sup>4</sup> This selection function can be thought of as analogous to the ordering source in Kratzerian double-base semantics for modals (Kratzer, 1991, 2012). We assume that the flow of time is held fixed in a rigid manner due to the doxastic accessibility relation's reliance on a *stereotypical* ordering source, whose assumptions adhere to a stereotypical flow of time across the relevant worlds. The intuition is that '*x thinks  $\phi$* ' does not make a statement about all time intervals in *x*'s life, but is rather defined over an implicitly contextual time interval. Furthermore, we propose that when  $\phi$  contains punctual *until*, the domain of quantification of *think* is further contextually restricted on the right temporal boundary to align with the time specified by *until*, (19).

(17) **Notations used in this paper:**

$S_{[m,n]}$  = selection function restricting a set of  $\langle w, t \rangle$ s to  $t \in [m, n]$

$R_x$  = doxastic accessibility relation

$\langle w_c, t_c \rangle$  = actual  $\langle w, t \rangle$

$g1, g2$  = contextual time points such that  $g1 | \dots | \mathbf{10pm} | \dots | g2$

$C_{[m,n]}$  = time interval concept (i.e.  $g(C_{[g1,g2]}) = \lambda w. \text{ a time interval in } w \text{ with } g1 \text{ and } g2 \text{ as its left and right boundaries}$ )

$\tau$  = event time span

- (18)  $\text{think}_x \phi = 1$  iff  $\forall \langle w, t \rangle \in S_{[g1,g2]}(R_x(\langle w_c, t_c \rangle)) : \langle w, t \rangle \in \phi$

- (19)  $\neg \text{think}_x \phi[\text{until } \mathbf{10pm}] = 1$  iff  $\neg \forall \langle w, t \rangle \in S_{[g1,\mathbf{10pm}]}(R_x(\langle w_c, t_c \rangle)) : \langle w, t \rangle \in \phi$

The intuition (19) captures is that *until* makes statements like '*x doesn't think  $\phi$* ' irrelevant after the time specified by *until* when the epistemic state is updated and the matter ( $\phi$  or  $\neg\phi$ ) is settled. More importantly, the additional domain restriction in (19) is necessary to ensure that CoSI is satisfied. (20) shows the assertion for (1a). (20a) states that it is not the case that in all speaker's belief  $\langle w, t \rangle$  pairs such that  $t$  spans  $[g1, \mathbf{10pm}]$ , John arrives before 10pm. (20b) spells out the CoSI of *until*, which says that from any speaker's belief  $\langle w, t \rangle$  with  $t \in [g1, \mathbf{10pm}]$ , there is an epistemically accessible  $\langle w', t' \rangle$  in which John arrives after 10pm. Crucially, if *think* ranges over  $\langle w, t \rangle$  with  $t \in [g1, g2]$  (unrestricted by *until*), CoSI cannot be satisfied since from

<sup>4</sup>We retain the standard term *accessibility relation* although strictly speaking we represent this relation as a function.



$\langle w, t \rangle$  with  $t \in [10pm, g2]$  where John arrived *before* 10pm,  $\langle w', t' \rangle$  where John arrives *after* 10pm is not accessible, (20c).

- (20) a. Assertion of (1a):  $\neg \forall \langle w, t \rangle \in S_{[g1, 10pm]}(R_{sp}(\langle w_c, t_c \rangle)) :$   
 $\langle w, t \rangle \in \exists \tau [\tau = C_{[t, 10pm]}(w) \wedge \exists e [\text{arrive}(e, j) \wedge e \subseteq \tau \wedge e < w]]$
- b. CoSI of *until*:  $\forall \langle w, t \rangle \in S_{[g1, 10pm]}(R_{sp}(\langle w_c, t_c \rangle)) \exists \langle w', t' \rangle \in R_{sp}(\langle w, t \rangle) :$   
 $\langle w', t' \rangle \in \exists \tau [\tau = C_{[10pm, g2]}(w') \wedge \exists e [\text{arrive}(e, j) \wedge e \subseteq \tau \wedge e < w']]$
- c. #CoSI of *until*:  $\forall \langle w, t \rangle \in S_{[g1, g2]}(R_{sp}(\langle w_c, t_c \rangle)) \exists \langle w', t' \rangle \in R_{sp}(\langle w, t \rangle) :$   
 $\langle w', t' \rangle \in \exists \tau [\tau = C_{[10pm, g2]}(w') \wedge \exists e [\text{arrive}(e, j) \wedge e \subseteq \tau \wedge e < w']]$

Informally, when an addressee contemplates the assertion of (1a), they contemplate the possibility of John arriving after 10pm. In order to do that, it must be the case that the contemplation is done *while there is something to contemplate* i.e. while the matter is not yet established. This is the work done by *until* - marking the point in time after which new information affects the ordering source in a way that makes further contemplating a matter which has already been settled, infelicitous.

### 3.3 *At all* as predicate modifiers that inhibit domain restriction

*At all*, in *John is \*(not) tired at all.*, is standardly analyzed as a property modifier that triggers domain widening (in the sense of Kadmon and Landman (1993)) by obligatorily introducing non-exhaustive alternatives, which explains its NPI status (Krifka, 1995; Chierchia, 2013, a.o.). In other words, *at all* requires to consider even minimal degrees of precision for tiredness, which are normally disregarded as pragmatically irrelevant. We propose that classical domain widening by *at all* can be re-conceptualized as a ban on (contextual) domain restriction. That is to say, instead of requiring to consider even the minimal degrees of tiredness, *at all* inhibits domain restriction to pragmatically relevant degrees of tiredness.

- (21) a. John is \*(not) tired at all.  
b. Alternatives for *at all*:  $\{ P \mid P \subset \mathbf{tired} \wedge \neg \min(P) \}$  (simplified)  
c. Non-exhaustivity:  $\cup \{ P \mid P \subset \mathbf{tired} \wedge \neg \min(P) \} \subset \mathbf{tired}$

With this re-conceptualization, it is easy to see how the paradigm in (1) is explained: (1a) has no *at all* and is shown in (20a,b). In (1b), *at all* requires that *think* ranges over all world-time pairs with an unrestricted time interval. This is unproblematic as there is no *until*. However, when both *until* and *at all* are present, (1c), the domain of quantification of *think* cannot be restricted by *until* for *at all* requires the widest possible domain and CoSI cannot be satisfied, (20c), which explains the infelicity of (1c). As *in weeks* does not have CoSI, we also explain the contrast in (14).

### 3.4 A note on three groups

Above, we proposed an explanation why in double NPI constructions, *until* is less acceptable than *in weeks*. We argued that the culprit is the non-cancellable Change of State inference of *until*. However, in Figure 3, we saw that there are actually three patterns, which we labelled ‘Down’, ‘Flat’, and ‘Up’.

We tentatively suggest that these three patterns represent distinct groups of population that differ with respect to their pragmatic strategies. Recently, Crnić et al. (2018) argue that the high error rate (>80%) in experiments concerning logical reasoning may be due to not



taking into account different reasoning strategies. They provide experimental evidence that at least three groups of people can be distinguished with respect to whether they compute scalar implicatures in premises, conclusions, or both in syllogistic arguments, see Table 1.

	<b>Logicians</b> don't compute SIs	<b>Validators</b> SIs for premises, but not conclusions	<b>Strengtheners</b> SIs for premises and conclusions
premise	weak	strong	strong
conclusion	weak	weak	strong

Table 1: Different reasoning strategies (Crnić et al., 2018)

Given these findings, the difference between ‘Up’-group, on the one hand, and ‘Flat’ and ‘Down’ groups, on the other hand, may be due to the fact that the former but not the latter compute CoSI for *until* when judging the naturalness of a sentence. More research needs to be done (including more experimental work) to confirm this suggestion as well as to get to the bottom of the difference between the ‘Flat’ and the ‘Down’ patterns.

## 4 Conclusion

This paper’s contribution to the literature on negation, via an examination of double NPI constructions, is both theoretical and empirical. Our glimpse into the empirical landscape yielded a much intricate picture than previously reported, favoring the semantic/pragmatic approach to Neg-raising over the syntactic one. Our theoretical account provided for a subset of double NPI constructions, in which both time and world indices play a role in epistemic accessibility relations, offers both a semantic-pragmatic way to account for these constructions as well as a re-conceptualization of NPI-like predicate modifiers as contextual domain restriction inhibitors rather than domain-wideners.

## References

- Renate Bartsch. Negative transportation gibt es nicht. *Linguistische Berichte*, 27:1–7, 1973.
- Gennaro Chierchia. *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press, Oxford, 2013.
- Chris Collins and Paul M. Postal. *Classical NEG raising*. The MIT Press, Cambridge MA, 2014.
- Cleo Condoravdi. Punctual until as a scalar NPI. In K. Hanson and S. Inkelas, editors, *The Nature of the Word*, pages 631–653. MIT Press, Cambridge, 2008.
- Luka Crnić, Andreas Haida, and Yosef Grodzinsky. Linguistic barriers to logical reasoning: a new perspective on aristotelian syllogisms. *Proceedings of Sinn und Bedeutung 22*, 2018.
- Paul Crowley. Neg-raising and neg movement. *Natural Language Semantics*, 27:1–17, 2019.
- Charles F. Fillmore. The position of embedding transformations in a grammar. *Word*, 19: 208–231, 1963.

- Jon Gajewski. *Neg-Raising: Polarity and presupposition*. PhD thesis, MIT, 2005.
- Jon Gajewski. Licensing strong NPIs. *Natural Language Semantics*, 19:109–148, 2011.
- Sabine Iatridou and Hedde Zeijlstra. The complex beauty of boundary adverbials: in years and until. *Linguistic Inquiry*, to appear.
- Nivit Kadmon and Fred Landman. Any. *Linguistics and Philosophy*, 16:353–422, 1993.
- Lauri Karttunen. Until. *Chicago Linguistic Society*, 10:283–297, 1974.
- Angelika Kratzer. Modality. In Arnim von Stechow and Dieter Wunderlich, editors, *Semantics: An international handbook of contemporary research*, pages 639–650. de Gruyter, Berlin, 1991.
- Angelika Kratzer. *Modals and conditionals*. Oxford University Press, New York, 2012.
- Manfred Krifka. The semantics and pragmatics of polarity items. *Linguistic Analysis*, 25: 209–257, 1995.
- Robin Lakoff. A syntactic argument for negative transportation. *Chicago Linguistic Society*, 5: 149–157, 1969.
- Ellen Prince. The syntax and semantics of neg-raising, with evidence from French. *Language*, 52:404–426, 1976.
- Jacopo Romoli. *Soft but Strong. Neg-Raising, Soft Triggers, and Exhaustification*. PhD thesis, Harvard University, 2012.
- Pieter A. M. Seuren. Negative’s travels. In Pieter A. M. Seuren, editor, *Semantic syntax*, pages 96–122. Oxford University Press, Oxford, 1974.
- Pieter A. M. Seuren. Negative’s travels. In *A View of Language*, chapter 8, pages 185–208. Oxford University Press, Oxford, 2001.
- Hedde Zeijlstra. Does Neg-Raising involve Neg-Raising. *Topoi*, 2017.

# Double Negation, Excluded Middle and Accessibility in Dynamic Semantics\*

Matthew Gotham

University of Oxford, Oxford, UK  
matthew.gotham@ling-phil.ox.ac.uk

## Abstract

This paper addresses a recalcitrant problem for dynamic semantics: the inaccessibility of discourse referents from under double negation or from a negated left disjunct into a right disjunct. I propose that these discourse referents are made accessible by the discourse being interpreted in the context of designated formulae, which validate double negation elimination in a controlled way.

## 1 Puzzles of Accessibility

First-generation dynamic semantic theories were developed, in part, in order to capture the anaphoric dependencies in discourses like (1).

(1) John owns a car. It is parked in a weird place.

As Karttunen (1976, 366) put it:

Let us say that the appearance of an indefinite noun phrase establishes a discourse referent just in case it justifies the occurrence of a corerential pronoun or a definite noun phrase later in the text.

So in (a natural interpretation of) (1), *a car* establishes a discourse referent, which is picked up by *it*. As Karttunen (1976) goes on to note, however, anaphoric dependencies created by indefinites don't always persist once introduced. To take just one example, negation closes off anaphoric dependencies. For example, (2) can't be interpreted with *it* dependent on *a car* if *a car* is in the scope of negation.

(2) John doesn't own a car. It is parked in a weird place.

Dynamic semantic theories generally have no problem accounting for these data. Concretely, let us look at the natural translations of (1) and (2) into dynamic predicate logic (DPL, Groenendijk and Stokhof (1991)), and their interpretations, in (3) and (4). DPL has the same syntax as classical predicate logic (PL) and its semantics is given in Figure 1.<sup>1</sup>

(3)  $\exists x(Cx \wedge Ojx) \wedge Px$

$\llbracket (3) \rrbracket_M^f = \{g \mid f[x]g \ \& \ g(x) \in \mathcal{I}(C) \ \& \ \langle \mathcal{I}(j), g(x) \rangle \in \mathcal{I}(O) \ \& \ g(x) \in \mathcal{I}(P)\}$

\*Thanks to Luisa Martí, Simon Charlow, Matt Mandelkern, the audience at the 2019 London Semantics Symposium at Queen Mary University of London and three anonymous AC2019 reviewers for helpful comments. This research is supported by an Early Career Fellowship from the Leverhulme Trust.

<sup>1</sup>This presentation is slightly idiosyncratic but is entirely equivalent to that in Groenendijk and Stokhof (1991).

$$\begin{aligned}
\llbracket Pt_1 \dots t_n \rrbracket_M^f &= \{g \mid f = g \ \& \ \langle \llbracket t_1 \rrbracket_M^g, \dots, \llbracket t_n \rrbracket_M^g \rangle \in \mathcal{I}(P)\} \\
\llbracket t_1 = t_2 \rrbracket_M^f &= \{g \mid f = g \ \& \ \llbracket t_1 \rrbracket_M^g = \llbracket t_2 \rrbracket_M^g\} \\
\llbracket \neg\phi \rrbracket_M^f &= \{g \mid f = g \ \& \ \llbracket \phi \rrbracket_M^g = \emptyset\} \\
\llbracket \phi \wedge \psi \rrbracket_M^f &= \left\{ h \mid \text{there's a } g : g \in \llbracket \phi \rrbracket_M^f \ \& \ h \in \llbracket \psi \rrbracket_M^g \right\} \\
\llbracket \phi \vee \psi \rrbracket_M^f &= \{g \mid f = g \ \& \ \llbracket \phi \rrbracket_M^g \cup \llbracket \psi \rrbracket_M^g \neq \emptyset\} \\
\llbracket \phi \rightarrow \psi \rrbracket_M^f &= \left\{ g \mid f = g \ \& \ \llbracket \phi \rrbracket_M^g \subseteq \left\{ h \mid \llbracket \psi \rrbracket_M^h \neq \emptyset \right\} \right\} \\
\llbracket \exists x\phi \rrbracket_M^f &= \{h \mid \text{there's a } g : f[x]g \ \& \ h \in \llbracket \phi \rrbracket_M^g\} \\
\llbracket \forall x\phi \rrbracket_M^f &= \left\{ g \mid f = g \ \& \ \{h \mid g[x]h\} \subseteq \left\{ h \mid \llbracket \phi \rrbracket_M^h \neq \emptyset \right\} \right\}
\end{aligned}$$

Figure 1: Semantics of DPL

$$(4) \quad \neg\exists x(Cx \wedge Ojx) \wedge Px$$

$$\llbracket (4) \rrbracket_M^f = \{g \mid f = g \ \& \ \{h \mid g[x]h \ \& \ h(x) \in \mathcal{I}(C) \ \& \ \langle \mathcal{I}(j), h(x) \rangle \in \mathcal{I}(O)\} = \emptyset \ \& \ g(x) \in \mathcal{I}(P)\}$$

By inspecting the semantic clauses in Figure 1 we can see that the closing off of anaphoric dependencies is tied to negation. In DPL, the negation of any formula is a **test**: an identity relation on some set of assignments. Pretty much all dynamic semantic theories treat negation like this,<sup>2</sup> and hence they all have the property that  $\neg\neg\phi$  is not generally equivalent to  $\phi$ :

[T]he law of double negation will not hold unconditionally. Consider a formula  $\phi$  that is not a test. Negating  $\phi$  results in the test  $\neg\phi$ , and a second negation, which gives  $\neg\neg\phi$ , does not reverse this effect [...]. Hence, double negation is not in general eliminable. (Groenendijk and Stokhof, 1991, 62)

As has repeatedly been noted,<sup>3</sup> this failure of double negation elimination is problematic in that there are several examples where it seems that we would like a doubly-negated existential statement to behave more like its un-negated counterpart than these theories predict.

## 1.1 Double Negation

One class of examples concerns straightforward double negations, such as (5).

$$(5) \quad \text{It's not true that John doesn't own a car. It's (just) parked in a weird place.}$$

The natural translation of (5) into DPL is given in (6). As the interpretation given shows, (6) does not capture the intended dependency, unlike (3).

$$(6) \quad \neg\neg\exists x(Cx \wedge Ojx) \wedge Px$$

$$\llbracket (6) \rrbracket_M^f = \{g \mid f = g \ \& \ \{h \mid g[x]h \ \& \ h(x) \in \mathcal{I}(C) \ \& \ \langle \mathcal{I}(j), h(x) \rangle \in \mathcal{I}(O)\} \neq \emptyset \ \& \ g(x) \in \mathcal{I}(P)\}$$

<sup>2</sup>An exception will be noted in Section 2.2.

<sup>3</sup>E.g. by Groenendijk and Stokhof (1990, 1991); Kamp and Reyle (1993); Krahmer and Muskens (1995).

## 1.2 Disjunction

Another class of examples concerns disjunctions like (7).

- (7) Either John doesn't own a car, or it is parked in a weird place.

Given the most natural translation of (7) as shown in (8), the intended anaphoric dependency is not captured. Once again, other dynamic semantic theories are essentially the same at this point.

- (8)  $\neg\exists x(Cx \wedge Ojx) \vee Px$

Note that in PL, (8) is equivalent to both (9) and (10).

- (9)  $\neg\exists x(Cx \wedge Ojx) \vee (\exists x(Cx \wedge Ojx) \wedge Px)$   
 (10)  $\neg\exists x(Cx \wedge Ojx) \vee (\neg\neg\exists x(Cx \wedge Ojx) \wedge Px)$

Meanwhile, in DPL (8) is equivalent to (10) but not (9); and (9) *would* capture the intended dependency when interpreted in DPL. So, apparently, we again have a situation where the PL equivalence based on double negation would be desirable.

## 1.3 Uniqueness

However, there are examples that seem to show that we don't want  $\phi$  to be *exactly* equivalent to  $\neg\neg\phi$ , such as (11)–(12).

- (11) ??It's not true that John doesn't own a shirt. It's in the wardrobe.  
 (12) ??Either John doesn't own a shirt, or it's in the wardrobe.

Examples (11) and (12) sound strange in a way that their counterparts (5) and (7) respectively don't. The reason seems to be that these examples carry the implication<sup>4</sup> that, if John owns a car/shirt, then he owns exactly one. While that is a plausible (though possibly false) assumption in the case of cars, it is much less plausible in the case of shirts. No such implication is present in (1), or (13).

- (13) John owns a shirt. It's in the wardrobe.

## 1.4 Plan

After very briefly reviewing the literature on this issue in Section 2, I will make a proposal for making discourse referents rendered inaccessible by (double) negation accessible again in Section 3. In Section 4 I will discuss how this proposal can be finessed in order to take account of the uniqueness effect. The proposal will be integrated into a compositional semantic system in Section 5. Section 6 concludes.

## 2 Previous Accounts

There have been a few attempts to address this issue in the literature, all of which approach it by doing something to the semantics of negation.

<sup>4</sup>I leave open the question of what exactly this 'implication' amounts to: whether entailment, presupposition, implicature, etc.

## 2.1 Separating Negation from Closure

Negation as defined in Figure 1 can be decomposed into three operators  $\bullet$ ,  $\sim$  and  $!$ , as defined in (14).  $!$  closes off anaphoric dependencies,  $\sim$  is (revised) negation and  $\bullet$  checks that its operand is a test. N.B.  $\simeq$  is DPL equivalence, i.e.  $\phi \simeq \psi \Leftrightarrow$  for all  $M$  and  $f$ ,  $\llbracket \phi \rrbracket_M^f = \llbracket \psi \rrbracket_M^f$ . An ancillary notion is *satisfaction equivalence*:  $\phi \simeq_s \psi \Leftrightarrow$  for all  $M$  and  $f$ ,  $\llbracket \phi \rrbracket_M^f \neq \emptyset$  just in case  $\llbracket \psi \rrbracket_M^f \neq \emptyset$  (Groenendijk and Stokhof, 1991, 56).

$$(14) \quad \begin{aligned} \llbracket \bullet\phi \rrbracket_M^f &= \{g \mid f = g \ \& \ g \in \llbracket \phi \rrbracket_M^f\} \\ \llbracket \sim\phi \rrbracket_M^f &= \{g \mid g \notin \llbracket \phi \rrbracket_M^f\} \\ \llbracket !\phi \rrbracket_M^f &= \{g \mid f = g \ \& \ \llbracket \phi \rrbracket_M^g \neq \emptyset\} \end{aligned}$$

Facts 1 and 2 then follow.

**Fact 1.**  $\sim\sim\phi \simeq \phi$

**Fact 2.**  $\bullet\sim!\phi \simeq \neg\phi$

Given the equivalence noted in Fact 1, we could imagine that *not* as expressed in (5) is translated as  $\sim$ , but as expressed in (2) is translated as  $\neg$  (or, equivalently,  $\sim$  augmented with  $\bullet$  and  $!$ ). Groenendijk and Stokhof (1990) and Rothschild (2017) both make suggestions somewhat like this, and in systems that are sufficiently different from DPL for the decomposition to be achieved with two operators rather than three.

Nevertheless, such an approach immediately raises the question of why there is no reading of, say, (2) in which *not* is translated just with the revised negation. In the system of Groenendijk and Stokhof (1990) the result would be an interpretation equivalent to ‘It’s not true that John owns a car which is parked in a weird place’, and in that of Rothschild (2017) the result would be an interpretation equivalent to ‘there is something which is not a car owned by John, and which is parked in a weird place’ (which is also what we’d get using just  $\sim$  as defined above). Needless to say, neither of these is a possible interpretation of (2). Both Groenendijk and Stokhof (1990) and Rothschild (2017) make suggestions<sup>5</sup> about how to avoid such interpretations, but these are incomplete. Furthermore, both theories require additional assumptions to account for the disjunction cases outlined in Section 1.2—in neither theory does the decomposition of negation achieve this alone. Nor do they have anything to say about the uniqueness effect.

## 2.2 Bilateralism

The approach adopted by Krahmer and Muskens (1995), when adapted from Discourse Representation Theory (DRT, Kamp and Reyle (1993)) to DPL, is to give formulae both positive (verifying) and negative (falsifying) extensions, with negation amounting to reversal of these. Double negation elimination then follows immediately. Examples like (7) are taken care of by a tweak to the semantics of  $\vee$  making  $\phi \vee \psi$  equivalent to  $\neg\phi \rightarrow \psi$ . This makes (8) equivalent to (15) which, given that the system now has double negation elimination, is equivalent to (16).

$$(15) \quad \neg\neg\exists x(Cx \wedge Ojx) \rightarrow Px$$

$$(16) \quad \exists x(Cx \wedge Ojx) \rightarrow Px$$

<sup>5</sup>Rather more developed in the case of Groenendijk and Stokhof (1990) than Rothschild (2017).

Formula (16) is not equivalent to (9) (or (10)): interpreted in DPL, (9) means ‘either John does not own a car, or he owns a car that is parked in a weird place’, while (16) means ‘every car John owns is parked in a weird place’. The difference is reminiscent of the difference between strong and weak readings of donkey sentences. Given the uniqueness effect it seems to be somewhat moot, however. Krahmer and Muskens (1995, 359) note the uniqueness effect but do not account for it. In any case, their semantics for  $\vee$  could be tweaked in a different way to get the ‘weak’ reading instead of the ‘strong’ one they do get (and defend).

In the following section I will present an account of pronoun accessibility in sentences like (5) and (7) in a unilateral semantics (in fact, without changing the semantics of DPL at all), which moreover accounts for the uniqueness effect.

### 3 Double Negation and Excluded Middle

The non-equivalence of  $\phi$  and  $\neg\neg\phi$  in DPL is reminiscent of the situation in intuitionistic logic (IL). Now, the parallel is by no means exact, since in IL this non-equivalence can be expressed as  $\phi \not\vdash \neg\neg\phi$ , whereas in DPL it can’t really be brought out directly in terms of entailment or derivability. Nevertheless, it’s worth looking at what one needs to add to IL in order to get the equivalence back.<sup>6</sup> Famously, adding any of (17)–(19) to IL gets you classical logic:

- (17)  $\neg\neg\phi \vdash \phi$  (double negation elimination)
- (18)  $\frac{\Gamma, \neg\phi \vdash \perp}{\Gamma \vdash \phi}$  (reductio ad absurdum)
- (19)  $\vdash \phi \vee \neg\phi$  (excluded middle)

This invites the following thought: could there be a way to achieve (something like) the double negation property for dynamic semantics by adding (something like) excluded middle? And could that help to resolve the issues we’ve identified with pronoun accessibility? The answer is *yes*, but it doesn’t involve the standard DPL disjunction. Rather, it involves ‘program disjunction’ (Groenendijk and Stokhof, 1991, 88), defined in (20).

- (20) Extend the language of DPL with the following clauses:
- If  $\phi$  and  $\psi$  are formulae, then  $\phi \cup \psi$  is a formula.
  - $\llbracket \phi \cup \psi \rrbracket_M^f = \llbracket \phi \rrbracket_M^f \cup \llbracket \psi \rrbracket_M^f$

Like  $\vee$ ,  $\cup$  is internally static, but unlike  $\vee$  it is externally dynamic. External dynamicity is crucial to the equivalence shown in Fact 3, which shows something of the extent to which (this form of) excluded middle gets us (something like) the double negation property in DPL.<sup>7</sup>

**Fact 3.** *If  $\phi \simeq \phi \wedge \phi$  then  $(\phi \cup \neg\phi) \wedge \neg\neg\phi \simeq \phi$*

In DPL  $\phi \cup \neg\phi$  is a tautology, but there are many semantically distinct tautologies in DPL. Consequently, DPL does not have the property that  $\phi$  is equivalent to  $T \wedge \phi$  for any DPL tautology  $T$  and formula  $\phi$ . So much can be seen from Fact 3. The relevance of this for us is that indefinites made inaccessible by double negation can be made accessible again on the assumption that the discourse is interpreted in the context of a specific tautology, namely, an appropriate instances of excluded middle with  $\cup$ .

<sup>6</sup>The connection is deeper than there is space to get into here. See (Ranta, 1994, 74–75) and Fernando (2001).

<sup>7</sup>Groenendijk and Stokhof (1991, 63–64) discuss the conditions under which  $\wedge$  is idempotent. In this paper we’re concerned about the case where  $\phi := \exists x(Cx \wedge Ojx)$ , and in that case  $\phi \simeq \phi \wedge \phi$ .

For example, if we assume that the discourse in (5) is interpreted in the context of  $\exists x(Cx \wedge Ojx) \cup \neg \exists x(Cx \wedge Ojx)$  and so augment (6) to (21), the interpretation we end up with is equivalent to (3), as shown below.

$$(21) \quad (\exists x(Cx \wedge Ojx) \cup \neg \exists x(Cx \wedge Ojx)) \wedge (\neg \neg \exists x(Cx \wedge Ojx) \wedge Px)$$

$$\begin{aligned} \llbracket (21) \rrbracket_M^f &= \left\{ g \mid g \in \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \cup \llbracket \neg \exists x(Cx \wedge Ojx) \rrbracket_M^f \ \& \ \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^g \neq \emptyset \right\} \\ &= \left\{ g \mid g \in \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \ \& \ g(x) \in \mathcal{I}(P) \right\} \\ &= \llbracket (3) \rrbracket_M^f \end{aligned}$$

In the same way, if we augment (8) to (22), the interpretation we end up with is satisfaction-equivalent to (9), as shown below.

$$(22) \quad (\exists x(Cx \wedge Ojx) \cup \neg \exists x(Cx \wedge Ojx)) \wedge (\neg \exists x(Cx \wedge Ojx) \vee Px)$$

$$\begin{aligned} \llbracket (22) \rrbracket_M^f &= \left\{ h \mid \text{there's a } g : g \in \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \cup \llbracket \neg \exists x(Cx \wedge Ojx) \rrbracket_M^f \ \& \ h \in \llbracket \neg \exists x(Cx \wedge Ojx) \vee Px \rrbracket_M^g \right\} \\ &= \left\{ g \mid (f = g \ \& \ \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^g = \emptyset) \ \text{or} \ (g \in \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \ \& \ g(x) \in \mathcal{I}(P)) \right\} \end{aligned}$$

Note that in simple positive cases like (3) and single-negation cases like (4), adding the instance of excluded middle does nothing bad, thanks to Facts 4 and 5.

**Fact 4.** *If  $\phi \simeq \phi \wedge \phi$  then  $(\phi \cup \neg \phi) \wedge (\phi \wedge \psi) \simeq \phi \wedge \psi$*

**Fact 5.** *If  $\phi \simeq \phi \wedge \phi$  then  $(\phi \cup \neg \phi) \wedge (\neg \phi \wedge \psi) \simeq \neg \phi \wedge \psi$*

Before we move on, I want to note two things about this treatment of disjunction. Firstly, binding is predicted to be symmetric, i.e. *either it's parked in a weird place, or John doesn't own a car* is predicted to be just as good as (7).

Secondly, in either case the interpretation amounts to ‘either John doesn't own a car, or some car he owns is parked in a weird place’. As noted above, this take on the truth conditions of (7) is disputed by [Krahmer and Muskens \(1995\)](#).

Both of these bugs/features follow from the (independently-given) semantics of  $\vee$  in DPL. Both would be changed on the assumption that *p or q* is translated into DPL not as  $p \vee q$  but as  $\neg p \rightarrow q$ . Nevertheless, as noted above, the strong vs. weak issue at least is somewhat moot given the uniqueness effect, to be discussed next.

## 4 Uniqueness

So far, the uniqueness effect is not accounted for. Intuitively, what it seems that this effect amounts to is a restriction on anaphoric licensing: while a simple indefinite *a P* licenses subsequent pronouns, one under double negation only does so on the assumption that there is exactly one *P*. With that in mind, let us reflect on what program disjunction does in cases of an existential statement and its negation.

$$\llbracket \exists xPx \cup \neg \exists xPx \rrbracket_M^f = \begin{cases} \text{the set of } x\text{-variants of } f \text{ mapping } x \text{ to a } P, \text{ if there are any} \\ \{f\} \text{ otherwise} \end{cases}$$



If we want the anaphoric dependency to be passed on only in the case of uniqueness, then the input context for our unaugmented formulae should look like this instead:

$$\begin{cases} \text{the (singleton) set of } x\text{-variants of } f \text{ mapping } x \text{ to a } P, \text{ if there's exactly one} \\ \{f\} \text{ otherwise} \end{cases}$$

That effect can be achieved by the introduction of an operator  $\uparrow$ , defined as follows.

(23) Extend the language of DPL with the following clauses:

- a. If  $\phi$  is a formula, then  $\uparrow\phi$  is a formula.
- b.  $\llbracket \uparrow\phi \rrbracket_M^f = \begin{cases} \llbracket \phi \rrbracket_M^f & \text{if } \left| \llbracket \phi \rrbracket_M^f \right| = 1 \\ \{f\} & \text{otherwise} \end{cases}$

Or, equivalently,

$$\llbracket \uparrow\phi \rrbracket_M^f = \left\{ g \mid g \in \llbracket \phi \rrbracket_M^f \ \& \ \left| \llbracket \phi \rrbracket_M^f \right| = 1 \right\} \cup \left\{ g \mid f = g \ \& \ \left| \llbracket \phi \rrbracket_M^g \right| \neq 1 \right\}$$

Note that  $\uparrow\phi$  is also a DPL tautology (for any  $\phi$ ). I will henceforth refer to formulae of the form  $\uparrow\phi$  as instances of ‘unique excluded middle’ (UEM).<sup>8</sup> If we now use UEM for our augmentations of (6) and (8) instead of excluded middle with  $\cup$ , we get the uniqueness effect, as seen below in (24) and (25) respectively.

(24)  $\uparrow\exists x(Cx \wedge Ojx) \wedge (\neg\neg\exists x(Cx \wedge Ojx) \wedge Px)$

$$\begin{aligned} \llbracket (24) \rrbracket_M^f &= \left\{ g \mid \begin{array}{l} g \in \left\{ h \mid h \in \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \ \& \ \left| \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \right| = 1 \right\} \\ \cup \left\{ h \mid f = h \ \& \ \left| \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^h \right| \neq 1 \right\} \\ \& \ \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \neq \emptyset \ \& \ g(x) \in \mathcal{I}(P) \end{array} \right\} \\ &= \left\{ g \mid \begin{array}{l} \left( \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f = \{g\} \ \& \ g(x) \in \mathcal{I}(P) \right) \\ \text{or } \left( f = g \ \& \ \left| \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \right| > 1 \ \& \ g(x) \in \mathcal{I}(P) \right) \end{array} \right\} \end{aligned}$$

‘Either John owns exactly one car, which is parked in a weird place, or John owns more than one car and  $x$  is parked in a weird place’ (with  $x$  free).

(25)  $\uparrow\exists x(Cx \wedge Ojx) \wedge (\neg\exists x(Cx \wedge Ojx) \vee Px)$

$$\begin{aligned} \llbracket (25) \rrbracket_M^f &= \left\{ g \mid \begin{array}{l} g \in \left\{ h \mid h \in \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \ \& \ \left| \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \right| = 1 \right\} \\ \cup \left\{ h \mid f = h \ \& \ \left| \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^h \right| \neq 1 \right\} \\ \& \ \llbracket \neg\exists x(Cx \wedge Ojx) \rrbracket_M^f \cup \llbracket Px \rrbracket_M^f \neq \emptyset \end{array} \right\} \\ &= \left\{ g \mid \begin{array}{l} \left( \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f = \{g\} \ \& \ g(x) \in \mathcal{I}(P) \right) \\ \text{or } \left( f = g \ \& \ \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^g = \emptyset \right) \\ \text{or } \left( f = g \ \& \ \left| \llbracket \exists x(Cx \wedge Ojx) \rrbracket_M^f \right| > 1 \ \& \ g(x) \in \mathcal{I}(P) \right) \end{array} \right\} \end{aligned}$$

‘Either John doesn’t own a car, or he owns exactly one car, which is parked in a weird place, or he owns more than one car and  $x$  is parked in a weird place’ (with  $x$  free). As noted above,

<sup>8</sup>In a previous version of this material, UEM was defined as a formula of the form  $\uparrow(\phi \cup \neg\phi)$ . That was before I realised that this is equivalent to  $\uparrow\phi$ .

this interpretation abolishes the distinction between weak and strong readings by making *a car* inaccessible as an antecedent to *it* if John owns more than one car.

As before, adding UEM to a simple positive formula like (3) doesn't change anything (Fact 6). Adding UEM to a single-negated sentence like (4) doesn't change anything either (Fact 7).

**Fact 6.** *If  $\phi \simeq \phi \wedge \phi$  then  $\uparrow\phi \wedge (\phi \wedge \psi) \simeq \phi \wedge \psi$*

**Fact 7.** *If  $\phi \simeq \phi \wedge \phi$  then  $\uparrow\phi \wedge (\neg\phi \wedge \psi) \simeq \neg\phi \wedge \psi$*

## 4.1 Uniqueness?

Matt Mandelkern (p.c.) has argued against building uniqueness into interpretations in this way, on the basis of examples like (26).

(26) Either Sue didn't have a drink last night, or she had a second drink right after it.

Clearly, (26) makes no sense if *it* is interpreted to mean the unique drink that Sue had last night, which is what the UEM-based analysis implies. On the other hand, if we stick with the analysis based on excluded middle with  $\cup$ , then (26) can mean 'either Sue didn't have a drink last night, or she did have a drink last night and had a second drink right after it', which is coherent.<sup>9</sup> The question remains, though, whether or not (26) can mean this.

Personally, I find (26) strange precisely because (I think) of the uniqueness implication, and so in what follows I will continue to use UEM. However, if this turns out not to be sustainable then we can switch to excluded middle with  $\cup$  and look for another way to treat the uniqueness effect in examples like (11)–(12).

## 5 Composition

The obvious question that this treatment of double negation etc. raises is where these instances of UEM might come from. Answering that question requires moving from DPL to a dynamic semantic system that permits compositionality below the level of the clause. To that end I will move to compositional discourse representation theory (CDRT, Muskens (1996)). Here, the boxes of DRT are taken to be abbreviations of type-logical expressions of type  $s \rightarrow s \rightarrow t$  (abbreviated  $T$ ), where  $s$  is the type of states and discourse referents  $u_n$  are of type  $s \rightarrow e$  (abbreviated  $E$ ). The abbreviations are summarized in Figure 2.<sup>10</sup>

I will treat the instances of UEM as being introduced lexically by negation as a kind of projective content. That is to say, in addition to introducing  $[|\neg D]$  in the standard dimension of meaning, instances of negation introduce  $\uparrow D$  in another dimension of meaning. The simplest way to do this is to make the result type of lexical entries not  $T$  but  $T \times T$ , with (some of) these instances of UEM in the first dimension and other content in the second.<sup>11</sup> In contrast to other kinds of projective content, however, I don't assume that these instances of UEM project very far, as can be seen from the lexical entries given to (*and*) and (*or*) in the lexicon shown in (27).<sup>12</sup>

<sup>9</sup>As Matt also pointed out, on the assumptions made by Krahmer and Muskens (1995) it has to mean that Sue had infinitely many drinks.

<sup>10</sup>Axioms ensure that discourse referents and states really do behave like variables and variable assignments.

<sup>11</sup>This can be done in a more principled way with a writer monad (Giorgolo and Asudeh, 2012). Thanks to Simon Charlow for pointing this out.

<sup>12</sup>(*and*) is a silent conjunction assumed to hold sentences together at the text level. Different projective behaviour might be required of overt *and*.

Conditions:	
$R(\delta_1, \dots, \delta_n)$	$\implies \lambda i. R(\delta_1(i), \dots, \delta_n(i))$
$\delta_m = \delta_n$	$\implies \lambda i. \delta_m(i) = \delta_n(i)$
$\neg D$	$\implies \lambda i. \neg \exists j. D(i)(j)$
$D \vee E$	$\implies \lambda i. \exists j. D(i)(j) \vee E(i)(j)$
DRSs:	
$[u_1 \dots u_n \mid C_1, \dots, C_n]$	$\implies \lambda i. \lambda j. (\forall u. (u \neq u_1 \wedge \dots \wedge u \neq u_n) \rightarrow u(i) = u(j))$ $\wedge C_1(j) \wedge \dots \wedge C_n(j)$
$D ; E$	$\implies \lambda i. \lambda k. \exists j. D(i)(j) \wedge E(j)(k)$
$D \cup E$	$\implies \lambda i. \lambda j. D(i)(j) \vee E(i)(j)$
$\uparrow D$	$\implies \lambda i. \lambda j. (D(i)(j) \wedge \exists! k. D(i)(k)) \vee (i = j \wedge \neg \exists! k. D(j)(k))$

Figure 2: Abbreviations for CDRT (Muskens, 1996, 157) augmented with  $\cup$  and  $\uparrow$ 

$$\begin{aligned}
(27) \quad & \text{John} \rightsquigarrow j : E \quad \dots \text{not} \dots, \text{doesn't} \rightsquigarrow \lambda p^{T \times T}. \langle \pi_1(p) ; \uparrow(\pi_2(p)), [ \mid \neg(\pi_2(p)) ] \rangle \\
& \text{own} \rightsquigarrow \lambda y^E. \lambda x^E. \langle [ \mid ], [ \mid \text{own}(x, y) ] \rangle \quad \text{is parked} \dots \rightsquigarrow \lambda x^E. \langle [ \mid ], [ \mid \text{parked}(x) ] \rangle \\
& a^1 \rightsquigarrow \lambda P^{E \rightarrow T \times T}. \lambda Q^{E \rightarrow T \times T}. \langle [ \mid ], [u_1 \mid ] ; \pi_1(P(u_1)) ; \pi_2(P(u_1)) ; \pi_1(Q(u_1)) ; \pi_2(Q(u_1)) \rangle \\
& \text{car} \rightsquigarrow \lambda x^E. \langle [ \mid ], [ \mid \text{car}(x) ] \rangle \quad (\text{and}) \rightsquigarrow \lambda q^{T \times T}. \lambda p^{T \times T}. \langle [ \mid ], \pi_1(p) ; \pi_2(p) ; \pi_1(q) ; \pi_2(q) \rangle \\
& \text{or} \rightsquigarrow \lambda q^{T \times T}. \lambda p^{T \times T}. \langle [ \mid ], \pi_1(p) ; \pi_1(q) ; [ \mid \pi_2(p) \vee \pi_2(q) ] \rangle \quad \text{it}_1 \rightsquigarrow u_1
\end{aligned}$$

Assuming the structures for (5) and (7) shown in (28) and (29) respectively, their interpretations are shown below.

$$\begin{aligned}
(28) \quad & [ [ \dots \text{not} \dots [ \text{John} [ i [ \text{doesn't} [ [ a^1 \text{ car} ] [ j [ t_i [ \text{own} t_j ] ] ] ] ] ] ] ] ] ] [ (\text{and}) [ \text{it}_1 \text{ is parked} \dots ] ] \\
& \rightsquigarrow \langle [ \mid ], \uparrow [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ; \uparrow [ \mid \neg [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ] ; \\
& \quad [ \mid \neg [ \mid \neg [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ] , \text{parked}(u_1) ] \rangle \\
& \equiv \langle [ \mid ], \uparrow [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ; [ \mid \neg [ \mid \neg [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ] , \text{parked}(u_1) ] \rangle \\
(29) \quad & [ [ \text{John} [ i [ \text{doesn't} [ [ a^1 \text{ car} ] [ j [ t_i [ \text{own} t_j ] ] ] ] ] ] ] ] [ \text{or} [ \text{it}_1 \text{ is parked} \dots ] ] \\
& \rightsquigarrow \langle [ \mid ], \uparrow [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ; [ \mid [ \mid \neg [u_1 \mid \text{car}(u_1), \text{own}(j, u_1)] ] \vee [ \mid \text{parked}(u_1) ] ] \rangle
\end{aligned}$$

Expanding the second projections of (28) and (29) according to the key in Figure 2 shows these interpretations to be equivalent to those given above for (24) and (25) respectively.

One thing to note about the lexicon in (27) is that the lexical entry for *or* predicts symmetric binding, as discussed at the end of Section 3. An alternative lexical entry predicting left-to-right binding only would be  $\lambda q^{T \times T}. \lambda p^{T \times T}. \langle [ \mid ], \pi_1(p) ; [ \mid \pi_2(p) \vee (\pi_1(q) ; \pi_2(q)) ] \rangle$ .

## 6 Conclusion

In this paper I have suggested a way for discourse referents that are problematically inaccessible in standard dynamic semantics to be rendered accessible again, without needing to change the semantic clauses for the existing connectives of any dynamic semantic theory at all. I proposed that these discourse referents are made accessible by the discourse being interpreted in the context of designated formulae, either excluded middle with  $\cup$  or UEM, which are tautologies in the sense of being always true according to the designated truth definition. The fact that

they can nevertheless be leveraged to help with this (in)accessibility issue is simply another example of meaning in dynamic semantics going beyond truth conditions.

I also proposed a compositional system for the introduction of these formulae, treating them as a kind of projective content introduced by the lexical semantics of negative operators. This part of the story is necessarily (even) less secure than the general idea. A reviewer notes that this seems to predict that other languages could behave differently to English in terms of pronoun accessibility from under double negation or from a negated left disjunct into a right disjunct. It certainly would be welcome if UEM could be introduced more systematically, but at the moment I don't see a way to do this.

## References

- Tim Fernando. A type reduction from proof-conditional to dynamic semantics. *Journal of Philosophical Logic*, 30(2):121–153, 2001. doi: 10.1023/A:1017541301458.
- Gianluca Giorgolo and Ash Asudeh.  $\langle M, \eta, \star \rangle$  Monads for conventional implicatures. In Ana Aguilar Guevara, Anna Chernilovskaya, and Rick Nouwen, editors, *Proceedings of Sinn und Bedeutung 16*, pages 265–278. MIT Working Papers in Linguistics, 2012.
- Jeroen Groenendijk and Martin Stokhof. Dynamic Montague grammar. Technical Report LP-1990-02, ILLC, University of Amsterdam, 1990. URL <https://eprints.illc.uva.nl/1148>.
- Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100, 1991.
- Hans Kamp and Uwe Reyle. *From Discourse to Logic*. Number 42 in Studies in Linguistics and Philosophy. Springer, Dordrecht, 1993.
- Lauri Karttunen. Discourse referents. In James McCawley, editor, *Syntax and Semantics*, volume 7, pages 363–385. Academic Press, New York, 1976.
- Emiel Krahmer and Reinhard Muskens. Negation and disjunction in discourse representation theory. *Journal of Semantics*, 12:357–376, 1995.
- Reinhard Muskens. Combining Montague semantics and discourse representation. *Linguistics and Philosophy*, 19(2):143–186, 1996.
- Aarne Ranta. *Type-Theoretical Grammar*. Oxford University Press, Oxford, 1994.
- Daniel Rothschild. A trivalent approach to anaphora and presupposition. In Alexandre Cremers, Thom van Gessel, and Floris Roelofsen, editors, *Proceedings of the 21st Amsterdam Colloquium*, pages 1–13. 2017. URL <http://events.illc.uva.nl/AC/AC2017/Proceedings/>.

# Dogwhistles, Trust and Ideology

Robert Henderson and Elin McCready

University of Arizona and Aoyama Gakuin University

## Abstract

Given that someone is consistently untruthful, why should we ever trust them? The question is not academic. Consider politicians and others who are known to consistently lie, but who are still listened to and voted back into office. This talk addresses this puzzle via three mechanisms: (i) a theory of source evaluation based on interactional histories and heuristics for judgments of reliability (McCready, 2015), (ii) a game-theoretic view of how speaker ideologies and political positions are communicated by linguistic acts (Burnett, 2018; Henderson and McCready, 2018), and (iii) a theory of how ideological considerations are valued alongside truth-conditional content. In the process, an analysis of *fake news* claims is provided.

## 1 Opening

In politics (and beyond), it is common to see people trusting others who are known to consistently not care much about the truth.<sup>1</sup> Donald Trump is the most prominent case. Many have devoted a great deal of time to exposing the false claims he routinely makes in his speeches and comments, but his supporters seem to trust him nonetheless. On standard views of the evaluation of information sources in epistemology, which usually involve the tracking of truth, this is a surprise. Given that someone is consistently untruthful, why should we ever trust them? This paper addresses this puzzle via three mechanisms. First is a theory of source evaluation based on interactional histories and heuristics for judgments of reliability (McCready, 2015), which is used to spell out the puzzle and as an exemplar of truth-based views on trust. The second is a game-theoretic analysis of how speaker ideologies and political positions are communicated by linguistic acts (Burnett, 2018), and an extension of that analysis to dogwhistles in political speech (Henderson and McCready, 2018). This latter contains the seeds of the third mechanism, one new to this paper: a theory of how ideological considerations are valued alongside truth-conditional content. We claim that such considerations are what makes people trust testimonial agents like Trump; in the process, an analysis of *fake news* claims is provided. We close with several conclusions, first the descriptive claim that trust in Trump-like agents can be rational depending on one's preferences, and second a set of prescriptions for how political discourse needs to proceed given the foundations of ideologically based trust.

## 2 Evaluating Information Sources

How can one determine whether content obtained from a particular agent, or other information source, should be believed? This is a longstanding and controversial issue in epistemology, and has been extensively addressed in thousands of works; our goal here is not to present a solution to this question, but simply to make it more precise. In order to do so, we make use of the work of McCready (2015), who proposes a two-factor theory of source reliability for both agent testimony and other evidence sources (as applied to dynamic update with evidential

---

<sup>1</sup>Acknowledgments: Audiences at Expressing Evidence and LENLS17, especially Regine Eckardt, Natasha Korotkova, Maribel Romero, Stephanie Solt.

constructions in formal semantics/pragmatics). Here we consider only the case of testimony, as the more general theory is not required to clarify the main question we are interested in: the relationship between reliability/truthfulness and trust.

On the view of [McCready \(2015\)](#), an agent first uses a set of heuristics to determine whether a particular testimonial source has properties associated with reliability. Sample properties that might be taken into account by these heuristics include the source’s profession, personal presentation, gender, race, etc. Of course, some of these properties are genuinely useful for determination of reliability and some are not. For example, an agent’s profession is relevant to deciding whether that agent is going to be reliably truth-tracking for assertions about content domains relevant to that profession: we might believe a linguist on issues of linguistic practice, for example. However, some choices of property lead to *epistemic injustice* ([Fricker, 2007](#)), such as gender and race, which aren’t in general relevant to evaluation of reliability, but are nonetheless taken into account by (some) agents in making judgements about whether to believe a piece of testimony (see [McCready and Winterstein 2017, 2019](#) for a detailed investigation of the case of gender, which shows via experimental methods that at least for some cases men are assigned higher baseline degrees of reliability than women). Still, application of these heuristics lead to a probability that the source is reliable. This is an initial determinant of whether it is a rational move to accept content proffered by the source.

The initial probability assigned using the heuristics is only a starting point which is subsequently modified by interaction with that source. Each observation of a discourse move by the source, together with verification of whether its content is truth-tracking, alters, via conditionalization, the probability the source is reliable. In particular, following [McCready 2015](#), suppose that each information-transmitting discourse move by a source of testimony contributes to a history of discourse moves, understood as iterations of a repeated game of information transmission. Such moves are modeled, simplifying slightly from the original system, as tuples  $\langle \varphi, V \rangle$ , where  $V$  is a value in the set  $\{T, F, ?\}$  indicating the truth value of the proffered content (‘?’ indicating that the move cannot be evaluated or that its value is otherwise indeterminate). The basic probability of reliability across this discourse sequence is then interpreted as the frequency of  $T$ -valued moves, so the result is a real number in  $[0,1]$  where each  $T$ -valued move induces an uptick in value (unless the frequency is already 1). The result is that each truth-tracking move raises (possibly very slightly) the perceived likelihood that the source is reliable, and each move that fails to track truth lowers it.

But how should reliability relate to belief? That is, how should an agent incorporate information from a variety of sources of varying reliability? Furthermore, when should an agent be willing to say they believe a proposition based on information from these various sources. Once we understand how reliability affects belief, we can model the fact that, on par, people discount unreliable sources, and begin to address the puzzle that some unreliable sources continue to be believed.

We link reliability and belief using a flavor of dynamic semantics ([Groenendijk and Stokhof, 1991](#)), following work on dynamic updates in Plausibility Models (e.g., [Baltag and Smets 2008](#)). In this literature, slightly simplified, a frame  $\sigma$  is a set of worlds ordered with a ‘plausibility ranking’ reflecting epistemic preferences on states, which is enriched into models in the usual way. In Reliability Dynamic Logic ([McCready, 2015](#)) the situation is slightly complicated. While we have a global model  $\sigma$ , we care more about indexed submodels (hereafter *information states*), each of which represents a source whose reliability we mean to track.

- (1)  $j \in \text{Source} \cup \mathcal{A}$ , where  $\text{Source}$  is the set of evidence sources and  $\mathcal{A}$  the set of agents.

These information states are ordered by the total ordering  $\preceq_a$  satisfying (2) for agent  $a$  who

is tracking reliability.

- (2)  $i \prec_a j$  iff  $P(\text{Rel}(i)) < P(\text{Rel}(j))$ , where  $P(\text{Rel}(i))$  is the probability that source  $i$  yields reliable information according to  $a$ .

Updates on information states are of the following form, where the subscript  $i$  marks the source of the evidence of  $\varphi$ . For cases of testimony, as we discuss here,  $i$  ranges over  $\mathcal{A}$ , the set of agents.

- (3)  $\sigma [E_i\varphi] = \sigma'$  where, for all  $\sigma_j \in \sigma$ ,  $\begin{cases} \sigma'_j = \sigma_j[\varphi] & \text{if } i = j \\ \sigma'_j = \sigma_j & \text{if } i \neq j \end{cases}$

Note that a sentence  $E_i\varphi$  always induces the standard update of state  $\sigma_i$ , namely  $\sigma_i[\varphi] = \{s \in \sigma_i \mid s \in \varphi\}$ . At this level, update with  $\varphi$  always takes place — but this is *not* the same as coming to believe  $\varphi$  at a global level.

To determine global beliefs, we unify the information of all tracked substates  $\sigma_i$  via lexicographic merge in (4).

- (4) Lexicographic merge.  
 $R_{a \circledast b} := R_a^{\prec} \cup (R_a^{\cong} \cap R_b) = R_a^{\prec} \cup (R_a \cap R_b) = R_a \cap (R_a^{\prec} \cup R_b)$

The core idea is that merging  $\sigma_i \circledast \sigma_j$  will result in state where all non-contradictory content survives, and in case of conflict, information from the higher-ranked source overrides the lower-ranked source—i.e.,  $i \prec_a j$  will privilege content from  $j$ . Recursively applied, the global state  $\sigma_T$  on which belief is defined will almost never exhibit conflicts.

With the link finally established between reliability (encoded in  $\prec_a$ ) and belief (propositions that follow from  $\sigma_T$ ), we are now in a position to state the puzzle. Each information-transmitting discourse move affects the perceived reliability of a source in a way dependent on the properties of the source and whether it is truth-tracking. Thus, an agent that is consistently non-truth-tracking will rank lower on  $\prec_a$  with each non-truth-tracking discourse move. As these agents fall on  $\prec_a$ -rank, the content this agent provides will be less likely to survive lexicographic merge and appear in  $\sigma_T$ . Thus, the content such an agent provides will not be believed, as belief is defined on the global information state. This seems right in the general case, but now the puzzle is clear: what about politicians who are known to consistently ignore the truth, but are still trusted? For the specific and obvious case, what about Trump? The remainder of this paper will address this puzzle, claiming in short that, in certain cases, ideology trumps reliability.<sup>2</sup>

### 3 Dogwhistles and Ideologies

Henderson and McCready (2017, 2018, 2019) build a novel theory of so-called *dogwhistle* communication, a communicative strategy familiar from the literature on political speech. This work is relevant to the question at hand because, as we will see, dogwhistle communication involves simultaneously sending truth-conditional information while at the same time sending information about the speaker's persona, which can include ideological orientation. The formal tools used to model dogwhistle communication can be, essentially, immediately extended to solve the puzzle developed in the previous section. The larger result is that we begin to see the emergence of a general theory of social meaning and how it interacts with truth-conditional

---

<sup>2</sup>Painful pun very much intended.



meaning, one that can handle *prima facie* disparate linguistic phenomena like dogwhistles and trust.

Dogwhistles can be roughly defined as messages which communicate aspects of the speaker's ideology to an ingroup in a way which is not accessible to an outgroup. Consider this line from George Bush's 2003 State of the Union address; the term 'wonder-working power' is used in an evangelical Christian hymn, and thus has associations that will be recognized only if the interpreter has the requisite background.

- (5) Yet there's power—wonder-working power—in the goodness and idealism and faith of the American people.

What is the covert message sent to listeners who recognize the phrase 'wonder-working power'? Henderson and McCready (2017, 2018, 2019) argue that it is only information about the speaker's conversational persona, and possible broader identity, as an evangelical Christian. This is backed by experimental literature in political science (e.g., Albertson 2015), where it is shown that evangelical listeners are more likely (than non-religious listeners) to recognize the phrase, more likely to approve of religious rhetoric in religious appeals, and more likely to identify the speaker as a religious conservative politician.

These *identifying* dogwhistles can be contrasted with *enriching* dogwhistles. With the latter sort of expression, the hidden message sent to savvy listeners involves the enrichment of truth-conditional content. We see a clear example in the following example from a 2014 radio program, where Representative Paul Ryan said the following.

- (6) We have got this tailspin of culture, in our inner cities in particular, of men not working and just generations of men not even thinking about working or learning the value and the culture of work.

He was criticized shortly after by fellow Representative Barbara Lee for making a "thinly veiled racial attack". This is because the phrase *inner-city* is code or euphemism for African American neighborhoods (especially stereotypically racialized views of such neighborhoods). We see, then, that this kind of dogwhistle does not only send information about the speaker's persona, but involves altering the truth-conditional content of the message. That is, one is committed to radically different propositional content when saying there is "a tailspin of culture, in our inner-cities", versus the enriched message that there is "a tailspin of culture, in our inner-city, African American neighborhoods."

The fact that dogwhistles have these two components—(i) identifying a speaker persona, and (ii) enriching truth-conditional meaning—motivates the account developed in (Henderson and McCready, 2017, 2018, 2019). We argue that identifying dogwhistles, which only do (i), are a kind of special case of enriching dogwhistle, which involve both (i) and (ii). In particular, we propose that recovering the enriched meaning of the dogwhistle is recovered *in virtue of* recognizing the persona of the speaker.

Given that the persona-signaling aspect of dogwhistles is core, we use the *social meaning games* of Burnett (2017) as the foundation for our analysis of dogwhistles. Such games involve utility assignment based on recovery of speaker persona/ideology and hearer evaluation of that persona/ideology. In this previous work, we modified Burnett's original definition of 'social utility' for use with dogwhistles in (7).<sup>3</sup> Here  $P_L$  gives the probability the speaker has persona  $p$  given the observed message  $m$ , which is just the RSA-style sociolinguistic literal listener

<sup>3</sup>Slightly simplified. We, for instance, are not using log probabilities here because we don't consider the issue of bounded rationality, which can be modeled with a temperature parameter.



(e.g., Goodman and Frank 2016; Franke and Jäger 2016; Franke and Degen 2016), while  $\nu_S$  and  $\nu_L$  (borrowed from Yoon et al. 2016), give the speaker and listener’s affective values of various personas weighted by the likelihood the speaker will be assigned that persona given their message.

$$(7) \quad U_S^{Soc}(m, L) = \sum_{p \in [m]} P_L(p|m) + \nu_S(p)P_L(p|m) + \nu_L(p)P_L(p|m)$$

Utility as in (7) is enough to model identifying dogwhistles. Speakers pick messages in the context of a listener (or group of listeners) weighing what persona they want to have and how listeners will judge that persona weighted by how likely it is that message will signal that persona. Dogwhistles arise when the speakers wants a persona  $p$ , listeners who don’t like  $p$  don’t realize that the  $m$  signals  $p$ , and listeners that do like  $p$  realize that  $m$  signals  $p$ .

To extend the account to handle enriching dogwhistles, which involve truth-conditional meaning, the account must allow for speakers and listeners to extract information from two dimensions. The core idea is to assuming signals with two possible meanings, one an enriched version of the other. Then, we let recovery of the enriched version be tied to recognition of the relevant persona. Thus, messages now denote pairs of truth-conditional meanings and social meanings:  $\langle [[m]], [m] \rangle$ . Given  $T$ , a set of states  $t$  (worlds), speaker strategies  $S_\sigma$  are now functions from pairs of states and personas to messages, and listener strategies  $L_\rho$  are functions from messages to such pairs. We then have the following utility function for information retrieval.

$$(8) \quad \begin{aligned} US(m, L) &= US_{Soc}(m, L) + EU(m, L), \text{ where} \\ EU(m, L) &= \sum_{t \in T} Pr'(t|p) \times U(t, m, L), \text{ where} \\ U(t, m, L) &> 0 \text{ if } t \in L_\rho(m) \text{ and else } = 0 \text{ (cf. van Rooij (2008)).} \end{aligned}$$

There are three critical aspects. First, the top line gives the general formula for the utility of a message given a listener—it is computer is a linear function of its social utility and its truth-conditional utility. This is critical for the puzzle at hand about truth and reliability, and we return to this in the following section. Second, note that in computing the utility of a message for signaling a state, we consider the (posterior) probability of the state given the speaker’s persona— $Pr'(t|p)$ . That is, we compute the speaker’s persona and then use that information to determine the probability that some meaning was transmitted given that the speaker has the persona they do. For most types of truth conditional meaning  $Pr'(t|p) = Pr'(t)$ . That is, we gain no information from knowing the speaker’s persona. When using enriching dogwhistles, though, the speaker’s persona is critical for catching the enriched meaning. Once we identify Paul Ryan’s persona, we are more likely to interpret *inner-city* as *African American neighborhoods*, rather than merely neighborhoods in the city. Finally, the third line (8) asserts that there is a payoff in the truth-conditional domain just in case the listener picks the correct state of the world. The bigger picture is that the social meaning is always recovered, but if the listener fails to recover the proper truth-conditional meaning, no value is extracted from this aspect of the communication.

With (8) we have a unified account of both identifying and enriching dogwhistles. The next section shows how a consequence of this analysis provides a solution to the puzzle we began with. That is, how is it that certain speakers are able to consistently speak falsely, while not being punished by certain listeners in terms of trust? The core insight that the account of dogwhistles provides is that messages package both social/ideological messages alongside truth-conditional content. This will allow a listener to track the utility of a speakers message along either dimension, which we will refer to as reliability and trust respectively.

## 4 Ideology and Trust

The analysis of dogwhistles and of social personas in the previous section (and also in [Burnett 2018](#)) assumes that we have a way to assign affective values to personas, as cast in the function  $\nu_{S/L}$ . But on what basis? Agents can have various reasons to assign positive (or negative) value to a given persona. There are obviously many options, including relatively high-level criteria, for instance the degree to which the persona instantiates some value independently held by the agent such as originality, rebelliousness, safety, or conformity to some social norm. Here, though, we will consider only a very simple metric, namely *similarity*: ‘I like people who are like me,’ or, in this context, ‘I like people who have social personas which are like mine.’ The intuition behind this criterion is that there is intrinsic value in discovering other individuals who share our core values, ideologies, and personal styles. Assuming this criterion means we can assign affective values on the basis of similarity metrics between speaker and hearer personas.

With even this simple metric we can enrich the analysis of reliability in [McCready 2015](#), which was predicated on truth-tracking alone. The core idea is to let hearers determine trust through a mix of truth-conditional and social meaning, with different mixes leading to different kinds of listener utilities. Fleshed out, speaker payoffs rest on two fundamental aspects of meaning: first, for truth-conditional meaning, the successful transmission of true information about the world, and, second, for social meaning, successfully transmitting information about speaker personas together with the hearer’s evaluation of that persona. We treat hearer payoffs identically because (for truth-conditional meaning) clearly the hearer cares about learning the truth, and (for social meaning) as social agents, we value finding individuals who share our values and social groupings, as already indicated in the previous paragraph.<sup>4</sup>

We follow [Henderson and McCready \(2018\)](#) and weight the two components of the utilities with values  $\delta, \gamma$ , giving (where  $U_S^{Soc}(m, L)$  is the utility of the social meaning and  $EU(L, Pr)$  is the value of the truth-conditional meaning, following [van Rooij 2008](#)):

$$U_S(m, L) = \delta U_S^{Soc}(m, L) + \gamma EU(L, Pr)$$

This  $\delta$  indexes the value placed on the social meaning, while  $\gamma$  indexes the value of the truth-conditional meaning. Setting  $\delta = 0$  gives a style of communication where social meaning is disregarded, for instance “science” as traditionally construed, i.e. as a completely objective enterprise where social aspects of the agents (scientists) involved are irrelevant. At the other extreme, setting  $\gamma = 0$  gives “post-truth”, a style of communication in which facts are irrelevant and only social persona matters, on the assumption that relevance is dependent on utility assignments.

This mechanism can be exploited for an analysis of reliability in the face of countervailing evidence as in the case of Donald Trump. When  $\delta \gg \gamma$ ,<sup>5</sup> we end up with Trump-voter-style confidence and trust, because considerations of truth are vastly undervalued compared to persona signaling. Indeed, enriching the system of [McCready \(2015\)](#) with social meanings means that perceived reliability can increase in repeated game settings in virtue of social signaling alone, given that social meaning is allowed to play a role in the ‘vetting’ process.<sup>6</sup> This kind of

<sup>4</sup>Stephanie Solt (p.c.) points out to us that this assumption is too simple: it can be useful for hearers to learn speaker personas even when they do not share them, for instance when the persona in question is highly objectionable or when it might otherwise impact judgements about the reliability of the speaker in a truth-conditional sense. The interaction between social and truth-conditional aspects of utility is complex and this paper is not the place to explore it; we will do this in future work.

<sup>5</sup>Where ‘ $\gg$ ’ indicates a vast difference in value.

<sup>6</sup>Space considerations preclude spelling this additional mechanism out in the present paper; but the basic idea is to allow, instead of simple truth-tracking as a vetting mechanism for discourse moves, either checking

strategy, while prima facie aberrant, can be rational in terms of maximizing utility in a variety of scenarios. One case would be in scenarios where social signals are more easily interpreted. In a political campaign where voters may not know much about certain policy domains, sending a true message about those policies may not actually help the listener pick the true state of the world. If these message are additionally uninformative in terms of social meaning, then it would be rational in terms of payoffs to pick an alternative message that is false, but which clearly sends a message about the speaker's persona / ideology. In the extreme case of this strategy, which we've been taking Trump to exemplify, over a conversational history interlocutors become ever more sure of each other's sociolinguistic personas and attendant ideologies, but do so by discounting truth-conditional considerations.

While  $\delta > \gamma$  discounts truth conditional information, note that this setting does not necessarily lead to perverse truth-conditional outcomes. Remember that on first interaction with a speaker, a listener must fix a prior for reliability based on, in part, the speaker's persona (indicated by non-linguistic social cues). We expect that one viable strategy to evaluate a speaker's reliability is to, early in the conversation, set  $\delta > \gamma$  so to better align on their persona and likely ideological commitments. Crucially, this can be truth-conditionally virtuous. Recall, for instance, that certain dogwhistles only get their enriched meaning once a listener has identified the persona of the speaker. This means that social meaning can affect the truth conditions of messages, and so having an accurate picture of the speaker is crucial. Focusing social meaning by setting  $\delta > \gamma$  early, and then switching to  $\gamma > \delta$  in order to monitor the truth-tracking of an agent whose persona/ideology has been sussed out is thus a viable strategy to maximize payoffs in both communicative domains.

We have shown how focusing on the social meaning domain is a viable strategy for maintaining high payoffs in repeated communication, and thus can be used to establish trust. We now turn to two communicative moves that can be used by speakers who are employing this strategy. The first, signals that, indeed they are pursuing this strategy. The second kneecaps alternative strategies.

In discussing Trump's communication style, multiple pundits and partisans have said that we must "take Trump seriously, not literally". What could this mean? In our account, the meaning is clear. It is an exhortation to set  $\delta > \gamma$ , or even to set  $\gamma = 0$ . The idea is that we should not pay attention to the literal truth-conditional content of his utterances, but to instead take him seriously. But take him seriously as what? We think it means to take his social persona as a billionaire, as a businessman, as a fighter, as a nationalist, etc. seriously. This rhetorical move immediately makes sense in the framework we have developed, and shows that social meaning and truth conditional meaning is separable, and that we can prioritize one or the other.

Moreover, with this setup, we are in a position to give an analysis of the *fake news* effect. In some conservative/right-wing discourse, taking a statement and calling it *fake news* is a signal to remove it from truth-relevant consideration. We see several different ways to model this within the present setup. The first is to view it as a call to set  $\gamma = 0$  in its payoff evaluation. The interpreter then considers only similarity of persona (ie. politics) in the utility calculation. With this, if the fake news move is accepted, considerations of truth become completely (though temporarily) irrelevant to political discourse. This is initially plausible, but it has several shortcomings. First, since it eliminates any utility stemming from the recovery of truth-conditional information, it leads to generally lower utilities for 'fake news' content; while

---

whether accurate information about social persona is transferred or whether the total utility of the discourse move exceeds a certain level, i.e. the communication is sufficiently useful that it's worth keeping on paying attention to the communicating agent. These two views give subtly different results which we plan to explore further in [Henderson and McCready \(2019\)](#).

this may be part of the intended effect, it is not clear to us that this way of computing utility properly tracks the way in which we value such content. Even if something is deemed to be fake news, there is a sense in which the truth-conditional content still provides some value to the interpreter, specifically that she comes to believe that the content is *false*, which is already useful in a way that should be reflected in utility calculation.

We thus think that *fake news* is best viewed as a kind of hedge which functions as a denial operator. Simultaneously, it has a pragmatic effect as an invitation to the hearer to ascribe a persona in ideological opposition to whatever agent made the original claim. Given the way in which we tie opposition and thus negative affect to lack of similarity, the ascription of low similarity will result in a negatively viewed persona. We use the following definition.<sup>7</sup>

$$(9) \quad FN(\varphi) = \begin{array}{l} (i) \text{ Deny}(\varphi) \\ (ii) \neg Sim(P_s, P_h) \end{array}$$

The utility of this operator can be motivated in various ways, but here is our preferred take. Imagine you're playing a Trump strategy in which the goal is to gain favor via the use of social meaning rather than being a reliable communicator in a truth-conditional sense. You would want to tell people that other news is fake so that they can't play a strategy which maximizes truth-conditional meaning: the explicit denial of a bit of content means that one won't be able to extract utility from it on the truth-conditional side of the utility calculation, so it will only be sensible to try to extract value from the social meaning side. As a result, the hearer (who accepts this discourse move) has to play the social meaning strategy-maximizing game you're playing, which is your preferred outcome, whether because you are better at the social meaning side in general, or because it is better for you to avoid strategies which pay attention to truth-conditional content in the particular case in question.

## 5 Conclusions

Trust is distinct from reliability. We can trust the unreliable, and not trust the reliable. Moreover, contrary to first impressions, this is even a rational way to behave, if we value social matching over truth—which we might in the case of political speech, if we care most about what kind of policies a politician might want to implement. This insight leads to a major takeaway. If we want to re-imbue politics with truth, pointing out the falsehoods of politicians is not a productive method in general. Instead, our only way forward is to show that the ideological presentation of those politicians is insincere, or that the ideology that they uphold is itself flawed.

## References

- Bethany L Albertson. Dog-whistle politics: Multivocal communication and religious appeals. *Political Behavior*, 37(1):3–26, 2015.
- Alexandru Baltag and Sonja Smets. A qualitative theory of dynamic belief revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory*, number 3 in Texts in Logic and Games, pages 13–60. Amsterdam University Press, 2008.

<sup>7</sup>As with the work of McCready (2015), the precise way in which hedges behave compositionally is left underdetermined here.

- Heather Burnett. Signalling games, sociolinguistic variation and the construction of style. *Linguistics and Philosophy*, 2018.
- Michael Franke and Judith Degen. Reasoning in reference games: Individual-vs. population-level probabilistic modeling. *PloS one*, 11(5):e0154854, 2016.
- Michael Franke and Gerhard Jäger. Probabilistic pragmatics, or why bayes??? rule is probably important for pragmatics. *Zeitschrift für sprachwissenschaft*, 35(1):3–44, 2016.
- Miranda Fricker. *Epistemic Injustice*. Oxford University Press, 2007.
- Noah D Goodman and Michael C Frank. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11):818–829, 2016.
- Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.
- Robert Henderson and Elin McCready. How dogwhistles work. In *Proceedings of LENLS 14*. JSAI, 2017.
- Robert Henderson and Elin McCready. Dog-whistles and the at-issue/not-at-issue distinction. In Daniel Gutzmann and Katherine Turgay, editors, *Secondary Content*, pages 191–210. Brill, 2018.
- Robert Henderson and Elin McCready. *Signaling without Saying: The Semantics and Pragmatics of Dogwhistles*. 2019. To appear from Oxford University Press.
- E. McCready. *Reliability in Pragmatics*. Oxford University Press, 2015.
- E. McCready and Grégoire Winterstein. Negotiating epistemic authority. In Kurahashi, Setsuya and Ohta, Yuiko and Arai, Sachiyo and Satoh, Ken and Bekki, Daisuke, editor, *New Frontiers in Artificial Intelligence (JSAI-isAI 2016 work-shops, LENLS, HAT-MASH, AI-Biz, JURISIN and SKL revised selected papers)*, volume LNAI 10247, pages 74–89. Springer, Berlin, 2017.
- Elin McCready and Grégoire Winterstein. Testing epistemic injustice. *Investigationes Linguisticae*, XXXVII, 2019.
- Robert van Rooij. Game theory and quantity implicatures. *Journal of Economic Methodology*, pages 261–274, 2008.
- Erica J Yoon, Michael Henry Tessler, Noah D Goodman, and Michael C Frank. Talking with tact: Polite language as a balance between kindness and informativity. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 2016.

# *Only* reconstruction and backwards association\*

Aron Hirsch<sup>1</sup> and Michael Wagner<sup>2</sup>

<sup>1</sup> The Hebrew University of Jerusalem, Israel

aronh@alum.mit.edu

<sup>2</sup> McGill University, Montréal, Canada

chael@mcgill.ca

## Abstract

Büring and Hartmann (2001) provide evidence that when the German exclusive particle *nur* linearly precedes a sentence-initial fronted DP that reconstructs, it cannot reconstruct along with the DP. They conclude with Jacobs (1983) that *nur* is always a propositional operator on the clausal spine. However, Smeets and Wagner (2018) show that *nur* can take reconstructed scope after all, and defend an analysis where *nur*+DP may be parsed together as a single quantifier. Yet, a comparison with English reveals that *only* cannot reconstruct, bearing out the pattern that Büring and Hartmann (2001) had suggested for German. We propose a new analysis that can account for both languages. We argue for our analysis based on a correlation between the apparent reconstruction of the exclusive with the DP and backwards association possibilities.

## 1 Introduction

As defined in (1), *only* composes with a proposition (‘prejacent’). Our aim is to assess whether *only* can also compose with a quantificational DP. Rooth (1985) and Wagner (2006) in different ways associate *only* with a meaning by which it can compose with a quantifier to output a new quantifier. In Rooth’s 1985 perspective, *only* is systematically ambiguous between (1) and further meanings of other semantic types. One is (2), taking a quantificational input.

$$(1) \quad \llbracket \text{only}_{st} \rrbracket^{ALT} = \lambda p_{st} . \lambda w : p(w) . \forall p' \in ALT [p'(w) \rightarrow p \subseteq p']$$

$$(2) \quad \llbracket \text{only}_Q \rrbracket^{ALT} = \lambda Q_{est,st} . \lambda f_{est} . \llbracket \text{only}_{st} \rrbracket^{ALT}(Q(f))$$

Our investigation will home in on one test environment: data where *only* is sentence-initial, and precedes a DP which has achieved its surface position by moving from a site lower in the clause, and can reconstruct to the low position. If *only* can compose with the DP, it should be able to reconstruct along with the DP to take low scope. We will insert an operator with respect to which *only* is scopally non-commutative into the structure between the landing site of movement and the origin, and test whether *only* can scope below the operator. Our primary test operator will be the adverbial *again*.

Testing for reconstructed scope in both German and English will yield conflicting results. In German, it appears that *nur* (‘only’) is able to scope below the additional operator, while in English, *only* will fixedly take wide scope over the operator. We will be concerned with reconciling the differential scope behavior in the two languages.

We will take the English data seriously as demonstrating that (2) is not an available meaning in any language, and propose an analysis for how *nur* takes low scope in German that does not rely on (2). When *nur* appears to reconstruct with a DP, the interpreted operator is a covert instance of (1), which attaches on the clausal spine under *again* and “backwards associates” with the fronted DP. Backwards association is independently blocked in English.

\*For their comments and judgments, we thank Bernhard Schwarz, Meghan Clayards, members of prosodylab, and three AC reviewers, among others. We, of course, bear sole responsibility for all errors.

## 2 Nur can take low scope in German

In (3), *nur* appears at the left edge of the sentence, preceding an object DP which has fronted to satisfy V2 requirements. The lower subject is a universal quantifier. [Büring and Hartmann \(2001, 260\)](#) report that *nur* is limited to take wide scope over the subject.

- (3) Nur das *Abstract* hat jeder gelesen.  
 only the abstract has everyone read  
 ‘Only the abstract is such that everyone has read it.’ (B&H: *only* >  $\forall$ ;  $\forall$  > *only*)

The fronted DP itself, by contrast, can reconstruct, as shown in (4), where the DP is existential. As a result, Büring & Hartmann conclude that (3) must involve a verb-third structure, with *nur* adjoined to the CP, interpreted as (1). More generally, this analysis, following [Jacobs \(1983\)](#), is based on the idea that *only* can never form a constituent with an (argument) DP.

- (4) Ein Abstract hat jeder gelesen.  
 an abstract has everyone read  
 ‘Everyone has read an abstract.’ ( $\exists$  >  $\forall$ ;  $\forall$  >  $\exists$ )

[Smeets and Wagner \(2018\)](#), however, find that inverse scope in (3) is detectable when the fronted DP is accented, and the remainder of the sentence deaccented. Moreover, they note that *nur* readily takes scope below *wieder* (‘again’) in (6). With low scope for *nur*, the prejacent of *wieder* conveys that Jan and no one else has failed in the present instance. Accordingly, the presupposition that *wieder* introduces is that there is a prior instance where Jan and no one else failed. The brief discourse in (5) offers a set up supporting the presupposition.

- (5) A: Last time, Jan failed the quiz, and everyone else passed.  
 B: And this time?
- (6) Nur *Jan* ist wieder durchgefallen.  
 only Jan is again flunked  
 ‘Only Jan has flunked again.’ (*only* > *again*; *again* > *only*)

In response, [Smeets and Wagner \(2018\)](#) claim that *nur* can compose with a quantifier, after all. As in (7), *Jan* Montague lifts to quantifier type to furnish a licit argument for *nur* with the meaning in (2). In turn, V2 can reconstruct, as in (8), to assign *nur*+DP low scope.

- (7)  $[[\text{Jan}^\uparrow]] = \lambda f_{est} . \lambda w . f(\text{Jan})(w)$   
 (8)  $[_{CP} \text{ist}_2 [_{TP} [\text{wieder} [_{vP} [\text{nur} [_{DP} \text{Jan}]] \text{durchgefallen}]]] t_2]]$

Importantly, note that narrow scope for *nur* is distinguishable from an alternative wide scope reading. If *nur* scoped over *wieder*, as in (9), a presupposition would be triggered that Jan has failed before. Moreover, each alternative would contain *wieder*, and thus carry a presupposition that *x had failed before*, for the given individual *x*. Assuming these presuppositions project, the global presupposition would require *people other than Jan* to have failed, as well (see [McKillen, 2016](#)). Surface scope is, of course, available, but crucially, inverse scope is, too.

- (9)  $[_{CP} [\text{nur} [_{DP} \text{Jan}]] \lambda 1 \text{ist}_2 [_{TP} t_1 [\text{wieder} [_{vP} t_1 \text{durchgefallen}]]] t_2]]$

[Smeets and Wagner \(2018\)](#)’s analysis makes a prediction: that the scope of *nur* should correlate with the reconstruction possibilities of fronted quantifiers. The possibility of narrow scope of *nur* under *wieder* does parallel (10), where a fronted existential can scope beneath *wieder* in



kind. On the reconstructed reading, (10) presupposes not that there is someone who has failed this time who also failed before, but rather just that someone or other failed before.

- (10) Jemand ist wieder durchgefallen.  
 someone is again flunked  
 ‘Someone flunked again.’ ( $\exists > \textit{again}$ ;  $\textit{again} > \exists$ )

Moreover, there are cases where quantifier reconstruction is blocked, and low scope for *nur* is unattested, too. For one, if we replace *wieder* with *zum zweiten Mal* (‘for the second time’), as in (11), then the quantifier obligatorily scopes wide. The fact with *nur* in (12) correlates: (12) seems infelicitous in the context in (5), indicating that *nur* cannot take narrow scope.

- (11) Jemand ist zum zweiten Mal durchgefallen.  
 someone is for the second time flunked  
 ‘Someone flunked for the second time.’ ( $\exists > \textit{second time}$ ;  $*\textit{second time} > \exists$ )
- (12) Nur Jan ist zum zweiten Mal durchgefallen.  
 only Jan is for the second time flunked  
 ‘Only Jan has again flunked.’ ( $\textit{only} > \textit{second time}$ ;  $*\textit{second time} > \textit{only}$ )

While quantifiers often can reconstruct from first position in a V2 sentence, reconstruction is blocked, including with *wieder*, when the quantifier remains in the ‘middle field’, as in (13). As observed in Smeets and Wagner (2018), low scope of *nur* is likewise impossible in (14). (See Frey (1993); Lechner (1998); Sternefeld (2001); Lechner (2018) for thorough discussions of reconstruction in the German middle field.)

- (13) Gestern ist jemand wieder durchgefallen.  
 gestern is someone again flunked  
 ‘Yesterday, someone flunked again.’ ( $\exists > \textit{again}$ ;  $*\textit{again} > \exists$ )
- (14) Gestern ist nur Jan wieder durchgefallen.  
 gestern is only Jan again flunked  
 ‘Yesterday, only Jan flunked again.’ ( $\textit{only} > \textit{again}$ ;  $*\textit{again} > \textit{only}$ )

Overall, an analysis of *nur* with (2) as an available meaning seems natural within German. That way, *nur* can compose with a following DP to form a quantifier and reconstruct to take narrow scope, except where reconstruction is independently blocked. In so far as logical vocabulary is universal, if (2) is attested in German, (2) should be available across languages. In this respect, a complication does, however, arise with English *only*.

### 3 English *only* cannot take low scope

McKillen (2016, 118) observed that when *only* is initial and *again* attaches at the *vP*, *only* takes fixed wide scope over *again*. We illustrate with (15). As noted by Smeets and Wagner (2018), this linear position of *again* is somewhat marked. Still, most speakers who accept the word order report a contrast between the two scope readings.

- (15) Only Sue has again failed. ( $\textit{only} > \textit{again}$ ,  $*\textit{again} > \textit{only}$ )

The context in (16) supports low scope for *only* and, as shown in (17-a), (15) seems infelicitous.



(17-b), where *only* takes transparent low scope, offers a point of comparison.<sup>1</sup>

(16) A: Last time, Sue failed the quiz, and everyone else passed. B: And this time?

(17) a. #Only Sue has again failed (the quiz). The others passed, like last week.<sup>2</sup>  
b. Again, only Sue has failed.

In (18), a baseline context is provided supporting surface scope, and (15) improves. The residual ‘?’ in (19) reflects some speakers’ general hesitance about the word order.

(18) A: Last time, Mary, John, and Sue failed the quiz. B: And this time?

(19) ?Only Sue has again failed (the quiz). The others passed this time.

While English does not exhibit V2, the initial DP is still in a derived position, as the subject A-moves from its thematic merger site in spec-*v*P to spec-TP, over adverbs such as *again*. In general, a quantifier undergoing such movement can reconstruct:

(20) Sue failed last week’s quiz. This week, someone has again failed (the quiz).

Like German (10), the second sentence in (20) does not necessarily presuppose that the person who failed this time has failed before, but rather just that someone or other failed before. This is clear, since the overall sequence in (20) need not convey that Sue failed twice. Hence, *someone* must be able to reconstruct under *again*. But then, if *only*+DP could be parsed as quantifier in (15), that should be able to reconstruct, too, to assign *only* low scope.

The correlation between quantifier reconstruction and the scope of the exclusive which seemed to hold *within* German breaks down *across* languages. English and German are alike in that a sentence-initial DP can reconstruct beneath *again*. But, the two languages come apart in the patterning of the exclusive: *nur* can take low scope; *only* cannot.

### 3.1 The paradox

To avoid over-generating unattested low scope in English, a parse where *only* reconstructs with the DP must be ruled out. This is directly predicted if (2) is simply not an available meaning, so that *only* cannot compose with a DP. Yet, how can German *nur* then take narrow scope? If *nur/only* directly encode (1), (6) and (15) must both be parsed with *nur/only* adjoined at the edge of the matrix clause, where they should take widest scope, as in Buring & Hartmann’s proposal. One possibility is that languages make use of different logical vocabulary, and that (2) is available in German, but not in English. We show, however, that the data receive a principled reconciliation in a system where just (1) is available in both languages.

## 4 Re-analyzing *only* and scope

Quek and Hirsch (2017) and Hirsch (2017) (hence QH) propose that the interpreted operator is always (1), but deny that overt forms like *nur* and *only* directly encode that operator. By

<sup>1</sup>As observed in McKillen (2016), the pattern of English *again* is unexpected in the account of ‘weakened projection’ in Sauerland (2013), where certain presupposition can be ignored in focus alternatives sets and therefore fail to project. This account could have offered an alternative interpretation of the German facts.

<sup>2</sup>A reviewer reports a conflicting judgment, and accepts (17-a), but notes adding a direct object might matter. McKillen’s example involved an object, maybe sharpening the contrast: *Only John has again forgotten his homework*. We have to leave this for future study. Note (6) allows inverse scope with or without an object.

dissociating the meaning from the overt morphology, they allow the propositional operator itself to attach at non-transparent sites in some cases. In this way, *only* can be perceived as taking different scopes, but without ever scoping together with a DP.

#### 4.1 Dissociating semantics and phonology

QH take *only* to be a front for *two* syntactic heads (for similar analyses, see Lee 2004; Barbiers 2014; Hole 2015; Bayer 2016). The higher head, ONLY, occurs on the clausal spine, and is interpreted as the propositional operator in (1). The lower head, F, is a focus marker, and attaches more local to the focus associate of ONLY. F is semantically inert. The basic sentences in (21) share the common underlying syntax in (22).

- (21) a. Mary *only* read one<sub>Foc</sub> book.                      b. Mary read *only* one<sub>Foc</sub> book.

- (22) [<sub>TP</sub> Mary [ONLY<sub>[iONLY]</sub> [<sub>vP</sub> t<sub>1</sub> read [<sub>FP</sub> [F<sub>[uONLY]</sub>] [<sub>DP</sub> one<sub>Foc</sub> book]]]]]]

The two heads are specified for an operator-specific feature, and stand in an Agree relation. In one conception, where probe-goal relationships are established through upwards Agree (e.g. Zeijlstra, 2012; Zeijlstra and Bjorkman, 2019), F probes and Agrees with ONLY above it. While the semantics interprets just the higher head, QH propose that the phonology can optionally realize the [ONLY] feature on either head. In (21-a), overt *only* realizes ONLY, while F is left covert. In (21-b), conversely, F is realized, and ONLY covert.

QH’s approach has an important consequence for DP-level *only*. Because ONLY is a propositional operator, when attached to a DP, *only* must realize F. In that case, then, the semantics and phonology dissociate, and scope ambiguity may arise. To illustrate, consider (23). Taglicht (1984) observed that *only* in (23) can take scope above or below *require*.

- (23) You are required to learn *only* Spanish.                      (*require* > *only*; *only* > *require*)

From QH’s perspective, pre-DP *only* in (23) realizes F. Because ONLY is covert, the string is compatible with different attachment sites for ONLY. One possible structure is (24), where ONLY attaches at the edge of the embedded *vP*. Another is (25), where ONLY attaches at the edge of the matrix *vP* instead. Because just ONLY is interpreted, its site fixes scope, and (24) and (25) yield the two observed readings for (23).

- (24) [<sub>TP</sub> you λ1 are [ONLY [<sub>vP</sub> required [<sub>TP</sub> t<sub>1</sub> to [<sub>vP</sub> t<sub>1</sub> learn [<sub>FP</sub> F [<sub>DP</sub> Spanish]<sub>Foc</sub>]]]]]]]]

- (25) [<sub>TP</sub> you λ1 are [<sub>vP</sub> required [<sub>TP</sub> t<sub>1</sub> to [ONLY [<sub>vP</sub> t<sub>1</sub> learn [<sub>FP</sub> F [<sub>DP</sub> Spanish]<sub>Foc</sub>]]]]]]]]

Hence, the route to scope variability is ambiguity in the attachment site of the interpreted ONLY head when it is covert. *Only* never semantically composes with a DP.

### 5 Back to initial *only*

Equipped with QH’s analysis, we return to sentence-initial *only*, and set about to reconcile the conflicting German and English facts. QH’s scope mechanism readily predicts the contrast, once we take into account independent generalizations about the placement of the propositional operator relative to its focus associate, which differ between the two languages.

## 5.1 Fixed scope in English

We first consider English (26), repeated from Section 3, where *only* takes obligatory wide scope over *only*. The structure in (27) straightforwardly derives the attested reading. The subject DP is focused, and F attaches local to the DP. The interpreted ONLY head attaches on the clausal spine, above the FP in spec-TP. Overt *only* can realize either ONLY or F and, one way or the other, the string in (26) results. Because ONLY is attached at the edge of the clause, it takes widest scope, above *again*.

(26) Only Sue has again failed (the quiz). (*only* > *again*; \**again* > *only*)

(27) [ONLY [<sub>TP</sub> [<sub>FP</sub> F [<sub>DP</sub> Sue]<sub>Foc</sub>] λ1 has [again [<sub>vP</sub> t<sub>1</sub> failed]]]]

Indeed, from the surface syntax in (27), there is no path to narrow scope for ONLY. The FP has undergone A-raising from spec-*vP* to spec-TP, and that movement could reconstruct to assign the FP narrow scope, as in (28). However, because F is semantically inert, reconstruction in (28) only diminishes the scope of the DP, as we noted was possible in Section 3. Being outside the raised constituent, the scope of the interpreted ONLY remains wide.

(28) [ONLY [<sub>TP</sub> has [again [<sub>vP</sub> [<sub>FP</sub> F [<sub>DP</sub> Sue]<sub>Foc</sub>] failed]]]]

To derive the unattested scope order, a different surface syntax would be required, one which merges ONLY lower, as in (29). As before, the FP raises from spec-*vP* to spec-TP, but now ONLY is beneath the landing site of movement, under *again*. Since initial *only* could be DP-level and realize F, the structure in (28) is compatible with the surface string.

(29) \* [<sub>TP</sub> [<sub>FP</sub> F [<sub>DP</sub> Sue]<sub>Foc</sub>] λ1 has [again [ONLY [<sub>vP</sub> t<sub>1</sub> failed]]]]

However, (29) runs afoul of a well-known constraint in English: *only* cannot associate with a focus outside its scope in the surface syntax. In other words, the propositional operator cannot “backwards associate” with a focus to its left. Jackendoff (1972) first observed the prohibition on backwards association. For the case at hand, the baseline is (30), where *only* is overtly realized at the putative position of ONLY in (29). Crucially, (30) is impossible on a reading where *Sue* is the associate of *only*.

(30) \*Sue<sub>Foc</sub> has again only failed.

From QH’s perspective, (30) is just another way of realizing the structure in (29). Whereas the string in (26) results if F is overt, the string in (30) results if ONLY is overt. What we learn from (30), then, is that the structure in (29) is ill-formed and, accordingly, it would also be ruled out as a possible parse of (26) as well. Just as overt ONLY cannot associate with a focus to its left in (30), covert ONLY could not associate with a focus to its left in (26).

It merits comment that the string in (30) is ungrammatical, while the string in (26) is grammatical, but lacks a narrow scope reading for *only*. Because *Sue* is focused, F must attach local to it and, in turn, overt *only* in (30) must be parsed as a realization of ONLY, as it is remote from *Sue*. The illicit structure in (29) is, therefore, the only parse compatible with the string in (30). On the other hand, as noted, the string in (26) is also compatible with high attachment of ONLY in (27), which does not involve backwards association.

In sum, there is no licit path to inverse scope in (26). The FP surfaces in spec-TP. Given that backwards association is prohibited, ONLY, whether overt or covert, must attach higher, and accordingly, ONLY fixedly takes widest scope.

## 5.2 Ambiguous scope in German

German differs from English in permitting backwards association. Consider again the example in (31), repeated from (6), where *nur* can take scope below, as well as above, *wieder*.

- (31) Nur *Jan* ist wieder durchgefallen.  
 only Jan is again flunked  
 ‘Only Jan has flunked again.’ (*only* > *again*; *again* > *only*)

Narrow scope for *nur* derives as in (32). The FP has undergone fronting to spec-CP for V2, and ONLY attaches at a position within the middle field below *wieder*, and backwards associates with the fronted DP to its left. The structure is analogous to (28) above, and compatible with the string in (31), if F is overtly realized, and ONLY covert.

- (32)  $[_{CP} [_{FP} F [_{DP} Jan]_{Foc}] \lambda 1 \text{ ist}_2 [_{TP} t_1 [\text{wieder } [_{\text{ONLY}} [_{vP} t_1 \text{ durchgefallen}]]] t_2]]$

The crucial difference from English emerges in the baseline in (30), where *nur* is pronounced at the site of ONLY. Unlike its English analog in (30), (33) is acceptable if *nur* associates with *Jan*, indicating that backwards association to a constituent in first position of a V2-clause is possible in German (cf. Jacobs, 1983).

- (33) *Jan*<sub>Foc</sub> ist wieder nur durchgefallen.  
 Jan is again only failed  
 ‘Only Jan has failed again.’

In QH’s system, (33) shows that (32) is licit, and (32) should thus be an available parse for (31) too, leading to a low scope reading for *nur*.

Wide scope for *nur* could derive as in (34), where ONLY attaches to the matrix CP, above the fronted FP, similar to English (27). However, (34) shares with Buring & Hartmann’s analysis that the derivation is verb-third, which would be exceptional for German.

- (34)  $[_{\text{ONLY}} [_{CP} [_{FP} F [_{DP} Jan]_{Foc}] \lambda 1 \text{ ist}_2 [_{TP} t_1 [\text{wieder } [_{vP} t_1 \text{ durchgefallen}]] t_2]]]$

Given that backwards association is licit, wide scope for *nur* can, in fact, derive without V3. The parse is (35). As in (32), ONLY attaches in the middle field and backwards associates with the FP, but now ONLY is at a site above *wieder*. Wide scope for ONLY results, and the correct surface string is again output if *nur* realizes F.

- (35)  $[_{CP} [_{FP} F [_{DP} Jan]_{Foc}] \lambda 1 \text{ ist}_2 [_{TP} t_1 [_{\text{ONLY}} [\text{wieder } [_{vP} t_1 \text{ durchgefallen}]]] t_2]]]$

Hence, with *nur* realizing F, scope ambiguity results from different options as to where covert ONLY is placed in the middle field, which are available because German, unlike English, allows for backwards association. ONLY can be placed at a site below, as well as above, *wieder*.

Of course, the question remains as to *why* English and German differ with respect to backwards association (see Erlewine 2014 for a proposal). Yet, what’s crucial for our purpose is that the baseline examples in (30) and (33) show that there is such a difference. Given the difference—regardless of its explanation—observed scope freezing in English and freedom in German “follows for free” from QH’s analysis, generalized to both languages.

### 5.3 Fixed scope in German

We noted in Section 2 that Smeets & Wagner predicted scope restrictions on exclusive operators should correlate with restrictions on quantifier reconstruction. Our approach makes a different prediction: that scope restrictions with exclusive operators should correlate with restrictions on backwards association. We've seen that differences *between* English and German follow from that prediction. Scope restrictions *within* German do, as well. In Section 2, we identified environments where *nur* cannot take low scope. In addition to quantifier reconstruction being blocked in those cases, backwards association is, too.<sup>3</sup>

First, we observed in (12) that *nur* takes obligatory wide scope over *zum zweiten Mal* ('for the second time'), unlike *wieder*. To derive narrow scope, we would require ONLY to attach low and backwards association with *Jan* across the adverbial. (36) shows that backwards association with the DP in first position is independently impossible across *zum zweiten Mal*.

- (36) \**Jan<sub>Foc</sub>* ist zum zweiten Mal nur durchgefallen.  
 Jan is for second time only flunked  
 'For the second time only Jan flunked.'

Likewise, we observed in (14) that *nur* cannot scope beneath *wieder* when it precedes a DP that is in the middle field. While backwards association across *wieder* to a DP in first position in V2 is possible, backwards association to a DP in the middle field is not. (37) is unacceptable if *nur* associated with *Jan*.

- (37) \*Gestern ist *Jan<sub>Foc</sub>* wieder nur durchgefallen.  
 gestern is Jan again only flunked  
 'Yesterday, only Jan again flunked.'

In addition to having the new empirical benefit of accounting for differences in scope of the exclusive relative to *wieder/again* between languages, our system thus maintains an account of scope freedom and fixing within German. Based on (36) and (37), we predict the impossibility of low scope for *nur* in (12) and (14) from Section 2.

## 6 A further prediction

In discussing English, we have focused entirely on data with *again*. Before concluding, we want to flag one additional correct prediction of our account for English. The prediction involves the perceived scope of sentence-initial *only* in subject raising configurations. The expectation is that *only* should take obligatory wide scope over the raising predicate:

- (38) Only Sue is certain to get an A. (*only* > *certain*; \**certain* > *only*)

This is supported in (38). The sentence naturally allows a reading where *only* scopes over *certain*, as paraphrased in (39-a). If *only* took narrow scope, the reading in (39-b) would result instead. The former, but not the latter, leaves open the possibility that students other than Sue might get an A, as well. The data in (40) demonstrate that *only* cannot take narrow scope: if we already know the outcome for every student, we cannot use (38) to convey what we are certain about. This stands in contrast to a baseline where *only* takes transparent low scope under *certain*, which is licit in the same context.

<sup>3</sup>This correlation might suggest that the analysis of backwards association depends on reconstruction of the associate into the scope of ONLY. We leave further consideration to a future occasion.

- (39) a. It's certain Sue will get an A, and not certain anyone else will. (*only* > *certain*)  
 b. It's certain Sue and no one else will get an A. (*certain* > *only*)
- (40) We already know exactly what will happen.  
 a. #Only Sue is certain to get an A.  
 b. It's certain that only Sue will get an A.

Convergent with conclusions from the *again* data in Section 3, scope freezing in (38) would be surprising if (2) were an available meaning for *only*. In general, A-raising can reconstruct (see e.g. Barss 1986, Romero 1998, Fox 1999, Sportiche 2006, among other works). For example, the following example shows a scope ambiguity, and there is a reading where no particular student is certain to get an A:

- (41) Some student or other is certain to get an A. ( $\exists$  > *certain*; *certain* >  $\exists$ )

If *only* could compose with the DP in (38), it should be able to reconstruct with the DP and take narrow scope in kind. On the other hand, our proposal predicts the pattern. Just the propositional operator is available, encoded in ONLY, and ONLY must take wide scope, as in (42), not narrow scope, as in (43), since backwards association is disallowed in English.

- (42) [ONLY [<sub>TP</sub> [<sub>FP</sub> F [<sub>DP</sub> Sue]<sub>Foc</sub>]  $\lambda$ 1 [<sub>vP</sub> is certain [<sub>TP</sub> t<sub>1</sub> to [<sub>vP</sub> t<sub>1</sub> get an A]]]]]]

- (43) [<sub>TP</sub> [<sub>FP</sub> F [<sub>DP</sub> Sue]<sub>Foc</sub>]  $\lambda$ 1 [<sub>vP</sub> is certain [<sub>TP</sub> t<sub>1</sub> to [ONLY [<sub>vP</sub> t<sub>1</sub> get an A]]]]]]

A broader range of quantifiers and operators should be looked at to see how general the pattern for *only* is. For example, our proposal is similar to an account for related missing reconstruction effects with negative quantifiers. Iatridou and Sichel (2011) noted that negation with a raised NegDP must scope over *certain* (*No one is certain to get an A*  $\neq$  *It's certain that no one will get an A*) (building on Lasnik 1999). To preclude negation reconstructing low, they took semantic negation to be separate from the DP, in a head on the clausal spine, above the subject (e.g. Penka, 2001; Zeijlstra, 2007; Penka, 2011). We extend this treatment to *only*.

## 7 Conclusion

We have presented evidence that exclusive operators are necessarily propositional operators. *Nur* and *only* can attach directly to a DP, but they then realize not the exclusive operator itself, but an inert focus head. The analysis avoids over-generation problems which would arise if English *only* could compose with a quantifier, and predicts a new correlation between backwards association and low scope for the exclusive within and across languages.

## References

- Barbiers, S. (2014). Syntactic doubling and deletion as a source of variation. In *Linguistic variation in the Minimalist framework*, pages 197–223. Oxford University Press.
- Barss, A. (1986). *Chains and anaphoric dependence*. PhD thesis, MIT.
- Bayer, J. (2016). Critical freezing in the syntax of particles. Ms. University of Konstanz.
- Büring, D. and Hartmann, K. (2001). The syntax and semantics of focus-sensitive particles in German. *Natural Language and Linguistic Theory*, 19:229–281.
- Erlewine, M. (2014). *Movement out of focus*. PhD thesis, MIT.

- Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry*, 30(2):157–196.
- Frey, W. (1993). *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus*, volume 35 of *Studia Grammatica*. Akademie Verlag, Berlin.
- Hirsch, A. (2017). *An inflexible semantics for cross-categorial operators*. PhD thesis, MIT.
- Hole, D. (2015). A distributed syntax for evaluative ‘only’ sentences. *Zeitschrift für Sprachwissenschaft*, 34(1):43–77.
- Iatridou, S. and Sichel, I. (2011). Negative DPs, A-movement, and scope diminishment. *Linguistic Inquiry*, 42(4):595–629.
- Jackendoff, R. S. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Ma.
- Jacobs, J. (1983). *Fokus und Skalen. Zur Syntax und Semantik der Gradpartikel im Deutschen*. Niemeyer, Tübingen.
- Lasnik, H. (1999). Chains of arguments. In *Current Studies In Linguistics Series*, volume 32, pages 189–216. MIT Press.
- Lechner, W. (1998). Two kinds of reconstruction. *Studia Linguistica*, 52(3):276–310.
- Lechner, W. (2018). A calculus for reconstruction and anti-reconstruction. In Schenner, M. and Krifka, M., editors, *Reconstruction Effects in Relative Clauses*, *Studia Grammatica*, pages 113–145. Akademie Verlag.
- Lee, Y. (2004). *The syntax and semantics of focus particles*. PhD thesis, MIT.
- McKillen, A. (2016). *On the interpretation of reflexive pronouns*. PhD thesis, McGill University.
- Penka, D. (2001). Kein muss kein rätsel sein. Magisterarbeit. Universität Tübingen.
- Penka, D. (2011). *Negative indefinites*. Oxford University Press.
- Quek, Y. and Hirsch, A. (2017). Severing focus form and meaning in Standard and Colloquial Singapore English. In *Proceedings of the North East Linguistic Society, UMass Amherst 2016*. GLSA, Amherst, Ma.
- Romero, M. (1998). Problems for a semantic account of scope reconstruction. In *Reconstruction: Proceedings of the Tübingen Workshop*. GLSA, Amherst, Ma.
- Rooth, M. (1985). *Association with Focus*. PhD thesis, University of Massachusetts, Amherst.
- Sauerland, U. (2013). Presuppositions and the alternative tier. In *Proceedings of SALT 23*, pages 156–173.
- Smeets, L. and Wagner, M. (2018). Reconstructing the syntax of focus operators. *Semantics & Pragmatics*, 11(6):1–27.
- Sportiche, D. (2006). Reconstruction, binding, and scope. *The Blackwell companion to syntax*, 4:35–93.
- Sternefeld, W. (2001). Semantic vs. syntactic reconstruction. In Kamp, H., Rossdeutscher, A., and Rhorer, C., editors, *Linguistic Form and its computation*, pages 145–182. Stanford: CSLI Publications.
- Taglicht, J. (1984). *Message and Emphasis. On focus and scope in English*, volume 15 of *English Language Series*. Longman, London and New York.
- Wagner, M. (2006). Association by movement. Evidence from NPI-licensing. *Natural Language Semantics*, 14(4):297–324.
- Zeijlstra, H. (2007). On the lexical status of negative indefinites in Dutch and German. *Journal of Comparative Germanic Linguistics*.
- Zeijlstra, H. (2012). There is only one way to agree. *The Linguistic Review*, 29:491–553.
- Zeijlstra, H. and Bjorkman, B. (2019). Checking up on ( $\phi$ )-agree. *Linguistic Inquiry*, 50(3):527–569.



# Loose Talk, Scale Presuppositions and QUD\*

Daniel Hoek

Princeton University  
dhoek@princeton.edu

## Abstract

I present a new pragmatic theory of loose talk, focussing on the loose use of numbers and measurement expressions. The account explains loose readings as arising from a pragmatic mechanism aimed at restoring relevance to the question under discussion (QUD), appealing to Krifka's notion of a measurement scale [7]. The core motivating observation is that the loose reading of a claim need not be weaker than its literal content, as almost all pragmatic treatments of loose talk have assumed (e.g. Lasersohn [10]). The loosening mechanism described here can be applied to a range of other linguistic phenomena as well.

## 1 Introduction

We often use precise measurements where only approximate quantities are intended. One can say “Ellen arrived at a quarter past six” without being dishonest or mistaken, even if Ellen in fact arrived just before 6.15. The reply “No you're wrong! She actually arrived at 6.14” is not just needlessly pedantic – it betrays a positive misunderstanding of the speaker's communicative intent. The same thing happens with numbers and measurements of every kind: in a typical context, “There were fifty thousand people at the rally”, “Billy's parrot is twenty inches tall”, and “The Earth is five billion years old” should be understood as *loose talk*, stating only the rough size of the crowd, height of the parrot, and age of the planet.

On the received view, speakers are not committed full, literal truth of their loose assertions, claiming only that the proposition they expressed is, in some sense, *close* to the truth. Below I build on observations by Sam Carter [2] that, I think, render this view untenable. I articulate a new, formal pragmatic account of loose talk, and show how it captures Carter's observation as well as a range of other tricky data points. While the loose use of numbers and measurements will be my primary focus here, the basic pragmatic mechanism described potentially has much wider application, and in §5 I briefly explore how it may be applied to other phenomena.

## 2 Loosening without Weakening

There is fairly wide agreement that loose talk is a pragmatic phenomenon [1][2][5][8][10][14][15] (but cf. [11]). There are a number of good reasons for that consensus, but the most important one is probably Lasersohn's observation of the following contrast [10]:

- (1) a. We popped the champagne around two o'clock, but after 2.02pm.
- b. # We popped the champagne at two o'clock, but after 2.02pm.

If the phrase “at two o'clock” literally *meant* ‘in the time interval around 2pm,’ then (1b) would have an unproblematic, consistent reading, just like (1a) does. But that is not what we find.

The entailment data also support a pragmatic approach. If Rob is 6'1.01", it follows that he is *over* 6'1". If Rob is 6'0.99", it follows that he is *under* 6'1". So taken literally, the claim

---

\*Thanks to Chris Barker, Sam Carter, Manuel Kriz, Philippe Schlenker and three anonymous reviewers at the Amsterdam Colloquium for helpful feedback and discussions about this material.



(2) Rob is six foot one.

apparently entails that Rob is in the tiny interval between 6'0.99" and 6'1.01".<sup>1</sup> But of course, a typical use of statement (2) conveys a weaker message, to the effect that *Rob's height is in the neighbourhood of the 6'1" mark*.

Most of the pragmatic accounts of loose talk just cited follow Lasersohn's classic treatment in assuming that the loose reading of a statement is always weaker than its literal content. While this assumption seems eminently plausible in light of simple examples like (1) and (2), it breaks down when we consider complex sentences.

## 2.1 Loose Talk under Embeddings

To begin with, Carter [2] noticed that the *negations* of such statements convey *stronger* messages than they literally express. The literal content of

(3) Rob is not six foot one

is very weak: (3) is true even if Rob's height is just the tiniest bit taller or shorter than 6'1". But in asserting (3), one ordinarily commits to something stronger, to the effect that *Rob is not even close to 6'1"*. The effective message of (3) is the negation of the weakened message conveyed by (2). Likewise, "Ellen did not arrive at 6.15" would ordinarily commit a speaker to something stronger than the literal content of that statement.

Carter's point about negation shows the widely held view of loosening as a form of weakening to be mistaken, and as such it is a counterexample to most extant approaches to loosening. Lasersohn and others view loosening as weakening and accordingly predict, incorrectly, that (3) weakens to triviality. The point extends to downward entailing environments generally:

(4) Everyone who arrives at 1 o'clock gets a free lunch

Again, any account on which loosening is a form of weakening predicts a loose reading for (4) that is even weaker than (4)'s literal content. But in fact, the loose reading of (4) is *Everyone who arrives around 1 o'clock gets a free lunch*, which is stronger than its literal content.

Relatedly, Lasersohn's particular theory also has trouble with conjunctions:

(5) Emma and Jack both weigh five stone.

On Lasersohn's account, the loose reading of (5) should be *Emma and Jack both have the same weight, which is close to five stone*. That is false if Emma weighs 71 pounds and Jack 68. But clearly (5) is loosely assertable in that situation. In fact, (5)'s loose meaning is the conjunction of the loose meanings of the conjuncts: *Emma weighs around five stone and so does Jack*.

## 2.2 Strict Comparatives

Strict comparatives are another example of strengthening through loosening.

(6) London is more than 30 miles away from here.

Again, Lasersohn and others predict (6) should have a weak loose reading. Something to the effect, say, that *London is at least 28 miles away from here*. But in fact it does not: (6) is true in no context if the distance to London is under 30 miles.

As observed by Solt [13], (6) can sometimes take a completely strict, literal reading, according to which it is true even if the distance to London is just slightly over 30 miles. But in addition, (6) can also take a strong loose reading. To see this, observe the following contrast:

<sup>1</sup>(2) also has a reading that *Rob is at least 6'1"*. Ignore such readings for present purposes.

- (7) A: London is 30 miles away from here.  
 B<sub>1</sub>: No, it's not. London is more than 30 miles away.  
 # B<sub>2</sub>: No, it's not. London is at least 30.1 miles away.

If A's remark is interpreted loosely, B<sub>2</sub>'s reply is infelicitous, presumably because it does not contradict the loose content of A's remark. But since B<sub>1</sub>'s reply is fine, it must have a loose reading that *does* contradict A, to the effect that *London is well over 30 miles away from here*.

### 3 Relevance and Scale Presuppositions

On the theory I want to propose, loosening always aims to restore *relevance*. It is a natural successor to the view that the loose reading of a statement is a relevant consequence of its literal reading [5][14][15]. On Stephen Yablo's articulation of that view [15], the loose reading of an assertion is basically the strongest proposition that (A) is entailed by its literal content and (B) is wholly relevant to the *question under discussion* (henceforth *QUD*).

As is common, I will model questions under discussion as partitions:

- (8) A *question*  $Q$  is a partition of the space  $\mathcal{W}$  of possible worlds. We may write  $w \sim_Q v$  to mean that  $Q$  groups the worlds  $w$  and  $v$  in the same partition cell.  
 (9) A proposition  $p \subseteq \mathcal{W}$  is *wholly relevant* to a question  $Q$  if and only if no partition cell of  $Q$  contains worlds where  $p$  is false and also worlds where  $p$  is true.

In other words, a proposition is wholly about a question just in case the truth value of that proposition is always settled by a complete answer to the question.

As we just saw, loosening need not be weakening, so we will revise condition (A). But I want to preserve the core idea behind (B). That idea is that, in a typical context where (2) or (3) is uttered, we are not interested in Rob's height to arbitrary levels of precision, so that the QUD is something like *What is Rob's height to the nearest inch?*. Consequently, (2) isn't wholly relevant because it is too specific: it specifies Rob's height to a greater level of precision than the QUD requires. And (3) isn't wholly relevant because it is not specific enough: it does not rule out any answer to the QUD. Loosening, then, is a pragmatic fix for irrelevance.

The gist of the present proposal is to keep (B), while replacing (A) with (C):

- (C) The loose reading of a claim involving measurements is *conditionally equivalent* to its literal reading, given the presupposition induced by the measurement scale.

(This formally resembles to Križ's treatment of plural indefinites; see [9], §3.A.) To flesh out this proposal, let me say a bit about scales and the presuppositions that attach to them.

#### 3.1 Scale Presuppositions

The proposed new condition (C) builds on Mandy Simons' idea that speakers sometimes make *contextual presuppositions* specifically in order to tie their literally non-relevant utterances to the QUD [12], and on Manfred Krifka's observation that measurement expressions are always used against the background of a *scale* [7][8]. A measurement scale is a 'menu' of conventionally associated measurement expressions. For instance, in English-speaking countries, specifications of personal height are typically selected from the feet-and-inches scale:

$$\{ \dots \text{"5 foot 11"}, \text{"6 feet"}, \text{"6 foot 1"}, \text{"6 foot 2"} \dots \}$$

How do measurement scales like this one induce presuppositions?

Well, note there are in fact many possible heights intermediate between the ones on the list: most likely, your height is missing, as is mine. After all, almost no-one is an exact number of inches tall: one is always a little over or under. When we use the feet-and-inches-scale, we nonetheless ignore the intermediate possibilities. So when we describe Rob's height using this scale, as (2) and (3) do, there is a good sense in which we (falsely) presuppose that

- (10) Rob's height is on the feet-and-inches scale. (Or: *Rob is an exact number of inches tall.*)

When we describe Rob's height using the feet-and-inches scale, the QUD is naturally taken to be *What is Rob's height to the nearest inch?*, which corresponds to this partition:

$$H = \{ \{ w : \text{at } w, \text{ Rob's height in inches is in the interval } [n - \frac{1}{2}, n + \frac{1}{2}) \} : n \in \mathbb{N} \}$$

As I'll explain in §3.2, there is a unique proposition that is wholly relevant to  $H$  (condition (B)) and also conditionally equivalent to (2) given the scale presupposition (10) (condition (C)). And that is the following proposition:

- (11) Rob is six foot one to the nearest inch (that is, he's between  $6' \frac{1}{2}''$  and  $6' 1 \frac{1}{2}''$ ).

Thus the present account of loose talk predicts that (11) is the loose reading of (2).

The unique wholly relevant proposition conditionally equivalent to (3) is (11)'s negation:

- (12) Rob is not six foot one to the nearest inch (that is, he's closer to  $6'$  or to  $6' 2''$ ).

And thus the loose reading of (3) is correctly predicted to be stronger than its literal reading. The desired readings for (4) and (5) can be derived in closely analogous ways, using appropriately modified QUDs and scale propositions. In §3.3 below, I show how the present account captures the general pattern that emerges from these embedded instances of loose talk.

On strict comparatives, the predictions are more subtle. If the question whether the distance to London is under or over the 30-miles mark is part of the QUD, (6) retains its strict, literal reading (it is already wholly relevant in such a context). But in a context where this is not so, like (7), (6) gets a strong loose reading to the effect that *London is well over 30 miles away (being closer to the next item on the scale)*.

### 3.2 Subject Matter Completion

So far we have stated the present account of loose talk as follows: the loose content of an utterance is the unique wholly relevant proposition that satisfies condition (C) (if there is such a proposition). This subsection offers a more elegant and suggestive reformulation of this view.

According to (C), the assertion a loose speaker makes is equivalent to the message they intend to communicate, but only conditional on the scale presupposition they employ. One way to understand this is that speakers (and hearers) attend only to the *incremental content* of the assertion over that scale presupposition, where this is understood as a partial proposition restricted to worlds where the scale presupposition is true:

- (13) A **partial proposition** is represented by an ordered pair  $\langle t, f \rangle$  of disjoint sets of possible worlds.  $\langle t, f \rangle$  is *true* at the worlds  $w \in t$  and *false* at the worlds  $w \in f$ . It has no defined truth-value at worlds outside of  $t \cup f$ . (In particular, the pair  $\langle p, \neg p \rangle$  should be taken to represent the full proposition  $p \subseteq \mathcal{W}$ , where  $z\neg p =_{\text{df}} \mathcal{W} \setminus p$ .)
- (14) The **incremental content** of a full proposition  $p$  over the full proposition  $s$ , or the **restriction** of  $p$  by  $s$ , written  $p \upharpoonright s$ , is the partial proposition  $\langle p \cap s, \neg p \cap s \rangle$ .

Loose utterances fit the following pattern. Let  $p$  be the full literal content of the speaker’s utterance,  $s$  the operative scale presupposition, and  $p \upharpoonright s$  the incremental content of  $p$  over  $s$ . Then, while there are cells of the QUD that intersect both  $p$  and  $\neg p$ , no cells intersect both  $p \cap s$  and  $\neg p \cap s$ . So the incremental content  $p \upharpoonright s$  of a loose assertion is *wholly relevant* to the QUD, in the sense defined above, even if its full semantic content  $p$  is not.

This is illustrated with three examples in the diagram below. In particular, the leftmost column shows that the literal content of (2), *Rob is 6’1’’*, is not wholly relevant to  $H$ : that is, (2) “colours outside the lines” of  $H$ , so to speak. But if we restrict (2)’s content by the scale presupposition (10), the resulting partial proposition (2)∣(10) is wholly relevant to  $H$  – it no longer colours outside the lines. In this way, the idea that listeners attend to the incremental content of the speaker’s assertion, rather than the full literal content, is a first step towards explaining how we are able to interpret the assertion as wholly relevant.

But the incremental content (2)∣(10) is not itself the communicated content of (2). Like most people, Rob is almost certainly not an exact number of inches tall: we’re all a little over or a little under. So (2)∣(10) is probably not truth-evaluable at the actual world, whereas the loose message of (2) is: (2) should still count as loosely true if Rob’s height is only close to 6’1’’. So to get at the communicated content, we need to “fill out” the incremental content somehow.

Here we use the QUD as our guide: to maintain relevance, the target message must take on a uniform truth value within each  $H$ -cell, and it should match the incremental content (2)∣(10) wherever the latter is defined. The fact that (2)∣(10) is wholly relevant to  $H$  guarantees that these constraints can be jointly satisfied, since the incremental content is never both true and false within the same  $H$ -cell. In fact, they pin down a unique full proposition, because every  $H$ -cell contains some worlds where (2)∣(10) has a truth-value.

In general, when the incremental content  $p \upharpoonright s$  is wholly relevant to the QUD  $Q$  (that is, no  $Q$ -cell intersects both  $p \cap s$  and  $\neg p \cap s$ ), we can consider its *completion* by that QUD.

- (15) Suppose a partial proposition  $\langle t, f \rangle$  is wholly relevant to a question  $Q$ . Then the *completion* of  $\langle t, f \rangle$  by  $Q$ , written  $Q(\langle t, f \rangle)$ , is defined as follows:

$$Q(\langle t, f \rangle) =_{\text{df}} \{ \{w : \text{for some } v \in t, w \sim_Q v\}, \{w : \text{for some } v \in f, w \sim_Q v\} \}$$

And provided  $s$  intersects every  $Q$ -cell,  $Q(p \upharpoonright s)$  is guaranteed to be a full proposition.

With this formalism in hand, we can restate the present account of loose talk in an attractively succinct way: when a speaker makes an utterance with semantic content  $p$  in a context where  $s$  is the scale presupposition and  $Q$  is the QUD, a loose reading  $Q(p \upharpoonright s)$  of the speaker’s utterance is available, provided  $Q(p \upharpoonright s)$  is well-defined.

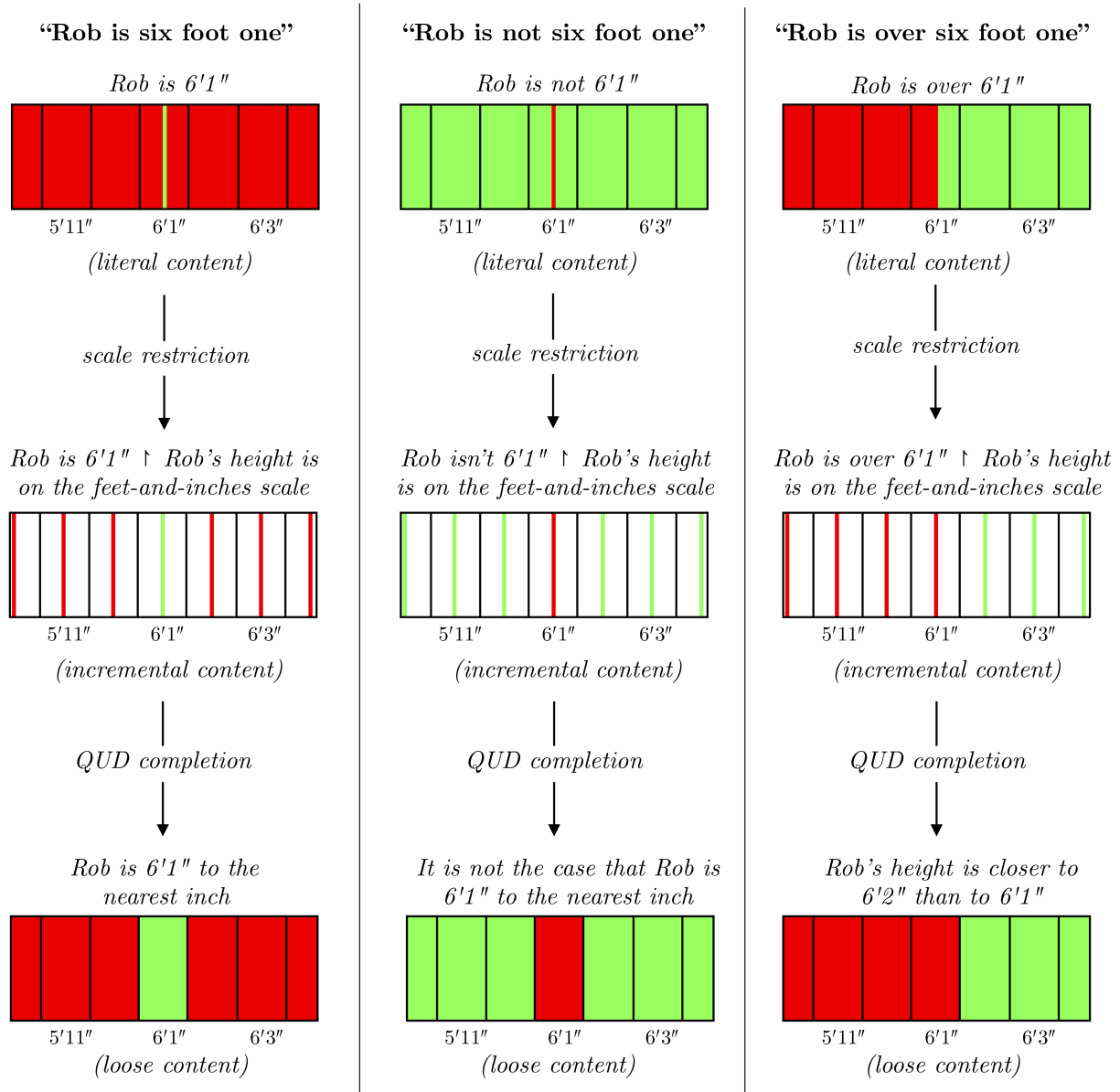
### 3.3 Boolean Transparency

Fixing a particular QUD  $Q$  and scale presupposition  $s$ , we can define a loosening operator “ $\circ$ ” as follows:  $p \mapsto Q(p \upharpoonright s)$ . This partial map takes suitable literal contents  $p$  to the corresponding loose content  $\circ p$ . This notation lets us compactly express two formal results that jointly capture all the patterns that emerge from the observations in §2:

- (16) **Boolean Transparency:**<sup>2</sup> Provided  $\circ p$  and  $\circ p'$  are both well-defined,  $\circ \neg p = \neg \circ p$ ,  $\circ (p \wedge p') = (\circ p \wedge \circ p')$ , and  $\circ (p \vee p') = (\circ p \vee \circ p')$

<sup>2</sup> *Proof:* Recall that  $\circ$  is the map  $Q(\cdot \upharpoonright s)$ . Now let  $r = \circ p$  and  $r' = \circ p'$ . First note that since  $r$  and  $r'$  are wholly relevant to  $Q$ , it is easy to see that  $Q(r \upharpoonright s) = r$  and  $Q(r' \upharpoonright s) = r'$ . Hence we have  $Q(r \upharpoonright s) = Q(p \upharpoonright s)$  whence  $r \upharpoonright s = p \upharpoonright s$ . Likewise,  $r' \upharpoonright s = p' \upharpoonright s$ .

Now it is easy to show that  $\circ \neg p = \neg \circ p$ . Since  $\neg r$  is wholly about  $Q$ ,  $Q(\neg r \upharpoonright s) = \neg r$ . And it follows from the fact that  $r \upharpoonright s = p \upharpoonright s$  that  $\neg r \upharpoonright s = \neg p \upharpoonright s$ . So completing both sides by  $Q$ , we have that  $\neg r = Q(\neg p \upharpoonright s) = \circ \neg p$ .



This diagram illustrates how, for three example sentences, the present account derives their loose reading from their literal contents. First we restrict the literal content to worlds in which the operative scale presupposition holds, in this case *Rob's height is on the feet-and-inches scale*. Then we complete that partial proposition in the only way possible to stay within the lines of the partition set by *H*, the QUD (*What is Rob's height to the nearest inch*).

Each of the nine subdiagrams represents the (partial) proposition in the caption above it. In each case, the large rectangle stands for the space of all possible worlds, and the smaller rectangles for cells of the QUD. The worlds are arranged from left to right according to Rob's height, as indicated by the ruler marks. **Green** marks the regions where the represented proposition is true. **Red** marks the regions where the represented proposition is false.

This result (16) captures the transparency observed in examples (3-5): the loose reading of a negation is the negation of the loose reading of the original claim, the loose reading of a conjunction is the conjunction of the loose readings of the conjuncts, etc. Now (16) can in turn be viewed as a consequence of (17), which captures a still wider pattern:

- (17) **Entailment Preservation:**<sup>3</sup> Provided  $\circlearrowleft p_i$  and  $\circlearrowleft c$  are well-defined, if  $p_1, p_2, \dots \models c$ , then also  $\circlearrowleft p_1, \circlearrowleft p_2, \dots \models \circlearrowleft c$

A direct consequence of (17) is the preservation of inconsistency under loosening observed in (1). In addition, (17) also explains the strengthening of strict comparatives. To see this, note that the literal content of “Rob is over six foot one” is inconsistent with (2). In order to preserve that inconsistency when both claims are read loosely, the comparative needs to have a strong loose reading to make up for the weak loose reading of (2). Or viewing it differently, “Rob is over six foot one” *entails* (3). To preserve that entailment after loosening, the content of the comparative must strengthen in order to entail the strengthened content of (3).

## 4 Further Observations

### 4.1 Round Numbers

As Krifka [7] emphasised, rounder numbers receive looser readings. The present treatment has a neat explanation for this phenomenon. Consider for instance (18a) and (18b):

- (18) a. This parrot is 22 inches tall.  
b. This parrot is 55.88 cm tall.

The semantic contents of (18a) and (18b) are truth-conditionally equivalent (by definition, an inch is equal to 2.54 cm). Nevertheless there is a clear contrast between them: (18a) gets a far looser reading than (18b). The reason is that while (18b) employs the extremely fine 100<sup>th</sup>-of-a-centimeter scale, the sentence (18a) uses the coarse inches scale, which is associated with a much stronger scale presupposition, and a more coarse-grained QUD; consequently, (18a) is correctly predicted to loosen more than (18b).

### 4.2 Scale Ambiguity and Slack Regulators

Another characteristic of round numbers is that they occur on multiple scales. For instance, the expression “700 miles” occurs on the 100-mile-scale, the 50-mile-scale, the 10-mile-scale, the 5-mile-scale and the 1-mile scale. This can lead to ambiguity:

- (19) Addis Ababa is 700 miles away from Mogadishu.

Depending on which scale one takes it to employ, the present account predicts looser or stricter readings for (19). And that ambiguity is indeed attested: is easy to construct stricter and looser

---

Likewise, since  $r \wedge r'$  is wholly about  $Q$ ,  $Q(r \wedge r' \uparrow s) = r \wedge r'$ . And we get from  $r \uparrow s = p \uparrow s$  and  $r' \uparrow s = p' \uparrow s$  that  $(r \wedge r') \uparrow s = (p \wedge p') \uparrow s$ . Completing both sides by  $Q$ , we have  $(r \wedge r') = Q((p \wedge p') \uparrow s) = \circlearrowleft (p \wedge p')$ .

Finally, using the previous two results,  $\circlearrowleft p \vee \circlearrowleft p' = \neg(\neg \circlearrowleft p \wedge \neg \circlearrowleft p') = \circlearrowleft \neg(\neg p \wedge \neg p') = \circlearrowleft (p \vee p')$ . ■

<sup>3</sup> *Proof:* Let  $r_i = \circlearrowleft p_i$  and  $d = \circlearrowleft c$ . Suppose  $p_1, p_2, \dots \models c$ , which is to say that  $\bigcap_i p_i \subseteq c$ . Now let  $w$  be any world where all the loosened premises  $\circlearrowleft p_i$  are true. We need to prove that  $d$  is also true at  $w$ . Since, for instance,  $\circlearrowleft p_1$  is true at  $w$ , there must be some world  $v \in s$  such that  $w \sim_Q v$ . Now all the  $\circlearrowleft p_i$  must have the same truth value at  $w$  and  $v$ ; so they are all true at  $v$  as well. Hence  $v \in \bigcap_i (p_i \cap s) \subseteq c \cap s$ . Consequently,  $d$  is true at  $w$ . And since  $d$  must also take the same truth-value at  $v$  and  $w$ ,  $d$  is true at  $w$ . ■

contexts in which the various readings are witnessed. If the context is ambiguous, the looser interpretations attached to coarser scales tend to be strongly preferred [1][8].

*Slack regulators* are operators like “exactly,” “precisely,” “roughly,” and “about,” which control the looseness with which claims are interpreted. On the present account, the natural treatment of slack regulators is to take them to be scale disambiguators: in case of ambiguity, “exactly” and “precisely” can be used to indicate the use of a fine scale, while “roughly” and “about” indicate the use of a coarse scale. One advantage of this treatment is that it explains why sentences like (20) are infelicitous:

(20) # There were about 27 476 people at the rally.

In (20), the word “about” purports to indicate the use of a coarse scale. But the number to which it attaches, “27 476,” only occurs on a maximally fine-grained integer scale.

### 4.3 Alternative Rounding Conventions

So far, I have assumed that loose talkers round off to the closest item on the scale, but this is not always the case. For instance, in a context where it is important that we do not underestimate Rob’s height, we might round upwards instead. In such a context, the question we are seeking to settle is not *What is Rob’s height to the nearest inch* but rather *What is the least height in inches that exceeds Rob’s*. That is to say, instead of  $H$ , the QUD is the partition  $H^*$ :

$$H^* = \{ \{ w : \text{at } w, \text{ Rob’s height in inches is in the interval } (n, n + 1] \} : n \in \mathbb{N} \}$$

The scale presupposition, meanwhile, is still the same, so that the loose reading of (2) in such special contexts will be  $H^*((2) \uparrow (10))$ , *Rob’s height is between 6’ and 6’1”*. Downward and intermediate rounding conventions can be handled analogously.

A more complex example is *age*. Typically, the convention is round off someone’s age downward to the nearest year. Except beginning at midnight on the day of their birthday, when we round upwards. This can be captured with a QUD of the following form:

$$A = \left\{ \left\{ w : \begin{array}{l} \text{at } w, \text{ Abby is fewer than } n + 1 \text{ years old, and her} \\ \text{age will exceed } n \text{ years some time today.} \end{array} \right\} : n \in \mathbb{N} \right\}$$

(Just because the present account *can* give a pragmatic account of the rounding convention for age does not mean that this *should* be captured pragmatically – arguably it is preferable in this case to just build it into the semantics of “age”.)

### 4.4 Looseness vs. Vagueness

It is sometimes assumed that looseness is always accompanied by vagueness [10][11]. I think that this is a mistake: looseness and vagueness are entirely separate phenomena, and loose readings often have perfectly sharp truth conditions. Consider for instance this claim:

(21) The molar mass of water is 18.015 grams.

In a scientific context, (21) is understood to convey a very precise and determinate piece of information, namely that *the molar mass of  $H_2O$  is at least 18.0145 and less than 18.0155 grams*. But this is still loose talk: a completely strict reading would be *the molar mass of  $H_2O$  is (exactly) 18.015000... grams*.

But I also think ordinary sentences like (2) have fairly determinate loose truth conditions, even outside of scientific contexts. That determinacy can be brought out by investigating what it takes to *contradict* the loose content of (2):



- (22) A: Rob is six foot one.  
 B<sub>1</sub>: No, he is closer to six foot two.  
 # B<sub>2</sub>: No, he is closer to six foot one and three quarters.

Here, B<sub>2</sub>'s response betrays precisely the kind of infelicity that is characteristic of rejections that contradict the literal but not the loose content of the original assertion. But B<sub>1</sub>'s response is perfectly felicitous. In general, for the response “No, he is closer to  $X$ ” to be felicitous,  $X$  must either be at least 6'2" or at most 6'. The reason for this, I suggest, is that only those responses contradict the loose content of A's assertion. It follows that loose content must really be *Rob is between 6'  $\frac{1}{2}$ " and 6'  $1\frac{1}{2}$ "*, just as the present account predicts.

## 4.5 Beyond Numbers

Some conventional scales are not quantitative. Colour scales for instance:

- (23) This carpet is burgundy.

In a well-stocked carpet shop, (23) may convey an extremely specific message about the carpet's colour. But describing the contents of my living room, (23) is naturally read more loosely. In ordinary conversation, we employ a relatively coarse colour scale, with a limited range of canonical, well-known colour names. In such a context, (23) gets the loose content *This carpet is closer to burgundy than to any other shade on the coarse scale*. But in carpet-shops, a more specialised, fine-grained colour vocabulary is employed, yielding a stricter reading for (23).

Another interesting application is to absolute adjectives:

- (24) The road is flat.

Arguably, the literal content of (24) is that *the road has no bumps whatsoever*. But (24) conveys something weaker. The smoothness of the road is described using a scale {“flat”, “bumpy”}, where “bumpy” means *containing a lot of bumps*. On these assumptions, the predicted loose content of (24) is *the road is closer to being flat than to being bumpy*.

## 5 Loose Talk as Conversational Exculpation

In paradigmatic cases of conversational implicature, a speaker is implicated in an additional commitment going beyond the literal content of their assertion. It is natural to wonder whether the reverse of this could also happen. Perhaps speakers are sometimes let off the hook for some of their literal commitments, for pragmatic reasons. One could call that conversational *exculpation* (to exculpate literally means ‘to release from blame’) [4].

The pragmatic mechanism described in §3 above yields a way to implement this idea. Speakers make a claim  $p$  that is based on a scale presupposition  $s$ , without thereby incurring any serious commitment to  $s$ . We have so far assumed that the ‘subtracted’ propositions are always a scale presupposition, but there is no reason other kinds of presuppositions could not be subtracted through the very same pragmatic loosening mechanism.

In fact, there is eminently good reason to think that this *is* possible. For scales and scale presuppositions sometimes appear “in the wild,” so to speak, in the absence of any conventional scale, numerical or otherwise. And that suggests that there is no sharp distinction between scale presuppositions and other kinds of contextual presupposition.

Here is an example to illustrate this (from Philippe Schlenker, p.c.). Imagine a wall with a vertical line of red pegs at different heights: a physically realised scale. The pegs need



not be evenly spaced. Some pegs have names attached to them, belonging to people who are approximately as tall as the peg indicates. I have been given a note labelled “Rob” and tasked to affix it to the appropriate peg. To help me out, you point to one of the pegs and say:

(25) This is the peg at Rob’s height.

The definite description in (25) induces the presupposition that *one of the pegs is at Rob’s height*. But I would not infer from this that you really believed this, any more than I would conclude from (2) that the speaker really thinks Rob is an exact number of inches tall. Instead (25) gets a loose reading that *this peg is closer to Rob’s height than the others*.

The phenomena here are precisely analogous to the numerical cases. The amount of loosening for (25) is dependent on the fineness of the scale: the sparser the pegs, the looser the talk. (25) also shows the same embedding behaviour as (2). In particular, (25)’s negation is has a strong loose reading, namely *some other peg is closer to Rob’s height than this one*.

This strongly suggests that presuppositions which are not associated with a conventional scale can be conversationally ‘exculpated’ using this same mechanism. And if that is granted, we have a very general pragmatic mechanism here, with the potential to explain a wide range of linguistic phenomena. To illustrate, I’ll briefly float two applications here (see also [4]).

First, there is application to ‘event-related’ readings: “Four thousand ships passed through the lock” has a reading *It happened 4000 times that some ship passed through the lock* [6]. This can be explained as the result of exculpating the contextual presupposition *No ship passed through the lock more than once*, with the QUD *How many ship-passings occurred?*

Second, rigidification of definite descriptions. “The man with the martini is a spy” can take a reading like *That man is a spy*, which may be true even if the man in question is holding a different drink [3]. This can be accounted for as the result of exculpating the presupposition *That man is the only guy in the room with a martini*, with the QUD *Who here is a spy?*

## References

- [1] Harald Bastiaanse. The rationality of round interpretation. In *Vagueness in Communication*. Springer, 2009.
- [2] Sam Carter. The dynamics of loose talk. *Noûs*, forthcoming.
- [3] Keith Donnellan. Reference and definite descriptions. *The Philosophical Review*, 75(3), 1966.
- [4] Daniel Hoek. Conversational exculpation. *The Philosophical Review*, 127(2), 2018.
- [5] Justine Kao, Jean Wu, Leon Bergen, and Nelson Goodman. Nonliteral understanding of number words. *PNAS*, 111(33), 2014.
- [6] Manfred Krifka. Four thousand ships passed through the lock. *Linguistics and Philosophy*, 1990.
- [7] Manfred Krifka. Be brief and vague! In *Sounds and Systems*. De Gruyter, 2002.
- [8] Manfred Krifka. Approximate interpretations of number words. In *Theory and Evidence in Semantics*. CSLI, 2009.
- [9] Manuel Križ. *Aspects of Homogeneity*. PhD thesis, University of Vienna, 2015.
- [10] Peter Lasnik. Pragmatic halos. *Language*, 75(3):522–551, 1999.
- [11] Uli Sauerland and Penka Stateva. Two types of vagueness. In *Vagueness and Language Use*. Palgrave MacMillan, 2011.
- [12] Mandy Simons. Presuppositions and relevance. In *Semantics vs. Pragmatics*. OUP, 2002.
- [13] Stephanie Solt. An alternative theory of imprecision. In *SALT 24*, 2014.
- [14] Dan Sperber and Deirdre Wilson. Loose talk. *Proc. of the Aristotelian Society*, 86, 1986.
- [15] Stephen Yablo. *Aboutness*, chapter 5, section 2: Loose Talk. Princeton University Press, 2014.

# The Anaphoric Potential of Indefinites under Negation and Disjunction\*

Lisa Hofmann

University of California Santa Cruz, Santa Cruz, California, U.S.A  
lihofman@ucsc.edu

## Abstract

This paper extends the analysis of modal subordination presented in Stone (1999); Stone and Hardt (1999) to other cases of non-veridically introduced discourse referents (drefs) with the goal of understanding the circumstances under which a dref is available for subsequent anaphoric reference. The central claim is that a dref introduced under negation can be the antecedent for a pronoun, only if the interpretation does not require that the assertion of its non-existence and the existence presupposition of the pronoun be true wrt the same set of worlds. The analysis is based on relativizing individual drefs to the worlds in which their referent exists (Stone (1999); Stone and Hardt (1999)), and sentential operators introducing propositional drefs that provide a local context for the interpretation of their prejacent (Karttunen (1973); Heim (1983)), and it is framed in intensional CDRT (Muskens (1996); Brasoveanu (2007, 2010)). This allows for a greater empirical coverage than Krahmer and Muskens’s (1995) account of drefs under negation and disjunction.

## 1 Introduction

Indefinite DPs under anti-veridical operators such as negation do not usually introduce a discourse referent (dref) that is available for subsequent reference:

- (1) a. *There is [no bathroom]<sup>v</sup> in this house.*  
b. *#It<sub>v</sub> is in a weird place.*

This generalization goes back to Karttunen (1969), who also points out the existence of counterexamples: Indefinites under negation can provide an antecedent for a pronoun when the antecedent is embedded under double negation (Karttunen (1969); Krahmer and Muskens (1995)), when the antecedent is in the first disjunct of a disjunction and the anaphor in the second one (Krahmer and Muskens (1995)), in modal subordination (Roberts (1989)), or when the utterance containing the anaphor rejects the utterance containing the antecedent.

- (2) a. Double negation:  
*It’s not true that there is [no bathroom]<sup>v1</sup> in this house. It<sub>v1</sub> is just in a weird place.*  
b. Disjunction:  
*Either there is [no bathroom]<sup>v2</sup> in this house, or it<sub>v2</sub> is in a weird place.*  
c. Modal subordination:  
*There is [no bathroom]<sup>v3</sup> in this house. It<sub>v3</sub> would be easier to find.*  
d. Disagreement:  
A: *There’s [no bathroom]<sup>v4</sup> in this house.*  
B: *(What are you talking about?) It<sub>v4</sub> is just in a weird place.*

---

\*This paper has benefited greatly from the comments of anonymous reviewers for the 2019 Amsterdam Colloquium and Simon Charlow, as well as from discussions with Adrian Brasoveanu, Jess Law, Morwenna Hooks, Jack Duff, and the audiences at a practice talk at UC Santa Cruz and at CUSP 2019.

In the examples in (2), *no bathroom* is the antecedent of *it*, although the former is in the scope of negation, and the latter is not. Krahmer and Muskens (1995) note that standard Discourse Representation Theory (DRT, Kamp (1981); Kamp and Reyle (1993)), and classic versions of File-Change Semantics (Heim (1982, 1983)) and Dynamic Predicate Logic (Groenendijk and Stokhof (1991)), were designed to capture Karttunen’s basic generalization illustrated in (1), and don’t allow drefs to outlive embedding under negation.

This paper presents an account that captures the above (1) and (2), in intensional Compositional DRT (CRDT, Muskens (1996); closely following Brasoveanu’s (2010) implementation). The account is based on the assumption that a pronoun may be co-referential with a preceding DP, only if the dref introduced by the DP exists in all the possible worlds in the local context of evaluation of the pronoun. This condition on anaphoric accessibility is captured by relativizing individual drefs to sets of worlds in which they exist (Stone (1999); Stone and Hardt (1999)). The paper therefore extends analyses of modal subordination with *would* based on simultaneous reference to sets of worlds and individuals (Stone (1999); Stone and Hardt (1999); Brasoveanu (2007, 2010)) to other cases of anaphors to non-verbally introduced individual drefs.

The presupposition of the pronoun is evaluated in its local context. The relation between local and global context sets is constrained by the semantics of sentential operators (Karttunen (1973); Heim (1983)), which may introduce drefs for sets of worlds in which their prejacent is interpreted. The global context set is constrained pragmatically to be compatible with the speaker’s commitments (Stalnaker (1978, 2002); Gunlogson (2004)). This allows for an account of the influence of the linguistic context, in particular, the veridicality of embedding operators, on the availability of an anaphoric dependency.

## 2 The analysis informally

The account relies on two basic analytical intuitions: (i) that the use of a pronoun presupposes the existence of a referent<sup>1</sup>, and (ii) an indefinite under negation introduces a counterfactual dref, i.e., a dref, s.t. the speaker is committed to its referent not existing in the global context. That explains why a discourse like (1) (repeated here) is infelicitous.

- (1) a. *There is [no bathroom]<sup>v</sup> in this house.*  
 b. *#It<sub>v</sub> is in a weird place.*

In (1), the speaker counterfactually introduces a dref *v* for a bathroom and is committed to it not existing. If the same speaker subsequently uses a pronoun in a veridical context, the existence of a referent is presupposed. Therefore, the pronoun couldn’t refer to *v*. If it did, the speaker would first assert the non-existence of a bathroom and then presuppose its existence, rendering the discourse inconsistent. This explanation is based on Stone’s (1999) analysis of modal subordination. Indefinites in the scope of modals also don’t provide antecedents for pronouns in veridical contexts, but they provide antecedents for pronouns in the scope of *would*:

- (3) *[A wolf]<sup>v5</sup> might walk in.* (Roberts (1989): 11)  
 a. *#It<sub>v5</sub> is gray.*  
 b. *It<sub>v5</sub> would eat you first.*

Stone suggests that the indefinite under *might* in (3) introduces *v*<sup>5</sup> as a hypothetical dref, i.e., a dref s.t. the speaker is not committed to the existence of its referent in the global context.

<sup>1</sup>I also assume that singular pronouns like *it* presuppose that the referent is a single individual, which is not further discussed in this paper.

As a result, the use of the pronoun  $It_{v_5}$  in a veridical context (3-a) is impossible, because the existence presupposition is not satisfied. In contrast,  $It_{v_5}$  in (3-b) can refer to a hypothetical dref, because *would* is assumed to be anaphoric to a hypothetical proposition, i.e., a proposition the speaker is not committed to, which provides a context of interpretation for the prejacent. Stone suggests that *would* can be anaphoric to the proposition in the scope of *might*, s.t. a wolf walks in, and  $It_{v_5}$  can be evaluated wrt the hypothetical worlds in which that wolf exists.

This paper extends an analysis along these lines to two new empirical cases: One of them is modal subordination with negative antecedents, such as (2-c) (repeated here):

(2-c) *There is [no bathroom]<sup>v3</sup> in this house.  $It_{v_3}$  would be easier to find.*

The availability of an anaphoric relation can be explained similarly to the classic modal subordination cases. A counterfactual proposition, like the one introduced by the prejacent of negation, is also a hypothetical proposition and may be an antecedent for *would*. The pronoun can be evaluated in the counterfactual context, where the referent introduced by  $[no\ bathroom]^{v3}$  exists, and the existence presupposition is satisfied locally. The other empirical case to which the analysis is extended involves cases where the discourse segment containing the antecedent and the discourse segment containing the pronoun are not required to be consistent with each other. This includes disjunction and inter-speaker disagreement ((2-b) + (2-d)), repeated here):

(2-b) Disjunction:  
*Either there is [no bathroom]<sup>v2</sup> in this house, or  $it_{v_2}$  is in a weird place.*

(2-d) Disagreement:  
A: *There's [no bathroom]<sup>v4</sup> in this house.*  
B: *(What are you talking about?)  $It_{v_4}$  is just in a weird place.*

Here, the pronouns are used veridically, which is inconsistent with a counterfactual antecedent. The discourse segments containing the antecedent and the pronoun are inconsistent with each other, but the contexts of disjunction and inter-speaker disagreement do not place a constraint on the two discourse segments to be consistent, which enables the veridical use of a pronoun with a counterfactual antecedent. This sets the cases in (2-b) and (2-d) apart from (1), where the propositions have to be compatible because they are assertions of the same speaker.

A formalization needs to keep track of the introduced drefs, the sets of worlds in which they exist, the local context sets of indefinites and anaphoric expressions, and who is committed to these propositions as being true, false, or neither. The formal system introduced in the following section is designed to do that.

### 3 An account of anaphoric accessibility

#### 3.1 The account

The system uses four basic types:  $t$  (truth-values),  $e$  (entities),  $w$  (possible worlds), and  $s$  (variable assignments). Following Stone (1999); Stone and Hardt (1999), individual drefs are formalized as drefs for individual concepts. An individual dref  $v$  is a function of type  $s(we)$  from assignments  $i_s$  and worlds  $w_w$  to individuals  $x_e$  (subscripts on terms indicate their type). The individual  $v_{s(we)}(i_s)(w_w)$  is the individual that the assignment  $i$  assigns to  $v$  in  $w$ . A propositional dref  $\phi$  is a function of type  $s(wt)$  from assignments  $i_s$  to sets of worlds ( $wt$ ). The proposition  $\phi_{s(wt)}(i_s)$  is the set of worlds that  $i$  assigns to  $\phi$ .

Natural language sentences are interpreted as DRSs, which are defined as binary relations

of type  $s(st)$  between input state  $i_s$  and output state  $j_s$ , where discourse states are variable assignments. A DRS contains a list of new drefs  $(v_1, \dots, v_n)$ , and a series of conditions of type  $st$ , i.e., properties of discourse states  $(C_1, \dots, C_n)$ , and is defined as (4):

$$(4) \quad [v_1, \dots, v_n \mid C_1, \dots, C_n] := \lambda i_s. \lambda j_s. i[v_1, \dots, v_n]j \wedge C_1(j) \wedge \dots \wedge C_n(j)$$

### 3.1.1 Variable updates relative to sets of worlds

The dynamic variable update  $(i[v]j)$  specifies a binary relation over variable assignments and works as random assignment of values to a variable  $v$ . This relation between discourse states  $i_s$  and  $j_s$  holds if  $i$  and  $j$  differ at most wrt the values assigned to  $v$  (see e.g. Groenendijk and Stokhof (1991); Muskens (1996); Brasoveanu (2007)).

Following Stone (1999); Stone and Hardt (1999), this system formalizes individual drefs as drefs for individual concepts, i.e., functions of type  $(we)$ , mapping possible worlds to individual referents. To characterize worlds in which no such referent exists, a dummy element  $\#$  is included in the domain of individuals.  $\#_e$  corresponds to the ‘indeterminate’ value in a trivalent logics, s.t. a term receives an indeterminate truth-value  $\#_t$  if any of its arguments evaluate to  $\#_e$ . This emulates partial functions in a system that uses only total functions (van den Berg (1996); Brasoveanu (2007, 2010)), but still captures that the use of a pronoun results in unacceptability if used in a context where a referent cannot be determined. Using individual concepts allows for the introduction of individual drefs along with the information about the set of worlds in which their referent exists (and does not):

(5) The relation  $i[\phi : v]j$  holds (for assignments  $i_s, j_s$ , and the drefs  $\phi \in \mathbf{Term}_{s(wt)}$ ,  $v \in \mathbf{Term}_{s(we)}$ ) iff the conjunction of the following holds:

- $i[v]j$
- $\forall w_w. (\phi(j)(w) \rightarrow v'(i)(w) \neq \#)$   
(an individual referent of  $v$  is determined in each world in  $\phi(j)$ )
- $\forall w_w. (\neg \phi(j)(w) \rightarrow v(j)(w) = \#)$   
(in each world not in  $\phi$ ,  $v$  points to the indeterminate value)

(5) guarantees that  $j$  is an update of  $i$  that differs at most wrt the value assigned to  $v$ . The second conjunct requires that for each world  $w$  in  $\phi(j)$ ,  $v(j)(w)$  doesn’t map to  $\#$  (but an actual individual). The third conjunct states that for each world  $w$  not in  $\phi(j)$ ,  $v(j)(w)$  maps to  $\#$ , ensuring that the referent of  $v$  exists only in the worlds in  $\phi(j)$ .

### 3.1.2 Dref accessibility, DRS conditions, and the local context

Relativizing individual drefs to sets of worlds in which they refer gives rise to an accessibility condition based on an existence condition for pronouns and the notion of a local context (Karttunen (1973); Heim (1983)). The latter involves the idea that expressions are interpreted wrt an intensional context set that is pragmatically constrained on a global level by the speaker’s commitments (Stalnaker (1978, 2002); Gunlogson (2004)), and locally constrained by the semantics of sentential operators. Based on this, a dref  $v$  is accessible for reference by a pronoun  $v'$ , only if the referent of  $v$  is determined in the local context of  $v'$ , i.e., if  $v$  maps to an individual (other than  $\#$ ), for each world in the local context.

In this system, the local context is defined in reference to the evaluation of DRS conditions and consists of an assignment  $i_s$ , of which the condition is predicated, and a dref for a set of possible worlds  $\phi$ , a compositionally supplied intensional argument. (6) illustrates how a

dynamic predicate is defined as an abbreviation for a condition involving the corresponding static predicate and the referents of the argument drefs given  $i$  and  $\phi$ .

- (6) Predicates with their arguments as basic conditions (type  $st$ ):  
 $R_\phi\{v\} := \lambda i_s. \forall w_w. \phi(i)(w) \rightarrow R(v(i)(w))(w)$ ,  
 for  $R \in \mathbf{Term}_{e(wt)}$ ,  $v \in \mathbf{Term}_{s(we)}$ ,  $\phi \in \mathbf{Term}_{s(wt)}$

The evaluation of a DRS condition in the context  $\langle i_s, \phi_{s(wt)} \rangle$  is mapped to a truth-value other than  $\#$  only if the argument dref is mapped to an individual other than  $\#$  for each world in  $\phi(i)$ . The general requirement that drefs refer in each world in their local context follows from the definition of DRS conditions. This is illustrated for the DRS in (7):

(7)  $S$ : *There is no bathroom.*  $\rightsquigarrow^2$

$\phi_1, \phi_2, \phi_2 : v_1$
$\phi_{DC_s} \subseteq \phi_1$
$\phi_1 = \overline{\phi_2}$
$\mathbf{bathroom}_{\phi_2}\{v_1\}$

An update with (7) three drefs: A proposition  $\phi_1$  (for the worlds where there is no bathroom), restricting the set of discourse commitments held by  $S$ , a proposition  $\phi_2$  (for the worlds where there is a bathroom), the complement of  $\phi_1$ , and an individual  $v_1$ , (for a bathroom in  $\phi_2$ ). The resulting assignment  $i$  maps  $v$  to an individual relative to each world in  $\phi_2(i)$ , and maps it to  $\#$  for each world not in  $\phi_2(i)$ . The condition  $\mathbf{bathroom}_{\phi_2}\{v_1\}$  requires that the referent of  $v_1$  is a bathroom in all  $\phi_2$ -worlds. For the truth-conditional import to be determined, the referent of  $v_1$  needs to be determined in all  $\phi_2$ -worlds. That is the case in (7) since the condition is interpreted wrt a variable assignment that is the result of updating an input state with  $[\phi_2 : v_1]$ .

As a special case of this general requirement, the condition on anaphoric accessibility also follows from the definition of DRS-conditions:

- (8) Accessibility of antecedents:  
 Given a relation  $R_{e(wt)}$ , an individual dref  $v_{s(we)}$ , a propositional dref  $\phi_{s(wt)}$ , and a discourse state  $i_s$ :
- A DRS-condition predicated of a discourse state  $R_\phi\{v\}(i)$  is defined as  $\forall w_w. \phi(i)(w) \rightarrow R(v(i)(w))(w)$
  - $\forall w_w. \phi(i)(w) \rightarrow R(v(i)(w))(w)$  can be determined only if  $v(i)(w)$  is determined for all worlds  $w$  s.t.  $\phi(i)(w)$ , therefore  $\forall w_w. \phi(i)(w) \rightarrow v'(i)(w) \neq \#$
  - As a result, a pronoun in the context of  $i$  and  $\phi$  can make reference to the dref  $v$  only if the value of  $v(i)(w)$  is determined for all worlds  $w$ , s.t.  $\phi(i)(w)$

Stone's (1999) mechanism introduces individual drefs relative to sets of worlds in which they exist, and gives rise to state this accessibility condition, that captures that pronouns presuppose the existence of a referent and are unacceptable otherwise. This is couched in an intensional version of CDRT, where sentential operators may introduce drefs for sets of worlds to provide a context in which their prejacent is interpreted.

<sup>2</sup> This representation can be derived compositionally based on the contributions of the following components:

- |     |   |   |
|-----|---|---|
| (i) | <ol style="list-style-type: none"> <li>there is <math>\rightsquigarrow \lambda \mathcal{P}_{wt}. \mathcal{P}</math></li> <li>no <math>\rightsquigarrow \lambda \mathcal{P}_{e(wt)}. \lambda \mathcal{P}'_{e(wt)}. \lambda \phi_w. [\phi_i, \phi_i : v_j];</math><br/> <math>[\phi = \overline{\phi_i}; \mathcal{P}(v_j)(\phi_i); \mathcal{P}'(v_j)(\phi_i)]</math></li> </ol> | <ol style="list-style-type: none"> <li>bathroom <math>\rightsquigarrow</math><br/> <math>\lambda v_e. \lambda \phi_w. [\mathbf{bathroom}_\phi\{v\}]</math></li> <li>Assertion by <math>S</math>:<br/> <math>\lambda \mathcal{P}_{wt}. [\phi_1 \mid \phi_{DC_s} \subseteq \phi_1]; \mathcal{P}(\phi_1)</math></li> </ol> |
|-----|---|---|

The bold-face subscripts denote CDRT 'meta-types' (see e.g. Brasoveanu (2007)), where  $e$  abbreviates the type for individual drefs  $s(we)$ ,  $w$  abbreviates the type for propositional drefs  $s(wt)$  and  $t$  abbreviates the type for dynamic sentence meanings  $s(st)$ .

## 3.2 Analysis

### 3.2.1 The counterfactual bathroom

The rest of this section illustrates how the analysis accounts for the above data in (1) and (2). First, let us see how the veridical use of a pronoun with a counterfactual antecedent where both discourse segments are uttered by the same speaker is ruled out.

(7) S: *There is no bathroom.*

$$\rightsquigarrow \begin{array}{|c|} \hline \phi_1, \phi_2, \phi_2 : \nu_1 \\ \hline \phi_{DC_S} \subseteq \phi_1 \\ \phi_1 = \phi_2 \\ \text{bathroom}_{\phi_2}\{\nu_1\} \\ \hline \end{array}$$

(9) # S: *It is in a weird place.*

$$\rightsquigarrow^3 \begin{array}{|c|} \hline \phi_3, \phi_3 : \nu_2 \\ \hline \phi_{DC_S} \subseteq \phi_3 \\ \text{place}_{\phi_3}\{\nu_2\} \\ \text{weird}_{\phi_3}\{\nu_2\} \\ \text{in}_{\phi_3}\{\underline{\nu_3}, \nu_2\} \\ \hline \end{array}$$

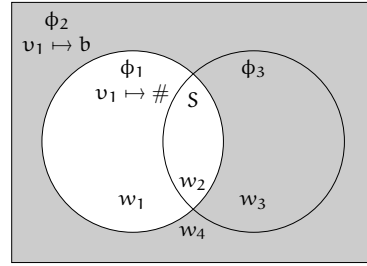
An update with (9) fails for the interpretation of  $\underline{\nu_3}$  (the underline indicates that it is anaphoric and requires an antecedent). Besides that, it specifies a dref  $\phi_3$  (the worlds where ‘it’ is in a weird place), and a dref  $\nu_2$  (mapping to a place in the  $\phi_3$ -worlds and to # in the non- $\phi_3$ -worlds). For illustration of the relationships between these drefs, consider a toy model  $M_1$  where  $D_w = \{w_1, w_2, w_3, w_4\}$ ,  $D_e = \{b, p, \#\}$ ,  $b$  is a bathroom and  $p$  is a place. The attempt of updating a discourse-initial discourse state with (7) and (9) wrt  $M_1$  results in an assignment  $i_1$  that assigns static referents to drefs as in (10). The relationship between these sets of worlds and the dref  $\nu_1$  storing the counterfactual bathroom are illustrated in (11).

(10) a.  $\phi_1(i_1) = \{w_1, w_2\}$   
 b.  $\phi_2(i_1) = \{w_3, w_4\}$   
 c.  $\phi_3(i_1) = \{w_2, w_3\}$

d.  $\nu_1(i_1) = \begin{cases} w_1 \mapsto \# \\ w_2 \mapsto \# \\ w_3 \mapsto b \\ w_4 \mapsto b \end{cases}$

e.  $\nu_2(i_1) = \begin{cases} w_1 \mapsto \# \\ w_2 \mapsto p \\ w_3 \mapsto p \\ w_4 \mapsto \# \end{cases}$

(11)



Pronominal reference fails because the potential antecedent  $\nu_1$  does not exist all the worlds in the local context of the anaphor  $\underline{\nu_3}$ , which is interpreted in the condition  $\text{in}_{\phi_3}\{\underline{\nu_3}, \nu_2\}$ . That guarantees that  $\underline{\nu_3}$  exists in all  $\phi_3$ -worlds. Therefore  $\nu_1$ , the dref introduced by *no bathroom*, can only be a possible antecedent for  $\underline{\nu_3}$ , if the referent of  $\nu_1$  exists in all the  $\phi_3$ -worlds.  $\nu_1$  is introduced as  $\phi_2 : \nu_1$  and its referent exists in all and only the  $\phi_2$ -worlds. As a result, the  $\phi_3$ -worlds have to be a subset of the  $\phi_2$ -worlds for this anaphoric relation to be available. Because all the worlds in  $\phi_{DC_S}$  are also ( $\phi_{DC_S} \subseteq (\phi_1 \cap \phi_3)$ ),  $\phi_1$  and  $\phi_3$  are interpreted in conjunction, they have to be compatible. Therefore, there are  $\phi_1$  worlds in  $\phi_3$ , where  $\nu_1$  does not exist ( $w_2$  in the above example). Resolving  $\underline{\nu_3}$  to  $\nu_1$  would result in undefinedness in the interpretation of the condition, so  $\nu_1$  is not an accessible antecedent for  $\underline{\nu_3}$ .

(ii)<sup>3</sup> a.  $\text{it} \rightsquigarrow \lambda P_{e(\text{wt})}.\lambda \phi_w.P(\nu)(\phi)$   
 b.  $\text{is} \rightsquigarrow \lambda P_{\text{wt}}.P$   
 c.  $\text{in} \rightsquigarrow \lambda Q_{((e(\text{wt}))(\text{wt}))}.\lambda \nu_e.\lambda \phi_w.Q(\lambda \nu'_e.\lambda \phi'_w.[\text{in}_{\phi}\{\nu, \nu'\}])(\phi)$

d.  $\text{a} \rightsquigarrow \lambda P_{e(\text{wt})}.\lambda P'_{e(\text{wt})}.\lambda \phi_w.[\phi_i, \phi_i : \nu_j];$   
 $[\phi = \overline{\phi_i}; P(\nu_j)(\phi_i); P'(\nu_j)(\phi_i)]$   
 e.  $\text{weird} \rightsquigarrow \lambda \nu_e.\lambda \phi_w.[\text{weird}_{\phi}\{\nu\}]$   
 f.  $\text{place} \rightsquigarrow \lambda \nu_e.\lambda \phi_w.[\text{place}_{\phi}\{\nu\}]$



### 3.2.2 The optional bathroom

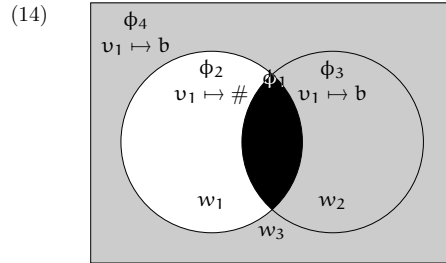
In contrast to propositions expressed by two assertions of the same speaker, the propositions expressed by the disjuncts of a disjunction are not required to be compatible with each other, which allows for reference to a counterfactual dref:

(12) S: *There is no bathroom or it is in a weird place.*  $\rightsquigarrow^4$

$\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_4 : v_1, \Phi_3 : v_2$ $\Phi_{DC_s} \subseteq \Phi_1$ $\Phi_1 = \Phi_2 \cup \Phi_3$ $\Phi_2 = \overline{\Phi_4}$ $\text{bathroom}_{\Phi_4}\{v_1\}$ $\text{place}_{\Phi_3}\{v_2\}$ $\text{weird}_{\Phi_3}\{v_2\}$ $\text{in}_{\Phi_3}\{v_3, v_2\}$
--

Consider a model  $M_2$ , where  $D_w = \{w_1, w_2, w_3\}$ ,  $D_e = \{b, p, \#\}$ ,  $b$  is a bathroom and  $p$  is a place. Updating an input state with (12) in  $M_2$  results in an output assignment  $i_2$  that assigns referents as in (13), depicted in (14).

- (13) a.  $\Phi_1(i_2) = \{w_1, w_2\}$   
 b.  $\Phi_2(i_2) = \{w_1\}$   
 c.  $\Phi_3(i_2) = \{w_2\}$   
 d.  $\Phi_4(i_2) = \{w_3\}$   
 e.  $v_1(i_2) = \left\{ \begin{array}{l} w_1 \mapsto \# \\ w_2 \mapsto b \\ w_3 \mapsto b \end{array} \right\}$   
 f.  $v_2(i_2) = \left\{ \begin{array}{l} w_1 \mapsto \# \\ w_2 \mapsto p \\ w_3 \mapsto \# \end{array} \right\}$



The anaphor  $v_3$  is interpreted in the condition  $\text{in}_{\Phi_3}\{v_3, v_2\}$ , which ensures that  $v_3$  exists in  $\Phi_3$ . For  $v_1$ , the dref introduced by *no bathroom*, to be a possible antecedent for  $v_3$ ,  $v_1$  needs to exist in all the  $\Phi_3$ -worlds.  $v_1$  is introduced as  $\Phi_4 : v_1$  and exists in all and only the  $\Phi_4$ -worlds,  $\Phi_3$  has to be a subset of  $\Phi_4$ .  $v_1$  exists in none of the worlds in  $\Phi_2$ , the complement of  $\Phi_4$ . Since  $\Phi_2$  and  $\Phi_3$  are not interpreted in conjunction, updating the context with (12) is compatible with an output state  $i_s$ , s.t.  $v_1$  exists in all  $\Phi_3$ -worlds, i.e. the ones where  $\Phi_2(i) \cap \Phi_3(i) = \emptyset$ , and  $v_3$  can be resolved as  $v_1$  (e.g. the above  $i_2$ ). Because disjunction places no requirement on its arguments to be compatible, a dref that is introduced counterfactually in the first disjunct is a possible antecedent for an anaphor in the second disjunct.

### 3.2.3 The hypothetical bathroom

Following Stone (1999), the modal subordination case is analyzed by assuming that the local context of the pronoun is an anaphorically provided hypothetical proposition, and the pronoun can therefore refer to a hypothetical individual.

(7) S: *There is no bathroom.*

(15) S: *It would be (right) there.*

$\rightsquigarrow$

$\Phi_1, \Phi_2, \Phi_2 : v_1$ $\Phi_{DC_s} \subseteq \Phi_1$ $\Phi_1 = \overline{\Phi_2}$ $\text{bathroom}_{\Phi_2}\{v_1\}$
---

$\rightsquigarrow^5$

$\Phi_3$ $\Phi_{DC_s} \subseteq \Phi_3$ $\text{would}_{\Phi_3}\{\Phi_4\}$ $\text{there}_{\Phi_4}\{v_2\}$
---

(iii)<sup>4</sup> or  $\rightsquigarrow \lambda P_{wt}. \lambda P'_{wt}. \lambda \Phi_w. [\Phi = \Phi' \cap \Phi'']; P(\Phi'); P'(\Phi'')$

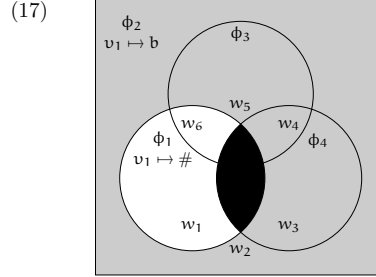
(iv)<sup>5</sup> a.  $\text{would} \rightsquigarrow \lambda P_{wt}. \lambda \Phi_w. [\text{would}_{\Phi} \{\Phi'\}]; P(\Phi')$

b.  $\text{there} \rightsquigarrow \lambda v_e. \lambda \Phi_w. [\text{there}_{\Phi} \{v\}]$

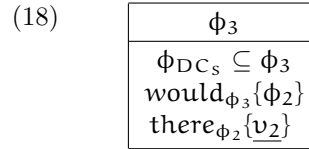


Consider a model  $M_3$ , where  $D_w = \{w_1, w_2, w_3, w_4, w_5, w_6\}$ ,  $D_e = \{b, p, \#\}$ ,  $b$  is a bathroom and  $p$  is a place. Updating an input state with (7) and then (15) in  $M_3$  results in an output assignment  $i_3$  that assigns static referents to drefs as in (16), depicted in (17).

- (16) a.  $\phi_1(i_3) = \{w_1, w_6\}$   
 b.  $\phi_2(i_3) = \{w_2, w_3, w_4, w_5, w_6\}$   
 c.  $\phi_3(i_3) = \{w_4, w_5, w_6\}$   
 d.  $\phi_4(i_3) = \{w_3, w_4\}$   
 e.  $v_1(i_3) = \left\{ \begin{array}{l} w_1 \mapsto \# \\ w_2 \mapsto b \\ w_3 \mapsto b \\ w_4 \mapsto b \\ w_5 \mapsto b \\ w_6 \mapsto \# \end{array} \right\}$



The interpretation of  $\text{would}_{\phi_3}\{\phi_4\}$  requires that  $\phi_4$  is a hypothetical proposition wrt  $\phi_3$ <sup>6</sup>  $\phi_2$ , the dref introduced under negation in (7), is a hypothetical proposition, and  $\phi_4$  can be resolved as  $\phi_2$ , as in (18).

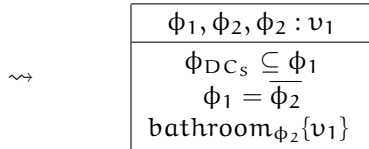


The pronoun  $v_2$  is interpreted in the condition  $\text{there}_{\phi_2}\{v_2\}$ . For  $v_1$ , the dref introduced by *no bathroom*, to be a possible antecedent for  $v_2$ , the referent of  $v_1$  needs to exist in  $\phi_2$ .  $v_1$  is introduced as  $\phi_2 : v_1$  and exists in all and only the  $\phi_2$  worlds, and  $v_2$  can be resolved as  $v_1$ .

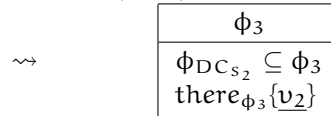
### 3.2.4 The contested bathroom

The case of inter-speaker disagreement is similar to the case of disjunction because the discourse segment containing the antecedent and the one containing the anaphor are not required to be compatible with each other.

(7) S: *There is no bathroom.*

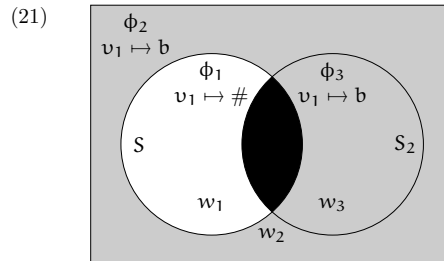


(19)  $S_2$ : *It is (right) there.*



Consider a model  $M_4$ , where  $D_w = \{w_1, w_2, w_3\}$ ,  $D_e = \{b, p, \#\}$ ,  $b$  is a bathroom and  $p$  is a place. Updating an input state with (7) and then (19) in  $M_4$  results in an output assignment  $i_4$  that assigns static referents to drefs as in (20), and illustrated in (21).

- (20) a.  $\phi_1(i_4) = \{w_1\}$   
 b.  $\phi_2(i_4) = \{w_2, w_3\}$   
 c.  $\phi_3(i_4) = \{w_3\}$   
 d.  $v_1(i_4) = \left\{ \begin{array}{l} w_1 \mapsto \# \\ w_2 \mapsto b \\ w_3 \mapsto b \end{array} \right\}$



<sup>6</sup>This is glossing over the specifics of modal semantics of *would* specifying the relationship between  $\phi_3$  and  $\phi_4$ . See Stone (1999); Stone and Hardt (1999) for discussion.

The anaphor  $\underline{v}_2$  is interpreted in the condition  $\text{there}_{\phi_3}\{\underline{v}_2\}$ .  $v_1$  can only be a possible antecedent for  $\underline{v}_2$ , if the referent of  $v_1$  exists in all  $\phi_3$ -worlds.  $v_1$  is introduced as  $\phi_2 : v_1$  and exists in all and only the  $\phi_2$  worlds, so  $\phi_3$  has to be a subset of  $\phi_2$ .  $\phi_1$  and  $\phi_3$  are asserted by different speakers, and therefore they need not be interpreted in conjunction. Updating the context with (7) and then with (19) is therefore compatible an output discourse state  $i_4$ , s.t.  $v_1$  exists in  $\phi_3$ , i.e. the one where  $\phi_1(i) \cap \phi_3(i) = \emptyset$ , and  $\underline{v}_2$  can be resolved to  $v_1$ . Because a successful resolution of the anaphoric dependency requires that  $\phi_1(i) \cap \phi_3(i) = \emptyset$ , we get an interpretation where B's utterance is incompatible with A's utterance.

## 4 Discussion

The presented analysis results in a dynamic semantics where all non-veridical operators are externally dynamic. Indefinites in their scope update the variable assignment globally, along with the information about the sets of worlds in which their referents exist. The analysis provides an understanding of when the surrounding context allows for an anaphoric relation between expressions introducing anaphora and potential antecedents. It constitutes a step forward from previous approaches to anaphoric accessibility in classical DRT (Kamp and Reyle (1993)), as well as analyses of modal subordination (Stone (1999)) and the double negation and disjunction cases (Krahmer and Muskens (1995)), by extending the empirical coverage.

Krahmer and Muskens's (1995) account for the disjunction and double negation cases within the framework of Double Negation DRT involves a semantics for negation that switches from the extension of an expression to its anti-extension, and a semantics for disjunction that analogizes it to conditionals, both in terms of their truth conditions and their dynamic potential. Accordingly, the sentences in (22), respectively, are taken to be equivalent to each other both dynamically and truth-conditionally.

- (22) a. Either there is no bathroom in this house, or it is in a weird place.  
 b. If it is not true that there is no bathroom in this house, it is in a weird place.  
 c. If there is a bathroom in this house, it is in a weird place.

Their analysis relies on the conventional semantics associated with negation and disjunction and therefore does not extend to cases without overt disjunction or double negation, like the modal subordination and inter-speaker disagreement cases. Further, it completely analogizes the semantics of disjunction to that of conditionals, which predicts that all anaphora pattern analogously in these two contexts. However, propositional anaphora exhibit a contrast between disjunctive and conditional contexts:

- (23) a. If Mary is sick, she knows that.  
 b. #Either Mary is not sick, or she knows that.<sup>7</sup>  
 c. Either Mary is not sick, or she is and knows that.

This asymmetry provides further evidence against a conditional analysis of disjunctions, besides not being able to account for modal subordination and inter-speaker disagreement. Although the account presented in this paper does not straightforwardly account for this data either, it does leave room for an explanation of this asymmetry, since disjunctions and conditionals may

---

<sup>7</sup>An anonymous reviewer points out that the acceptability of this sentence somewhat improves when replacing *that* with *it*. Observations of this kind provide an interesting angle for future research on the anaphoric availability of propositional drefs.

have distinct semantic representations. Therefore, the account provides a vantage point over asymmetries between individual and propositional anaphora, to be explored in future research.

## References

- Adrian Brasoveanu. Structured Nominal and Modal Reference. *Doctoral dissertation, Rutgers University*, 2007.
- Adrian Brasoveanu. Decomposing Modal Quantification. *Journal of Semantics*, 27(4):437–527, November 2010.
- Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100, February 1991.
- Christine Gunlogson. *True to Form : Rising and Falling Declaratives as Questions in English*. Routledge, June 2004.
- Irene Heim. *The semantics of definite and indefinite noun phrases*. Doctoral dissertation, University of Massachusetts, Amherst, 1982.
- Irene Heim. On the Projection Problem for Presuppositions. In *Proceedings of the Second West Coast Conference on Formal Linguistics*, pages 114–125, 1983.
- Hans Kamp. A theory of truth and semantic representation. *Formal semantics-the essential readings*, pages 189–222, 1981.
- Hans Kamp and Uwe Reyle. *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers, 1993.
- Lauri Karttunen. Discourse Referents. 1969.
- Lauri Karttunen. Presuppositions of compound sentences. *Linguistic Inquiry*, 4:169–93, 1973.
- Emiel Krahmer and Reinhard Muskens. Negation and Disjunction in Discourse Representation Theory. *Journal of Semantics*, 12(4):357–376, 1995.
- Reinhard Muskens. Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy*, 19(2):143–186, 1996.
- Craige Roberts. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12(6):683–721, December 1989.
- Robert Stalnaker. Assertion. In P. Cole, editor, *Pragmatics*, number 9 in *Syntax and Semantics*, pages 315–332. London: Academic Press, 1978.
- Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25:701–721, 2002.
- Matthew Stone. Reference to possible worlds. *RuCCS Report*, 46, 1999.
- Matthew Stone and Daniel Hardt. Dynamic discourse referents for tense and modals. In *Computing meaning*, pages 301–319. Springer, 1999.
- Martin van den Berg. *Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora*. PhD Thesis, dissertation, Universiteit van Amsterdam, 1996.

# Evidence Acquisition Time as Belief-State Change A View from Mvskoke (Creek)\*

Kimberly Johnson

University of Massachusetts Amherst, Amherst, Massachusetts, U.S.A  
kcjohnson@umass.edu

## Abstract

In some languages a single morpheme appears to encode both tense/aspect meaning and the evidential source for the assertion [7, 4, 5, 3, 9, 13]. I provide evidence from fieldwork that Creek (Muskogean) past tenses also encode evidentiality and show how the Creek data supports an approach which refers to *Evidence Acquisition Time* (EAT) [9, 13]. I then provide a novel formalization of EAT as the time the speaker's belief-state changes and demonstrate how this can be extended to Bulgarian, Korean and Matses.

## 1 Introduction

In recent years, two competing approaches have emerged regarding the puzzling relationship between evidentiality and tense/aspect in certain languages. In these languages, a single morpheme appears to encode both tense or aspect information and the evidential source for the assertion. This type of morpheme has been labeled the 'perfect of evidentiality' [7] and 'non-experienced past' [4]. The two approaches have centered on accounts of the Bulgarian evidential participle [7, 13] and the Korean evidential tense *-te* [3, 9].

Faller [4] and Chung [3] propose that evidential-tense morphemes are spatial deictic tenses which refer to a location at the topic time (TT). Both authors use spatio-temporal traces to model the intuition that direct evidence means the speaker was present as the event took place. Faller's account of Quechua non-experienced past *-sqa* uses two spatio-temporal trace functions. The *e-trace* function identifies the run-time and run-space of the event. The *P-trace* function maps the speaker to the locations she perceives at each point in her lifetime. Indirect evidentials, like *-sqa*, encode that speaker's perceptual field does not overlap with the spatio-temporal event trace. Chung [3] extends the spatial approach to Korean evidential morphemes, arguing that they require the speaker's perceptual field to cross paths with the spatio-temporal trace of the evidence for the event. Direct evidentials on the spatial approach encode that the speaker's perceptual field and the (evidence for the) event overlap in both location and time.

On the other hand, Lee [9] and Smirnova [13] propose that tense-evidential morphemes encode a relation between the *evidence acquisition time* (EAT) and another time. This approach relies on tense (relating EAT and TT) to derive the direct-indirect evidential distinction. Present tense indicates an overlap between EAT and TT; this means evidence was direct. Past and future introduce a precedence relation between TT and EAT so that they do not overlap; this means that evidence must be indirect. Under this view, a direct evidential encodes that

---

\*I owe a debt of gratitude to Jack Martin, Jennifer Johnson, Judy Montiel and Colleen Fitzgerald for paving the way for me to work in the Muscogee (Creek) and Seminole communities. *Mvto* to my language consultants Linda Sulphur Bear, Mary Ann Emarthla, Paul Fixico, Richard Harjo, Juanita Walker Harris, Rosemary McCombs Maxey, Ina Ann Micco Hickey, De Lois Roulston and Lucy Tiger. Thanks also to Seth Cable, Daniel Altshuler, John Kingston, Jack Martin, Ana Arregui, Vincent Homer and audiences at the UMass Semantics Workshop for helpful feedback and discussion. Any errors are my own.

evidence was acquired simultaneously to the event, but doesn't require the speaker to be in the same location as the event. This is not predicted by a spatial approach.

In what follows, I present original data from fieldwork that the tense system of Creek (mus; also Mvskoke, Seminole) also makes an indirect-direct evidential distinction. The contexts that I discuss include simultaneous indirect evidence and direct evidence after the fact. The behavior of Creek tenses in these contexts supports the temporal approach and proves problematic for the spatial approach. I propose that Creek tenses identify TT with EAT, which is in turn related to the event time (ET) through viewpoint aspect. The direct-indirect evidential distinction is derived entirely through the interaction of tense with aspect. I provide a novel formalization of the concept of EAT as the time when the belief state of the speaker changes. This formalization extends to other languages discussed in the literature and reduces cross-linguistic variation to the temporal restriction (if any) encoded in the morphemes.

## 2 Evidential Tenses in Creek

Creek is a severely endangered Muskogean language spoken by a little over 600 people in Oklahoma and Florida. Creek has four (previously five) past tenses [6, 10, a.o.]. One of these is a general, unrestricted past marker ('P4'), while the others ('P1-P3') cover increasingly broad intervals of past time (P1 = today; P2 = up to 5 years ago; P3 = up to 70 years ago). In matrix clauses, Gricean competition between the restricted past tenses results in their covering disjoint intervals of time, (1) and (2).<sup>1</sup> The restricted tenses are exemplified below. Imperfective and perfective aspect differ in the tonal pattern of the word. Note that P1 can appear as an infix, especially when combining with perfective aspect.<sup>2</sup>

(1) Perfective Past Tenses	(2) Imperfective Past Tenses
a. Mary lat <i>éyk</i> -is Mary fall.P1.PFV-IND 'Mary fell.' (today/last night)	a. ni:s- <i>éy</i> -s buy.IMP-P1-IND 'He/she was buying it.'
b. Mary lâ:tk- <i>ánk</i> -s Mary fall.PFV-P2-IND 'Mary fell.' (yesterday-5 yrs ago)	b. ni:s- <i>ánk</i> -s buy.IMP-P2-IND 'He/she was buying it.'
c. Mary lâ:tk- <i>imát</i> -s Mary fall.PFV-P3-IND 'Mary fell.' (long ago) (PF-Mus-06/2018)	c. ni:s- <i>imát</i> -s buy.IMP-P3-IND 'He/she was buying it.' [11, 263-66]

I claim P4 is not restricted to a specific interval of past time, but pragmatic competition with P1-P3 results in P4 usually triggering an inference that the event happened in the remote past, (3).

<sup>1</sup>I argue elsewhere based on their behavior in questions and remoteness indeterminacy contexts that the tenses cover increasingly broad intervals of time instead of disjoint intervals [8].

<sup>2</sup>Examples are in a phonemic transcription (following Martin [11]). Personal field-notes are cited by speaker code, dialect (Mus-Muskogee; Sem-Seminole), and date of elicitation. The abbreviations used are as follows: ACC accusative; CAUS causative; COMP complementizer; DAT dative; DEM demonstrative; DS different subject; DUR durative; IMP imperfective aspect; IND indicative; IP medio-passive; NOM nominative; NZL nominalizer; P1 recent past; P2 intermediate past; P3 distant past; P4 remote past (P5 in the Muskogean tradition); PASS impersonal passive; PAT patient; PFV perfective aspect; PL plural; RECP reciprocal; SG singular; SS same subject; TPL triplural.

- (3) Pòkkicc-a:k-**atí**:-s.  
 play.ball-PL.IMP-**P4**-IND  
 ‘They played stickball.’ (long ago) (RH-Sem-07/2019)

P4 typically combines with imperfective aspect (according to [11]) - a puzzle for future research.

**Evidentiality** The graded past tense system also encodes a direct-indirect evidential distinction. The sentences in examples (1) and (2) are most naturally used in direct witness contexts. The sentence in (3) generally imply that the speaker did not witness the ball game. Likewise in reported contexts, speakers reject the restricted tenses (P1-P3) and instead accept P4. Speaker comments for each of the tenses point to the direct witness meaning associated with P1-P3. Example (4) establishes a reported evidence context within the day of utterance (an interval compatible with P1). In this context, a sentence marked with P4 is felicitous, but a sentence marked with P1 is infelicitous.

- (4) Context: Imagine your friend Mary tells you that she talked to the chief today.
- a. Mary mocá nítta míkko im-ponây-**atí**:-s.  
 Mary this day chief 3.DAT-talk.PFV-**P4**-IND  
 ‘Mary talked to the chief today.’
- b. #Mary míkko im-pona:yêyy-s.  
 Mary chief 3.DAT-talk.**P1**.PFV-IND  
 ‘Mary talked with the chief.’ (today)  
 Speaker Comment: That would be if you saw her talking to him. (DLR-Mus-06/2019)

Example (5) establishes a reported evidence context within a temporal interval compatible with P2. In this context, a P4-marked sentence is felicitous, but a P2-marked sentence is not.

- (5) Context: Last year, a woman you know, Mary, spoke with Chief Leonard Harjo and she told you about it.
- a. Mary ohłolopí: hank-ánk-i: míkko Leonard im-ponâ:y-**atí**:-s.  
 Mary year one-P2-NLZ chief Leonard 3.DAT-speak.PFV-**P4**-IND  
 ‘Last year, Mary spoke with Chief Leonard.’
- b. #Mary ohłolopí: hank-ánk-i: míkko Leonard im-ponâ:y-**ánk**-s.  
 Mary year one-P2-NLZ chief Leonard 3.DAT-speak.PFV-**P2**-IND  
 ‘Last year, Mary spoke with Chief Leonard.’  
 Speaker Comment: That’s like I saw Mary talking to him. (JWH-Sem-07/2018)

In example (6), the reported evidence is established in a subordinate chained clause “It was written in the newspaper” and the event in question is 30-40 years ago. Although this falls within P3 interval, P3 is unacceptable. Instead, the speaker corrected the sentence to (6-a).

- (6) Prompt: Imagine you read a newspaper story about the chief in which you learned that he frequented a certain church when he was young (30-40 years ago).
- a. Coka-talámi: hocéyhoc-â:t, míkko má mi:kosapka-cóko  
 paper-daily write.IMP-CAUS.PASS-COMP.SS chief DEM prayer-house  
 a:r-í:-t o:w-**atí**:-s.  
 go.SG.IMP-DUR-SS be.IMP-**P4**-IND  
 ‘It was written in the newspaper, the chief went to that church.’

- b. #Coka-talámi: hocéyhoc-â:t, míkko má mi:kosapka-cóko  
 paper-daily write.IMP-CAUS.PASS-COMP.SS chief DEM prayer-house  
 a:r-í:t ô:w-i:mát-s.  
 go.SG.IMP-DUR-SS be.PFV-P3-IND  
 ‘It was written in the newspaper, the chief went to that church.’  
 Speaker Comment: No, if it’s according to the paper it would be [(6-a)]  
 (MAE-Sem-07/2018)

These three examples have demonstrated that P1-P3 are infelicitous in indirect evidence contexts. Instead, P4 is used for an utterance based on reported evidence at any past time. Normally, P4 is infelicitous in direct evidence contexts.

- (7) a. Context: Imagine you’ve been telling your brother there’s a girl who wants to see him. Then last week you drove by the diner and saw them together. To tell me that they saw each other, could you say...  
 b. #Iti-hi:c-â:k-ati:s.  
 RECP-see-PL.PFV-P4-IND  
 ‘They saw each other.’  
 Speaker Comment: Not if I saw them at the diner. (LSB-Mus-06/2017)

One explanation could be that P4 is an indirect evidential, as proposed in previous literature [2, 12]. Under this view, P1-P3 are not necessarily evidential, since their use in direct evidence contexts could be an implicature arising from competition with the indirect P4. A strong argument against this explanation is that P4 can be used in direct witness contexts when it appears in a downward entailing environment, such as in the restrictor of a universal quantifier.

- (8) a. Context: Imagine you and your preacher were going around town together this month. Every time he invited someone to your next church meeting you were with him and saw him do it. When your Sunday meeting rolled around, everyone he invited came. On Monday you want to tell me about it.  
 b. Iġkanáka istí: im-pohátt-ati: omâlk-a:t, foll-â:k-ânk-s.  
 preacher people 3.DAT-invite-P4 all.PFV-COMP go.about.TPL-PL.PFV-P2-IND  
 ‘All the people the preacher invited showed up/came.’ (LSB-Mus-11/2019)<sup>3</sup>

Implicatures do not arise in downward entailing environments, but semantic meaning is preserved. If indirect evidence was part of the semantic meaning of P4, then it is predicted to be infelicitous in a direct evidence context even when it is in the restrictor of the universal quantifier *omalka* ‘all, every’. I take this as evidence that the indirect meaning associated with P4 is a scalar implicature.

I argue instead that P1-P3 are semantically restricted to direct evidence contexts as well as to certain time intervals. Based on speaker comments, an initial hypothesis would be that P1-P3 are direct witness evidentials. If a speaker directly witnesses an event, it follows that there was overlap both in times (of the event and of evidence acquisition) and in locations (the location of the event and the speaker’s perceptual field). As such, the contexts in (4), (5), and (6) do not allow us to choose between the two approaches to evidential tense. There are, however, two challenges to analyzing these morphemes as encoding overlap in location.

<sup>3</sup>I would like to thank Jack B. Martin for eliciting this example for me.

**Simultaneous Indirect Evidence** First, P1-P3 can be used in indirect contexts if the speaker learned of the event as it took place. For example, P2 is acceptable in contexts where the speaker got evidence over the phone (9) as the event was happening. In this context, the speaker acquires evidence simultaneously to the event without the event being inside their perceptual field. There is overlap in times, but not in locations.

- (9) a. Context: Your friend called yesterday and said, “I’m at the Salon right now and my wife is getting her hair cut.” Today I ask you what your friend’s wife did yesterday.  
 b. An-híssi: i:héywa-n iké-yssi:-n ín-tonhô:w-**ank-s**.  
 1.SG.DAT-friend 3.PAT.wife-ACC hair-ACC 3.DAT-cut.PFV.PASS-**P2-IND**  
 ‘My friend’s wife got her hair cut.’ (RMM-Mus-06/2019,PF-Mus-11/2018)

The use of P2 in (9) is unexplained on the spatial approach. As such, this supports the temporal approach which derives direct evidentiality through temporal overlap and not necessarily spatial overlap.

**Direct Evidence After the Fact** Furthermore, P1-P3 can be used in contexts where the speaker learned of the event after the fact, so long as the morpheme appears on a special verbal auxiliary. In the indirect context below, P1 is unacceptable on the main verb. Instead it must occur on the auxiliary verb. In a direct witness context, speakers report opposite judgments: (10-b) is unacceptable and P1 must be marked on the main verb.

- (10) P1 Context: Imagine your friend Bill called you on the phone just now and told you he just cut his hair. Would it sound okay to say the following sentence?  
 a. #An-híssi-t an-hôyheyk-it ika-éyssi: wá:hł-is.  
 1.SG.DAT-friend-NOM 1.SG.DAT-call.PFV-SS head-hair cut.**P1**.PFV-IND  
 Speaker Comment: No, you’d say (10-b).  
 b. An-híssi-t an-hôyheyk-it ika-éyssi: wał-î:p-at  
 1.SG.DAT-friend-NOM 1.SG.DAT-call.PFV-SS head-hair cut-IP.PFV-P4  
**hâ:k-éy-s**.  
**become.PFV-P1-IND**  
 ‘My friend called, he cut his hair.’ (AM-Mus-Elic11/2018)

P2 also cannot appear on the main verb in an indirect witness context like in (11). It is only acceptable if it marks the auxiliary. Again, in a direct witness context, the (a) example is acceptable and the (b) example is not. There is parallel data for P3.

- (11) P2 Context: A month ago, you weren’t home when Sam called, but you saw it on the answering machine.  
 a. #Hasí: hank-ank-í: Sam ahokkolá-n a:-an-hôyhk-**ank-s**.  
 month one-P2-NLZ Sam twice-ACC DIR-1.SG.DAT-call-**P2-IND**  
 ‘Sam called me twice, one month ago.’  
 Speaker Comment: No, that would be if you were home when your phone rang.  
 b. Hasí: hank-ank-í: Sam ahokkolá-n a:-an-hôyhk-at **hâ:k-ánk-s**.  
 month one-P2-NLZ Sam twice-ACC DIR-1.SG.DAT-call-P4 **become.PFV-P2-IND**  
 ‘Sam called me twice, one month ago.’ (MAE-Sem-07/2018)



- (12) P3 Context: Two years ago, Sam called you on the phone. You weren't in your house when he called, but you saw his message on the answering machine.
- a. #Ohłolopí: hokkol-ánk-i: mâ:h-in Sam an-hôyhk-**i:mát-s**.  
 year two-P2-NLZ about.PFV-DS Sam 1.SG.DAT-call.PFV-**P3**-IND  
 'About two years ago, Sam called me.'  
 Speaker Comment: This would mean that you were home when he called and saw your phone.
- b. Ohłolopí: hókkol-ánk-i: mâ:h-in Sam an-hôyhk-at  
 year two-P2-NLZ about.PFV-DS Sam 1.SG.DAT-call.PFV-P4  
**hâ:k-imát-s**  
**become.PFV-P3-IND**  
 'About two years ago, Sam called me.' (MAE-Sem-08/2018)

The generalization that emerges from these data is that when the restricted tenses appear on the main verb, the time of the event and the time of evidence acquisition must coincide. When the restricted tenses appear on the auxiliary *hak-*, the time of the event must precede the time of learning about the event. The spatial approach incorrectly predicts that P1-P3 should be unacceptable in *simultaneous* and *after-the-fact* contexts. I argue that when P1-P3 appear on the main verb, ET and EAT must overlap, but when P1-P3 appear on the special auxiliary *hak-*, ET must precede EAT.

### 3 Formalizing Evidence Acquisition Time

Similarly to Smirnova's account for Bulgarian and Lee's for Korean, I propose that Creek evidential tenses locate the EAT with respect to another time. Whereas Smirnova and Lee augmented the Neo-Reichenbachian system of times with EAT as distinct from ET, TT and UT, I propose instead that P1-P3 identify the TT as the time at which the speaker 'came to believe' the proposition - the Belief-State Change time. The difference between monoverbal and auxiliary constructions with P1-P3 is derived through the tenses' interaction with viewpoint aspect.

In monoverbal constructions EAT and ET overlap. The main verb is inflected for either imperfective or perfective aspect, both of which require ET to overlap with TT. P1-P3 then restrict the TT in two ways. First, they place it within an interval of time that is a certain distance from UT. Second, they restrict it to the time the speaker learned of the prejacent (the EAT). This derives the direct witness contexts as well as the simultaneous contexts. In auxiliary constructions ET precedes EAT. I take auxiliary constructions to be complex tense constructions equivalent in meaning to past perfect. The main verb bears both (im)perfective aspect and P4; aspect places the ET overlapping with an interval of time contributed by P4, which precedes another interval, contributed by tense on the auxiliary verb. The evidential tense on the auxiliary then relates that interval to the UT and identifies it as the time of learning. In essence, auxiliary constructions communicate that the speaker learned of a past event at a subsequent past time. This derives the use of P1-P3 on the auxiliary in the indirect contexts in (10) - (12). Monoverbal constructions with P4, on the other hand, quite simply locate an event in the past of the utterance. Thus under this approach, P4 is an unrestricted past tense similar to English past.

**EAT as Belief-State-Change** I formalize EAT as the time at which someone comes to believe a proposition, and define a meta-language predicate COME-TO-BELIEVE.  $\text{COME-TO-BEL}(x,t,p) = \text{T}$  iff at all times prior to  $t$ , not all  $x$ 's belief worlds were compatible with  $p$ , but at all times following  $t$ , all  $x$ 's belief worlds are  $p$ -worlds.

$$(13) \quad \lambda x_e \lambda t_i \lambda P_{\langle st \rangle} : \text{COME-TO-BEL}(x,t,P) = \lambda x_e \lambda t_i \lambda P_{\langle st \rangle} [\forall t' : t' \prec t. \neg \forall w' \in \text{BEL}(x,w,t') : P(w') = \text{T} \ \& \ \forall t'' : t \preceq t''. \forall w'' \in \text{BEL}(x,w,t'') : P(w'') = \text{T}]$$

I propose that P1-P3 are temporal operators that introduce a temporal index  $i$  and restrict that time to the appropriate interval associated with P1-P3 and to the time the speaker 'came to believe' a given proposition. That is, P1-P3 state that the speaker 'came to believe' the proposition at the TT which overlaps with a particular time interval. I abstract away from the precise semantics of the time intervals and refer to them as the P1-interval, P2-interval, and P3-interval. I illustrate with P2.

$$(14) \quad \llbracket \text{P2}_i \rrbracket^{c,w,g,t} = [\lambda P_{\langle i,t \rangle}. g(i) \text{ o P2-interval}(t) \ \& \ \text{COME-TO-BEL}(\text{sp}(c),g(i),P(g(i)))].$$

This denotation for P2 states that  $g(i)$  lies within the P2-interval and the speaker came to believe - at  $g(i)$  - that the proposition was true at  $g(i)$ . The semantics in (14) correctly predicts that the event of learning must also take place in the P2-interval.

**Interactions with Viewpoint Aspect** Given the proposed semantics for P1-P3, the puzzling facts in (9) - (12) follow from their interaction with aspectual morphology. When P1-P3 are affixed to the main verb, I argue the verb is inflected for perfective or imperfective aspect. Imperfective aspect requires ET to contain TT; perfective aspect requires TT to contain ET. The semantics are found in (15).

$$(15) \quad \begin{array}{l} \text{a. } \llbracket \text{Perfective} \rrbracket^{w,g,t} = [\lambda P_{\langle et \rangle}. [\lambda t'_i. \exists e. \tau(e) \subseteq t' \ \& \ P(e) = \text{T}]] \\ \text{b. } \llbracket \text{Imperfective} \rrbracket^{w,g,t} = [\lambda P_{\langle et \rangle}. [\lambda t'_i. \exists e. t' \subseteq \tau(e) \ \& \ P(e) = \text{T}]] \end{array}$$

Given the contribution of P1-P3, the ET will end up overlapping with the EAT, so that the speaker 'comes to believe' the proposition as the event is taking place. (16) shows the LF and truth conditions for the monoverbal sentence (5).

$$(16) \quad \begin{array}{l} \text{a. LF for (5): } [\text{TP } \text{P2}_i [\text{AspP } \text{PFV} [\text{vP } \text{Mary } v [\text{vP } \text{Chief Leonard} [\text{v } \text{talk.to} ] ] ] ] ] \\ \text{b. Truth conditions: } \llbracket (5) \rrbracket^{c,w,g,t} = \text{T iff } g(i) \text{ o P2-interval}(t) \ \& \ \text{COME-TO-BEL}(\text{sp}(c), \\ g(i), \exists e. \underline{\tau(e)} \subseteq g(i) \ \& \ \text{talk.to}(e, \text{Mary}, \text{Chief-Leonard})) = \text{T} \end{array}$$

Since the semantics require the time of belief-state change to coincide with the run-time of the event, these truth conditions of (5) are satisfied in direct witness contexts as well as simultaneous contexts. This accounts for the judgments in (4), (5), (6) and (9).

I further argue that the special auxiliary construction in (10) - (12) is a complex tense construction with a sequence of two past tenses - P4 on the main verb and P1-P3 on the auxiliary. (Im)perfective aspect on the main verb requires that the ET overlap with the interval contributed by P4. P4 in turn encodes a precedence relation between that time and the TT. The semantics I assume for P4 are given below.

$$(17) \quad \llbracket \text{P4}_i \rrbracket^{w,g,t} = [\lambda P_{\langle i,t \rangle}. [\lambda t'_i. g(i) \prec t' \ \& \ P(g(i)) = \text{T}]]$$

Consequently an auxiliary construction combining with P1-P3 will require that the ET precede the EAT. Thus sentences like (11) have the following LF and truth conditions.



- (22) Truth Conditions for (20-b) and (21-b):  $\llbracket \text{Maria pišela kniga} \rrbracket^{c,w,g,t} = \text{T}$  iff  
 COME-TO-BEL(sp(c),g(i), $\exists e. g(i) \subseteq \tau(e) \ \& \ \text{write}(e,\text{Maria},\text{book}) = \text{T}$ )

The temporal pronoun  $g(i)$  will receive its value from a contextually determined variable assignment function  $g$ , so that in a past context it refer to the contextually salient past time and in a present tense contexts it refers to the time of utterance.

Arregui et al. [1] characterize the perfective aspect form of the participle as enforcing a strict precedence relation between ET and TT. In this way, their intuition about the temporal contribution of the perfective participle aligns with the past tense semantics that Smirnova gives it. I model this with the semantics in (23).

- (23) Bulgarian Perfective Semantics:  
 $\llbracket \text{PFV} \rrbracket^{w,g,t} = [\lambda P_{\langle \epsilon, t \rangle} . [\lambda t'_i . \exists e. \tau(e) \prec t' \ \& \ P(e) = T]]$

When the evidential participle combines with perfective aspect, its truth conditions are satisfied if the speaker learned of an event after it is finished. In past contexts, this will mean that the speaker learned at some past time of a completed event, (24). In present tense contexts, my approach predicts it will mean that the speaker just learned of a past event or state. Smirnova, unfortunately, does not provide a good example of a perfective evidential participle in a present tense context.

- (24) Perfective Bulgarian Evidential Participle
- a. *Past Tense Context-Learning after-the-fact*: Last month Ivan told you that Maria your former classmate, spend last year writing a book and that it has just been published. You believe Ivan. Today, your friend asks what Maria was doing last year. You say:
  - b. Maria pisala kniga.  
 Maria write.PFV.EV book  
 ‘Maria was writing/wrote a book, [I heard].’ [13, 481, glossing adapted]

For the above example, the evidential meaning of the participle combines with perfective aspect to yield the truth conditions in (25).

- (25) Truth Conditions for (24-b):  $\llbracket \text{Maria pisala kniga} \rrbracket^{c,w,g,t} = \text{T}$  iff  
 COME-TO-BEL(sp(c),g(i), $\exists e. \tau(e) \prec g(i) \ \& \ \text{write}(e,\text{Maria},\text{book}) = \text{T}$ )

Smirnova spends a great deal of time accounting for the exclamatory intonation which is obligatory in present tense contexts. Under the Belief-State Change account, this follows quite naturally. In present tense contexts, the use of the evidential participle conveys that the speaker is learning of the event/state at the present time. This accounts for the exclamatory intonation that must accompany the evidential participle in such contexts and why the evidential participle is unacceptable if the speaker was previously aware of the truth of  $p$ , (26).

- (26) a. *Previous knowledge context*: Stojan sustained an injury and was unable to play four games. You know that he has fully recovered and is scheduled to play the game today. When you go to the game and see Stojan scoring the goal, you say:  
 b. #Stojan igrael!  
 Stojan play.IMP.EV  
 ‘Stojan is playing, [I see]!’ [13, 509-10, glossing adapted]

In conclusion, I have shown how the Belief-State Change system can be extended to account

for the Bulgarian evidential participle. Just as in Creek, the evidential interacts with viewpoint aspect. However I proposed that the evidential in Bulgarian does not place a temporal restriction on the Belief-State Change time, but that it receives its value from context. Furthermore, formalizing EAT as the time one's beliefs change provides an explanation of the Bulgarian evidential participle's mirative-like behavior in present tense contexts.

## 5 Conclusion

I have argued that Creek restricted past tenses encode the time a speaker came to believe a proposition. We saw evidence from Creek that adjudicated in favor of a temporal and not a spatial approach to evidential tenses. I formalized *evidence acquisition time* in terms of time of belief-state change and was able to account for the Creek data and extend my approach to Bulgarian. In ongoing work, I have explored how this system extends to Korean [9] and Matses [5]. My account reduces cross-linguistic variation in evidential tense systems to the temporal restriction the marker imposes on the topic time, identified as the time of coming to believe. The direct and indirect evidential meanings are derived through the tenses' interaction with lower tense and aspectual morphology.

## References

- [1] Ana Arregui, María Luisa Rivero, and Andrés Salanova. Aspect and tense in evidentials. In Ana Arregui, María Luisa Rivero, and Andrés Salanova, editors, *Modality across syntactic categories*, pages 211–234. Oxford University Press, first edition, 2017.
- [2] D. G. Brinton. Contributions to a grammar of the Muskokee language. In *Proceedings of the American Philosophical Society*, volume 11, pages 301–309, 1870.
- [3] Kyung-Sook Chung. Spatial deictic tense and evidentials in Korean. *Natural Language Semantics*, 15(3):187–219, 2007.
- [4] Martina Faller. The deictic core of "non-experienced past" in Cuzco Quechua. *Journal of Semantics*, 21:45–85, 2004.
- [5] David W. Fleck. Evidentiality and double tense in Matses. *Language*, 83(3):589–614, 2007.
- [6] Mary R. Haas. Ablaut and Its Function in Muskogee. *Language*, 16(2):141, 1940.
- [7] Roumyana Izvorski. The Present Perfect as an Epistemic Modal. In Aaron Lawson, editor, *SALT VII*, pages 222–239, Ithaca, NY: Cornell University, 1997.
- [8] Kimberly Johnson. Graded and Evidential Tenses in Mvskoke (Creek). 2019.
- [9] Jungmee Lee. Temporal constraints on the meaning of evidentiality. *Natural Language Semantics*, 21(1):1–41, 2013.
- [10] Jack B. Martin. How to Tell a Creek Story in Five Past Tenses. *International Journal of American Linguistics*, 76(1):43–70, 2010.
- [11] Jack B. Martin. *A grammar of Creek (Muskogee)*. University of Nebraska Press, 2011.
- [12] Michelle Nathan. *Grammatical description of the Florida Seminole dialect of Creek*. Ph.D. dissertation, Tulane University, New Orleans, 1977.
- [13] A. Smirnova. Evidentiality in Bulgarian: Temporality, Epistemic Modality, and Information Source. *Journal of Semantics*, 30(4):479–532, 2013.

# Pair-List Questions from Dependent Plurals \*

William Johnston

McGill University, Montréal, Québec, Canada  
william.johnston4@mcgill.ca

## Abstract

Questions with plural definites may receive apparent pair-list answers. The dominant analysis in the literature holds that these questions can only receive weak cumulative Hamblin answers, and that these cumulative answers may be elaborated into pair-list responses for pragmatic reasons. Here, I provide evidence that the “cumulation-only” account is insufficient, and that in certain contexts questions containing plural definites receive genuine pair-list Hamblin answers. I argue that these answers arise when the question nucleus contains an anaphoric dependency function between the *wh*-item and a plural definite, parallel to the proposal for declaratives given by Winter (2000). I conclude that two distinct parses must be available for questions with plural definites: a cumulation-and-elaboration strategy, and a dependency strategy.

## 1 Introduction

It has been observed that questions with plural definites (QPDs) may receive pair-list responses.<sup>1</sup> For example, (1a) may elicit (1b).

1. (a) Who do these men like? (Dayal, 2016)  
(b) John likes Sue and Bill likes Jane.

While this response is felicitous in many contexts, is (1b) truly the semantic answer to (1a)? That is, is (1b) the strongest true member of (1a)’s Hamblin set?<sup>2</sup> Dayal (1992, 1996) and Krifka (1992) hold that Hamblin answers to QPDs can only take the form of weak cumulative propositions, and thus there is no member of (1a)’s Hamblin set that takes the form of (1b). Instead, (1b) is held to be a pragmatically-motivated elaboration on the strongest true (cumulative) Hamblin answer. In Section 2, I review the Dayal-Krifka account (which I will refer to as the “cumulation-only” account), drawing attention to two key pieces of supporting evidence.

In Section 3, I present new data that shows that the cumulation-only account is insufficient, and that pair-list answers to QPDs must be derivable as Hamblin answers in the semantics. In Sections 4 and 5 I argue that plural *wh*-phrases in QPDs can have a semantics parallel to Winter’s proposal for so-called dependent definites, and I explore some predictions that this structure makes. I conclude with a brief discussion of an unresolved question raised by this approach.

---

\*Thanks to Bernhard Schwarz for his tremendous assistance with this project. Thanks as well to three anonymous reviewers and to the Syntax/Semantics reading group at McGill University for their feedback and suggestions. This research was supported by the Social Sciences and Humanities Research Council of Canada (Insight Grant #435-2016-1448, PI: Bernhard Schwarz).

<sup>1</sup>See e.g. Groenendijk and Stokhof (1984) and Pritchett (1990).

<sup>2</sup>I.e., the set of all possible answers to (1a). (Hamblin, 1973)

## 2 Pair-list answers via cumulation

### 2.1 The cumulation-only account.

Under the cumulation-only account, the Hamblin answers to a QPD must receive a so-called cumulative interpretation of the main predicate. This is described by the  $**$ -operator of Krifka (1986) and Sternefeld (1998). “ $**P(X, Y)$ ” expresses the weak proposition that for each  $x$  in  $X$ , there is some  $y$  in  $Y$  such that  $P(x, y)$  holds, and for each  $y$  in  $Y$ , there is some  $x$  in  $X$  such that  $P(x, y)$  holds. That is, each part of the first group must be related to some part of the second, and vice-versa.

In the case of (1), the cumulation-only account holds that the strongest true Hamblin answer to (1a) is equivalent to the proposition that “these men (between them) like Sue and Jane”. This can be expressed using the  $**$ -operator as in (2a), which is equivalent to the first-order logical expression in (2b).

2. (a)  $**\text{like}(\text{these men}, \text{Sue}+\text{Jane})$   
 (b)  $\forall x[\text{man}(x) \rightarrow \exists y[y \in \{\text{Sue}, \text{Jane}\} \wedge \text{like}(x, y)]] \wedge$   
 $\forall y[y \in \{\text{Sue}, \text{Jane}\} \rightarrow \exists x[\text{man}(x) \wedge \text{like}(x, y)]]$

This answer is the strongest true member of the Hamblin set shown in (3), which contains all possible cumulative *liking* relations from the relevant men to women.

3. Hamblin set:  $\{**\text{like}(\text{these men}, x) : x \in *\{\text{Sue}, \text{Jane}, \text{Lisa}, \text{Amy}, \dots\}\}$

Note that this Hamblin set does not contain propositions that encode mappings from particular men to particular women. Thus, the pair-list response (1b) cannot, under this account, be derived in the semantics.

Instead, (1b) is analyzed as a pragmatically-motivated elaborative response—one that expresses the strongest true Hamblin set member (2a), plus some amount of additional information. The truth of (1b) entails the truth of the Hamblin answer (2a), so no information expressed by that Hamblin answer is lost. (However, (2a) is equally compatible with a situation in which John likes Jane and Bill likes Sue, or in which John likes Sue and Bill likes both Jane and Sue, etc.)

### 2.2 Review: Evidence for cumulation-only

**Cumulation is a possible source for pair-list responses.** QPDs have been shown to lack the subject-object asymmetries that are characteristic of canonical pair-list questions. Compare the questions with quantifiers in (4-5) with the QPDs in (6-7). The pair-list answer is felicitous in (4), when *wh*-moving the object, but infelicitous when *wh*-moving the subject in (5).

4. (a) Which professor does every student like? (Dayal, 2016)  
 (b) Bill likes Professor Carl and Alice likes Professor Dan.  
 5. (a) Which student likes every professor? (Dayal, 2016)  
 (b) #Bill likes Professor Smith and Alice likes Professor Jones.

The contrast between (4) and (5) is discussed by Chierchia (1993), who holds it to be a matter of quantifier scope. In (4), the universal quantifier attached to the subject takes wide scope relative to the base position of the *wh*-item, enabling the pair-list reading. In (5), the



quantifier attaches to the object, which remains below the base position of the *wh*-item. This means there is no way to derive the necessary scope ordering for a pair-list reading without incurring a weak crossover violation. The pair-list reading is therefore unavailable in (5).

Compare questions with quantifiers to QPDs, where a pair-list response is felicitous regardless of which argument is *wh*-moved.

6. (a) Who do these men like? (Dayal, 2016),  
 (b) John likes Sue, and Bill likes Jane. repeated from (1a)
7. (a) Who likes these men? (Dayal, 2016)  
 (b) Sue likes John, and Jane likes Bill.

This lack of asymmetry has been taken as evidence that QPDs do not receive their pair-list responses from the same source as questions with quantifiers. Instead, these pair-list responses must derive from a source that is available in both (6) and (7), and that is not sensitive to the relative scope of the two arguments. This source is argued to be cumulation. (Dayal, 1992; Krifka, 1992)

However, while this is persuasive evidence that a cumulation-and-elaboration strategy must be available for (6-7), this does not preclude the availability of other strategies.

**Pair-list responses *only* derive from cumulation.** It has also been shown that pair-list responses to QPDs appear to depend on the plurality of both arguments. For example, when the *wh*-item is singular, the pair-list response becomes infelicitous.

8. (a) Which women do these men like?  
 (b) John likes Sue, and Bill likes Jane.
9. (a) Which woman do these men like?  
 (b) #John likes Sue, and Bill likes Jane.

As Dayal and Krifka point out, we would not expect a quantifier scope-based account of the pair-list reading to require plurality of the *wh*-item, since questions that uncontroversially receive their pair-list readings in this way do not abide by such a requirement.<sup>3</sup>

On the other hand, the contrast between (8) and (9) is quite natural under the cumulation-only account, since the only cumulative relations possible with a singular argument are trivial. We can certainly apply the *\*\**-operator to a predicate with one singular and one plural argument; however, there can be no variation in which part of the singular argument the various parts of the plural argument are related to.<sup>4</sup> In the question context, this effectively gives us an individual answer.

In summary, the lack of subject-object asymmetries in QPDs provides evidence that pair-list responses can be generated by a cumulation-and-elaboration strategy, and the unavailability of pair-list responses in precisely the cases where cumulation is trivial provides evidence that no other strategy is available.

<sup>3</sup>E.g., example (6) in the preceding section.

<sup>4</sup>To spell this out more clearly: the answer to (9a) will be of the form shown below. This spells out the proposition that every atomic part of *the men* likes some atomic part of *the woman*, and since there is only one atomic woman in this scenario, every man must like that one woman.

i.  $\forall x[\text{man}(x) \rightarrow \exists y[\text{woman}(y) \wedge \text{likes}(x, y)]] \wedge \forall y[\text{woman}(y) \rightarrow \exists x[\text{man}(x) \wedge \text{likes}(x, y)]]$



### 3 Cumulative answers are not strong enough

**Cumulation-only undergenerates pair-list responses.** While the arguments reviewed in the preceding section are persuasive, insofar as they establish that a cumulation-and-elaboration strategy is necessary, several new observations show that cumulation cannot be the only source of pair-list responses to QPDs. To demonstrate, I will consider the example in (10).

10. (a) Which numbers did the players pick?  
 (b) Ann picked 1, Ben picked 2, Chris picked 3, Dan picked 4, Emma picked 5.

This is a felicitous question-answer exchange in context (11). The assistant coach in this example could quite naturally ask (10a), eliciting a response of (10b) from the head coach.

11. The head coach of a basketball team had five jerseys made, numbered 1-5, for the five players on the team. Each player chose a jersey. The assistant coach knows all five players on the team, knows the numbers that were available, and believes that each of those player chose exactly one of those jerseys. However, the assistant coach was not present for the choosing, and so doesn't yet know which player selected which jersey.

Under the cumulation-only account, (10a) is assumed to have a Hamblin set of the form shown in (12a), of which (12b) is the strongest true member. In fact, if *picked* is interpreted cumulatively, (12b) is the *unique* true Hamblin answer in this context.

12. (a) Hamblin set:  $\{**\text{picked}(\text{the players}, x) : x \in \{1, 2, 3, 4, 5, \dots\}\}$   
 (b) Strongest true answer:  $**\text{picked}(\text{the players}, 1+2+3+4+5)$

Any proposition in the Hamblin set (12a) that contains more than five numbers will be false, since only five numbers are available. Any proposition that contains fewer than five numbers will also be false, since in this context, all players must have picked a jersey and two players cannot choose the same jersey. And any proposition that contains a number not available to be picked, e.g. 6, will also be false. All true answers must relate the players to exactly the five numbers available in context, and since *picked* is interpreted cumulatively, there is exactly one proposition in the Hamblin set that does so.

That (10a) is a felicitous question in this context poses a problem under the cumulation-only analysis. (11) clearly establishes that the assistant coach already believes the unique true cumulative answer (12b) to be the case (i.e., the assistant coach believes the numbers that the players chose were 1–5). But information-seeking questions are generally infelicitous when the strongest true answer is known to the questioner (and when the answerer knows that the questioner knows). For example, the assistant coach in this context could not ask (13), since the players' identities are already common knowledge.

13. #Who are the players?

In short, if the cumulation-only analysis is correct, we expect the asking of (10a) to be as infelicitous as the asking of (13). Yet it is perfectly natural for the assistant coach to ask (10a), eliciting (10b). This suggests that (10a) can have a true Hamblin answer that is stronger than the weak cumulative proposition in (12b).

I hypothesize that this stronger answer is a true pair-list answer, encoding the mappings stated in (10b). To test this, we can modify (10a) by adding the adverbial expression *between them*, which will disambiguate the question towards a cumulative reading. Under the

cumulation-only account, this should produce no change in interpretation, since (10a) is necessarily already cumulative. If my hypothesis is correct, and (10a) must receive a pair-list answer in order to be a felicitous information-seeking question in this context, then forcing a cumulative reading should result in infelicity. This prediction is borne out.

14. #Which numbers did the players pick between them?

As (14) shows, forcing cumulation causes the question to become infelicitous in context (11). (14) can only be understood as a question about the identity of the numbers chosen, which is already known to the assistant coach.

These observations suggest that a QPD can receive a true pair-list answer in the semantics. However, one might argue that these facts can be explained away in the pragmatics. Perhaps questioners have learned to exploit the tendency of answerers to give elaborative responses, and this has become conventionalized to such a degree that prior knowledge of the strongest true answer is immaterial. And perhaps the addition of *between them* somehow precludes this conventionalized usage. As we will see in the next section, however, data from question-embedding poses a significant challenge to proposals of this type.

**Question-embedding removes pragmatic strategies.** We can isolate these problems from pragmatic effects by considering cases where the question is embedded. Certain question-embedding verbs, such as *wonder* or *discover*, only produce a true and felicitous proposition if their subject lacks complete knowledge of the embedded question's strongest true answer. (Roelofsen and Uegaki, 2016; Spector and Egré, 2015) When that answer is known, infelicity or falsity results. For example, neither proposition in (15) is true and felicitous in (11), since the assistant coach already has complete knowledge of the players' identities.

15. (a) #The assistant coach wonders who the players are.  
(b) #The assistant coach will discover who the players are.

Both examples in (16), then, imply ignorance of the strongest true answer to (10a). If the cumulation-only account holds, this answer is already known to the assistant, which should result in the falsity or infelicity of the examples in (16). Yet both of these can be true and felicitous in context!

16. (a) The assistant coach wonders which numbers the players picked.  
(b) The assistant coach will discover which numbers the players picked.

Crucially, the truth or falsity of (16a) and (16b) does not depend on the available pragmatic strategies; it is determined by the semantics of the embedded question. In essence, this embedding data presents the same prior knowledge problem discussed in the preceding section; however, in this case a pragmatic rescue is not possible.

This shows that the semantic answer to (10a) in context (11) must be stronger than the cumulative proposition expressed in (12b). I take this to confirm the hypothesis that the strongest true answer to (10a) is, in fact, the pair-list answer (10b).

## 4 Pair-list answers via dependent plurals.

**Functional dependencies between plurals.** How, then, is the pair-list reading derived? It is established that questions can receive functional answers, and pair-list answers have been

argued to form a subset of these. (Engdahl, 1980, 1986; Chierchia, 1993) This functional reading arises, in cases like (4) from Section 2.2, from a quantifier taking wide scope over the *wh*-item. This makes examples like (10a) seem slightly puzzling. After all, there is no overt quantifier in this case. But as I will argue, the interaction of the two plurals gives us the quantification we need.

I propose that pair-list readings of QPDs arise from an anaphoric dependency between the plural definite in the question nucleus and the plural *wh*-item, analogous to the model of dependent definites developed by Winter (2000). In Winter’s account, a structurally lower definite contains a covert syntactic variable bound by a distributive operator that accompanies a higher definite. This dependency is represented in (17) by  $T$ , “a contextually salient function from individuals to individuals mapping each soldier to a target”. (Winter, 2000)

17. (a) The soldiers hit the targets. (Winter, 2000)  
 (b)  $\forall x[\text{soldier}(x) \rightarrow \text{hit}(x, T(x))]$

Note that this representation does contain a universal quantifier, which originates from a distributivity operator  $D$  attached to the subject. This operator binds the lower definite, as illustrated in (18).

18. [[the soldiers]  $D_1$ ] [hit [the targets( $x_1$ )]] (Winter, 2000)

Winter deals with declaratives, but there is no reason *prima facie* this structure cannot occur in questions as well. In QPDs, then, the *wh*-marked object behaves as Winter’s dependent definites do, and is bound by the distributivity operator attached to the plural definite subject. *Wh*-movement then yields abstraction over the dependency function, resulting in a Hamblin set like (19a) that varies over functions from individuals to individuals. Pair-list answers are thereby derived in the semantics. In context (11), this Hamblin set contains exactly one true proposition, (19b), which is based on the dependency function in (19c). This unique true member is equivalent to the pair-list answer (10b).

19. (a) Hamblin set:  $\{\forall x[\text{player}(x) \rightarrow \text{picked}(x, F(x))] \mid F : D \rightarrow \{1, 2, 3, 4, 5 \dots\}\}$   
 (b) Strongest true answer:  $\forall x[\text{player}(x) \rightarrow \text{picked}(x, F'(x))]$   
 (c)  $F' = \{\langle \text{Ann}, 1 \rangle, \langle \text{Ben}, 2 \rangle, \langle \text{Chris}, 3 \rangle, \langle \text{Dan}, 4 \rangle, \langle \text{Emma}, 5 \rangle\}$

This produces a structure that is, in fact, quite similar to questions with overt quantifiers. Ultimately, pair-list answers in both cases are derived by placing the universal quantifier and the *wh*-item in the same relative scope.<sup>5</sup>

**Two strategies are needed for QPDs.** Note that the evidence presented here does not do away with the cumulation-and-elaboration strategy! Cumulative answers are obviously still possible, and answerers may certainly give elaborative responses that communicate additional information not encoded in the strongest true Hamblin answer. Examples (6-7) in Section

<sup>5</sup>As this system ultimately depends on the relative scope of, and a binding relationship between, a quantifier and a structurally low DP, it inherits some issues found in other scope/binding analyses of similar phenomena. One of these, as a reviewer has brought to my attention, is the lack of pair-list readings under negative quantification. For example, the following question cannot mean “give me a list containing each student and the corresponding exam which that student cannot fail”:

- i. Which exam(s) must no student(s) fail?

2, which cannot be produced by a dependency parse, clearly demonstrate the need for the cumulation-and-elaboration strategy.

However, this strategy must coexist with another, one that is capable of deriving pair-list Hamblin answers. This is the role of the dependency strategy.<sup>6</sup> Absent context, any arbitrary QPD will have two parses, one deriving a pair-list answer, and the other deriving a cumulative answer (which can then be elaborated upon).

## 5 Confirmed predictions

**Emergence of subject-object asymmetries.** Subject-object asymmetries are, as noted in Section 2, a hallmark of canonical pair-list questions. To reiterate: pair-list questions require the universal quantifier to *c-command* the base position of the *wh*-item. When this is not possible (i.e., when the *wh*-item is the subject and the universal quantifier is contained within the object) pair-list readings become unavailable due to weak crossover effects.

The dependency model, like any other scope/binding analysis, should display these asymmetries. This prediction is borne out—compare (20a), which in context (11) lacks a pair-list reading, with example (10a) from Section (3).

20. (a) Which players picked the numbers?  
 (b) #Ann picked 1, Ben picked 2, Chris picked 3, Dan picked 4, Emma picked 5.

The *wh*-marked object in (10a) may be bound at its base position by the distributive operator accompanying the subject, but no such binding relationship is possible when, as in (20a), the distributive operator does not take scope above the *wh*-item.

The reason that these subject object asymmetries in QPDs have so far gone unnoticed, I ascribe to the many contexts in which a QPD could conceivably support either the cumulation-and-elaboration parse or the dependency parse. The asymmetries that arise from the dependency strategy can often be masked by the more permissive cumulation-and-elaboration strategy.

**Restricted interpretation of numerals.** The cumulation-and-elaboration strategy and the dependency strategy diverge in the possible interpretations they allow for numeral modifiers. In contexts such as (11) where the cumulation-and-elaboration strategy is unavailable, numerals are interpreted as expected under the dependency strategy.

In the cumulation-and-elaboration strategy, numeral modifiers constrain the cardinality of the range of the entire cumulative relation. However, they do not directly constrain the possible elaborative responses that can be generated in the pragmatics. For example, (21a) has a Hamblin set of the form in (21b), where *x* must be a plural individual of women with exactly two atomic parts. Thus, the strongest true answer must map the two men to a group of two women. The answerer is then free to elaborate this cumulative answer into a pair-list response such as (21d), which maintains a range of cardinality two.<sup>7</sup>

<sup>6</sup>See Beck and Sauerland (2000) for arguments that cumulation and dependency strategies coexist in declaratives as well.

<sup>7</sup>In many contexts, questions like (21a) may be ambiguous as to which parse is intended. For example, (21a) could also receive a scope/binding parse, in which case the answer is a pair-list mapping each man to two women:

- i. John likes Sue and Jane, and Bill likes Lisa and Amy.

However, it is not possible to give a cumulative answer, within which each man likes two women, for a total of

21. (a) Which two women do these men like?  
 (b) Hamblin set:  $\{\text{**like}(\text{these men}, x) : x \in \text{*}\{\text{Sue, Jane, Lisa, Amy, ...}\} \wedge x \text{ has two atomic parts}\}$   
 (c) Strongest true answer:  $\text{**like}(\text{these men}, \text{Sue+Jane})$   
 (d) Elaborative response: John likes Sue and Bill likes Jane.

However, the dependency strategy differs in the interpretation it allows. In this case, the numeral constrains the cardinality of *each individual* in the range. Returning to the basketball team example, replacing “which numbers” in (10a) with “which five numbers” only permits the pair-list answer to map each player to a group of five numbers. If evaluated within context (11), (22) is in fact infelicitous, since it contradicts the presupposition that the players chose one number each.

22. *In context (11):*  
 #Which five numbers did the players pick?

Evaluated in a neutral context, we see that this same question can only receive a response like (23b), that maps each player to a group of five numbers.

23. *In a neutral context:*
- (a) Which five numbers did the players pick?  
 (b) Ann picked 1-5, Ben picked 6-10. . .  
 (c) #Ann picked 1, Ben picked 2, Chris picked 3, Dan picked 4, Emma picked 5.

## 6 An unresolved puzzle

Modifying (10a) by making *numbers* singular instead of plural, as shown in (24), causes the pair-list answer to disappear. Setting aside any context, (24) can only be felicitously answered by a proposition that maps the group of players to a single number. As discussed in Section 2.2, this has been interpreted as a signature of cumulation.

24. (a) Which number did the players pick?  
 (b) #Ann picked 1, Ben picked 2, Chris picked 3, Dan picked 4, Emma picked 5.  
 (c) The players picked 1.

Although I cannot yet offer a complete answer, I will make two observations about this fact. First, this phenomenon is not limited to the domain of questions. A parallel constraint appears with demonstratives. In a suitable context, (25a) can be read as expressing a proposition equivalent to the pair-list answer (10b)<sup>8</sup>, while (25b) cannot.

25. (a) The players picked these/those numbers.

---

four women between them:

- ii. #These men like Sue, Jane, Lisa, and Amy.

<sup>8</sup>(25a) could even serve as an answer to the assistant coach’s question in context (11); for example, if the head coach uttered (25a) as she was in the process of handing the assistant coach a written list of the player-number pairs.

(b) #The players picked this/that number.

The second observation I offer is that there appears to be some inter-speaker variation with respect to the judgments expressed in (24). The prevailing view agrees with the literature: that (24c) is a felicitous answer to (24a), while (24b) is not. However, two of the approximately ten native speakers of North American English with whom I have discussed this data find both answers in (24) to be felicitous. That is, for a (presumed) minority of speakers, pair-list answers to (24a) are readily available.

How do we make sense of this data? It's possible that some kind of dynamic plural logic might resolve this puzzle.<sup>9</sup> For example, the inter-speaker contrast with respect to (24b) might be described in the PCDR system of Brasoveanu (2008) as a contrast in whether the singular morphology is interpreted as discourse-level or domain-level atomicity. In the former case, there must exist only one atomic number in the total *picking* relation, giving rise to the infelicity of (24b). In the latter case, there must exist only one atomic number in each mapping within the *picking* relation, in which case (24b) is a felicitous response. At present, however, I cannot offer a full account of these facts.

## 7 Conclusion

Pair-list answers to QPDs may be available even when cumulative answers are contextually precluded. This contradicts the cumulation-only analysis, lending support instead to a model of QPDs wherein two distinct pathways for deriving pair-list responses are available. One is the cumulation-and-elaboration approach of Dayal (1992, 1996) and Krifka (1992). The other, I argue, arises from an anaphoric dependency, akin to the system described by Winter (2000). Both cumulation and dependency strategies are necessary to reach a complete description of questions with plural definites.

## References

- Sigrid Beck and Uli Sauerland. Cumulation is needed: A reply to Winter (2000). *Natural Language Semantics*, 8(4):349–371, 2000.
- Adrian Brasoveanu. Donkey pluralities: plural information states versus non-atomic individuals. *Linguistics and philosophy*, 31(2):129–209, 2008.
- Gennaro Chierchia. Questions with quantifiers. *Natural Language Semantics*, 1(2):181–234, 1993.
- Veneeta Dayal, 1992. See Srivastav (1992).
- Veneeta Dayal. *Locality in WH Quantification: Questions and Relative Clauses in Hindi*. Dordrecht, Kluwer Academic Publishers, 1996.
- Veneeta Dayal. *Questions*. Oxford University Press, Oxford, 2016.
- Elisabet Engdahl. *The syntax and semantics of questions in Swedish*. PhD thesis, University of Massachusetts, Amherst, 1980.

---

<sup>9</sup>Thank you to my reviewers for suggesting this approach.

- Elisabet Engdahl. *Constituent Questions: The Syntax and Semantics of Questions with Special Reference to Swedish*. Reidel, Dordrecht, 1986.
- Jeroen Groenendijk and Martin Stokhof. *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, Universiteit van Amsterdam, 1984.
- Charles L. Hamblin. Questions in montague english. *Foundations of Language*, 10:41–53, 1973.
- Manfred Krifka. *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. PhD thesis, Universität München, 1986. (Published by Wilhelm Finck, Munich, 1989.).
- Manfred Krifka. Definite NPs aren't quantifiers. *Linguistic Inquiry*, 23:156–63, 1992.
- Bradley L. Pritchett. A note on scope interaction with definite plural nps. *Linguistic Inquiry*, 21(4):646–54, 1990.
- Floris Roelofsen and Wataru Uegaki. The distributive ignorance puzzle. In *Proceedings of Sinn und Bedeutung 21*, volume 21, pages 999–1016, 2016.
- Benjamin Spector and Paul Egré. A uniform semantics for embedded interrogatives: an answer, not necessarily the answer. *Synthese*, 192(6):1729–1784, 2015.
- Veneeta Srivastav. Two types of universal terms in questions. In *Proceedings of the North Eastern Linguistics Society (NELS) 22*, Amherst, 1992. GLSA Publications. Reprinted 2003 in Javier Gutiérrez-Rexach (ed.), *Semantics: Critical Concepts in Linguistics*. London/New York: Routledge, 373–85.
- Wolfgang Sternefeld. Reciprocity and cumulative predication. *Natural Language Semantics*, 6(3):303–337, 1998.
- Yoad Winter. Distributivity and dependency. *Natural Language Semantics*, 8:27–69, 2000.

# The Asymmetry of Past Tense \*

Hadil Karawani<sup>1</sup>, Carina Kauf<sup>2</sup>, and Hedde Zeijlstra<sup>3</sup>

<sup>1</sup> Leibniz-ZAS, Berlin & University of Konstanz [hadil.karawani@gmail.com](mailto:hadil.karawani@gmail.com)

<sup>2</sup> Georg-August-Universität Göttingen & Massachusetts Institute of Technology  
[carina.kauf@stud.uni-goettingen.de](mailto:carina.kauf@stud.uni-goettingen.de),

<sup>3</sup> Georg-August-Universität Göttingen [hzeijls@uni-goettingen.de](mailto:hzeijls@uni-goettingen.de)

## Abstract

In this paper, we propose a semantics for (the highest instance) of past tense in a syntactic domain that is essentially modal and not strictly temporal. Given this asymmetry we are able to account for the fact that, once embedded under another modal, past tense morphology can receive a modal interpretation and is not an inherent time shifter. This naturally derives the syntax of counterfactual *if* - and *wish* clauses. Overgeneration of modal readings in other modal contexts is ruled out by means of pragmatic competition with present tense morphology.

## 1 Background

One of the central questions in the studies of counterfactual conditionals (and similar constructions, such as *wish*-clauses, (cf. [Iat00])) concerns the usage of past tense morphology. In many languages, including English, past tense morphology in such constructions does not give rise to an anteriority effect, but rather conveys that the predicate holds in a different world than the actual one. Why is it that in (1) the inclusion of past tense morphology conveys that the antecedent does not hold in the actual world?

- (1) If Mary didn't speak English, she would be helpless.

Two approaches have been formulated to account for the semantic contribution of such past tense morphology: (i) a fake tense approach where past tense morphology expresses exclusion from either the time of utterance or the actual world (see [Pal01, Iat00, Sch05, Sch14, Kar14], a.o.); and (ii) a real tense approach where past tense is inherently temporal, but the locus of tense can be shifted outside the conditional. In the latter, the past tense morpheme refers to a point back in time at which the antecedent was still a possibility: the conditional means that in the worlds at which the antecedent holds (and which are closest to the actual world), the consequent holds (cf. [Dud83, Dud84, Ipp03, Arr09, GVS09]).

The real tense approach may provide the null hypothesis in the sense that it keeps past tense morphology purely temporal, but is challenged by (i) the fact that the locus of past tense interpretation does not match its surface position—*if*-clauses form an island for movement (see [Rom14])—and (ii) the fact that not every counterfactual conditional denotes a possibility that was available at an earlier point in time—in other words, some conditionals may be counterfactual in earlier times, or throughout time. For instance, for counteridenticals like (2), arguably, at no time in history was its antecedent inferred not to be false.

---

\*We thank Sabine Iatridou, and Manfred Krifka, as well as audiences at the semantics reading group at ZAS, the philosophy colloquium at Konstanz, and the English linguistics reading group in Goettingen for generous and helpful feedback. All remaining mistakes are, of course, our own. We also thank Floris Roelofsen, Julian Schlöder and the anonymous reviewers of the Amsterdam Colloquium.



- (2) If I were you, I would be thrilled.

At the same time, the fake tense (exclusion) approach also suffers from non-trivial problems. For one, there is a clear asymmetry between the availability of the temporal and the modal reading of the past tense morpheme: A sentence like (3) can only mean that John’s sleeping took place, not that he may sleep now. What blocks a modal reading in such cases?

- (3) John slept.

[Iat00] argues that when access to alternative worlds is given, the modal interpretation becomes available, but does not formalize this requirement; the semantics of her exclusion feature, which past tense morphology denotes, does not have an in-built preference for a temporal interpretation. Moreover, not in every modal context does past tense morphology trigger a modal interpretation. Under modal predicates or adverbs, or under intentional predicates, past tense morphology is interpreted in a purely temporal way, as is the case in the examples below.

- (4) a. It is possible that John was happy.  
 b. Probably, Mary was leaving.  
 c. Mary believes John was leaving.

Hence, the question remains open as to what blocks a modal interpretation for plain past tense constructions. We aim to solve this problem, and present a viable alternative to the temporal approach, by lexically encoding this asymmetry into the semantics of past tense morphology. However, in doing so, we explicitly make the past tense morpheme a modal indexical and not a temporal one. For us, a past tense morpheme makes reference to world-time pairs involving the the actual world, but not the time of utterance.

Before continuing, let us point out one potential caveat. In some languages, like Greek, Italian, and Zulu, a.o., imperfective aspect accompanies past tense in constructions akin to the example in (1). While some authors (cf. [Ipp02, Fer14]) argue that imperfective aspectual semantics plays a role in the composition of hypothetical and/or counterfactual readings, we follow [HK12] in their analysis, which takes imperfective morphology to fulfil a morpho-syntactic role, rather than a semantic one: being default, imperfective aspect ‘comes along for the ride’ to fulfil a syntactic requirement for aspect. Hence, we will, therefore, not further discuss the role of (fake) aspect here.

## 2 Our Proposal

### 2.1 Assumptions

Our proposal is based on the following assumptions. First, we note that past tense makes reference to a local evaluation time, not to the time of utterance. Evidence for this comes from examples like (5), where the time of hiding takes place prior to the time of thinking, not prior to the time of utterance.

- (5) Alan will think that everybody hid.

This means that past tense morphology is inherently relative to a local evaluation time. However, this also entails that the locus evaluation time, in cases where the past tense morpheme is the highest instantiation of tense morphology, must be set to a default time-of-utterance index,  $t_u$ , that enters the derivation as a last resort, as in (6).

(6) Mary left.

$$\begin{array}{c}
 \lambda w. \exists t [t < t_u. \text{Leave}(\text{Mary}) \text{ at } t] \\
 \swarrow \quad \searrow \\
 \text{PAST Leave}(\text{Mary}) \quad t_u \\
 \lambda t^*. \lambda w. \exists t [t < t^*. \text{Leave}(\text{Mary}) \text{ at } t]
 \end{array}$$

In the same vein, an actual world index  $w_0$  is applied to every proposition that is not further modally anchored. Uttering (6) will ultimately receive an interpretation about the actual world, so an additional  $w_0$  must be added to the derivation as a last resort as well, as in (7):

(7) Mary left.

$$\begin{array}{c}
 \exists t [t < t_u. \text{Leave}(\text{Mary}) \text{ at } t \text{ in } w_0] \\
 \swarrow \quad \searrow \\
 \lambda w. \exists t [t < t_u. \text{Leave}(\text{Mary}) \text{ at } t \text{ in } w] \quad w_0 \\
 \swarrow \quad \searrow \\
 \text{PAST Leave}(\text{Mary}) \quad t_u \\
 \lambda t^*. \lambda w. \exists t [t < t^*. \text{Leave}(\text{Mary}) \text{ at } t]
 \end{array}$$

With this in mind, we take past tense to have the following semantics, which we present as an operation (for the sake of exposition), but which can easily be recast in presuppositional terms (see [Kar19]):

$$(8) \quad \llbracket \text{PAST} \rrbracket = \lambda P. \lambda t^*. \lambda w. \exists t [ \langle w, t \rangle \neq \langle w_0, t^* \rangle \ \& \ P \text{ holds at } t \text{ in } w ]$$

Under this semantics for past tense morphology, the asymmetry between the two usages follows. In plain past tense constructions, the temporal interpretation immediately follows: Since the proposition is applied to  $t_u$  and  $w_0$  at the final stage of the derivation, both world-time pairs will include  $w_0$ , but only one of them  $t_u$ , and therefore the two tense variables must receive a distinct interpretation; taking the future to be a modal and not a tense [Iat00, Cop09, GM17], only a past tense reading thus emerges.

$$\begin{aligned}
 (9) \quad & [ \text{PAST} (\text{John sleep})(t_u)(w_0) ] \\
 & = \lambda P. \lambda t^*. \lambda w. \exists t [ \langle w, t \rangle \neq \langle w_0, t^* \rangle \ \& \ P \text{ holds at } t \text{ in } w ] (\text{John sleep})(t_u)(w_0) \\
 & = \lambda t^*. \lambda w. \exists t [ \langle w, t \rangle \neq \langle w_0, t^* \rangle \ \& \ \text{John sleep holds at } t \text{ in } w ] (t_u)(w_0) \\
 & = \lambda w. \exists t [ \langle w, t \rangle \neq \langle w_0, t_u \rangle \ \& \ \text{John sleep holds at } t \text{ in } w ] (w_0) \\
 & = \exists t [ \langle w_0, t \rangle \neq \langle w_0, t_u \rangle \ \& \ \text{John sleep holds at } t \text{ in } w_0 ] \\
 & = \exists t [ t \neq t_u \ \& \ \text{John sleep holds at } t \text{ in } w_0 ]
 \end{aligned}$$

Only if the entire proposition is embedded under a modal quantifier, as is the case in a conditional, the world argument is not set to  $w_0$ , and a modal interpretation may arise. Assuming an (oversimplified) semantics for the conditional  $\lambda p. \lambda q. \forall w [ p(w) \longrightarrow q(w) ]$  ([Sta68, Vel86]), the interpretation of the antecedent of the conditional *If John slept, ...* is as in (10).

$$\begin{aligned}
 (10) \quad & [ \text{IF} (\text{PAST} (\text{John sleep})(t_u)) ] \\
 & = [ \lambda p. \lambda q. \forall w [ p(w) \longrightarrow q(w) ] ] (\lambda P. \lambda t^*. \lambda w. \exists t [ \langle w, t \rangle \neq \langle w_0, t^* \rangle \ \& \ P \text{ holds at } t \text{ in } w ] (\text{John sleep})(t_u)) \\
 & = \lambda q. \forall w [ \exists t [ \langle w, t \rangle \neq \langle w_0, t_u \rangle \ \& \ \text{John sleep holds at } t \text{ in } w ] \longrightarrow q(w) ]
 \end{aligned}$$

What past tense morphology does here is make a non-actual veridical contribution. It says of any predicate that it holds at a world-time pair that is different from the local evaluation time and the actual world. The fact that the past tense morpheme is ultimately a modal indexical and not a temporal indexical derives the desired asymmetry.

The idea that past tense is inherently modal might lead to over generalization — so what blocks the modal reading from arising in every modal, or intentional, environment, for that matter?

## 2.2 Preventing overgeneration

Our proposal states that once a past tense is modally embedded, a non-temporal reading may emerge. The idea that past tense is inherently modal might therefore lead to overgeneralization. To see this, take for instance the example in (11) below.

(11) It is possible that Mary was leaving.

Under our proposal, this sentence is assigned the following meaning: there is a world  $w'$  accessible from the actual world  $w_0$ , and there is a time  $t$ , such that  $\langle w', t \rangle \neq \langle w_0, t_u \rangle$ , and Mary's leaving takes place at  $t$ .

(12) [POSSIBLE (PAST(Mary leave)) ( $t_u$ )( $w_0$ )]  
 $= \exists w' [ \text{ACC}(w', w_0) \ \& \ \exists t [ \langle w', t \rangle \neq \langle w_0, t_u \rangle \ \& \ \text{Leave}(\text{Mary}) \text{ holds at } t \text{ in } w' ] ]$

The meaning in (12) is compatible with a scenario in which it is possible for Mary to be leaving *now*, and where we know that Mary was not leaving *before*. Clearly, this is not the meaning that (11) ought to have: it is intended to read as past and to not be compatible with a present reading. What is the mechanism that blocks the non-temporal reading under a modal when only a temporal reading is intended?

A similar challenge arises for sentence like (13), which is predicted, under our proposal, to have the following meaning in (14).

(13) John believes Mary was leaving.

(14)  $\forall w [ w \in \text{BEL}(\text{John}, t_u) \longrightarrow [ \exists t. \langle w, t \rangle \neq \langle w_0, t_u \rangle \ \& \ \text{Leave}(\text{Mary}) \text{ at } t \text{ in } w ] ]$

According to (14), in all worlds compatible with John's beliefs at the time of utterance  $t_u$ , there is a time  $t$ , such that  $\langle w, t \rangle \neq \langle w_0, t_u \rangle$ , and Mary's leaving takes place at  $t$ . That means that Mary, in principle, could be leaving at the time of the believing, at least in every world where John's beliefs do not correspond with the actual world. These readings are thus too weak.

However, rather than strengthening our original proposal, we conjecture that these readings result from pragmatic competition with the present tense. This competition is most likely an instance of Maximize Presupposition (originally postulated by [Hei91]) — provided that the contributions that tense morphology make are actually presuppositional in nature. To see this, take (15a)–(15b).

(15) a. It is possible that Mary is leaving.  
 b. John believes Mary is leaving.

As shown in (16a) and (16b) below, the meanings of (15a)–(15b) are stronger than (12)–(13):

- (16) a.  $\exists w'. \text{ACC}(w', w_0) \ \& \ \text{Mary leave holds at } t_u \text{ in } w'$   
 b.  $\forall w. w \in \text{BEL}(\text{John}, t_u) \longrightarrow [\text{Leave}(\text{Mary}) \text{ at } t_u \text{ in } w]$

Hence, by uttering (11)–(12), instead of (15a)–(15b), it can be inferred that the speaker does not believe (15a)–(15b) to be true. As (15a)–(15b) only allow a simultaneous reading, i.e., a reading where Mary’s leaving takes place in the local context time, (11)–(12) can only be uttered with a temporally shifted reading. The modal contribution that past tense morphology makes in comparison to present tense morphology gets annihilated by the modal embedding under competition with the present tense.

The reason for us to opt for this pragmatic competition instead of altering our denotation of past tense morphology is that the simultaneous reading comes about in exactly those environments where a present tense alternative is absent. This is the case in counterfactual *if*- and *wish*-clauses. To see this, consider (17) and (18):

- (17) a. If Mary was leaving, John would be happy.  
 b. \*If Mary is leaving, John would be happy.
- (18) a. John wishes Mary was leaving.  
 b. \*John wishes Mary is leaving.

Hence, the restriction, according to which non-temporal interpretations of past tense morphology are preserved to those grammatical contexts that are known to give rise to counterfactuality inferences, is warranted. Naturally, the question as to why (17b) and (18b) are ungrammatical arises, though. We do not have a concrete answer to this question, but we presume that the natural answer here seems to be that the lexical semantics of *wish* does not allow it to reach within the set containing the actual world and the time of utterance.

### 3 Sequence of Tense

One might wonder to what extent our analysis of fake past applies to Sequence of Tense (SoT) contexts as the temporal contribution of past tense morphology in these environments appears to be redundant — just like the temporal contribution in fake past is also absent. This depends on how you treat SoT; for us, what counts is the highest instance of past tense in the chain.

That our semantics for the past tense only necessarily applies to the highest instance of past tense in a sentence means that for multiple embedded past tenses, lower past tenses are either only phonologically marked for past tense (cf. [Kra98, Abu97, Sto95], a.o.), or agreement markers with respect to a higher covert tense operator (cf. [KZ18]), which should carry the semantics in (8). Our proposal is thus fully compatible with the existence of Sequence-of-Tense readings.

As a case study, consider the approach by [KZ18]. They argue that the perceived ambiguity of past-under-past embeddings is not the result of ambiguity but rather of lexical underspecification, which they cast in terms of two ingredients. Every past tense morpheme for them denotes a relative non-future (*RNF*) with respect to its local evaluation time; and every past tense morpheme (*-ed*) is assumed to be equipped with a past tense feature that needs to be checked by a past tense operator (*Op-PAST*) higher up in the structure. Given the fact that a past tense operator can check the features of all past tense morphemes in its syntactic domain via multiple agree, no second operator is allowed if all past tense morphemes are part of the same syntactic domain (cf. [Zei12]). The logical form of *John said Mary was ill* is then as follows:

- (19) John said Mary was ill.
- $[Op-PAST_{[iPAST]} [John [say-ed_{[uPAST]} [Mary [be-ed_{[uPAST]} ill.]]]]]$   
 $\exists t' < t_u \quad \exists t^2 \leq t' \quad \exists t^3 \leq t^2$
  - $\exists t' < t_u \ \& \ [\exists t^2 \leq t' \ \& \ say(John, t^2, [\exists t^3 \leq t^2 \ \& \ be-ill(Mary, t^3)])]$
  - John's saying is strictly before the utterance time  $t_u$  and Mary's being ill starts out no later than at the time of John's saying.*

Combining this approach with the generalized past-tense meaning proposed in this paper can be achieved straightforwardly, by replacing the past tense operator with PAST as proposed in (8), but keeping the not-later-than meaning of past tense morphology as well as their feature checking relation the same.

- (20)  $[[PAST]] = \lambda P. \lambda t^*. \lambda w. \exists t [ < w, t > \neq < w_0, t^* > \ \& \ P \text{ holds at } t \text{ in } w ]$

- (21) John said Mary was ill.
- $PAST(John \ say-ed \ Mary \ be-ed \ ill)(t_u)(w_0)$
  - $\lambda P. \lambda t^*. \lambda w. \exists t [ < w, t > \neq < w_0, t^* > \ \& \ P \text{ holds at } t \text{ in } w ](John \ say-ed \ Mary \ be-ed \ ill)(t_u)(w_0)$
  - $\exists t [ < w_0, t > \neq < w_0, t_u > \ \& \ [\exists t^2 \leq t \ \& \ say(John, t^2, [\exists t^3 \leq t^2 \ \& \ be-ill(Mary, t^3)])]$
  - There exists a time  $t$  such that the ordered pair  $< w_0, t >$  is not the same as  $< w_0, t_u >$  and John said no later than this time  $t$  that Mary was ill at a time no later than his saying time.*

Even though these semantics look quite complex, it can be easily confirmed that the worlds in which this statement is true are the same as those in (19b).

## 4 Conclusions

In this paper, we have proposed a semantics for (the highest instance) of past tense in a syntactic domain that is essentially modal and not strictly temporal. Given this asymmetry we are able to account for the fact that, once embedded under another modal, past tense morphology can receive a modal interpretation and is not an inherent time shifter. This naturally derives the syntax of counterfactual *if*- and *wish* clauses. Overgeneration of modal readings in other modal contexts is ruled out by means of pragmatic competition with present tense morphology.

## References

- [Abu97] D. Abusch. Sequence of tense and temporal de re. *Linguistics and Philosophy*, 20(1):1–50, 1997.
- [Arr09] Ana Arregui. On similarity in counterfactuals. *Linguistics and Philosophy*, 32(3):245–278, 2009.
- [Cop09] Bridget Copley. *The semantics of the future*. Routledge, 2009.
- [Dud83] Victor H Dudman. Tense and time in english verb clusters of the primary pattern. *Australian Journal of Linguistics*, 3(1):25–44, 1983.
- [Dud84] Victor H Dudman. Parsing 'if'-sentences. *Analysis*, 44(4):145–153, 1984.

- [Fer14] Marcelo Ferreira. Displaced aspect in counterfactuals: towards a more unified theory of imperfectivity. *The art and craft of semantics: A festschrift for Irene Heim*, ed. L Crnic and U. Sauerland, pages 147–164, 2014.
- [GM17] Anastasia Giannakidou and Alda Mari. Epistemic future and epistemic must: nonveridicality, evidence, and partial knowledge. *Mood, Aspect, Modality Revisited: New Answers to Old Questions*, ed. J. Blaszczak, A. Giannakidou, D. Klimek-Jankowska, and K. Migdalski, pages 75–117, 2017.
- [GVS09] Atle Grønn and Arnim Von Stechow. Temporal interpretation and organization of subjunctive conditionals. *Ms. U. Oslo*, 2009.
- [Hei91] Irene Heim. Artikel und definitheit (articles and definiteness). *Semantics: An International Handbook of Contemporary Research*, ed. A. von Stechow and D. Wunderlich, pages 487–535, 1991.
- [HK12] Claire Halpert and Hadil Karawani. Aspect in counterfactuals from a (rabic) to z (ulu). In *Proceedings of WCCFL*, volume 29, pages 99–107, 2012.
- [Iat00] Sabine Iatridou. The grammatical ingredients of counterfactuality. *Linguistic inquiry*, 31(2):231–270, 2000.
- [Ipp02] Michela Ippolito. On the semantic composition of subjunctive conditionals. *ms., MIT/Tübingen University*, 2002.
- [Ipp03] Michela Ippolito. Presuppositions and implicatures in counterfactuals. *Natural language semantics*, 11(2):145–186, 2003.
- [Kar14] Hadil Karawani. *The real, the fake, and the fake fake: In counterfactual conditionals, crosslinguistically*. LOT Publications, 2014.
- [Kar19] Hadil Karawani. The past is rewritten. In *Making and meaning: Snippets for Uli*, ed. A. Nicolae, P. Elliott & Y. Sudo, pages 52–53, 2019.
- [Kra98] A. Kratzer. More structural analogies between pronouns and tenses. In *Proceedings of SALT VIII. (Cornell Working Papers in Linguistics)*, volume 8, pages 92–110, 1998.
- [KZ18] Carina Kauf and Hedde Zeijlstra. Towards a new explanation of sequence of tense. In *Proceedings of Semantics and Linguistic Theory*, volume 28, pages 59–77, 2018.
- [Pal01] Frank Robert Palmer. *Mood and modality*. Cambridge University Press, 2001.
- [Rom14] Maribel Romero. ‘fake tense’ in counterfactuals: A temporal remoteness approach. *The art and craft of semantics: A festschrift for Irene Heim*, ed. L Crnic and U. Sauerland, 2:47–63, 2014.
- [Sch05] Philippe Schlenker. The lazy frenchman’s approach to the subjunctive. *Proceedings of Going Romance XVII*, pages 269–309, 2005.
- [Sch14] Katrin Schulz. Fake tense in conditional sentences: a modal approach. *Natural language semantics*, 22(2):117–144, 2014.
- [Sta68] Robert C Stalnaker. A theory of conditionals. In *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, ed. N. Rescher, pages 98–112. Springer, 1968.
- [Sto95] Tim Stowell. What is the meaning of the present and past tenses. *Temporal reference, aspect, and actionality. Vol. I: Semantic and syntactic perspectives*, ed. P.-M. Bertinetto, V. Bianchi, and M. Squartini, pages 381–396, 1995.
- [Vel86] Frank Veltman. Data semantics and the pragmatics of indicative conditionals. *On conditionals*, ed. E. C. Traugott, A. ter Meulen, J. S. Reilly, and C. A. Ferguson, pages 147–168, 1986.
- [Zei12] Hedde Zeijlstra. There is only one way to agree. *The linguistic review*, 29(3):491–539, 2012.

# Dynamic Unioning Plural Logic

Ezra Keshet

University of Michigan

## Abstract

Popular plural logics based on van den Berg (1996) require complex machinery: a *structured inclusion* relation to properly maintain dependencies (e.g., ‘ $\sqsubseteq$ ’ in Brasoveanu (2013)), and *maximization* and *distributivity* operators to correctly analyze quantification, among others. Systems following van den Berg (1996) introduce even further machinery (see Nouwen, 2003; Brasoveanu, 2008). Here instead I propose the new Dynamic Unioning Plural Logic (DUPL) a simpler system that replicates the van den Berg system (and addresses some of its empirical issues) with only one new operator and one new term type.

## 1 Defining Dynamic Unioning Plural Logic

Following van den Berg (1996), I assume an interpretation relation  $\llbracket \cdot \rrbracket :: \langle \mathbf{G}, \mathbf{G} \rangle \rightarrow \mathbf{t}$  over pairs of states  $\langle G, H \rangle$ , each consisting of sets of assignments.  $\llbracket \cdot \rrbracket$  is interpreted relative to a model  $\langle D, I \rangle$  comprising a domain of individuals  $D :: \mathbf{e}$  and a predicate interpretation function  $I :: \mathbf{p}_n \rightarrow \{\mathbf{e}\}^n$ . The domain contains only singular individuals, but the predicates expect plural individuals (sets of individuals) as their arguments, where singleton sets represent singular arguments. Specifically,  $I$  maps each predicate of arity  $n$  to an  $n$ -tuple of plural individuals, those intuitively making the predicate true. The assignments  $g :: \mathbf{v} \rightarrow \mathbf{e} \mid \star$  in each state  $G :: \mathbf{G}$  map variables to  $DU\{\star\}$ , individuals in the model plus the dummy value  $\star$ , representing missing values in the assignment. Finally, we will use two varieties of terms (type  $\tau$ ) in predicate literals: plain variables like “ $x$ ” and augmented variables preceded by a plus-sign, like “ $+x$ .” See Figure 1 for a summary of these types and their conventional meta-language notations.

Type	Meta-language notation	Description
$\mathbf{t}$	$\mathbf{T}, \mathbf{F}$	Truth values
$\mathbf{e}$	$A_1 \dots A_n, \dots, Z_1 \dots Z_n$	Individuals in the model
$\star$	$\star$	Dummy value
$\mathbf{v}$	$\nu, a \dots z, a' \dots z', \dots$	Variables
$\tau$	$\tau$	Terms: “ $\nu$ ” or “ $+\nu$ ”
$\mathbf{g} = \mathbf{v} \rightarrow \mathbf{e} \mid \star$	$g, h$	Assignments
$\mathbf{G} = \{\mathbf{g}\}$	$G, H, K, L$	Sets of assignments
$\mathbf{f}$	$\phi, \psi$	Formulas
$\mathbf{p}_n$	$P$	Predicates of arity $n$

Figure 1: Domains in DUPL

Some useful abbreviations are shown in Figure 2. Namely, “ $g[x]h$ ” indicates that assignments  $g$  and  $h$  differ at most in their value for variable  $x$ . Relatedly, “ $G[x]H$ ” indicates that  $H$  is a superset of  $G$  where each assignment in  $H$  is related to an assignment in  $G$  via  $[x]$ . Next, “ $G_{x,+y}$ ” (for example) represents that subset of state  $G$  where assignments all have defined values for the variables  $x$  and  $y$  in the subscript (ignoring the plus signs of terms). “ $G(x)$ ” represents a plural individual drawn from all the values in component assignments  $g \in G$ . This

is the only source of plurality, as assignments map variables to singular individuals. Sequences of terms “ $(\tau_1, \dots, \tau_n)$ ” may be abbreviated as “ $\vec{\tau}_n$ .” Finally,  $\star$  is the “all-star” assignment, mapping any variable to  $\star$ .

Notation	:: Type	Expansion
$g[x]h$	:: <b>t</b>	$g \setminus \{x, g(x)\} = h \setminus \{x, h(x)\}$
$G[x]H$	:: <b>t</b>	$G \subseteq H \subseteq \{h : \exists g \in G (g[x]h)\}$
$G_{(+x, (+y), \dots)}$	:: <b>G</b>	$\{g \in G : g(x) \neq \star \ \& \ g(y) \neq \star \ \& \ \dots\}$
$G(x)$	:: <b>{e}</b>	$\{g(x) : g \in G_x\}$
$\vec{\tau}_n$	:: <b><math>\tau^n</math></b>	$(\tau_1, \dots, \tau_n)$
$\star$	:: <b>g</b>	$\lambda x. \star$

Figure 2: Abbreviations

The full definition of DUPL, summarized in Figure 3, defines the following kinds of formulas:

- Random assignment “ $[x]$ ” relates an input state  $G$  to any superset  $H$  of  $G$  whose assignments differ from assignments in  $G$  only in their values for variable  $x$ . Thus, random assignment introduces states with new plural values for  $x$  and new dependencies between these values and existing values in  $G$ . This (along with the definitions of conjunction “ $\&$ ” and the union operator “ $\cup_x(\dots)$ ” below) ensures that states always get larger as formulas progress.
- Predicate literals “ $P \vec{\tau}_n$ ” are tests over the values of their argument terms  $\tau_1 \dots \tau_n$ . For plain variable terms  $x$ , this is the set of values for  $x$  just in the substate where all the argument variables are defined (i.e.,  $\neq \star$ ). For instance,  $P(x, y)$  evaluated in a state  $G$  will only examine  $G_{x,y}$ . This allows testing a predicate literal to ignore values defined outside the current context. When global testing is required, augmented terms  $+x$  are used, since their interpretation comprises all values for  $x$  throughout the current state.
- Dynamic conjunction is defined, as usual, via relation composition, and negation is a test ensuring that  $\llbracket \ ]$  does not relate the input  $G$  to any output.
- The **union** operator “ $\cup_x \phi$ ” is the only new operator in DUPL. It relates an input  $G$  to the union of certain outputs for  $G$  in  $\llbracket [x]; \phi \rrbracket$ , namely those outputs  $K$  that only differ from  $G$  in assignments where  $x$  is defined. The expression  $K \setminus K_x$  returns the set of assignments

$G \llbracket [x] \rrbracket H$ iff $G[x]H$ $G \llbracket P \vec{\tau}_n \rrbracket H$ iff $G=H \ \& \ \langle G \vec{\tau}_n(\tau_1), \dots, G \vec{\tau}_n(\tau_n) \rangle \in I(P)$ $G \llbracket \phi; \psi \rrbracket H$ iff $G(\llbracket \phi \rrbracket \circ \llbracket \psi \rrbracket)H$ $G \llbracket \neg \phi \rrbracket H$ iff $G=H \ \& \ G \notin \text{dom}(\llbracket \phi \rrbracket)$ $G \llbracket \cup_x \phi \rrbracket H$ iff $H = \bigcup \{K : G \llbracket [x]; \phi \rrbracket K \ \& \ K \setminus K_x = G\} \ \& \ H \neq \emptyset$	$G \vec{\tau}_n(x) = G \vec{\tau}_n(x)$ $G \vec{\tau}_n(+x) = G(x)$
---	--

Figure 4: Terms

Figure 3: The DUPL System



in  $K$  where  $x=\star$  (i.e.,  $x$  is undefined). This set must be identical to the input  $G$ . The overall output state  $H$  will always be a superset of  $G$ , since each  $K$  will also be a superset of  $G$ . This definition also effectively requires  $x$  to be undefined before  $\cup_x$ ; otherwise, each  $K \setminus K_x$  cannot be equal to  $G$  (if  $G$  had values for  $x$  they would be removed by “ $\setminus K_x$ ”).

This condition on the states  $K$  essentially maintains the input state as “read-only”: the output state  $H$  may add assignments, but never remove assignments from  $G$  or alter members of  $G$  without also introducing a value for  $x$ . In practice, this feature serves to limit the operation of random assignment. For instance, inside a union clause  $\cup_x(\dots[y];\dots)$ , a clause “[ $y$ ]” will add values for  $y$  only to those assignments that also have a value for  $x$ .

Finally, we will call a formula  $\phi$  true in DUPL iff  $\{\star\} \in \text{dom}(\llbracket\phi\rrbracket)$ , where  $\star$ , again, is the “all-star” assignment that maps every variable to  $\star$ .

## 2 Using DUPL

### 2.1 Preliminaries

Without “ $\cup_x$ ,” and restricting ourselves to singleton values, DUPL roughly replicates Dynamic Predicate Logic (DPL, [Groenendijk and Stokhof, 1991](#)), since each state contains a single contentful assignment, as shown in (1):<sup>1</sup>

- (1) A woman entered. She sat.  $\rightsquigarrow [w]; 1(w); \text{woman}(w); \text{entered}(w); \text{sat}(w)$   
 $\forall h : \{\star\} \llbracket(1)\rrbracket \{h\}$  iff  $h(w)$  is a woman who entered and sat, and  $\forall \nu \neq w (h(\nu) = \star)$ .

The clause “[ $w$ ]” introduces all states whose assignments differ from  $\star$  at most in their values for  $w$ . For instance,  $\{\star\}$ ,  $\{[w \rightarrow A_1]\}$ ,  $\{\star, [w \rightarrow A_1]\}$ ,  $\{[w \rightarrow W_7]\}$ ,  $\{[w \rightarrow A_1], [w \rightarrow W_7]\}$ , and  $\{\star, [w \rightarrow A_1], [w \rightarrow W_7]\}$  are all possible output states. The predicate “1” is meant to be true of only singleton values, and therefore “1( $w$ )” eliminates those states  $G$  with non-singleton values for  $G(w)$ . The remainder of the formula eliminates those states  $G$  in which  $G(w)$  is not a woman who entered and sat. This still may leave several states, each of whose singular value for  $G(w)$  satisfies these requirements.

Such variation among states in DUPL serves roughly the same purpose as in DPL: it models indeterminacy about the value of variables (sometimes explained as listener uncertainty about speaker reference). In order to introduce plurality, though, we will need variation within a DUPL state, among assignments, as shown in (2):

- (2) Three women entered. They sat (together).  $\rightsquigarrow [w]; 3(w); \text{women}(w); \text{entered}(w); \text{sat}(w)$   
 $\forall H : \{\star\} \llbracket(2)\rrbracket H$  iff  $H(w)$  is 3 women who entered and sat, and  $\forall \nu \neq w (H(\nu) = \emptyset)$ .

Here, the predicate literal  $3(w)$  allows only states  $G$  where  $|G(w)| = 3$ . Assuming the predicates *women*, *entered*, and *sat* all apply to (non-singleton) sets of individuals, the output states for  $\{\star\}$  generated by (2) will all contain three women who entered and sat together in the model. Depending on the definitions for the predicates, these sets may even overlap: for instance, if five women entered and sat together, the outputs of this formula will contain all size-three subsets of this group of five.

<sup>1</sup>The output states  $H$  may also contain  $\star$ , but this will not affect the value for  $H(w)$  since  $\star(w) = \star$ .

## 2.2 Introducing Unions

The “ $\cup_x$ ” operator converts such state-external variation into state-internal variation, achieving a sort of maximization. For instance, a formula including “ $\cup_w(\dots)$ ” can represent a maximal version of plural “some,” as shown in (3). The “ $\cup_w$ ” flattens several states exhibiting indeterminacy (i.e., which women are  $w$ ?) into one state exhibiting plurality (i.e.,  $G(w)$  represents all women who entered).

- (3) Some women entered. They sat together.  $\rightsquigarrow \cup_w(\text{women}(w); \text{entered}(w)); \text{sat}(w)$   
 $\{\star\} \llbracket (3) \rrbracket H$  for the  $H$  s.t.  $\star \in H$  and  $H(w)$  comprises all women who entered,  $H(w)$  all sat together, and  $\forall \nu \neq w (H(\nu) = \emptyset)$ .

Note that the formula “ $[w]; \text{women}(w); \text{entered}(w)$ ” alone always relates its input  $G$  to outputs containing  $G$ . Therefore, the definition of  $\cup_w$  requires that the input  $\{\star\}$  be a subset of the output state in (3), since each for each  $K$  being unioned inside  $\cup_w$ ,  $K$  will always be a superset of its input. Since each  $K$  must therefore contain  $\star$ , the union of all such  $K$  will also include  $\star$  and the output of the union clause will be a superset of its input. In fact, the outputs of a union clause  $\cup_x$  will always be supersets of their inputs. This is reflected in Fact 1:

**Fact 1.** For any  $G$  and  $H$  such that  $G \llbracket \cup_x(\dots) \rrbracket H$ ,  $G \subseteq H$ .

This fact prevents assignments in the input from being lost in the output of “ $\cup_x(\dots)$ ”.

Building on (3), we can introduce a formula for a maximal reading of the indefinite “three women,” as shown in (4). The clause “ $3(w)$ ” appears outside the union, and therefore it simply acts as a test on states. Specifically, it will only be true in models where three women total entered, since  $w$  has already been maximized to all the women in the model who entered.

- (4) Three women entered. They sat together.  $\rightsquigarrow \cup_w(\text{women}(w); \text{entered}(w)); 3(w); \text{sat}(w)$

## 2.3 Collective Quantifiers

Generalized quantifiers use the union operator once for their restrictor and once for their nuclear scope. This process allows later reference to a restrictor set and a nuclear scope / reference set. For instance, in outputs  $G$  of the formula in (5),  $G(s)$  will comprise all the students, and  $G(s')$  all the students who gathered on the quad. As shown in (6), such plural values may be referenced by future pronouns.

Fact 1 ensures that the embedded union clause of (5) does not erase any of the values for  $s$ , guaranteeing that  $G(s)$  will still hold all the students by the end of the formula. The subclause “ $s'=s$ ” requires that  $K(s') \subseteq K(s)$  in each of its output states  $K$ . Recall that  $[s']$  can introduce any size plural value for  $s'$  in its output state. Two-variable literals<sup>2</sup> like  $s'=s$  only check their truth in the substate where both variables are defined, though. Thus, for the output  $H$  of (5), the values for  $s$  where  $s'$  is defined, namely  $H_{s,s'}(s)$ , must equal the values for  $s'$ , namely  $H_{s,s'}(s')$ . This only requires  $H(s')$  to be a subset of  $H(s)$ , though, since there can be other values for  $s$ , in assignments where  $s'$  is undefined.

Finally, *MOST* represents an appropriate predicate over sets. Notice that the literal  $MOST(+s, +s')$  is the first we have seen to use the augmented terms  $+x$ . This type of term is necessary here, because we want to compare all values for  $s$ , not just those where  $s'$  is also defined. An illustration of states used to calculate this sentence is shown in Figure 5, assuming that  $S_1 \dots S_4$  are the students in the model and  $S_1 \dots S_3$  gathered on the quad. The values for

<sup>2</sup>For simplicity, I will assume that each model provides an appropriate equality predicate.

$s$  and  $s'$  do not need to line up in any given assignment, as long as they match up correctly overall.

- (5) Most students gathered on the quad.  
 $\rightsquigarrow \cup_s(students(s); \cup_{s'}(s'=s; gathered(s'))); MOST(+s, +s')$
- (6) They waved signs.  $\rightsquigarrow waved-signs(s')$

$s$	$s'$
*	*
$S_1$	*
$S_2$	*
$S_3$	*
$S_4$	*
$S_1$	$S_2$
$S_2$	$S_1$
$S_3$	$S_3$

$s$	$s'$
*	*
$S_1$	*
$S_2$	*
$S_3$	*
$S_4$	*
$S_1$	$S_1$
$S_2$	$S_2$
$S_2$	$S_3$

Figure 5: Two possible outputs of (5), ignoring unused variables

One more feature of the nested union operators in (5) is that there are no assignments  $g$  in any output state such that  $g(s)=*$  but  $g(s')\neq*$ . Such an assignment might be part of an output of the inner union clause, since  $\cup_{s'}(\dots)$  does not remove outputs with defined values for  $s'$ . However, the outer union clause will not let such assignments pass, since  $\cup_s(\dots)$  will not allow assignments lacking a value for  $s$  to add a value for  $s'$ . In general, the following fact can be deduced:

**Fact 2.** For any  $G$  such that  $G_y=\emptyset$  and  $H$  such that  $G[\cup_x(\dots[y]; \dots)]H$ ,  $H_y \setminus H_x = \emptyset$ .

This is a relativized form of opacity: variables  $y$  introduced inside unions  $\cup_x(\dots)$  will not be accessible outside of  $G_x$  for output states  $G$ .

### 2.4 Distributive Quantifiers

Distributivity can be achieved in DUPL by ensuring that states within a union clause contain at most one value for a particular variable, i.e.,  $\cup_x(1(x); \dots)$ . This is illustrated in (7). Internal to the distributive “ $\cup_{w'}(1(w'); \dots)$ ”, representing the nuclear scope of (7), “ $w'=w; entered(w')$ ” might produce outputs  $K_1$  and  $K_2$  in Figure 6, assuming only  $W_1$  and  $W_2$  smiled in the model. Note that  $K_1(w')$  and  $K_2(w')$  both contain a single woman who smiled.<sup>3</sup> Crucially, though,  $K_1(w)$  and  $K_2(w)$  both still contain all women in the domain. Therefore, the union of these sets creates a state  $H$  such that  $H(w)$  is all women (here  $W_1 \dots W_3$ ) while  $H(w')$  is only those who smiled.

- (7) Most women smiled.  
 $\rightsquigarrow \cup_w(1(w); woman(w); \cup_{w'}(1(w'); w'=w; smiled(w'))); MOST(+w, +w')$

Donkey anaphora follows easily from this set-up, as shown in (8). For example, Figure 7 shows a component nuclear scope state representing a woman  $W_1$  who bought two books  $B_1$  and

<sup>3</sup>Technically, distribution would not be necessary on the nuclear scope if the restrictor is already distributed, as in this case. Distribution is shown for consistency with other cases and the final version of the system presented below.

$B_2$ , but only read  $B_1$ , capturing (weak) donkey anaphora. Notice that the other women who bought books are also included in this state (along with their books), but the interpretation of predicate literals like “ $read(w', b)$ ” only considers the substate where all their argument variables are defined. In this state, that substate would be the single assignment represented by the last row of Figure 7.

For strong donkey anaphora, we can simply replace the nuclear scope with the formula “ $\cup_{w'}(1(w'); w'=w; \cup_{b'}(b'=b); read(w', b'))$ ” which stores in the variable  $b'$  all books read by the woman  $w'$  and then tests whether she read all these books.<sup>4</sup> The reason  $b'$  will only store values for the current woman  $w'$  is that the outer union  $\cup_{w'}(1(w'); \dots)$  will filter out states that alter assignments where  $w'=\star$ . Since there is only one value for  $w'$  in each state, the only books stored in  $b'$  in each state that passes this filter will be those bought by this particular woman  $w'$ . The same process could be repeated separately for donkey pronouns other than books, and therefore mixed cases of weak and strong donkey anaphora can also be handled (van der Does, 1992; Brasoveanu, 2008).

Quantificational subordination is the same process as donkey anaphora, just across sentences, as shown in (9), where *SOME* is also a set relation. Discourse plurals are also simple in DUPL, as shown, e.g., by  $w$  and  $w'$  in  $MOST(w, w')$ ;  $b$  in the same context would be all woman-bought and -read books. For instance a following clause “*on-table(b)*” could assert that the books the women bought and read are on the table.

- (8) Most women who bought a book read it.  
 $\rightsquigarrow \cup_w(1(w); woman(w); [b]; book(b); bought(w, b); \cup_{w'}(1(w'); w'=w; read(w', b)))$ ;  
 $MOST(+w, +w')$
- (9) Some of them loved it.  
 $\rightsquigarrow \cup_{w^2}(1(w^2); w^2=w'; \cup_{w^3}(1(w^3); w^3=w^2; loved(w^3, b)))$ ;  $SOME(+w^2, +w^3)$

$K_1$		$K_2$	
$w$	$w'$	$w$	$w'$
*	*	*	*
$W_1$	*	$W_1$	*
$W_2$	*	$W_2$	*
$W_3$	*	$W_3$	*
$W_1$	$W_1$	$W_2$	$W_2$

$H$	
$w$	$w'$
*	*
$W_1$	*
$W_2$	*
$W_3$	*
$W_1$	$W_1$
$W_2$	$W_2$

$w$	$b$	$w'$
*	*	*
$W_1$	$B_1$	*
$W_1$	$B_2$	*
$W_2$	$B_3$	*
$W_3$	$B_4$	*
$W_1$	$B_1$	$W_1$

Figure 6: Possible states in (7)

Figure 7: State in the calculation of (8)

### 3 Discussion

#### 3.1 Multiple Maximal Outputs

Referential indeterminacy may extend to seemingly quantified variables, not just those introduced by indefinites, when collective predicates are involved. For instance, (10) may be true

<sup>4</sup>This formula assumes *read* can apply collectively to a set of books; an additional distribution would be required otherwise.

where there are multiple separate groups of students who sat together, each comprising one-third of the students total. This reading is not captured by a straight-forward generalized quantifier translation like the first translation in (10):

- (10) One third of the students sat together.  
 $\rightsquigarrow_1 \cup_s(\text{students}(s); \cup_{s'}(s'=s; \text{sat-together}(s'))); \frac{1}{3}(+s, +s')$   
 $\rightsquigarrow_2 \cup_s(\text{students}(s); [s']; s'=s; \frac{1}{3}(+s, +s'); \text{sat-together}(s'))$

The “multiple thirds” reading seems to be captured better by an indefinite-like definition for “one-third”, as shown in the second translation in (10). In general, then, it seems as though certain determiners are ambiguous between a maximal reading and an indefinite reading. This distinction is not captured in the current system, although the indefinite reading seems to be easier with the more numerically precise determiners. For instance, *they* in (11) cannot refer to a set of most students that sat together that is not the maximal one:

- (11) Most students sat together. They discussed politics.

### 3.2 Reference to the Restrictor Set inside the Nuclear Scope

Nouwen (2003) points out cases where within a distributive nuclear scope, a plural pronoun seems to refer to a value established in the restrictor, as shown in (12). Failure to capture such cases is one empirical shortcoming of the original van den Berg (1996) system. DUPL can handle this case without further modification, though, since even in a distributive context, the entire previous state is available for reference. (See, for instance, states  $K_1$  and  $K_2$  in Figure 6 above.)

If we use plain  $x$  variable terms, the only reading available is one where the fathers each give a pep talk only to their own daughters, not all the daughters. This is because plain terms in predicate literals are always evaluated at the substate where all their arguments are defined; so, “*gave-peptalk*( $f', d$ )” for instance would exclude any value for  $d$  other than  $f'$ ’s daughter.

Thus, here is another case where we will use augmented variable terms “ $+x$ ,” which are evaluated in the full state, rather than some substate thereof. Using these terms, “ $+d$ ” in the literal “*gave-peptalk*( $f', +d$ )” in (12) will refer to all the daughters, rather than only the local value of  $d$  where  $f'$  is defined:

- (12) Each father with a daughter at the meet gave them (all) a pep talk.  
 $\rightsquigarrow \cup_f(\cup_f(1(f); \text{father}(f); [d]; \text{daughter}(d, f); \text{at-meet}(d));$   
 $\cup_{f'}(1(f'); f'=f; \text{gave-peptalk}(f', +d)); \text{ALL}(+f, +f'))$

Notice that the translation here assumes a second, embedded union term  $\cup_f(1(f); \dots)$  distributing over the restrictor. This is so that the full collected / unioned set of fathers and daughters from the restrictor is available in the calculation of the nuclear scope. The outer, surrounding union  $\cup_f(\dots)$  merely serves to restrict where the nuclear scope variable  $f'$  is defined, as discussed when Fact 2 was introduced. This outer union will have no other effect, though, since the inner union over  $f$  prevents its input state from having any defined value for  $f$ . This is therefore required to be a case where random assignment of  $f$  in the outer union does nothing.

## 4 Conclusion and Comparisons

A few small features in the definition of DUPL conspire to allow the behaviors described above. First, the progression of discourse states in DUPL is monotonically increasing, in the sense that every output state (where there is one) is a superset of its input. This means that previous discourse states are always available for reference in later clauses of a formula. For instance, even when distributing over the nuclear scope of a quantifier, while the nuclear scope variable, say  $x'$ , might only have one value, all previous values for the restrictor variable  $x$  will always be available. This allows reference to the restrictor set inside the nuclear scope. This monotonicity is due to random assignment always outputting a superset of its input (and no other clause type interfering with this feature). Although predicate literals still must refer sometimes to a substate of the current state, this is achieved via plain  $x$  variable terms, whose values are calculated within such a substate.

Second, the union clauses in DUPL provide the maximization necessary for quantification, but limit changes as follows: the union clauses introduce a new reference variable  $x$  and filter out any states that make changes outside the “scope” of  $x$  (where the scope is the substate where  $x$  is defined). This constraint prevents spurious values for a subordinate variable from projecting beyond the union clause. Instead, any variable  $y$  introduced inside the scope of  $x$  will only be available in future contexts where  $x$  is defined, as in quantificational subordination.

Further operators, such as complex inclusion relations and distributive operators are not necessary, given this initial set-up.

### 4.1 Comparison to previous systems

Random assignment in [van den Berg \(1996\)](#) introduces a new plural value for a variable  $x$ , but does not introduce any dependencies between the new variable and other variables. His definition also explicitly maintains the dependencies between other variables in the input state. [Nouwen \(2003\)](#) uses a very similar definition for his plural logic. [Brasoveanu \(2007\)](#), following earlier work of van den Berg ([van den Berg, 1994](#)), introduces both a value and dependencies for a new variable  $x$ , but still explicitly maintains the dependencies between other variables.

Random assignment in DUPL is most similar to Brasoveanu’s version, introducing any value or dependencies for a new variable. The major difference, as mentioned above, is that DUPL ensures that the output of random assignment is a superset of its input. This guarantees that we preserve the input as read-only, a fact which was crucial for correctly capturing generalized quantifiers and reference to the restrictor set inside the nuclear scope.

Predicate literals in van den Berg’s system (and Nouwen’s) always take the entire state-wide value for their argument variables. Brasoveanu introduces both a distributive and a state-wide version of predicate literal interpretation, building on the “abstraction” operation of [Kamp and Reyle \(1993\)](#).

DUPL uses a different definition, restricting variable values in a predicate literal to those in the substate where all the predicate’s argument variables are defined. In addition, “ $+x$ ” style variable terms allow reference to the complete state. This small change allows reference to local contexts via plain terms and global contexts via augmented terms.

Existing plural logics all help themselves to further operators defined as relations over states, rather than as abbreviations for more basic operations. DUPL is a reaction against this, an attempt to derive as much as possible from a smaller basic set of operations.

The other systems all include a distributive operator. Distribution in van den Berg’s system is over a variable  $x$ , and explicitly applies a given formula to all substates where  $x$  has only one value, before unioning / stitching these values back together. Nouwen updates this definition

only to allow reference to the restrictor set. Brasoveanu distributes down to the individual assignment level, rather than relativizing the distribution to the value of a particular variable.

Existing systems also include a maximization operator, somewhat akin to  $\cup_x(\dots)$  above. For van den Berg and Brasoveanu, maximizing a formula with respect to a variable  $x$  outputs only those states whose values for  $x$  is not smaller than any other output's value for  $x$ . Nouwen's version of maximization ( $\zeta$ ) is closer to DUPL, since his is a combination distributive and maximization operator.

DUPL, instead, takes advantage of the fact that the union operation inherently stitches together states. Then, distribution can be defined without a new operator, simply by requiring substates within a union clause to contain at most one value for a given variable.

The systems due to van den Berg and Brasoveanu define a special inclusion operator " $\subseteq$ " or " $\sqsubseteq$ " to ensure the correct relationship between the restrictor set variable (e.g.,  $x$ ) and the nuclear scope set variable (e.g.,  $x'$ ). (Nouwen has a different view on projection of the restrictor set.) Clauses like " $w'=w$ " in DUPL replicate van den Berg's inclusion operator, ensuring that if  $w'$  has a non- $\star$  value, this value will match  $w$ . Since random assignment always carries along the full input state, even if  $w'$  does not comprise all values in  $w$ , all values in  $w$  will nevertheless be available for later use. No special inclusion operator is required.

There are other differences. Brasoveanu allows individual assignments to return plural values. Nouwen's system eschews variables and assignments in preference to indexed stacks of values. These choices are largely orthogonal to the points in this paper, though. In addition, the other systems introduce further operators that cannot be defined as abbreviations of multiple simpler operations (for instance, Brasoveanu (2007) has a relativized "atom" operator to capture certain scope effects). Although DUPL aims to avoid the introduction of further such operators, I leave to future work the examination of whether DUPL already captures the empirical cases that drove the creation of each such operator.

## References

- Martin H. van den Berg. A direct definition of generalized dynamic quantifiers. In MJB Stokhof and P Dekker, editors, *Proceedings 9th Amsterdam Colloquium*, pages 121–140. ILLC, 1994.
- Martin H. van den Berg. *Some aspects of the internal structure of discourse. The dynamics of nominal anaphora*. PhD thesis, University of Amsterdam, 1996.
- Adrian Brasoveanu. *Structured nominal and modal reference*. PhD thesis, Rutgers, The State University of New Jersey, 2007.
- Adrian Brasoveanu. Donkey pluralities: plural information states versus non-atomic individuals. *Linguistics and philosophy*, 31(2):129–209, 2008.
- Adrian Brasoveanu. The grammar of quantification and the fine structure of interpretation contexts. *Synthese*, 190(15):3001–3051, 2013.
- Jacob Maarten van der Does. *Applied quantifier logics*. PhD thesis, University of Amsterdam, 1992.
- Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100, 1991.

Hans Kamp and Uwe Reyle. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*, volume 42. Kluwer Academic Dordrecht,, The Netherlands, 1993. ISBN 0792310276.

Rick Nouwen. *Plural pronominal anaphora in context*. PhD thesis, University of Utrecht, 2003.



# Coordinating Complete Answers: The Case of *Tanto-Quanto* Conjunction\*

Filipe Hisao Kobayashi and Enrico Flor

Massachusetts Institute of Technology  
filipek@mit.edu  
eflor@mit.edu

## Abstract

We discuss a coordination strategy found in Portuguese and Italian which we call *Tanto-Quanto Conjunction* (TQC). The semantic properties that distinguish this construction from run-of-the-mill *and*-conjunction are the focus of this paper. TQC imposes a discourse related requirement on its conjuncts, namely that they each be a complete answer to a question raised in the discourse. We propose an analysis of TQC where each of its conjuncts falls under the scope of a focus sensitive operator which, by means of an answerhood operator, checks that its prejacent satisfies this requirement.

## 1 Introduction

In this paper we investigate the semantic properties of a coordinating construction found in Portuguese and Italian, which we name *Tanto-Quanto Conjunction* (henceforth: TQC) after the pieces that compose it. TQC in both languages is illustrated in (1):<sup>1</sup>

- (1) a. Tanto a Sara quanto a Maria trabalham em Paris. Portuguese  
TANTO the Sara QUANTO the Maria work in Paris  
*Sara and Maria work in Paris.*
- b. Tanto Sara quanto Maria lavorano a Parigi. Italian  
TANTO Sara QUANTO Maria work in Paris  
*Sara and Maria work in Paris.*

*Tanto* and *quanto* are also found in equative constructions, as shown in (2a). However, the TQC in (2b) lacks an equative meaning altogether: rather than implying that Sara and Maria like cars to the same degree, its truth conditional import is simply that they both like cars. For example, as opposed to (2a), (2b) could be truthfully uttered in a context in which one of Sara and Maria likes cars much more than the other.

- (2) a. A Sara gosta tanto de carro quanto a Maria.  
the Sara like TANTO of car QUANTO the Maria  
*Sara likes cars as much as Maria.*
- b. Tanto a Sara quanto a Maria gostam de carro.  
TANTO the Sara QUANTO the Maria like of car  
*Sara and Maria like cars.*

---

\*For valuable comments and discussion, we thank Athulya Aravind, Itai Bassi, Keny Chatain, Sherry Yong Chen, Patrick Elliott, Émile Enguehard, Kai von Fintel, Maša Močnik, Patrick Niedzielski, Vincent Rouillard, Viola Schmitt, Roger Schwarzschild, and Stanislao Zoppi. All mistakes are our own.

<sup>1</sup>For conciseness, we use only Portuguese examples in the rest of the paper, but all facts we observe and discuss hold in both languages in the same way.

Translating TQC into English *and* (i.e., via *and*-conjunction), as done above, preserves the truth conditional meaning of the utterance: that both conjuncts are true. However, we argue that TQC imposes a discourse related requirement on its conjuncts that *and*-conjunction (expressed in Italian and Portuguese as *e*) does not. We propose that the distribution of TQC is constrained by the following appropriateness condition:

(3) **TQC's appropriateness condition**

⌈Tanto A quanto B VP⌋ is felicitously uttered in a context *c* only if each of  $\llbracket A \text{ VP} \rrbracket$  and  $\llbracket B \text{ VP} \rrbracket$  are a **complete answer** to a question raised in *c*.

The core data supporting this claim come from observations concerning the kind of explicit questions that TQC is a felicitous answer to (section 2). We propose an analysis of such facts in terms of recent accounts of exhaustivity in questions [3] (section 3.1) and [13]'s alternative semantics for focus (section 3.2). In our proposal, each conjunct of a TQC is taken to fall under the scope of a focus sensitive operator which, by means of an answerhood operator, checks that its prejacent is a complete answer to a question raised in the discourse (section 4). Our analysis relies on TQC being an instance of sentential coordination, and we present independent motivation for this assumption (section 5).

## 2 Tanto-Quanto Conjunction in answers to questions

Questions usually demand exhaustive answers ([5] and much subsequent work). For example, an appropriate answer to (4) is one in which all people with a PhD are specified.<sup>2</sup> *A'* is inappropriate because it is interpreted exhaustively, which renders it incompatible with the facts.

- (4) CONTEXT: Only Sara and Maria have a PhD.  
 Q: Who has a PhD? A: Sara and Maria.  
 A': #Sara.

The question-answer pairs in (4) contrasts with those in (5), in which the question has the possibility modal *can*. As observed by [5], such questions do not seem to require an exhaustive answer: both *A* and *A'* are an appropriate answer to the question in (5).

- (5) CONTEXT: Sara can drive people back and Maria can drive people back. No one else can.  
 Q: Who can drive us back? A: Sara and Maria.  
 A': Sara.

Questions that demand an exhaustive answer, like the one in (4), are referred to as **mention-all** (MA) questions, whereas those that do not, like the one (5), are referred to as **mention-some** (MS) questions. From now on, we refer to answers from appropriate question-answer pairs as **complete answers**. MA and MS questions differ in that the former has a single complete answer, whereas the latter allows for the possibility of there being more than one.

The MA/MS distinction is relevant here because TQC is sensitive to it, and therefore it is an important tool to probe into the meaning of this construction. Consider the question-answer pairs in (6)-(7) where TQCs appear in the answers, coordinating the NP correlates of the *wh*-item in the question. The crucial fact here is that, different from *and*-conjunction, TQC cannot be used as an answer to an MA question, as seen in (6). Nonetheless, TQC can be felicitously used as an answer to MS questions, as seen in (7).

<sup>2</sup>In (4) we assume all answers are pronounced with a falling tone. This is important because certain prosodic contours can be used to indicate that the answer is not to be interpreted exhaustively.

(6) Q: Who has a PhD?

a. #Tanto a Maria quanto a Sara.  
TANTO the Maria QUANTO the Sara

b. A Maria e a Sara.  $\leadsto$ Only Maria and Sara have a PhD  
the Maria and the Sara

(7) Q: Who can drive us back?

a. Tanto a Maria quanto a Sara.  $\not\leadsto$ Only Maria and Sara can drive us back  
TANTO the Maria QUANTO the Sara

b. A Maria e a Sara.  $\leadsto$ Only Maria and Sara can drive us back  
the Maria and the Sara

The question-answer pairs in (6) and (7) show that TQC can be an answer to MS questions, but not to MA questions. This paradigm can be accounted for by the appropriateness conditions in (3), which states that each of TQC's conjuncts must be a complete answer to a question raised in the context. In (6), an MA question is raised, and therefore, answers must be interpreted exhaustively to be appropriate. Answering such question with a TQC is unavoidably infelicitous: interpreting its conjuncts non-exhaustively makes them an infelicitous answer, whereas interpreting them exhaustively makes them contradictory. This contrasts with contexts in which an MS question is raised, since these questions may have more than one complete answer. For example, in (7a) each of the TQC's conjuncts is individually a complete answer to the MS question.

Note that there is still a difference between TQC and *and*-conjunction in answers to MS questions: TQC does not lead to an exhaustivity inference, while *and*-conjunction does. The latter was noted by [14], who argues that MS questions are more precisely described as "mention-one": an MS answer specifies exactly one of the possible options. This difference between TQC and *and*-conjunction follows from the fact that the answer with *and*-conjunction in (7b) is interpreted as a whole as an answer to the question, whereas only the individual conjuncts of TQC are interpreted as such, as the appropriateness condition stated in (3) indicates.

Even with TQC one can still observe the effects that MS is "mention-one". Take (8) to be uttered in a context in which there are exactly three people who can drive us back (Maria, Paula and Sara). Although it is possible to answer this question with *and*-conjunction of all three entities as in (8a), the TQC in (8b), whose first conjunct is an *and*-conjunction, is infelicitous.

(8) Q: Who can drive us back?

a. A Maria, a Paula, e a Sara.  
the Maria the Paula and the Sara

b. #Tanto [a Maria e a Paula] quanto a Sara.  
TANTO the Maria and the Paula QUANTO the Sara

c. #A Maria e a Paula.  
the Maria and the Paula

In order for the TQC in (8b) to be appropriate, each conjunct has to be a complete answer to the question Q. Although this is the case for the second conjunct (*a Sara*), it is not so for the first conjunct: as seen in (8c), *Maria and Paula* is not an appropriate answer to Q.

Note that the appropriateness condition in (3) does not require that the conjuncts of a TQC be a complete answer to the same question. Observe the dialogue in (9), in which two questions are raised. TQC's appropriateness condition is satisfied because *Maria drives a Ferrari* is a

complete answer to *What does Maria drive?* and *Sara drives a Ferrari* is a complete answer to *What does Sara drive?*. Crucially, if the context were such that Maria drives both a Ferrari and an Lamborghini, the TQC in (9) would be infelicitous.

- (9) a. Q: What does Maria drive? What about Paula, what does she drive?  
 b. Tanto a Maria quanto a Paula dirigem Ferrari.  
 TANTO the Maria QUANTO the Paula drive Ferrari

In this section, we used question-answer pairs to investigate the appropriateness conditions of TQC. Of course, TQC can be appropriately used in utterances that are not answers to explicit questions. We however believe that the appropriateness condition in (3) can be extended to such cases as well. One way to do so would be to follow [12] in modeling discourse as always involving questions being raised, be they explicit or implicit. If such a model were adopted, (3) would make reference to salient Questions Under Discussions (QUDs). For the remaining of this paper, we remain agnostic to an actual implementation along these lines, and the account of TQC we propose can be straightforwardly implemented in different models of discourse.

### 3 Theoretical background

Here we discuss our assumptions concerning exhaustivity in questions and focus semantics.

#### 3.1 A working analysis of exhaustivity in questions

Following [7], we take questions to denote the set of their answers, as shown in (10).<sup>3</sup>

$$(10) \quad \llbracket \text{Who has a PhD?} \rrbracket^w = \{\lambda w'. \text{has-a-PhD}_{w'}(x) \mid \text{human}_w(x)\}$$

Under this assumption, exhaustivity in questions can be accounted for by an answerhood operator *Ans* [8]. A variety of such operators have been proposed in the literature. Here, we adopt [3]’s, presented in (11), which is a weakened version of [2]’s: while [2]’s *Ans* returns the maximally informative true answer to a question, [3]’s returns the set of maximally informative true answers (i.e., the set of true answers that are not entailed by any other true answer).<sup>4</sup> This weakening of *Ans* is motivated by the need of accounting for MS questions, which, as we have seen, allow for more than one complete answer. Furthermore, we assume the answerhood condition in (12), which specifies whether a given proposition is a complete answer to a question.

$$(11) \quad \text{Ans}(Q) := \lambda w. \{p \in Q \mid p(w) \wedge \forall q \in Q[q(w) \rightarrow q \not\subseteq p]\}$$

#### (12) The Answerhood Condition

A proposition  $p$  is a complete answer to a question  $Q$  in a world  $w$  iff  $p \in \text{Ans}(Q)(w)$ .

Here’s an illustration of how MA readings of *wh*-questions are derived. Assume that there are only three human—Sara, Maria, and Paula—and only the first two have a PhD. The question in (10) thus denotes the question set in (13),<sup>5</sup> and its true answers are  $\mathbf{s}$ ,  $\mathbf{m}$ , and  $\mathbf{s} \oplus \mathbf{m}$ , underlined in (13). As shown in (14), when *Ans* applies to this set, it only returns  $\mathbf{s} \oplus \mathbf{m}$ : although  $\mathbf{s}$  and  $\mathbf{m}$  are true, they are both entailed by  $\mathbf{s} \oplus \mathbf{m}$  (since *to have a PhD* is a distributive predicate). Therefore, given (12), only *Sara and Maria* will be an appropriate answer to (10).

<sup>3</sup>We assume with [2] that *who* is number neutral—the predicate *human* in (10) thus ranges over both singularities and pluralities.

<sup>4</sup>Both [2]’s and [3]’s entries for *Ans* are presuppositional, a matter we leave aside in the present discussion.

<sup>5</sup>From now on, given an individual constant  $\alpha$ , we abbreviate a proposition  $\lambda w. \phi(\alpha)(w)$  to simply  $\alpha$ .

$$(13) \quad Q = \{\underline{s}, \underline{m}, \underline{p}, \underline{s \oplus m}, \underline{s \oplus p}, \underline{m \oplus p}, \underline{s \oplus m \oplus p}\}$$

$$(14) \quad \text{Ans}(Q)(@) = \{s \oplus m\} \quad (\text{'@' stands for the world of evaluation})$$

[3] assumes questions with possibility modals, such as (15), to be ambiguous between an MS and an MA reading. The proposed ambiguity is due to the scope of distributivity in these sentences: if it is above the modal, as in (15b), we get an MA reading of the question, but if it is below the modal, as in (15c), we get an MS reading.

(15) a. Who can drive us back?

$$b. \quad \{\lambda w. \forall y \leq_{at} x : \diamond_w(\lambda w'. \text{drive-us-back}_{w'}(y)) \mid \text{human}_@(x)\} \quad \text{MENTION-ALL}$$

$$c. \quad \{\lambda w. \diamond_w(\lambda w'. \forall y \leq_{at} x : \text{drive-us-back}_{w'}(y)) \mid \text{human}_@(x)\} \quad \text{MENTION-SOME}$$

To see this, assume there are four humans—Maria, Sarah, Téó and Léo—and only the first three can be the person who will drive people back (assuming that driving is a one-person job). The question set of (15a) when distributivity takes wide scope is (16), where the true answers are underlined. All the true answers are entailed by  $\diamond m \wedge \diamond s \wedge \diamond t$ , and, therefore, this will be the sole member of the set *Ans* outputs (17).

$$(16) \quad Q = \left\{ \begin{array}{c} \underline{\diamond m}, \underline{\diamond s}, \underline{\diamond t}, \diamond l, \\ \underline{\diamond m \wedge \diamond s}, \underline{\diamond m \wedge \diamond t}, \underline{\diamond m \wedge \diamond l}, \underline{\diamond s \wedge \diamond t}, \underline{\diamond s \wedge \diamond l}, \underline{\diamond t \wedge \diamond l} \\ \underline{\diamond m \wedge \diamond s \wedge \diamond t}, \underline{\diamond m \wedge \diamond s \wedge \diamond l}, \underline{\diamond s \wedge \diamond t \wedge \diamond l} \\ \underline{\diamond m \wedge \diamond s \wedge \diamond t \wedge \diamond l} \end{array} \right\}$$

$$(17) \quad \text{Ans}(Q)(@) = \{\diamond m \wedge \diamond s \wedge \diamond t\}$$

When distributivity scopes below the modal, the question set is quite different. This is illustrated in (18). First, because driving is a one person job, all members of (18) which involve disjunction below the possibility modal are incompatible with the context at hand. Therefore, the only true answers in this set are  $\diamond m$ ,  $\diamond s$ , and  $\diamond t$ . Since these propositions are logically independent, they will all be in the set of answers that *Ans* outputs (19). Therefore, (15c) has more than one complete answer.

$$(18) \quad Q = \left\{ \begin{array}{c} \underline{\diamond m}, \underline{\diamond s}, \underline{\diamond t}, \diamond l, \\ \diamond(m \wedge s), \diamond(m \wedge t), \underline{\diamond(m \wedge l)}, \diamond(s \wedge t), \diamond(s \wedge l), \diamond(t \wedge l) \\ \diamond(m \wedge s \wedge t), \diamond(m \wedge s \wedge l), \diamond(s \wedge t \wedge l) \\ \underline{\diamond(m \wedge s \wedge t \wedge l)} \end{array} \right\}$$

$$(19) \quad \text{Ans}(Q)(@) = \{\diamond m, \diamond s, \diamond t\}$$

The fact that MS readings do not accept conjunctive answers is accounted for in this approach. In the MS question set, there are no viable conjunctive alternatives, therefore, a conjunctive answer could only be complete in the MA reading of this (15). However, in the MA reading of (15), there is a single complete answer.<sup>6</sup>

### 3.2 Alternative Semantics for Focus

In the alternative semantics account of focus [13], linguistic expressions are assigned two different semantic values: an ordinary semantic value ( $\llbracket \cdot \rrbracket$ ) and a focus value ( $\{\!\{ \cdot \}\!\}$ ). Any expression mapped to an ordinary value of type  $\sigma$  is mapped to a focus value of type  $\sigma t$ . Lexical items are

<sup>6</sup>[14] notes some problems with [3]'s approach to MS readings and offers an alternative analysis which still relies on the entry for *Ans* in (11). The choice between these two proposals does not affect our analysis of TQC.

in general assigned trivial focus values (i.e., the singleton containing their ordinary value (20a)) and focus values are combined via Pointwise Functional Application (20b).

$$(20) \quad \text{a. } \{\{\alpha\}\} = \{\llbracket \alpha \rrbracket\}$$

$$\text{b. } \{\{\alpha \beta\}\} = \{f(x) \mid f \in \llbracket \alpha \rrbracket \wedge x \in \llbracket \beta \rrbracket\}$$

Focused constituents are syntactically represented by having a morpheme *F* attach to them, whose lexical entry is specified in (21). Semantically, *F* is responsible for introducing multiple alternatives in the focus dimension of meaning. As illustrated in (22), if the proper name *Sara* is focused, its focus semantic value will be the set of all individuals.<sup>7</sup>

$$(21) \quad \text{a. } \llbracket F \rrbracket = \lambda x_\sigma. x$$

$$\text{b. } \{\{F\}\} = \{\lambda x_\sigma. y \mid y \in D_\sigma\}$$

$$(22) \quad \text{a. } \llbracket \text{Sara } F \rrbracket = s$$

$$\text{b. } \{\{\text{Sara } F\}\} = \{y \mid y \in D_e\}$$

[13] offers a uniform way to account for a whole range of focus-related phenomena by proposing that natural language has a single operator that is capable of manipulating focus values of expressions:  $\sim$ . Its semantics is stated in terms of the syncategorematic rule in (23).

$$(23) \quad \llbracket \sim_Q \phi \rrbracket = \lambda w : Q \subseteq \{\{\phi\}\} \wedge \llbracket \phi \rrbracket \in Q \wedge |Q| > 1. \llbracket \phi \rrbracket(w)$$

One of the phenomena  $\sim$  was proposed to account for is **Question-Answer Congruence** (QAC). This is illustrated in (24), where focus prosody is indicated with small caps. Although the ordinary values of both *A* and *A'* are good candidates for complete answers to *Q*, the latter is infelicitous. The contrast is due to the placement of focus: for QAC to be satisfied, the constituent that is the correlate of the *wh*-item of the question must be the one to bear focus.

$$(24) \quad \text{Q: What did Alex read?} \quad \text{A: Alex read WAR AND PEACE.}$$

$$\text{A': } \# \text{ALEX read War and Peace.}$$

In [13], QAC follows from the presupposition of  $\sim$ . Under the assumption that answers fall under the scope of  $\sim$  with a restrictor that matches the denotation of the question being asked, as in (25a), the presupposition of  $\sim$  in (25b) will guarantee that QAC is satisfied.

$$(25) \quad \text{a. } \sim_Q [\phi \text{ Alex read } [ \text{War and Peace } F ]]$$

$$\text{b. } Q \subseteq \{\{\phi\}\} \wedge \{w \mid \text{read}_w(a, w \& p)\} \in Q \wedge |Q| > 1$$

$$\text{where } Q = \{\{\phi\}\} = \{\lambda w. \text{read}_w(a, x) \mid x \in D_e\}$$

## 4 Analysis

We propose that TQC is an instance of what [10] refers to as Focus Sensitive Coordination (FSC), i.e., coordinations in which at least one of the coordinates is embedded under a Focus Sensitive Operator (FSO). In the case of TQC, we assume that both *tanto* and *quanto* are FSOs that locally check whether their prejacent is a complete answer to a question raised in the discourse. Their lexical entries are defined in (26): they each combine with a proposition *p* and introduce the presupposition that *p* is a complete answer to a question *Q*, as determined by *Ans*. A covert  $\&$  with the meaning in (27) conjoins the *tanto* and *quanto* clauses.

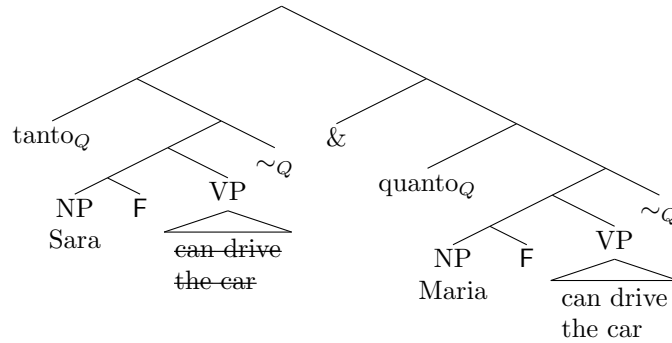
$$(26) \quad \llbracket \text{tanto}_Q \rrbracket = \llbracket \text{quanto}_Q \rrbracket = \lambda p. \lambda w : p \in \text{Ans}(Q)(w). p(w)$$

<sup>7</sup>In [13]'s actual system, a special lexical rule is given to focused constituents.

$$(27) \quad \llbracket \& \rrbracket = \lambda p.\lambda q.\lambda w. p(w) \wedge q(w)$$

In (28b), we provide the syntactic structure of the TQC in (28a). We assume that TQC involves sentential conjunction, even though, on the surface, (28a) seems to be an instance of NP conjunction.<sup>8</sup> This is so because *tanto* and *quanto* take a propositional argument: under our current assumptions, being a complete answer to a question is a property of propositions. We therefore assume that in (28a), the VP of the first clause is elided (see [9] for a recent proposal in which *and*-conjunction is completely reduced to sentence conjunction).

- (28) a. Tanto a Sara quanto a Maria podem dirigir o carro.  
 TANTO the Pedro QUANTO the Sara can drive the car
- b.



Due to the ungrammaticality of (29), we must assume that ellipsis is obligatorily.

- (29) \*Tanto o Pedro comeu peixe quanto a Sara comeu peixe.  
 TANTO the Pedro ate fish QUANTO the Sara ate fish

Finally, note that neither *tanto* nor *quanto* directly access the focus values of their prejacent. As in [13], this is done by having a  $\sim$  with a matching restrictor immediately dominated by the sister node of the FSO. Among other things,  $\sim$  also guarantees that QAC is satisfied.

We now show how this analysis accounts for the data discussed in section 2, starting with MS questions. Assuming that the value of  $Q$  in (28b) is the question in (30a) and that only Maria, Sara and Léo can drive the car, the set of complete answers is the one in (30b).<sup>9</sup> Ignoring the presuppositions triggered by  $\sim$ , we predict the appropriateness conditions of the TQC in question to be (30c),<sup>10</sup> which in this case is satisfied.

- (30) a.  $Q = \{\lambda w. \diamond_w(\lambda w'. \forall y \leq_{at} x : \text{drive-the-car}_{w'}(y)) \mid x \in D_e\}$   
 b.  $\text{Ans}(Q)(@) = \{\diamond s, \diamond m, \diamond l\}$   
 c.  $\diamond s \in \text{Ans}(Q)(@) \wedge \diamond m \in \text{Ans}(Q)(@)$

If the first conjunct is itself an *and*-conjunction, the sentence is correctly predicted to be bad (31a). No matter whether the *and*-conjunction scopes below (31b) or above (31c) the modal, the appropriateness condition of TQC cannot be satisfied. Even if the distributivity operator scopes above the modal in the question, (31b) and (31a) will still be false: neither proposition denoted by the conjuncts would be a complete answer to  $Q$ .

<sup>8</sup>In fact, according to [10], Focus Sensitive Coordination always involves sentential coordination.

<sup>9</sup>For QAC to be satisfied, we must assume that singular DPs may combine with distributive quantifiers.

<sup>10</sup>Actually, under certain assumptions of presupposition projection out of conjuncts, a conditional presupposition would be predicted to arise. In many cases, these presuppositions are strengthened, however (the “proviso problem”, see [4]). Here we assume this to be the case.



- (31) a. # [Tanto a Sara e o Léo] quanto a Maria podem dirigir o carro.  
 TANTO the Pedro and the Léo QUANTO the Sara can drive the car  
 b.  $\diamond(s \wedge l) \in \text{Ans}(Q)(@) \wedge \diamond m \in \text{Ans}(Q)(@)$   
 c.  $(\diamond s \wedge \diamond l) \in \text{Ans}(Q)(@) \wedge \diamond m \in \text{Ans}(Q)(@)$

When the value of  $Q$  in both conjuncts is the same MA question, the requirements are once again unsatisfiable. If the question in (32a) is raised and only Sara and Maria have a PhD, the set of complete answers to  $Q$  will be the singleton in (32b). Consequently, (32c) is false.

- (32) a.  $Q = \{\lambda w. \text{has-a-PhD}_w(x) \mid x \in D_e\}$   
 b.  $\text{Ans}(Q)(@) = \{\lambda w. \text{has-a-PhD}_w(s \oplus m)\}$   
 c.  $\{w \mid \text{has-a-PhD}_w(s)\} \in \text{Ans}(Q)(@) \wedge \{w \mid \text{has-a-PhD}_w(m)\} \in \text{Ans}(Q)(@)$

The final point to be addressed are the cases in which the conjuncts of the TQC do not answer the same question. Example (9) is repeated below in (33). The two questions being raised are *What does Maria drive?* (34a) and *What does Paula drive?* (34b). For QAC to be satisfied, *Ferrari* must be focused, so we assume that (33b) has the LF in (34c). To avoid claiming that a focused constituent is elided in (33b), we assume that *Ferrari* is moved in an Across-the-Board fashion in surface syntax but is reconstructed at LF. The appropriateness conditions our analysis derives for (33b) are stated in (34), which can be satisfied, as one can easily verify.

- (33) a. Q: What does Maria drive? What about Sara, what does she drive?  
 b. Tanto a Maria quanto a Paula dirigem Ferrari.  
 TANTO the Maria QUANTO the Paula drive Ferrari
- (34) a.  $Q_1 = \{\lambda w. \text{drive}_w(m, x) \mid x \in D_e\}$   
 b.  $Q_2 = \{\lambda w. \text{drive}_w(p, x) \mid x \in D_e\}$   
 c.  $[[\text{tanto}_{Q_1} \sim_{Q_1} \text{Maria drives [Ferrari F]}] \& [\text{quanto}_{Q_2} \sim_{Q_2} \text{Paula drives [Ferrari F]}]]$   
 d.  $\{w \mid \text{drive}_w(m, f)\} \in \text{Ans}(Q_1)(@) \wedge \{w \mid \text{drive}_w(p, f)\} \in \text{Ans}(Q_2)(@)$

## 5 Additional considerations

In this section we present additional and independent evidence corroborating the assumption that TQC is an instance of sentential conjunction.

NPs coordinated by TQC can never combine with a predicate and give rise to a non-distributive reading, a phenomenon we dub **anti-collectivity**, following [6]. When combined with mixed predicates like *earn \$100.000*, TQC only gives rise to a distributive reading, while *and*-conjunction results in ambiguity between distributive and a non-distributive reading.

- (35) a. Tanto a Sara quanto a Maria recebem \$100.000.  
 TANTO the Sara QUANTO the Maria receive \$100.000  
*Sara and Maria earn \$100.000 each. / \*Sara and Maria earn \$100.000 between them.*  
 b. A Sara e a Maria recebem \$100.000.  
 the Sara and the Maria receive \$100.000  
*Sara and Maria earn \$100.000 each. / Sara and Maria earn \$100.000 between them.*

Similarly, TQC is incompatible with collective predicates:



- (36) a. \*Tanto a Maria quanto a Sara se encontraram.  
 TANTO the Maria QUANTO the Sara SE met
- b. A Maria e a Sara se encontraram.  
 the Maria and the Sara SE met  
*Maria and Sara met.*

Although the facts in (35)-(36) seem to suggest that TQC is some kind of distributive conjunction, it also lacks characteristic properties of distributive constructions. For example, TQC does not licence so-called sentence-internal readings of *diferente* ‘different’ ([1]). Thus, while (37b), with *and*-conjunction, can be understood as saying that the book Maria read is different from the one Sara read, (37a) can only mean that Sara and Maria bought books different from another book made salient in the discourse (the sentence external reading).<sup>11</sup>

- (37) a. Tanto a Maria quanto a Sara leram um livro diferente. \*SENT-INTERNAL  
 TANTO the Maria QUANTO the Sara read one book different
- b. A Maria e a Sara leram um livro diferente. ✓SENT-INTERNAL  
 the Maria and the Sara read one book different

Sentential *and*-conjunction patterns with TQC in all these cases (to the exclusion of other forms of conjunction one could assume to exist). In fact, as (38) shows, sentential *and*-conjunction has anti-collectivity properties and does not license an internal reading of *different*, i.e., (38c) cannot be uttered out of the blue to mean that the books Sara and Maria read were different:

- (38) a. Sara earns \$100.00 and Maria earns \$100.000. \*COLLECTIVE/CUMULATIVE
- b. \*Sara met and Maria met.
- c. Sara read a different book and Maria read a different book. \*INTERNAL

A fact we do not account for is that TQC of the subject NP controls plural verbal agreement, as shown in (39). Under the assumption that TQC is an instance of sentence coordination coupled with ellipsis, this is unexpected.

- (39) Tanto a Maria quanto o Vicente am-**am**/\*-**a** peixe.  
 TANTO the Maria QUANTO the Vicente love-3PL/-3SG fish  
*Maria and Vicente love fish.*

Similar facts have been observed in Right Node Raising constructions (see [6] and references therein). However, in those cases plural agreement is optional, whereas here it is obligatory. We leave for future research whether TQC is better analyzed as an instance of ellipsis (as done here) or Right Node Raising. In any case, the compositional account defended here should be implementable in either case.

## 6 Conclusion

In this paper we have described and analyzed the semantic properties of a type of coordination found in Portuguese and Italian, TQC. While its truth conditional import is that of propositional conjunction, TQC has the feature of imposing certain discourse related requirements on its

<sup>11</sup>[11], and much subsequent literature on *different*, claims that English *and*-conjunction does not license sentence internal readings of singular *different* (\**Mary and Sara read a different book*). According to our judgements, and of other speakers we consulted with, this is not the case in either Italian or Portuguese.

conjuncts. This case study is thus yet another illustration of the different ways in which a propositional logic connector, in our case conjunction, is enriched in natural language to convey more than its run-of-the-mill at-issue meaning (as in other focus sensitive coordinators). Furthermore, we can conclude that a semantic definition of complete answers is needed in order to capture TQC’s semantic properties in a compositional fashion. Since TQC locally checks that its conjuncts are complete answers, the property of being so cannot be a pragmatic one.

Finally, a remaining question in this study is the connection between TQC and equative constructions. As mentioned in the introduction, the two constructions are built from the same pieces (*tanto* and *quanto*), however there does not seem to be any obvious common semantic property from which to derive this morphological identity. We speculate that at least at a very abstract level of analysis, TQC does involve a notion of equality between two propositions with respect to their relevance in the discourse. We leave this matter for future research.

## References

- [1] Greg Carlson. Same and different: Some consequences for syntax and semantics. *Linguistics and Philosophy*, 10(4):531–565, 1987.
- [2] Veneeta Dayal. *Locality in WH quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer, 1996.
- [3] Danny Fox. Mention-some readings of questions. MIT Seminar Notes, 2013.
- [4] Bart Geurts. Local satisfaction guaranteed: A presupposition theory and its problems. *Linguistics and Philosophy*, 19(3):259–294, 1996.
- [5] Jeroen Groenendijk and Martin Stokhof. *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam, 1984.
- [6] Patrick Georg Grosz. Movement and agreement in right-node-raising constructions. *Syntax*, 18(1):1–38, 2015.
- [7] Charles L. Hamblin. Questions in Montague English. *Foundations of Language*, 10(1):41–53, 1973.
- [8] Irene Heim. Interrogative semantics and Karttunen’s semantics for “know”. In R. Buhalla and A. Mittwoch, editors, *Proceedings of the Israeli Association for Theoretical Linguistics*, pages 128–144, 1994.
- [9] Aron Hirsch. *An inflexible semantic for cross-categorical operators*. PhD thesis, MIT, 2017.
- [10] Sarah Hulsey. *Focus Sensitive Coordination*. PhD thesis, MIT, 2008.
- [11] Friederike Moltmann. Reciprocals and Same/Different: Towards a semantic analysis. *Linguistics and Philosophy*, 15(4):411–462, 1992.
- [12] Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69, 2012.
- [13] Mats Rooth. A theory of focus interpretation. *Natural language semantics*, 1(1):75–116, 1992.
- [14] Yimei Xiang. *Interpreting Questions with Non-exhaustive Answers*. PhD thesis, Harvard University, 2016.

# Hyperlogic: A System for Talking about Logics\*

Alexander W. Kocurek<sup>1</sup>

Cornell University, Ithaca, NY, USA  
awk78@cornell.edu

## Abstract

Sentences about logic are often used to show that certain embedding expressions, including attitude verbs, conditionals, and epistemic modals, are hyperintensional. Yet it is not clear how to regiment “logic talk” in the object language so that it can be compositionally embedded under such expressions. This paper does two things. First, it argues against a standard account of logic talk, viz., the impossible worlds semantics [2]. It is shown that this semantics does not easily extend to a language with propositional quantifiers, which are necessary for regimenting some logic talk. Second, it develops an alternative framework based on logical expressivism, which explains logic talk using shifting conventions [6]. When combined with the standard **S5** $\pi$ + semantics for propositional quantifiers, this framework results in a well-behaved system that does not face the problems of the impossible worlds semantics. It can also be naturally extended with hybrid operators [1] to regiment a broader range of logic talk, e.g., claims about what laws hold according to other logics. The resulting system, called *hyperlogic*, is therefore a better framework for modeling logic talk than previous accounts.

## 1 Introduction

Sentences like (1)–(3) suggest that attitude verbs, conditionals, and epistemic modals are hyperintensional, i.e., they do not validate the replacement of necessary (or even logical) equivalents.

- (1) Inej believes intuitionistic logic is the correct logic.
- (2) If the Liar were both true and not true, the law of non-contradiction would fail.
- (3) Classical logic might not be correct.

In order to develop a semantics for these expressions that captures their hyperintensionality, we first need a way of regimenting sentences like ‘intuitionistic logic is the correct logic’ and ‘the law of non-contradiction fails’ in the object language so that they may be meaningfully and non-trivially compositionally embedded. Yet it is not entirely clear how this can be done. We cannot, for instance, simply regiment (2) as  $(l \wedge \neg l) \Box \rightarrow \neg \vdash \neg(p \wedge \neg p)$ , since this illicitly imports notation from the metalanguage into the object language. We might instead try to regiment (2) as  $(l \wedge \neg l) \Box \rightarrow \neg lnc$ , where *lnc* is some primitive atomic standing for the law of non-contradiction. But this is unsatisfactory, as atomic formulas are generally assumed to be logically contingent, whereas the law of non-contradiction, intuitively, is not.

This paper does two things. First, it presents and argues against a standard hyperintensional framework for modeling logic talk, viz., the *impossible worlds semantics*, which introduces logically impossible worlds to accommodate logic talk [2]. The main problem with this approach is that it cannot be easily extended to a language with propositional quantifiers [4]. Propositional quantifiers are useful, and even necessary in some cases, for regimenting laws

---

\*Thanks to Yifeng Ding, Melissa Fusco, Wes Holliday, Ethan Jerzak, and Rachel Rudolph, as well as the participants of the 11th Semantics and Philosophy in Europe conference for helpful discussion.

of logic in the object language. But it turns out that when we try to interpret propositional quantifiers in the presence of logically impossible worlds, we run into serious troubles.

Second, this paper presents an alternative to the impossible worlds semantics. This system, which I call *hyperlogic*, is inspired by a philosophical view on the nature of logic known as *logical expressivism*, which explains logic talk by appealing to shifting conventions rather than impossible worlds [6]. It turns out that logical expressivism, when combined with the standard **S5** $\pi$ + semantics for propositional quantifiers, does not face any of the problems that plagued the impossible worlds semantics, and generally results in a nice, well-behaved system. What's more, this system can be naturally extended to a more expressive language that can accommodate a broader range of logic talk. In particular, we can introduce operators borrowed and modified from hybrid logic [1] to regiment claims about what laws hold *according to* other logics and to regiment the distinction between axioms and rules in the object language.

## 2 The Impossible Worlds Semantics

To begin, let's review the impossible worlds semantics and its account of logic talk. To simplify the discussion, I will only focus on counterfactuals. The points made about counterfactuals easily extend to attitude verbs and modals more generally, but I leave that to future work.

We start with a simple base language  $\mathcal{L}_0$  consisting of an infinite stock of propositional variables  $\text{Prop} = \{p_1, p_2, p_3, \dots\}$ , all the standard boolean connectives ( $\neg, \wedge, \vee, \rightarrow$ ), a pair of modalities ( $\Box$  and  $\Diamond$ ), and a counterfactual operator ( $\Box\rightarrow$ ). The syntax of  $\mathcal{L}_0$  is summarized in Backus-Naur form as follows:

$$\phi ::= p \mid \neg \phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid (\phi \rightarrow \phi) \mid \Box \phi \mid \Diamond \phi \mid (\phi \Box\rightarrow \phi).$$

On the standard, intensional semantics for counterfactuals,  $\phi \Box\rightarrow \psi$  is true iff all of the closest possible  $\phi$ -worlds are  $\psi$ -worlds [8, 10]. The impossible worlds semantics takes this idea and adds a twist:  $\phi \Box\rightarrow \psi$  is true iff all of the closest  $\phi$ -worlds, whether or not those worlds are possible, are  $\psi$ -worlds [2]. In other words, the impossible worlds semantics differs from the standard semantics in allowing the closest  $\phi$ -worlds to include impossible worlds.

What is an “impossible” world? First, think of a world as a kind of ersatz entity: a world (possible or not) is just a set of formulas. The members of an ersatz world intuitively represent what is true according to that world—that is,  $\phi$  is true at an ersatz world  $w$  iff  $\phi \in w$ . On this understanding of a world, a possible world is just a special kind of set, viz., one that is maximally compossible, whereas an impossible world is just a set that is *not* maximally compossible. Thus, impossible worlds are not a wholly new or alien kind of entity: we are already committed to them if we accept ersatz possible worlds [9].

Using this conception of impossible worlds, here is the impossible worlds semantics for  $\mathcal{L}_0$ . First, an *impossible worlds model* is a quadruple of the form  $\mathcal{I} = \langle W, P, f, V \rangle$ , where:

- $W \neq \emptyset$  is the set of *worlds*;
- $\emptyset \neq P \subseteq W$  is the set of *possible worlds*;
- $f: \wp W \times W \rightarrow \wp W$  is the *selection function*;
- $V$  is the *valuation function*, where  $V: \text{Prop} \times P \rightarrow \{0, 1\}$  and  $V: \mathcal{L}_0 \times \overline{P} \rightarrow \{0, 1\}$ .

Intuitively,  $f(X, w)$  is the set of worlds “closest” to  $w$  in  $X$ . Various constraints may be placed on  $f$  if desired; e.g., many authors require  $f(X, w) \subseteq X$  (corresponding to  $\phi \Box\rightarrow \phi$ ) or  $w \in f(X, w)$  if  $w \in X$  (corresponding to  $(\phi \Box\rightarrow \psi) \rightarrow (\phi \rightarrow \psi)$ ). Even simple constraints like

these, however, are controversial in the context of counterlogicals [3, 9]. For our purposes, it does not particularly matter what constraints we impose so long as  $f(X, w)$  can contain some impossible worlds, i.e., so long as we don't require that  $f(X, w) \subseteq P$  for all  $X$  and  $w$ .

Valuation functions determine the truth of *every* formula at impossible worlds. This is because, in order to model logic talk, the impossible worlds semantics needs to appeal to *logically* impossible worlds. In general, though, there is no single rule for determining the truth of complex formulas from the truth of atomics that applies to every logically impossible world. When it comes to the logically impossible, anything goes: there are impossible worlds governed by logics where conjunction is equivalent to disjunction, where negations are redundant, or even where everything is true. Any collection of formulas constitute a world and can be said to conform to some wacky logic or other. Such impossible worlds are *strange*, for sure, but nothing in the intuitive conception of an impossible world rules them out. So for impossible worlds, truth must be determined by fiat via the valuation function.

Given an impossible worlds model  $\mathcal{I} = \langle W, P, f, V \rangle$  and a  $w \in W$ , we define *satisfaction*  $\Vdash_{\mathbf{I}}$  as follows. First, if  $w \in P$ , then  $\mathcal{I}, w \Vdash_{\mathbf{I}} \phi$  iff  $V(\phi, w) = 1$ . Second, if  $w \in P$ , then satisfaction is defined recursively (where  $\llbracket \phi \rrbracket^{\mathcal{I}} = \{u \in W \mid \mathcal{I}, u \Vdash_{\mathbf{I}} \phi\}$ ):

$$\begin{aligned}
\mathcal{I}, w \Vdash_{\mathbf{I}} p &\quad \Leftrightarrow \quad V(p, w) = 1 \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \neg \phi &\quad \Leftrightarrow \quad \mathcal{I}, w \not\Vdash_{\mathbf{I}} \phi \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \phi \wedge \psi &\quad \Leftrightarrow \quad \mathcal{I}, w \Vdash_{\mathbf{I}} \phi \text{ and } \mathcal{I}, w \Vdash_{\mathbf{I}} \psi \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \phi \vee \psi &\quad \Leftrightarrow \quad \mathcal{I}, w \Vdash_{\mathbf{I}} \phi \text{ or } \mathcal{I}, w \Vdash_{\mathbf{I}} \psi \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \phi \rightarrow \psi &\quad \Leftrightarrow \quad \mathcal{I}, w \Vdash_{\mathbf{I}} \phi \text{ only if } \mathcal{I}, w \Vdash_{\mathbf{I}} \psi \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \Box \phi &\quad \Leftrightarrow \quad \text{for all } v \in P: \mathcal{I}, v \Vdash_{\mathbf{I}} \phi \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \Diamond \phi &\quad \Leftrightarrow \quad \text{for some } v \in P: \mathcal{I}, v \Vdash_{\mathbf{I}} \phi \\
\mathcal{I}, w \Vdash_{\mathbf{I}} \phi \Box \rightarrow \psi &\quad \Leftrightarrow \quad f(\llbracket \phi \rrbracket^{\mathcal{I}}, w) \subseteq \llbracket \psi \rrbracket^{\mathcal{I}}.
\end{aligned}$$

Finally, given a set  $\Gamma \subseteq \mathcal{L}_0$  and a  $\phi \in \mathcal{L}_0$ , we say  $\Gamma$  *I-entails*  $\phi$ , or  $\Gamma \models_{\mathbf{I}} \phi$ , if for every impossible worlds model  $\mathcal{I} = \langle W, P, f, V \rangle$  and every  $w \in P$ , if  $\mathcal{I}, w \Vdash_{\mathbf{I}} \Gamma$ , then  $\mathcal{I}, w \Vdash_{\mathbf{I}} \phi$ . In other words, consequence is satisfaction-preservation over *possible* worlds. This ensures that  $\models_{\mathbf{I}}$  is an extension of classical logic even though the constituents of counterfactuals may behave nonclassically. Thus,  $\models_{\mathbf{I}} p \vee \neg p$ , even though  $\not\models_{\mathbf{I}} p \Box \rightarrow (p \vee \neg p)$ , since the closest  $\phi$ -worlds may include some impossible worlds  $w$  where  $V(p \vee \neg p, w) = 0$ .

### 3 Adding Propositional Quantifiers

Now we will consider extending our base language  $\mathcal{L}_0$  with propositional quantifiers. Why? Abstractly, of course, it would be unfortunate for the impossible worlds semantics if it could not be extended with propositional quantifiers. But also, there are several reasons specific to logic talk for introducing propositional quantifiers into the language. I will mention two.

First, propositional quantifiers are useful for regimenting *laws* of logic. As it stands, the best we can do in  $\mathcal{L}_0$  is pick an arbitrary propositional variable  $l$  for each law and simply stipulate, in the metalanguage, that  $l$  stands for that law. This is less than ideal. We would like to capture the sense in which, say, the law of excluded middle is a valid principle by saying not just that it is *true* but that it is *logically necessary*. So where *lem* stands for the law of excluded middle, we would like to say that  $\models_{\mathbf{I}} \text{lem}$ . Propositional variables are logically contingent, however, in that there are no constraints on what truth values we can assign to them. Unless we impose further *ad hoc* constraints on our class of models, it won't be the case that  $\models_{\mathbf{I}} \text{lem}$ .

Second, some logic talk seems to require propositional quantifiers. Here are some examples:

- (4) No contradiction is true.
- (5) Everything that is intuitionistically valid is classically valid.
- (6) Some propositions are neither true nor not true.
- (7) If the Liar were both true and not true, everything would be true.

What's more, we need propositional quantifiers to make familiar distinctions between *de dicto* and *de re* counterfactuals. Contrast the following:

- (8) a. There is a contradiction such that, if it were true, everything would be true and not true.
- b. If there were a true contradiction, everything would be true and not true.

These do not seem equivalent. Intuitively, (8-a) seems true: e.g., the conjunction of “Everything would be true and not true” and its negation would be one such contradiction. But (8-b) seems false (or at least could be false): if there were a true contradiction, paraconsistent logic would be correct, so not everything would be true and not true. Regimenting laws as brute atomic formulas leaves us ill-equipped for distinguishing such counterfactuals.

Propositional quantifiers allows us to avoid these problems. Instead of choosing an arbitrary propositional variable to represent the law of excluded middle, we could regiment it as a universally quantified claim, viz.,  $p \vee \neg p$  necessarily holds *for any proposition*  $p$ :

$$\forall p \Box(p \vee \neg p).$$

Assuming we define the semantics correctly, we will not need to stipulate that this formula is valid: it will simply fall out of the semantics for propositional quantifiers that  $\models \forall p \Box(p \vee \neg p)$ . This can be so even if counterfactuals are hyperintensional, and so non-trivially embed in counterfactuals. So writing the law of non-contradiction as  $\forall p \Box \neg(p \wedge \neg p)$ , (2) becomes:

$$(l \wedge \neg l) \Box \rightarrow \neg \forall p \Box \neg(p \wedge \neg p).$$

Propositional quantifiers also make it easier to regiment the quantified examples above. For instance, (4) and (7) could be regimented respectively as:

$$\begin{aligned} & \neg \exists p(p \wedge \neg p) \\ (l \wedge \neg l) \Box & \rightarrow \forall p p. \end{aligned}$$

Moreover, the distinction between (8-a) and (8-b) can be captured using scope:

$$\begin{aligned} & \exists p((p \wedge \neg p) \Box \rightarrow \forall q(q \wedge \neg q)) \\ & (\exists p(p \wedge \neg p) \Box \rightarrow \forall q(q \wedge \neg q)). \end{aligned}$$

Thus, it makes sense to consider extending  $\mathcal{L}_0$  with propositional quantifiers. So let's extend  $\mathcal{L}_0$  to a language  $\mathcal{L}_Q$  with propositional quantifiers  $\forall p$  and  $\exists p$  binding into sentence position:

$$\phi ::= p \mid \neg \phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid (\phi \rightarrow \phi) \mid \Box \phi \mid \Diamond \phi \mid (\phi \Box \rightarrow \phi) \mid \forall p \phi \mid \exists p \phi.$$

The simplest semantics for modal logic with propositional quantifiers (**S5** $\pi$ +) interprets the quantifiers as ranging over arbitrary *sets of worlds* [4]. The semantics looks something like this:

$$\begin{aligned} \mathcal{M}, w \Vdash \forall p \phi & \Leftrightarrow \text{for all } X \subseteq W: \mathcal{M}_X^p, w \Vdash \phi \\ \mathcal{M}, w \Vdash \exists p \phi & \Leftrightarrow \text{for some } X \subseteq W: \mathcal{M}_X^p, w \Vdash \phi \end{aligned}$$

where  $\mathcal{M}_X^p = \langle W, R, V_X^p \rangle$  and  $V_X^p$  is exactly like  $V$  except that  $V_X^p(p) = X$ . In words,  $\forall p\phi$  is true just in case  $\phi$  comes out true on any way of interpreting  $p$ , where “a way of interpreting  $p$ ” is just an assignment of  $p$  to a possible worlds proposition (i.e., a set of worlds).

A natural strategy for extending the impossible worlds semantics to  $\mathcal{L}_Q$  is to simply import the **S5** $\pi$ + semantic entries directly. Thus, where  $w \in P$ , where  $\mathcal{I}_X^p = \langle W, P, f, V_X^p \rangle$ , and where  $V_X^p$  is exactly like  $V$  except that for all  $w \in W$ ,  $V_X^p(p, w) = 1$  iff  $w \in X$ :

$$\begin{aligned} \mathcal{I}, w \Vdash_{\mathbf{I}} \forall p\phi &\Leftrightarrow \text{for all } X \subseteq W: \mathcal{I}_X^p, v \Vdash_{\mathbf{I}} \phi \\ \mathcal{I}, w \Vdash_{\mathbf{I}} \exists p\phi &\Leftrightarrow \text{for some } X \subseteq W: \mathcal{I}_X^p, v \Vdash_{\mathbf{I}} \phi \end{aligned}$$

Notice, however, that as it’s defined,  $V_X^p$  only differs from  $V$  in the interpretation of  $p$ .  $V_X^p$  does not differ from  $V$  on *complex* formulas involving  $p$ , at least at impossible worlds: for instance, if  $w \in \bar{P}$ , then  $V_X^p(\neg p, w) = V(\neg p, w)$  regardless of  $X$ . In the **S5** $\pi$ + semantics, this doesn’t matter since the interpretation of a complex formula involving  $p$  is recursively determined from the interpretation of  $p$ . But in the impossible worlds semantics, truth at impossible worlds is not determined in a recursive manner: it’s determined *by fiat* by the valuation function. So we cannot, in general, determine how to change the interpretation of, say,  $\neg p$  or  $p \vee q$  at an arbitrary impossible world when we change the interpretation of  $p$ .

Not only is this counterintuitive, it leads to formal difficulties. For instance, the following is valid on the current semantics for  $\mathcal{L}_Q$ :

$$\exists p((p \wedge \neg p) \Box \rightarrow q) \rightarrow \forall p((p \wedge \neg p) \Box \rightarrow q).$$

Since  $\llbracket p \wedge \neg p \rrbracket^{\mathcal{I}} \subseteq \bar{P}$  for any  $\mathcal{I}$  and since  $V_X^p(p \wedge \neg p, w) = V(p \wedge \neg p, w)$  for any  $w \in \bar{P}$ , it follows that  $\llbracket p \wedge \neg p \rrbracket^{\mathcal{I}_X^p} = \llbracket p \wedge \neg p \rrbracket^{\mathcal{I}}$ , and so  $f(\llbracket p \wedge \neg p \rrbracket^{\mathcal{I}_X^p}, w) = f(\llbracket p \wedge \neg p \rrbracket^{\mathcal{I}}, w)$  for any  $X \subseteq W$ . Hence, if  $f(\llbracket p \wedge \neg p \rrbracket^{\mathcal{I}_X^p}, w) \subseteq \llbracket q \rrbracket^{\mathcal{I}_X^p} = \llbracket q \rrbracket^{\mathcal{I}}$  for one  $X \subseteq W$ , it holds for all  $X$ . But this principle is implausible: just because *one* contradiction counterfactually implies  $q$ , it does not follow that *all* contradictions do.

Clearly, the solution to this problem will involve defining  $V_X^p$  so that it differs from  $V$  not just on the interpretation  $p$  but also on the interpretation of complex formulas involving  $p$ . The trouble is that it is not clear how to do this. Again, when it comes to the logically impossible, anything goes. Even if  $p$  is true and  $\neg p$  is false at an impossible world, making  $p$  false at that world does not automatically mean we must make  $\neg p$  true at that world: we might be at a world that allows a sentence and its negation to be false.

Here is one promising line of thought. The problem seems to stem from our conception of impossible worlds as fixed sets of formulas. Perhaps impossible worlds need to also be equipped with a *rule* for how truth is determined, which we could model as a function from an interpretation of the propositional variables to an interpretation of complex formulas. Such rules need not be “natural”; an impossible world governed by a wacky logic will have a wacky rule for determining truth. But however wacky, such a rule will help us interpret propositional quantifiers by making sure we are not at a loss for how to reinterpret complex formulas when we reinterpret propositional variables.

More precisely, define a **variable assignment** over  $W$  to be a function  $g: \text{Prop} \rightarrow \wp W$  mapping each propositional variable to a proposition. Where  $\text{VA}_W$  is the set of variable assignments over  $W$ , let’s redefine a valuation function to be a map  $V: \mathcal{L}_Q \times \bar{P} \times \text{VA}_W \rightarrow \{0, 1\}$  such that  $V(p, w, g) = 1$  iff  $w \in g(p)$ . In other words,  $V$  is a rule for determining what complex formulas are true at an impossible world *given* an interpretation of the propositional variables. Likewise, let’s redefine impossible worlds models as quadruples  $\mathcal{I} = \langle W, P, f, V \rangle$  with our new valuation functions. We now relativize truth to a world and a variable assignment. So if  $w \in \bar{P}$ ,

then  $\mathcal{I}, w, g \Vdash_{\mathbf{I}} \phi$  iff  $V(\phi, w, g) = 1$ . Otherwise, the truth conditions are as before (except relativized to variable assignments) with the following amendments:

$$\begin{aligned} \mathcal{I}, w, g \Vdash_{\mathbf{I}} p &\Leftrightarrow w \in g(p) \\ \mathcal{I}, w, g \Vdash_{\mathbf{I}} \forall p\phi &\Leftrightarrow \text{for all } X \subseteq W: \mathcal{I}, w, g_X^p \Vdash \phi \\ \mathcal{I}, w, g \Vdash_{\mathbf{I}} \exists p\phi &\Leftrightarrow \text{for some } X \subseteq W: \mathcal{I}, w, g_X^p \Vdash \phi. \end{aligned}$$

This semantics does not validate  $\exists p((p \wedge \neg p) \Box \rightarrow q) \rightarrow \forall p((p \wedge \neg p) \Box \rightarrow q)$ , since  $V(p \wedge \neg p, w, g_X^p)$  need not be the same for every  $X \subseteq W$ .

Unfortunately, this proposal faces further problems. Whereas before the impossible worlds semantics validated too much, now it validates too little. For example, it does not validate existential introduction ( $\phi(\chi) \models \exists p\phi(p)$ , subject to the usual restrictions), variable exchange ( $\exists p\phi(p) \models \exists q\phi(q)$ , subject to the usual restrictions), or vacuous quantification ( $\phi \models \forall p\phi$  where  $p$  does not occur free in  $\phi$ ). To illustrate, none of the following are valid:

$$\begin{aligned} ((q \wedge \neg q) \Box \rightarrow r) &\rightarrow \exists p((q \wedge p) \Box \rightarrow r) \\ \exists p(p \Box \rightarrow r) &\rightarrow \exists q(q \Box \rightarrow r) \\ (q \Box \rightarrow r) &\rightarrow \forall p(q \Box \rightarrow r). \end{aligned}$$

To see why, consider the first principle.  $\exists p((q \wedge p) \Box \rightarrow r)$  is true at  $w$  iff for some  $X \subseteq W$ ,  $f(\llbracket q \wedge p \rrbracket^{\mathcal{I}, g_X^p}, w) \subseteq \llbracket r \rrbracket^{\mathcal{I}, g_X^p}$ . But even if  $X = \llbracket \neg q \rrbracket^{\mathcal{I}, g}$ , we cannot be sure that  $f(\llbracket q \wedge p \rrbracket^{\mathcal{I}, g_X^p}, w) = f(\llbracket q \wedge \neg q \rrbracket^{\mathcal{I}, g}, w)$  since it need not be that  $V(q \wedge \neg q, w, g) = V(q \wedge p, w, g_X^p)$ . And this gap can be exploited to construct a counterexample to the first principle.

We could try to impose various constraints on  $V$  to avoid these problems. But it is not obvious how to do this in a systematic fashion. For each such problem, we would need to impose an additional constraint on  $V$  to block that specific problem. Such a gerrymandered approach seems undesirable, to say the least. What's more, it is not clear what would conceptually motivate such constraints apart from the fact that they help avoid these technical problems.

## 4 Logical Expressivism

The previous section outlined a problem with the impossible worlds semantics: the presence of logically impossible worlds makes it difficult to interpret propositional quantifiers. I will now show how a new approach, viz., logical expressivism, does better.

Logical expressivism is motivated by the thought that counterlogicals seem to involve shifts in the meaning of the logical connectives [6]. For example, consider the following inference:

- (9) a. If intuitionistic logic were correct, the continuum hypothesis would not not be true.  
b.  $\therefore$  If intuitionistic logic were correct, the continuum hypothesis would be true.

Intuitively, this inference seems invalid because in the consequent of (9-a), we are interpreting ‘not’ according to an intuitionistic interpretation, which does not validate the law of double negation elimination. If we held fixed the actual, classical meaning of ‘not’, then the inference would be valid. The antecedent ‘if intuitionistic logic were correct’ seems to be triggering a shift in the meaning of ‘not’ from a classical to an intuitionistic one.

This sort of phenomenon arises in ordinary, non-logical examples, too [7]. There is an old joke: if a dog’s tail were called a “leg”, how many legs would a dog have? The answer is supposed to be four—calling a tail a leg doesn’t make it one! At the risk of ruining a (bad) joke by explaining it, the reason this is a joke is that there are two natural readings of (10).



(10) If a dog's tail were called a "leg", a dog would have five legs.

On one reading—call it the *shifty* reading—(10) is true because we interpret 'leg' in the consequent according to the conventions described in the antecedent. On another reading—call it the *rigid* reading—(10) is false because we interpret 'leg' in the consequent according to our actual conventions, on which a tail does not count as a leg. Both readings are available, though the shifty reading seems more salient in the context of the joke; hence why the joke is "funny".

Logical expressivism holds that the same thing is happening in (9). On this view, logic just is a convention governing logical vocabulary. Just as speakers may adopt any number of conventions for how to talk, so too, they may adopt any number of conventions for how to use words like 'not', 'and', 'or', and so on. There is no such thing as the "correct" or "one true" logic, just as there is no such thing as the "correct" or "one true" language. By adopting a non-classical logic, one is not thereby describing things inaccurately. Logic is more a matter of *decision* than *discovery*. But when speakers interpret counterlogicals, they interpret the logical connectives according to nonclassical conventions, even if they adopt classical logic. That is, counterlogicals, on their most natural interpretation, are really *counterconventionals* [6].

We can develop a formal semantics for logical expressivism by adapting another well-known expressivist semantics, viz., the hyperplan semantics for normative discourse due to [5]. Informally, a hyperplan can be thought of as a maximally specific plan, specifying what actions to take in every conceivable situation. Formally, a hyperplan is just a total function from worlds to sets of permissible actions. Gibbard's proposal was to think of normative vocabulary ('ought', 'may', etc.) as being sensitive to hyperplans. Thus, ' $\alpha$  ought to  $\phi$ ' is true relative a world-hyperplan pair  $\langle w, h \rangle$  iff  $\phi$ ing is included amongst the actions in  $h(w)$ .

Since logical expressivism thinks of a logic as a kind of convention, where a convention can be thought of as a kind of plan for how to use words, logics are, effectively, just special kinds of plans. Thus, we can define a *hyperconvention* as a maximally specific plan for how to use words. Formally, we can model hyperconventions as interpretation functions: functions mapping propositional variables, connectives, and so on to intensions. Then a sentence of the form ' $\phi$  is valid' is true relative to a world-hyperconvention pair  $\langle w, c \rangle$  iff  $\phi$  as interpreted according to  $c$  holds in every situation.

To make this more precise, define a *hyperconvention* over  $W$  to be a function  $c$  where:

- for each  $p \in \text{Prop}$ ,  $c(p) \subseteq W$
- for each  $n$ -place  $\Delta \in \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \Box, \Diamond\}$ ,  $c(\Delta): \wp W^n \rightarrow \wp W$ .

In other words, if we think of a proposition as just a set of worlds, hyperconventions map each propositional variable to a proposition and each  $n$ -place connective to an  $n$ -ary operation on propositions.<sup>1</sup> An *index* over  $W$  is a pair  $\langle w, c \rangle$  where  $w \in W$  and  $c$  is a hyperconvention over  $W$ . The set of indices over  $W$  is denoted  $\text{Ind}_W$ .

Now for the semantics. An *expressivist model* is a pair of the form  $\mathcal{E} = \langle W, f \rangle$ , where:

- $W \neq \emptyset$  is the set of *possible worlds*;
- $f: \wp \text{Ind}_W \times \text{Ind}_W \rightarrow \wp \text{Ind}_W$  is the *selection function*.

Notice that expressivist models do not include a valuation function since hyperconventions can play that role. Notice also that  $W$  is described as the set of *possible* worlds. This is because the role of impossible worlds is effectively played by possible worlds-under-descriptions.

<sup>1</sup>Note that this set excludes  $\Box \rightarrow$ . Since  $\Box \rightarrow$  denotes an operation on sets of indices, which themselves contain hyperconventions, we cannot also have hyperconventions interpret  $\Box \rightarrow$  without circularity.

Satisfaction is defined relative to indices, i.e., world-hyperconvention pairs. Thus, given an expressivist model  $\mathcal{E} = \langle W, f \rangle$  and a  $\langle w, c \rangle \in \text{Ind}_W$ , we define *satisfaction*  $\Vdash_{\mathbf{E}}$  as follows (where  $\Delta \in \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \Box, \Diamond\}$ ):

$$\begin{aligned} \mathcal{E}, w, c \Vdash_{\mathbf{E}} p &\Leftrightarrow w \in c(p) \\ \mathcal{E}, w, c \Vdash_{\mathbf{E}} \Delta(\phi_1, \dots, \phi_n) &\Leftrightarrow w \in c(\Delta)(\llbracket \phi_1 \rrbracket^{\mathcal{E}, c}, \dots, \llbracket \phi_n \rrbracket^{\mathcal{E}, c}) \\ \mathcal{E}, w, c \Vdash_{\mathbf{E}} \phi \Box \rightarrow \psi &\Leftrightarrow f(\llbracket \phi \rrbracket^{\mathcal{E}}, w, c) \subseteq \llbracket \psi \rrbracket^{\mathcal{E}}, \end{aligned}$$

where  $\llbracket \phi \rrbracket^{\mathcal{E}} := \{\langle w, c \rangle \in \text{Ind}_W \mid \mathcal{E}, w, c \Vdash_{\mathbf{E}} \phi\}$  and  $\llbracket \phi \rrbracket^{\mathcal{E}, c} := \{w \in W \mid \mathcal{E}, w, c \Vdash_{\mathbf{E}} \phi\}$ .

Finally, consequence is defined as satisfaction-preservation over “classical” indices, i.e., indices that interpret the connectives in the ordinary classical way. More precisely, let’s say a hyperconvention  $c$  is *classical* if the following conditions are met for all  $X, Y \subseteq W$ :

$$\begin{aligned} c(\neg)(X) &= \overline{X} & c(\rightarrow)(X, Y) &= \overline{X} \cup Y \\ c(\wedge)(X, Y) &= X \cap Y & c(\Box)(X) &= \{w \in W \mid X = W\} \\ c(\vee)(X, Y) &= X \cup Y & c(\Diamond)(X) &= \{w \in W \mid X \neq \emptyset\}. \end{aligned}$$

An index is *classical* if its hyperconvention is classical. We let  $\text{CInd}_W$  be the set of classical indices over  $W$ . Then we say  $\Gamma$  ***E**-entails*  $\phi$ , or  $\Gamma \Vdash_{\mathbf{E}} \phi$ , iff for every expressivist model  $\mathcal{E}$  and every *classical*  $\langle w, c \rangle \in \text{CInd}_W$ , if  $\mathcal{E}, w, c \Vdash_{\mathbf{E}} \Gamma$ , then  $\mathcal{E}, w, c \Vdash_{\mathbf{E}} \phi$ . Restricting consequence to truth-preservation over classical indices ensures that  $\Vdash_{\mathbf{E}}$  is classical. In fact, over  $\mathcal{L}_0$ , the logic of logical expressivism exactly matches the logic of the impossible worlds semantics (see [6]):

**Theorem 1.** *For all  $\Gamma \subseteq \mathcal{L}_0$  and  $\phi \in \mathcal{L}_0$ ,  $\Gamma \Vdash_{\mathbf{I}} \phi$  iff  $\Gamma \Vdash_{\mathbf{E}} \phi$ .*

In particular, this means that (i)  $\Vdash_{\mathbf{E}}$  is an extension of classical logic, and (ii) the logical expressivist semantics is hyperintensional, i.e.,  $\neg \Diamond \phi \not\equiv_{\mathbf{E}} \phi \Box \rightarrow \psi$ . Thus, when it comes to  $\mathcal{L}_0$ , the impossible worlds semantics and logical expressivist semantics are on a par: anything one can do, the other can do as well.

But the logical expressivist semantics does significantly better when extended to  $\mathcal{L}_Q$ . Let us say an *expressivist variable assignment* over  $W$  is a function  $g: \text{Prop} \rightarrow \wp \text{Ind}_W$ . Here, then, are the semantic clauses for propositional quantifiers in the logical expressivist framework:

$$\begin{aligned} \mathcal{E}, w, c, g \Vdash_{\mathbf{E}} p &\Leftrightarrow \langle w, c \rangle \in g(p) \\ \mathcal{E}, w, c, g \Vdash_{\mathbf{E}} \forall p \phi &\Leftrightarrow \text{for all } X \subseteq \text{Ind}_W: \mathcal{E}, w, c, g_X^p \Vdash \phi \\ \mathcal{E}, w, c, g \Vdash_{\mathbf{E}} \exists p \phi &\Leftrightarrow \text{for some } X \subseteq \text{Ind}_W: \mathcal{E}, w, c, g_X^p \Vdash \phi. \end{aligned}$$

Let’s say  $\langle w, c, g \rangle$  is *initialized* if  $\langle w, c \rangle \in g(p)$  iff  $w \in c(p)$ . If we define consequence as satisfaction-preservation over initialized (classical) points, we do not get any of the counterintuitive consequences that plagued the impossible worlds semantics. On the one hand,  $\exists p((p \wedge \neg p) \Box \rightarrow q)$  does not entail  $\forall p((p \wedge \neg p) \Box \rightarrow q)$ , since  $\llbracket p \wedge \neg p \rrbracket^{\mathcal{E}, g_X^p}$  need not be the same for all  $X \subseteq \text{Ind}_W$ . On the other hand, this semantics does validate the normal principles governing universal and existential propositional quantifiers, such as existential introduction, variable exchange, and vacuous quantification. So we do validate the principles from the end of § 3 (e.g.,  $\exists p(p \Box \rightarrow r) \rightarrow \exists q(q \Box \rightarrow r)$ ). Thus, the logical expressivist framework already does significantly better than the impossible worlds semantics at interpreting propositional quantifiers.

## 5 Hybrid Logic

In the previous section, we saw that logical expressivism fared better than the impossible worlds semantics when we extended  $\mathcal{L}_0$  to  $\mathcal{L}_Q$ . However, there are two expressive limitations of  $\mathcal{L}_Q$

that affect its ability to regiment logic talk. First,  $\mathcal{L}_Q$  has no way of expressing claims about whole logics, such as ‘intuitionistic logic is correct’. Second,  $\mathcal{L}_Q$  cannot regiment claims about what holds *according to* a particular logic. For instance, even though classical and intuitionistic logicians disagree over (11-a), both parties agree that (11-b) is true:

- (11) a. The law of excluded middle holds.  
b. According to classical logic, the law of excluded middle holds.

Third, while  $\mathcal{L}_Q$  has the resources to regiment *axioms*, it does not have the resources to regiment *rules of inference*. For instance, consider the rule of modus ponens:

$$\phi, \phi \rightarrow \psi \therefore \psi$$

The only way to represent this rule in  $\mathcal{L}_Q$  is as the propositionally quantified sentence:

$$\forall p \forall q \Box((p \wedge (p \rightarrow q)) \rightarrow q).$$

But this treats modus ponens as an axiom, not as a rule. The distinction matters since some logics accept one but not the other. For example, the strong Kleene logic **K3** rejects modus ponens as an axiom ( $\not\vdash_{\mathbf{K3}} (\phi \wedge (\phi \rightarrow \psi)) \rightarrow \psi$ ) but accepts it as a rule ( $\phi, \phi \rightarrow \psi \vdash_{\mathbf{K3}} \psi$ ).

To overcome this last problem, we need a way to distinguish in the object language between *necessarily,  $(p \wedge (p \rightarrow q)) \rightarrow q$  is true* and *necessarily, if  $p$  and  $p \rightarrow q$  are true, then  $q$  is true*. The key difference is that we are using the *if...then...* construction of the (classical) *metalanguage* to state the rule, whereas we are using the *object language*  $\rightarrow$  to state the axiom. Thus, what we would like is access to the classical interpretation of the connectives in the object language—written, say, as  $\sim$ ,  $\&$ ,  $\supset$ , etc.—and then to regiment the rule of modus ponens as:

$$\forall p \forall q \Box((p \& (p \rightarrow q)) \supset q).$$

Intuitively, we want this to say that for any propositions  $p$  and  $q$ , necessarily, if  $p$  and  $p \rightarrow q$  are true, then  $q$  is true—rather than say for any  $p$  and  $q$ , necessarily,  $(p \wedge (p \rightarrow q)) \rightarrow q$  is true.

Fortunately, all three limitations can be overcome by extending to a language with hybrid operators [1]. In short, we’ll extend  $\mathcal{L}_Q$  with an infinite stock of *interpretation variables*  $\mathbf{IVar} = \{i_1, i_2, i_3, \dots\}$  as new atomic formulas (intuitively standing for particular logics), and with two new hybrid operators  $@$  (“according to”) and  $\downarrow$  (a variable binding operator). The resulting syntax  $\mathcal{L}_{QH}$  is summarized as follows:

$$\phi ::= p \mid \neg \phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid (\phi \rightarrow \phi) \mid \Box \phi \mid \Diamond \phi \mid (\phi \Box \rightarrow \phi) \mid \forall p \phi \mid \exists p \phi \mid i \mid @_i \phi \mid \downarrow i. \phi$$

Very roughly, we can think of  $i$  as standing for “ $i$  is the correct logic”,  $@_i \phi$  for “according to  $i$ ,  $\phi$ ”, and  $\downarrow i. \phi$  for “where  $i$  is the current logic,  $\phi$ ”.

It is easy to extend the expressivist semantics to  $\mathcal{L}_{QH}$ . The models are as before. The main difference is that now we will extend variable assignments  $g$  to not only assign each  $p \in \mathbf{Prop}$  to a set of indices, but also each  $i \in \mathbf{IVar}$  to a hyperconvention. The semantics for the new hybrid vocabulary then becomes:

$$\begin{aligned} \mathcal{E}, w, c, g \Vdash_{\mathbf{E}} i &\iff g(i) = c \\ \mathcal{E}, w, c, g \Vdash_{\mathbf{E}} @_i \phi &\iff \mathcal{E}, w, g(i), g \Vdash_{\mathbf{E}} \phi \\ \mathcal{E}, w, c, g \Vdash_{\mathbf{E}} \downarrow i. \phi &\iff \mathcal{E}, w, c, g_c^i \Vdash_{\mathbf{E}} \phi \end{aligned}$$

As before, consequence is defined as satisfaction preservation over initialized classical points. Call the resulting semantics *hyperlogic*.

Hyperlogic has the expressive resources to overcome the limitations mentioned above. For instance, suppose we single out an interpretation variable  $k$  to stand for our initial, classical hyperconvention (we can always rewrite formulas with  $k$  into equivalent formulas without it). Then we can regiment the difference between (11-a) and (11-b) as:

$$\begin{aligned} & \forall p \Box (p \vee \neg p) \\ & @_k \forall p \Box (p \vee \neg p). \end{aligned}$$

To account for the difference between axioms and rules, we can use the binder  $\downarrow$  to grant us access, in the object language, to the classical interpretation of connectives in the scope of operators (such as  $@$ ) that might shift the underlying logic. So we can define  $\sim$ ,  $\&$ ,  $\supset$ , and so on as follows (where  $i$  does not occur in  $\phi$  or  $\psi$ ):

$$\begin{aligned} \sim \phi & := \downarrow i. @_k \neg @_i \phi \\ (\phi \& \psi) & := \downarrow i. @_k (@_i \phi \wedge @_i \psi) \\ (\phi \supset \psi) & := \downarrow i. @_k (@_i \phi \rightarrow @_i \psi). \end{aligned}$$

This means we actually can just regiment the rule of modus ponens as  $\forall p \forall q \Box ((p \& (p \rightarrow q)) \supset q)$ . After a little simplifying, this becomes:

$$\forall p \forall q \Box \downarrow i. ((@_i p \wedge @_i (p \rightarrow q)) \rightarrow @_i q).$$

Thus, hyperlogic offers a promising new framework for regimenting a broad range of logic talk, including claims about axioms, rules, and even entire logics, so that such talk can be compositionally embedded. By contrast, it is unclear how the impossible worlds semantics can be extended to such expressions, given that it simply appeals to impossible worlds. Somewhere, the role of logic has to be made explicit. Hyperlogic is able to achieve its success in large part because it makes the logic governing the connectives play a central role in the semantics.

## References

- [1] Carlos Areces and Balder ten Cate. Hybrid Logics. In Patrick Blackburn, Frank Wolter, and Johan van Benthem, editors, *Handbook of Modal Logic*, pages 821–868. Elsevier, 2006.
- [2] Francesco Berto, Rohan French, Graham Priest, and David Ripley. Williamson on Counterpossibles. *Journal of Philosophical Logic*, 47:693–713, 2018.
- [3] Daniel H Cohen. On What Cannot Be. In *Truth or Consequences*, pages 123–132. Springer Netherlands, Dordrecht, 1990.
- [4] Kit Fine. Propositional quantifiers in modal logic. *Theoria*, 36(3):336–346, 1970.
- [5] Allan Gibbard. *Thinking How to Live*. Harvard University Press, Cambridge, MA, June 2003.
- [6] Alexander W. Kocurek and Ethan Jerzak. Counterlogicals as counterconventionals. Manuscript, 2019.
- [7] Alexander W. Kocurek, Ethan Jerzak, and Rachel Rudolph. Against conventional wisdom. Manuscript, 2019.
- [8] David K Lewis. *Counterfactuals*. Blackwell Publishing, 1973.
- [9] Daniel Nolan. Impossible Worlds: A Modest Approach. *Notre Dame Journal of Formal Logic*, 38(4):535–572, 1997.
- [10] Robert C Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, pages 98–112. Basil Blackwell Publishers, Oxford, 1968.

# Indicative and Subjunctive Conditionals in Commitment Spaces

Manfred Krifka<sup>1</sup>

Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS)  
and Humboldt Universität zu Berlin  
krifka@leibniz-zas.de

## Abstract

The paper argues for a treatment of conditional sentences as conditional speech acts. It provides a formal implementation in the framework of commitment spaces, arguing that this approach has advantages over the conditional proposition account, as it motivates the known restrictions for embedded conditionals. It also introduces an extended model of commitment spaces for subjunctive conditionals, it shows how they affect revisionary updates, and it indicates ways to deal with the problem for the propositional account pointed out by Alonso-Ovalle and Tichý.

## 1 Conditional Propositions or Conditional Speech Acts?

There are two general approaches to the interpretation of conditional sentences:

- (1) *If Fred was at the party, the party was fun.*

One approach analyzes indicative conditionals like (1) as conditional propositions (CP). For example, Stalnaker (1968) interprets *if  $\varphi$  then  $\psi$* , where  $\varphi$ ,  $\psi$  denote propositions  $\varphi$ ,  $\psi$  (functions from indices  $i$  to truth values), as a proposition  $\lambda i[\psi(\max(i, \varphi))]$ , where  $\max(i, \varphi)$  is the index that is maximally similar to  $i$  such that  $\varphi$  is true at  $i$ . Hence (1) is true at an index  $i$  iff for the index  $i'$  that is maximally similar to  $i$  except that Fred was at the party at  $i'$ , it holds that the party was fun at  $i'$ . There are further developments of this view, e.g. Lewis (1973), Kratzer (1981) and much subsequent work, especially in linguistic semantics. This tradition can explain why conditionals occur as embedded clauses in positions like propositional attitude predicates:

- (2) *Wilma believes that if Fred was at the party, the party was fun.*

However, there are syntactic slots for propositional expressions where conditionals cannot occur, or are at least very hard to interpret, e.g. in the protasis of another conditional, cf. Gibbard (1981):

- (3) *#If Kripke was there if Strawson was (there), then Anscombe was there.*

The other approach takes (1) to be a conditional assertion, and conditionals in general as conditional speech acts (CA): “An affirmation of the form ‘if  $p$ , then  $q$ ’ is commonly felt less

---

<sup>1</sup>Parts of this paper were presented at ZAS, as the Annual Logic Lecture at the University of Connecticut and at the Workshop of Forschergruppe 1783 at the University of Frankfurt in 2017, furthermore at the University of British Columbia and at the University of Nantes in 2018 at at EHESS Paris in 2019. The author is grateful for many discussions at these venues, Clemens Steiner-Mayr, Stefan Kaufmann, Magda Kaufmann, Cleo Condoravdi, Sabine Iatridou, Frank Veltman, François Recanatì, Hedde Zeijlstra, Carina Kauf, Hadil Karawani and to the three anonymous reviewers of the Amsterdam Colloquium 2019. Research on this topic was supported by DFG project PaTMO “Past Tense Morphology in Tense and Modality” and the ERC 787929 SPAGAD “Speech Acts in Grammar and Discourse”. I dedicate this paper to the memory of Arthur Merin and Susan Rothstein.

as an affirmation of a conditional than as a conditional affirmation of the consequent.” (Quine 1950). This view is popular in philosophy of language, cf. Barker (1995), Edgington (1995) and subsequent work. One point for the CA analysis is that conditionals are hard to interpret in certain positions where other propositions are fine, cf. (3). Another is that the apodosis of conditionals can be filled by speech acts other than assertions, like questions, exclamatives and directives, cf. (4).

- (4) *If Fred was at the party, was it fun? / how fun it must have been! / tell me more about it!*

The CA analysis realizes the insight of Peirce / Ramsey that conditionals involve temporary assumptions allowing for conditional assertions and other argumentative moves. And it appears to have a lot of intuitive appeal, even for proponents of the CP analysis.

“...[T]he consequent of a conditional proposition asserts what is true, not throughout the whole universe of possibilities considered, but in a subordinate universe marked off by the antecedent.” (Peirce in the Grand Logic [1893-4]; Collected Papers 4.435)

“While there are some complex constructions with indicative conditionals as constituents, the embedding possibilities seem, intuitively, to be highly constrained. For example, simple disjunctions of indicative conditionals with different antecedents, and conditionals with conditional antecedents are difficult to make sense of. The proponent of a non-truth-conditional account needs to explain what embeddings there are, but the proponent of a truth-conditional account must explain why embedded conditionals don’t seem to be interpretable in full generality.” (Stalnaker 2011).

The current paper proposes a semantic representation for the CA view, and argue that it should be considered a viable option in linguistic semantics. This will be done using Commitment Spaces as developed by Krifka (2015), a format for representing different kinds of speech acts.

## 2 Commitment Spaces

The current paper will make use of a somewhat simplified version of the framework of Krifka (2015), who introduced as basic notion “commitment states” as sets of propositions; here I will work with context sets  $c$  in the sense of Stalnaker (1974), i.e. sets of indices that represent the information considered to be shared by the interlocutors. In addition, the ways how this shared information  $c$  can develop at a particular point in conversation to other context sets  $c'$ , with  $c' \subset c$ , will be represented as well. We will assume sets of context states  $C$ , called “commitment spaces” (CS). The context sets  $c$  in  $C$  with minimal information, those for which there is no  $c' \in C$  such that  $c \subset c'$ , are special insofar as they represent the shared factual information, or “root” of  $C$ , written  $\sqrt{C}$ , cf. (4.a). Ideally,  $\sqrt{C}$  is a singleton set containing the information that is the classical common ground, cf. the illustration in Figure 1. Clauses are interpreted as propositions  $\phi$ , which are sets of indices like context sets  $c$ . Propositions can be turned to assertive updates of context spaces

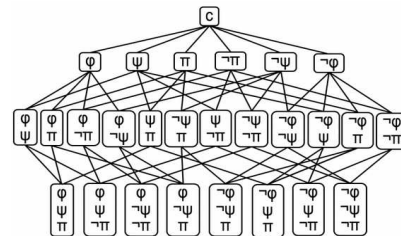


Figure 1. Commitment space  $C$  with root  $\sqrt{C} = \{c\}$ , where  $\phi, \psi, \pi$  are logically independent propositions;  $\boxed{\phi}$  stands for  $c \cap \phi$ .

by a function “ $\cdot$ ” as in (4.b), cf. the illustration in Figure 2. Notice that assertive updates of  $C$  by  $\cdot \varphi$  restricts  $C$  to those context sets  $c$  for which the proposition  $\varphi$  holds. Update by functions  $A$  will be written as in (4.c).

- (4) a.  $\sqrt{C} := \{c \in C \mid \exists c' \in C [c \subseteq c']\}$
- b.  $\cdot \varphi := \lambda C \{c \in C \mid c \subseteq \varphi\}$
- c.  $C + A := A(C)$

Update functions in general are closed under the operations of dynamic and Boolean conjunction, of disjunction and of denegation  $;$ ,  $\&$ ,  $\vee$ ,  $\sim$  as defined by functional and set-theoretic operations, cf. (5.a,b,c,d), and illustrated in Figures 3, 4, 5. Notice that a disjunction of assertive updates leads to a context space with a multiple root.

- (5) a.  $[A ; B] := \lambda C. B(A(C))$
- b.  $[A \& B] := \lambda C [A(C) \cap B(C)]$
- c.  $[A \vee B] := \lambda C [A(C) \cup B(C)]$
- d.  $\sim A := \lambda C [C - A(C)]$

Interrogative updates, e.g. the question if  $\varphi$  is true, are defined as in (6.a). In contrast to assertive updates, they do not change the root of the input CS but reduce the continuations. The alternative question whether  $\varphi$  or  $\psi$  is true can be rendered as in (b), and the question whether  $\varphi$  or not  $\varphi$  is true as in (c), cf. Figure 6.

- (6) a.  $?\varphi := \sqrt{C} \cup C + \cdot \varphi$
- b.  $[?\varphi \vee ?\psi] = \sqrt{C} \cup C + \cdot \varphi \cup C + \cdot \psi$
- c.  $[?\varphi \vee ?\neg\varphi] = \sqrt{C} \cup C + \cdot \varphi \cup C + \cdot \neg\varphi$

The disjunction of two interrogative updates does not result in a multiply-rooted CS; rather, the root of the input CS does not change at all, reflecting the fact that questions do not add information but restrict the possible continuations of the conversation. Krifka (2015) assumes in addition that in assertions the speaker  $s$  declares commitment for a proposition  $\varphi$ , resulting in an update with a proposition  $\cdot s \vdash \varphi$ , whereas in a question the speaker requests the addressee  $a$  to declare commitment for a proposition, resulting in updates like  $?a \vdash \varphi$  for monopolar questions or  $[?a \vdash \varphi \vee ?a \vdash \neg\varphi]$  for bipolar questions. The expected move is that the addressee performs one of these commitments, or rejects the proposal with some other move.

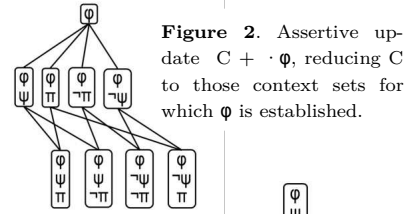


Figure 2. Assertive update  $C + \cdot \varphi$ , reducing  $C$  to those context sets for which  $\varphi$  is established.

Figure 3.  $C + [\cdot \varphi ; \cdot \psi]$ ,  $C + [\cdot \varphi \& \cdot \psi]$

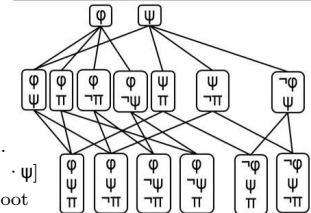


Figure 4.  $C + [\cdot \varphi \vee \cdot \psi]$  multiple root

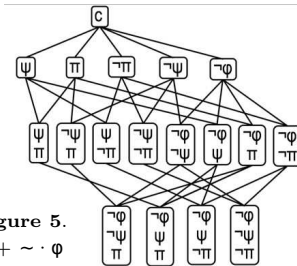


Figure 5.  $C + \sim \cdot \varphi$

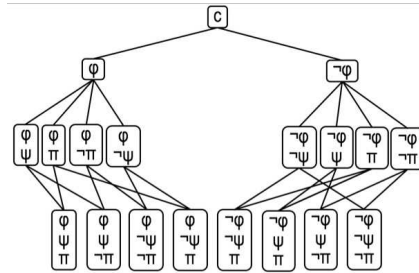


Figure 6.  $C + [?\varphi \vee ?\neg\varphi]$

### 3 Conditionals in Commitment Spaces

Under the CA analysis of conditionals, the apodosis is an update. We define the notion of a conditional update  $C + [A \Rightarrow B]$  as an update by  $B$  that involves only the part  $C+A$ . This

can be expressed as in (7.a) or, disregarding anaphoric bindings from A to B, (7.b); cf. Figure 7.

- (7) a.  $[A \Rightarrow B] := [\sim A \vee A; B]$
- b.  $[A \Rightarrow B] := [\sim A \vee B]$

Conditional update uses the denegation operator  $\sim$  to deal with the protasis of the conditional. However, the protasis of natural-language conditionals is not a speech act but a proposition. Notice that the protasis cannot accommodate speechact-related adverbs, cf. *If Fred (\*presumably) was at the party, the party was fun*. Also, in German the protasis has the verb-final word order characteristic for embedded propositions, cf. *Wenn Fred da war, dann ...hat die Party Spaß gemacht* ‘If Fred was there, the party was fun’. Hence we assume that the protasis is a proposition, and the apodosis is a speech act. This calls for the a definition for update as in (9).

- (8)  $[\text{if } \varphi, A] := \lambda C [\{c \in C \mid c \not\subseteq \varphi\} \cup C + A],$   
 $= [\sim \cdot \varphi \vee A]$

The conditional assertion like *if Fred was at the party, it was fun* restricts the commitment space  $C$  in such a way that whenever the proposition ‘Fred was at the party’ is established, the speaker is committed to the proposition ‘the party was fun’, cf. Figure 7. The conditional question *if Fred was at the party, was it fun or not?* restricts  $C$  in such a way that whenever ‘Fred as at the party’ is established, the only continuations are that the addressee commits to ‘the party was fun’ or to its negation, cf. Figure 8.

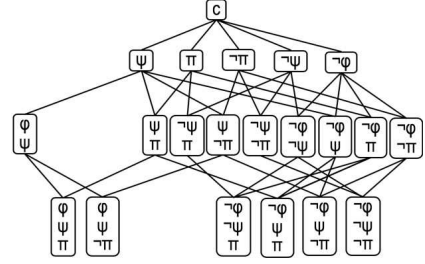


Figure 7.  $C + [\cdot \varphi \Rightarrow \cdot \psi].$   
 $C + [\text{if } \varphi, \cdot \psi]$

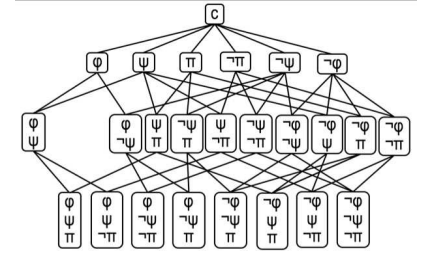


Figure 8.  $C + [\cdot \psi \Rightarrow [? \psi \vee ? \neg \varphi]]$

## 4 Embedding of conditional assertions

In this section we will discuss the embedding of conditional sentences in larger constructions, which is sometimes possible, and restricted at other times (cf. Stalnaker 2011). We will compare how the analysis as conditional assertions fares in comparison with conditional propositions.

### 4.1 Conjunction and Disjunction

Conjunction of conditionals is straightforward and can be modelled by dynamic or Boolean conjunction on updates. This predicts transitivity for conditional assertions, cf. (9). Let  $C$  be updated to  $C'$  by the Boolean conjunction of  $[\text{if } \varphi, \cdot \psi]$  and  $[\text{if } \psi, \cdot \pi]$ , then it also holds that  $[\text{if } \varphi, \cdot \pi]$  is established in  $C'$ , that is,  $C' + [\text{if } \varphi, \cdot \pi] = C'$ . For the CP analysis we need a stipulation for transitivity: We have  $[\varphi > \psi] \wedge [\psi > \pi] = \lambda i [\psi(\max(i, \varphi)) \wedge \pi(\max(i, \psi))]$  and  $[\varphi > \pi] = \lambda i [\pi(\max(i, \varphi))]$ ; transitivity  $[[\varphi > \psi] \wedge [\psi > \pi]] \subseteq [\varphi > \pi]$  is guaranteed only if  $\max(i, \varphi) = \max(i, \psi)$ .

- (9)  $C + [[\text{if } \varphi, \cdot \psi] \& [\text{if } \psi, \cdot \pi]] \subseteq C + [\text{if } \varphi, \cdot \pi]$



Disjunction of conditionals is known to be problematic, as the results are often hard to make sense of (cf. Barker 1995, Abbott 2004, Stalnaker 2011). Take the example by Edgington (1995):

- (10) *If you open Box A you will get ten pounds, or if you open Box B you will get a button.*

Under the CA analysis as developed here we find that  $[[\text{if } \varphi, \cdot \psi] \vee [\text{if } \varphi', \cdot \psi']]$  is equivalent to  $[[\text{if } \varphi, \cdot \psi'] \vee [\text{if } \varphi', \cdot \psi]]^2$ , that is, the protases can be swapped. This is confusing, as the particular grouping of clauses should be informative. The CP analysis should not have a problem with (10),  $[\varphi > \psi] \vee [\varphi' > \psi']$  is straightforwardly interpreted as  $\lambda i[\psi(\max(i, \varphi)) \vee \psi'(\max(i, \varphi'))]$ .

However, under certain conditions disjunctive conditionals are interpretable easily, as in (11), cf. Barker (1995). Notice that this sentence states unconditionally that the check will arrive today or tomorrow. It then gives the additional information that if George has put it into the mail, it will arrive today, and that if he hasn't, it will arrive tomorrow. Hence (11) is not a disjunction of conditionals, but rather has the structure  $\cdot [\psi \vee \psi'] \& [\text{if } \varphi; \cdot \psi] \& [\text{if } \neg \varphi; \cdot \psi']$ . The prosodic realization, with deaccented conditional clauses, helps to create this interpretation.

- (11) *The check will arrive today, if George has put it into the mail, or it will come with him tomorrow, if he hasn't.*

The problem of (10) should also not arise when the apodosis is the same for both protases; this predicts that (12.a) should be fine. However, (12.a) turns out to be equivalent to the shorter (12.b), cf. (13.a), which appears to disfavour (12.a).

- (12) a. *If you open Box A you get ten pounds or if you open Box B you get ten pounds.*  
 b. *If you open Box A and you open Box B you get ten pounds.*  
 c. *If you open Box A or if you open Box B, you get ten pounds.*  
 d. *If you open Box A you get ten pounds and if you open Box B you get ten pounds.*

Due to  $[\cdot \varphi \& \cdot \psi] = \cdot [\varphi \wedge \psi]$  and  $[\cdot \varphi \vee \cdot \psi] \subseteq \cdot [\varphi \vee \psi]^3$ , we have the logical relationships in (13). Due to (13.a), (12.a) is equivalent to (12.b) and (12.c). Due to (13.b), (12.c) has (12.d) as a close paraphrase under a propositional interpretation of the disjunction, though not as an equivalence (here,  $A \subseteq B$  holds iff for all C, for  $A(C) \subseteq B(C)$ ).

We have true equivalence if (12.c) is interpreted following the scheme  $[\sim[\cdot \varphi \vee \cdot \psi] \vee \cdot \pi]$ .

- (13) a.  $[\text{if } [\varphi \wedge \psi], \cdot \pi] = [[\text{if } \varphi, \cdot \pi] \vee [\text{if } \psi, \cdot \pi]] = [[\sim \cdot \varphi] \vee [\sim \cdot \psi]] \vee \cdot \pi^4$   
 b.  $[\text{if } [\varphi \vee \psi], \cdot \pi] \subseteq [[\text{if } \varphi, \cdot \pi] \& [\text{if } \psi, \cdot \pi]] = [\sim[\cdot \varphi \vee \cdot \psi] \vee \cdot \pi]^5$

DP coordination like *you open Box A or/and Box B* might express narrow-scope propositional or wider-scope speech-act coordination. e.g. *If you open Box A and Box B...* may be interpreted as  $[[\text{if } \varphi, \cdot \pi] \& [\text{if } \psi, \cdot \pi]]$ . Several issues remain to be investigated that will not be pursued in this paper, e.g. the role of scalar implicature, but cf. (30.a,b) for subjunctive conditionals.

<sup>2</sup>Due to commutativity and associativity of disjunction,  $[[\sim \cdot \varphi \vee \cdot \psi] \vee [\sim \cdot \varphi' \vee \psi']] = [[\sim \cdot \varphi' \vee \cdot \psi] \vee [\sim \cdot \varphi \vee \psi']]$

<sup>3</sup>Note that  $[\cdot \varphi \vee \cdot \psi]$  may have a multiple root, cf. Figure 4, whereas  $\cdot [\varphi \vee \psi]$  includes nodes above this root.

<sup>4</sup>Due to  $[[\sim \cdot \varphi \vee \cdot \pi] \vee [\sim \cdot \psi \vee \cdot \pi]] = [[\sim \cdot \varphi \vee \sim \cdot \psi] \vee \cdot \pi] = [\sim[\cdot \varphi \& \cdot \psi] \vee \cdot \pi] = [\sim \cdot [\varphi \wedge \psi] \vee \cdot \pi]$ .

<sup>5</sup>Due to  $[\sim[\cdot \varphi \vee \psi] \vee \cdot \pi] \subseteq [\sim[\cdot \varphi \vee \cdot \psi] \vee \cdot \pi] = [[\sim \cdot \varphi \& \sim \cdot \psi] \vee \cdot \pi] = [[\sim \cdot \varphi \vee \cdot \pi] \& [\sim \cdot \psi \vee \cdot \pi]]$ .

## 4.2 Negation

Another semantic operation on conditionals that is notoriously difficult to grasp is negation (cf. Barker 1995, Edgington 1995). In (14.a), negation does not scope over the whole sentence in contrast to (b), which shows that in principle negation can take wide scope over a dependent clause.

- (14) a. *The party was not fun if Fred was there.*  
 b. *The party was not fun because Fred was there (but because there was no beer.)*

This begs for explanation in the CP view, as  $[\Phi > \Psi]$  can be negated, resulting in  $\lambda i \neg[\Psi(\text{ms}(i, \Phi))]$ . The CA view predicts lack of propositional negation, as this negation could not take scope over an update. However, cases with wide-scope negation have been discussed by Barker (1995):

- (15) a. *It is not the case that if God is dead, then everything is permitted.*  
 b. *If God is dead, then everything is NOT permitted.*

Barker suggests an analysis in terms of a metalinguistic negation that rejects the claim made by the non-negated assertion. This negation cannot be expressed by denegation, as  $\sim[\text{if } \Phi, \cdot \Psi]$  is equivalent to  $[\cdot \Phi \ \& \ \sim \cdot \Psi]$ , hence (15) would mean that God is dead and it is ruled out that everything is permitted. Rather, we assume the weak negation that Punčochář (2015) has proposed for inquisitive semantics. It can be expressed by a combination of dynamic possibility and denegation, cf. (17), where  $C + \diamond A$  returns  $C$  iff  $C + A$  is defined, cf. (16). Independent evidence for this type of negation comes from interactions like  $S_1$ : *The number 37753 is certainly prime.*  $S_2$ : *No, it might just have very high prime factors,* where *no* expresses possible falsehood of the antecedent.

$$(16) \ \diamond A := \lambda C. C + A \neq \emptyset [C]$$

$$(17) \ C + \diamond \sim[\text{if } \Phi, \cdot \Psi] = C \text{ if } \exists c \in C[\Phi \subseteq c \wedge \Psi \not\subseteq c], \text{ else undefined}$$

Egré & Politzer (2013), in an experimental study of conditionals that are rejected by *No*, distinguish between three kinds of negation within the CP framework. However, we can work with just one negation,  $\diamond \sim$ , with different kinds of explanation why the negation holds.

- |   |  |
|---|--|
| (18) $S_1$ : <i>If it is a square chip, it will be black.</i>     | $[\text{if } s(c), \cdot b(c)], = A$                                 |
| $S_2$ : <i>No. (i) There are square chips that are not black.</i> | $\diamond \sim A ; \cdot \exists x[s(x) \wedge \neg b(x)]$           |
| (ii) <i>all square chips are not black.</i>                       | $\diamond \sim A ; \cdot \forall x[s(x) \rightarrow \neg b(x)]$      |
| (iii) <i>square chips are not necessarily black.</i>              | $\diamond \sim A ; \cdot \neg \Box \forall x[s(x) \rightarrow b(x)]$ |

## 4.3 Conditional apodosis and conditional protasis

Conditional apodosis clauses are unremarkable, cf. (19), and can be easily modelled within the CA approach, as  $[\text{if } \Phi, [\text{if } \Psi, \cdot \Pi]]$ .

- (19) *If Fred was at the party, then if there was beer at the party, the party was fun.*

We have  $[\text{if } \Phi, [\text{if } \Psi, \cdot \Pi]] = [\text{if } [\Phi \wedge \Psi]; \cdot \Pi]^6$ , as it should be. For the CP account, observe that  $[\Phi > [\Psi > \Pi]] = \lambda i[\Pi(\text{max}(\text{max}(i, \Phi), \Psi))]$  and  $[[\Phi \wedge \Psi] > \Pi] = \lambda i[\Pi(\text{max}(i, [\Phi \wedge \Psi]))]$ , so to get equality of the two terms we have to stipulate  $\text{max}(\text{max}(i, \Phi), \Psi) = \text{max}(i, [\Phi \wedge \Psi])$ . There are apparent counterexamples of this rule like (20) by Barker (1995). However, here *even* scopes over the embedded conditional sentence, preventing a conjunction with the first protasis.

- (20) *If Fred is a millionaire, then even if he fails the entry requirement,*

<sup>6</sup>Due to  $[\sim \cdot \Phi \vee [\sim \cdot \Psi \vee \Pi]] = [[\sim \cdot \Phi \vee \sim \cdot \Psi] \vee \Pi] = [\sim[\cdot \Phi \ \& \ \cdot \Psi] \vee \Pi] = [\sim[\Phi \wedge \Psi] \vee \Pi]$

*he would still get the job.*

In contrast, conditional sentences cannot occur in the protasis of another conditional, cf. (3). This is because *if* in the protasis selects proposition, hence conditionals are of the wrong semantic type. However, there are cases in which conditionals in protasis position are fine, as in (21). But notice that this example is naturally read with accent on *broke* and deaccented *if it was dropped*, making this *if*-clause the topic of the whole clause, leading to the interpretation spelled out in (b).

- (21) a. *If the glass broke if it was dropped, it was fragile.*  
 b. ‘If the glass was dropped, then if it broke, it was fragile.’

#### 4.4 Conditionals in propositional attitude contexts

We have seen in (2) that conditional clauses occur in propositional attitude contexts, cf. also (22) for a different set of predicates. This constitutes a strong argument for the CP approach.

- (22) *Fred knows / thinks / assumes / hopes / doubts that if Wilma applies, she will get the job.*

However, there is a line of defense for the CA approach here. Similar to other lexical predicates that come with sortal requirements, propositional attitude contexts can lead to a coercion of an update to a proposition. This is similar to *drink the whole bottle*, which is understood as drink the whole content of the bottle. The coercion of an update A would be to the proposition that A is assertable, where a simple assertion  $\cdot \varphi$  is assertable at an index  $i$  iff  $\varphi$  is true at  $i$ . This means that  $\cdot \varphi$ , short for  $\lambda C\{c \in C \mid c \subseteq \varphi\}$ , is coerced to  $\varphi$ , the function  $\lambda i[\varphi(i)]$ . For conditional updates like  $[\text{if } \varphi, \psi]$  the assertability condition would result in a proposition that is close to one of the CP accounts of propositions, like Stalnaker’s  $\lambda i[\psi(\max(i, \varphi))]$ . This coercion approach would have to be worked out in greater detail, which is not the focus of this paper.

## 5 Subjunctive Conditionals and Generalized CSs

### 5.1 The interpretation of subjunctive conditionals

Indicative conditionals have a pragmatic requirement that their protasis can be asserted at the current CS, as otherwise the update would be uninformative (cf. Veltman 1985: p.181): If  $C + \cdot \varphi = \emptyset$ , then  $C + [\text{if } \varphi, A] = [C \cup C+A] = C$ . Subjunctive conditionals like (23) violate this requirement, as they are uttered felicitously under the assumption that Fred was not at the party.

- (23) *If Fred had been at the party, it would have been fun.*

Classical approaches to subjunctive conditionals assume that they denote propositions that have a truth value, which is defined via a relation of closeness of worlds (cf. Lewis 1973). But just as indicative conditionals, subjunctive conditionals can have other speech acts as their apodosis, and resist certain kinds of embeddings.

- (24) *If Fred had been at the party, would it had been fun? / how fun it would have been!*

(25) #If Kripke would have been there if Strawson had been, then Anscombe was there.<sup>7</sup>

How can we extend the current representation framework to accommodate subjunctive conditionals? A subjunctive conditional [if  $\phi$ , A] should be interpretable at an input commitment space  $C$  even if  $C + \cdot \phi = \emptyset$ . The idea that will be pursued here is that this can be done by relaxing  $C$  to a  $C'$ ,  $C \subset C'$ , such that  $C' + \cdot \phi \neq \emptyset$ . Relaxing should be minimal, that is,  $C'$  should be as similar to  $C$  as possible. This  $C'$  is a hypothetical commitment space that is entertained in case  $\phi$  were true, after which we return to  $C$ . Nevertheless, the hypothetical commitment space might actually become relevant in case  $\phi$  turns out to be true, necessitating a revisionary update.

To work out this idea, we introduce the notion of a “generalized CS” as a pair of an actual CS and a background CS,  $\langle C_a, C_b \rangle$ , where  $C_a$  is a sub-CS of  $C_b$  as defined in (26).

$$(26) C' \leq_{CS} C'' :\Leftrightarrow C' \subseteq C'' \text{ and } \forall c \in C'' [\exists c' [c' \in C' \wedge c \subseteq c'] \rightarrow c \in C']$$

For example, updating the CS  $C$  of Figure 1 by  $[\cdot \phi; \cdot \psi]$  leads to the generalized CS Figure 9, with the actual CS  $C_a$  rendered in bold, and the background CS  $C_b$ , identical to the original  $C$ .

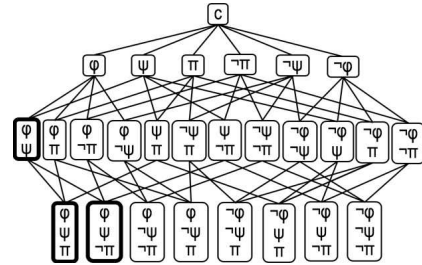


Figure 9.  $C + [\cdot \phi; \cdot \psi] = \langle C_a, C_b \rangle$   
actual and background CS

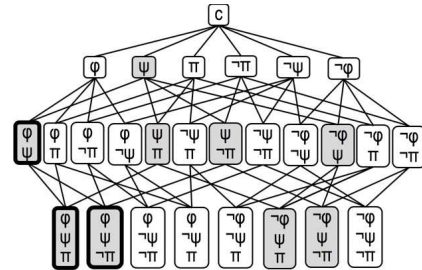


Figure 10. Smallest hypothetical  $C_a'$  such that  $C_a \subseteq C_a' \subseteq C$  and  $C_a' + [\text{if } \neg\phi, \cdot \pi] \neq C_a'$

When we want to update  $\langle C_a, C_b \rangle$  by  $[\text{if } \neg\phi, \cdot \pi]$ , we fail, as updating  $C_a$  with  $\cdot \neg\phi$  would result in the empty actual CS. Hence we assume an hypothetical CS  $C_a'$  that differs from  $C_a$  minimally such  $\cdot \neg\phi$  can be interpreted. This  $C_a'$  is defined as  $\min(C_a, \cdot \neg\phi, C_b)$  by (27). In our example, this is the gray CS in Figure 10.

$$(27) \min(C_a, A, C_b) = \text{the smallest } C \text{ such that } C_a \leq_{CS} C \leq_{CS} C_b \text{ and } C + A \neq \emptyset$$

Updating  $C_a'$  by  $[\text{if } \neg\phi, \cdot \pi]$  leads to the removal of all context sets in which  $\neg\phi$  but not  $\pi$  are established. This does not affect the actual CS  $C_a$  but only the background CS  $C_b$ . After the update with the subjunctive conditional, the resulting generalized CS is as in Figure 11.

In general, we can assume that regular indicative update affects primarily the actual CS and only secondarily the background CS, as it must be guaranteed that the actual CS is a sub-CS of the background CS. This is achieved by (28).

$$(28) \langle C_a, C_b \rangle + A = \langle C_a + A, [C_b - C_a] \cup C_a + A \rangle$$

Subjunctive update, on the other hand, affects primarily the background CS, which can be expressed as in (29). The output background CS  $C$  for a conditional update is defined via the CS  $C^*$  that is the smallest CS between  $C_a$  and  $C_b$  such that the protasis  $\phi$  can be asserted:

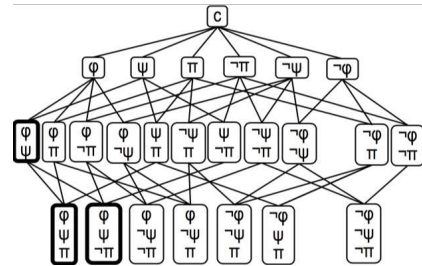


Figure 11. Generalized CS after update with  $[\text{if } \neg\phi, \cdot \pi]$

<sup>7</sup>Gibbard (1981) considers this better than with the indicative case, “Delphic but not incomprehensible”.

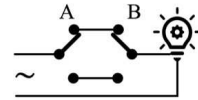
$$(29) \langle C_a, C_b \rangle + [\text{if } \varphi, A] = \langle C_a \cap C, C \rangle$$

where  $C = [C_b - C^*] \cup C^* + [\text{if } \varphi, A]$ , and  $C^* = \min(C_a, \cdot \varphi, C_b)$

Notice that (29) can be taken as the general rule for conditional updates. In case  $C_a + \cdot \varphi \neq \emptyset$ , it holds that  $C^* = C_a$ , and we get the same result as under rule (28), as only the actual input CS  $C_a$  is affected. The use of indicative vs. subjunctive mood indicates whether  $C^* = C_a$  or  $C_a <_{CS} C^*$ . Hence indicative is a morphological index that expresses coreference with the actual CS, whereas subjunctive expresses disjointness with the actual CS.

The current account can explain the experimental findings by Ciardelli et al. (2018) that (30.a) is often judged true whereas (b) is often judged false in the given scenario.

- (30) a. *If Switch A or Switch B was down, the light would be off.*  
 b. *If Switch A and Switch B were not both up, the light would be off.*



Assume that (30.a) is interpreted following the scheme  $[\sim[\cdot \varphi \vee \cdot \psi] \vee \cdot \pi]$ , it has an interpretation in which it is equivalent to a conjunction of two conditionals,  $[\sim \cdot \varphi \vee \cdot \pi]$  &  $[\sim \cdot \psi \vee \cdot \pi]$ , cf. (13.b). Interpreted independently of each other (cf. Alonso-Ovalle 2009), the first conjunct would require a minimal hypothetical CS in which ‘A down’ can be assumed, for which case the light would be off (and similarly for the second conjunct). Hence the judgement that (30.a) is true. On the other hand, (30.b) is interpreted as  $[\sim \cdot \neg[\varphi \wedge \psi] \vee \cdot \pi]$ , equivalent to  $[\sim \cdot [\varphi \vee \psi] \vee \cdot \pi]$ , and requires a minimal hypothetical CS in which the negation of ‘A and B up’ can be assumed, with one prominent option a CS in which ‘A and B down’, for which the light would be on. Hence the judgement that (30.b) is false. Note that one crucial step was the difference in the understanding of the protasis, as  $\sim[\cdot \varphi \vee \cdot \psi]$  or as  $[\sim \cdot [\varphi \vee \psi]]$ ; this is structurally similar to inquisitive lifting in Ciardelli et al. (2018).

The indicative / subjunctive distinction is reminiscent of temporal reference: Just as present tense refers to the actual time of utterance and past tense shifts to some prior time, indicative refers to the actual assumptions of the common ground  $C_a$  and subjunctive shifts to a stage of the common ground development in which certain assumptions are not made. This motivates the observation that subjunctive is often expressed with past-like morphology (cf. Iatridou 2000, Karawani 2014):

- (31) *If Fred was at the party right now, the party would be fun.*

The current proposal leads to a straightforward explanation of the relation between subjunctive and past tense than theories based on closeness of possible worlds: If tense morphology expresses temporal or modal distance from the actual point of reference, then it is not clear why it is past tense and not, for example, future tense is used to express counterfactuality.

## 5.2 Revisionary Updates

In the current setup, subjunctive conditionals only affect the background CS. As communication typically develops in the actual CS, the question arises what subjunctive conditionals contribute to the communication. Intuitively, they express general rules that, due to the subjunctive that requires  $C_a \neq C^*$ , do not have an effect on the part of the common ground that describes the way how the world is. For example, (31) implicates that Fred is not at the party, and nothing follows concerning whether the party was fun. However, if it turns out that Fred is, in fact, at the party, we can conclude that the it is fun. Subjunctive conditionals unfold their inferential power after a revisionary update.

Revisionary update can be seen as a rescue strategy if  $C+A$  results in the empty set. In this case, the input CS  $C$  may be changed minimally to a  $C'$  for which  $C'+A$  is defined. In a

generalized CS framework, revisionary update can be specified as in (32). For example, revisionary update of  $\langle C_a, C_b \rangle$  by  $\cdot \phi$  in Figure 9 results in the generalized CS of Figure 12.

$$(32) \langle C_a, C_b \rangle +_{\text{rev}} A = \langle \min(C_a, A, C_b) + A, C_b \rangle$$

Revisionary update after the subjunctive conditional [if  $\neg\phi, \cdot \pi$ ] of Figure 11 leads to the generalized CS in Figure 13, showing that the subjunctive conditional affects the new actual CS.

### 5.3 A Solution to Tichý’s problem

The current proposal suggests that conditionals do not express propositions about the world but rather statements concerning the assumptions made in a conversation. In this, it is an example of the premise semantics approach to conditionals, cf. Kratzer 1989, Veltman 2005, and Starr 2019 for an overview. According to this approach, subjunctive assertive conditionals adjust a body of premises with the protasis, and assert that the apodosis is a consequence of this revised premise set.

Tichý (1976) pointed out a problem for the modal similarity analysis. Assume that Jones wears a hat if it is raining, and otherwise wears a hat or not at random. Assume furthermore that it is in fact raining (hence Jones wears a hat). Now, is the subjunctive conditional *If it were not raining, Jones would wear a hat* true? Intuitively, it is not true, but the modal similarity analysis asks us to consider a world that is maximally close to the current one except that it is not raining; as wearing a hat is compatible with there being no rain, in this world Jones would wear a hat (see Starr 2019 for discussion).

As it stands, the current model of generalized CSs would run into the same problem. This is because it does not record the way how the actual CS developed. As an example, take the generalized CS  $\langle C, C \rangle$ , where  $C$  is the CS of Figure 1. Update with  $[[\text{if } \phi, \cdot \psi] \ \& \ [\text{if } \neg\phi, [\cdot \psi \vee \neg\psi]]]$  and further update with  $\cdot \phi$  leads to the generalized CS in Figure 14. Notice that this entails the subjunctive update [if  $\neg\phi, \cdot \psi$ ], as the conditional will be interpreted at the CS with root  $\psi$ , as this is the closest CS for which  $\neg\phi$  can be interpreted, and in this CS the conditional [if  $\neg\phi, \cdot \psi$ ] holds.

What is necessary is that by asserting [if  $\phi, \cdot \psi$ ] at  $\langle C, C \rangle$ , the root of the resulting CS,  $C+[\text{if } \phi, \cdot \psi]$ , i.e. the node  $\boxed{\phi\psi}$ , gets as immediate predecessor the node  $\boxed{C}$  itself, the root of the input CS  $C$ . This results in the generalized CS of Figure 15. Notice that the subjunctive update [if  $\neg\phi, \cdot \psi$ ] is not already established here, as the root of the

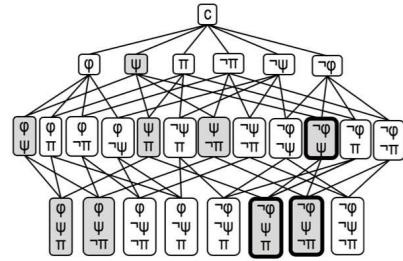


Figure 12. Revisionary update of generalized CS in Fig. 9 with  $\cdot \neg\phi, \max(C_a, \cdot \neg\phi, C_b)$  grayed out

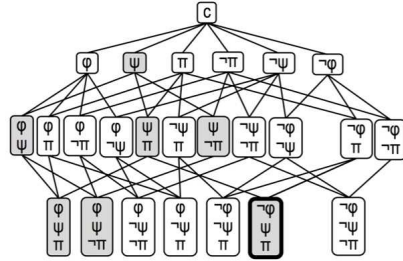


Figure 13. Revisionary update of generalized CS in Fig. 11 with  $\cdot \neg\phi$

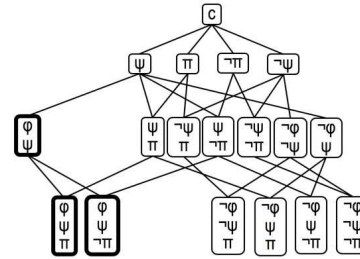


Figure 14.

$$\langle C, C \rangle + [[\text{if } \phi, \cdot \psi] \ \& \ [\text{if } \neg\phi, [\cdot \psi \vee \neg\psi]]] + \cdot \phi$$

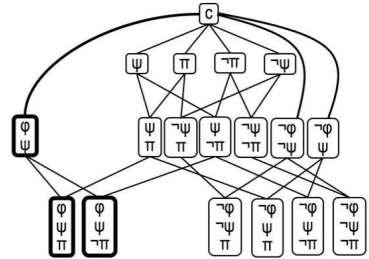


Figure 15.

$$\langle C, C \rangle + [[\text{if } \phi, \cdot \psi] \ \& \ [\text{if } \neg\phi, [\cdot \psi \vee \neg\psi]]] + \cdot \phi$$

with constructive predecessor rules

minimal CS at which  $\neg\Phi$  can be assumed is the node  $\boxed{\Phi}$ . In this account, the relation between the context sets of a CS do not just follow from the inclusion relation, but in addition by an accessibility relation that is determined by how the conversation actually moves forward.

## References

- Abbott, Barbara. 2004. Some remarks on indicative conditionals. *SALT* 14. 1-19.
- Alonso-Ovalle, Luis. 2009. Counterfactuals, correlatives, and disjunction. *L&P* 32: 207-244.
- Barker, Stephen J. 1995. Towards a pragmatic theory of 'if'. *Philosophical Studies* 79: 185-211.
- Ciardelli, Ivano, Linmin Zhang & Lucas Champollion. 2018. Two switches in the theory of counterfactuals: A study of truth conditionality and minimal change. *L&P* 41: 577-621.
- Edgington, Dorothy. 1995. On conditionals. *Mind* 104: 235-329.
- Égré, Paul & Guy Politzer. 2013. On the negation of indicative conditionals. *19th Amsterdam Colloquium*. 10-18.
- Gibbard, A. 1981. Two recent theories of conditionals. In: Harper, W. et al.(eds), *Ifs: Conditionals, Belief, Decision, Change and Time*. Dordrecht: Reidel,
- Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31: 231-270.
- Karawani, Hadil. 2014. *The Real, the Fake, and the Fake Fake in Counterfactual Conditionals, Crosslinguistically*. Doctoral dissertation, Utrecht.
- Kratzer, Angelika. 1981. The notional category of modality. In: Eikmeyer, H.-J. & H. Rieser, (eds), *Words, Worlds, and Contexts*. Berlin: de Gruyter, 38-74.
- Kratzer, Angelika. 1989. An investigation of the lumps of thought. *L&P*. 12: 607-653.
- Krifka, Manfred. 2015. Bias in Commitment Space Semantics: Declarative questions, negated questions, and question tags. *SALT* 25, 328-345.
- Lewis, David. 1973. *Counterfactuals*. Cambridge, Mass.: Harvard University Press.
- Punčochář, Vít. 2015. Weak negation in inquisitive semantics. *JLLI* 24: 323-355.
- Stalnaker, Robert. 1968. A theory of conditionals. In: Rescher, N., (ed), *Studies in logical theory*. Oxford: Basil Blackwell, 98-112.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. In: Munitz, M. K. & P. K. Unger, (eds), *Semantics and Philosophy*. New York: New York University Press, 197-214.
- Stalnaker, Robert. 2011. Conditional propositions and conditional assertions. In: Egan, A. & E. Weatherson, (eds), *Epistemic modality*. Oxford University Press, 227-248.
- Starr, William B. 2019. Counterfactuals. In; Zalta, E. (ed.), *Stanford Encyclopedia of Philosophy*.
- Tichý, Pavel. 1976. A counterexample to the Stalnaker-Lewis analysis of counterfactuals. *Philosophical Studies* 29: 271-273.
- Veltman, Frank. 1985. *Logics for conditionals*. Doctoral dissertation. University of Amsterdam.
- Veltman, Frank. 2005. Making counterfactual assumptions. *Journal of Semantics* 22: 159-180.

# DPs and CPs in Depiction Complements\*

Kristina Liefke

Goethe University Frankfurt, Frankfurt am Main, Germany

Liefke@lingua.uni-frankfurt.de

## Abstract

We present a compositional semantics for DP/CP-neutral depiction verbs (e.g. *paint*, *imagine*, *visualize*, *write (about)*) that interprets the DP and CP complements of these verbs in the same semantic type,  $\langle s, t \rangle$  (i.e. as partial sets of situations). Our semantics answers several challenges for semantics for depiction reports (e.g. [29, 32, 33]) – in particular, the difficulty of interpreting declarative and coordinated DP/CP complements of these reports, of explaining the selectional restrictions (*qua* different kinds of CPs) of many depiction verbs, and of accounting for entailment relations between DP- and CP-complemented depiction reports. At the same time, it preserves the merits of existing semantics for depiction verbs.

## 1 Introduction

Depiction verbs are verbs like *paint* and *imagine*<sup>1</sup> (see [6, Ch. 7], [33]) whose complements can be used to describe the content of pictures or mental images (see (1), (2)):

- (1) Paul is painting  $[_{DP} \text{a penguin}]$  ( $\equiv$  Paul is pictorially representing a penguin)
- (2) Uli is imagining  $[_{DP} \text{a unicorn}]$  ( $\equiv$  Uli is mentally depicting a unicorn)

To capture the semantic properties of depiction verbs (esp. their missing readings and entailment pattern; see [29, 32, 33]), the object DPs of such verbs are commonly interpreted as (type- $\langle s, \langle e, t \rangle \rangle$ ) properties. Depending on the particular account, these are the properties that are denoted by the DP's restrictor (for the *de dicto*-reading of (2): the property 'is a unicorn', see (3); cf. [3, 8, 29]) or existentially quantified sub-properties of these properties (see (4); cf. [32]):

- (3)  $imagine_i(uli, unicorn)$ , where *unicorn* is a non-logical constant of type  $\langle s, \langle e, t \rangle \rangle$
- (4)  $(\exists P^{\langle s, \langle e, t \rangle \rangle})[P \sqsubseteq unicorn \wedge imagine_i(uli, P)]$ , with  $P \sqsubseteq Q :=$  '*P* is more specific than *Q*'

Property-type accounts like the above are fully compositional and capture the intensional behavior of depiction complements. These advantages notwithstanding, all current incarnations of such accounts face the following challenges:

**Challenge (i): CP complements.** Property-type semantics have traditionally focused on DP complements of depiction reports (see [3, 8, 32, 33]). However, depiction verbs also license (certain kinds of) CP complements (see (5), (6)):

- (5) Uli is imagining  $[_{NON-FIN} [_{DP} \text{a girl}] \text{ riding } [_{DP} \text{a unicorn}]]$
- (6) Uli is imagining  $[_{FIN} \text{that } [_{DP} \text{a girl}] \text{ is riding } [_{DP} \text{a unicorn}]]$

Property-type accounts suggest – but do not explicitly claim – that clausal complements of depiction reports are also interpreted as properties. However, for CPs with *multiple* non-specific indefinites (e.g. (6)), this strategy fails<sup>2</sup> to yield a unique interpretation (see [31]). For example,

\*I thank three anonymous referees for AC 2019 for valuable comments on an earlier version of this paper. The paper has profited from discussions with Kai von Fintel, Markus Werning, and Ede Zimmermann. The research for this paper is supported by the German Research Foundation (via Ede Zimmermann's grant ZI 683/13-1).

<sup>1</sup>Other members of this class include *conceive*, *visualize*, *portray*, *sculpt*, *write (about)*, and *draw*.

<sup>2</sup>This failure may be remedied by interpreting the complement of (6) as a *pair* of properties. However, since



this strategy interprets the doubly non-specific reading of (6) as (7a) and/or (7b):

- (7) a.  $imagine_i(uli, \lambda j^s \lambda x^e. unicorn_j(x) \wedge (\exists y^e)[girl_j(y) \wedge ride_j(y, x)])$   
 b.  $imagine_i(uli, \lambda j^s \lambda y^e. girl_j(y) \wedge (\exists x^e)[unicorn_j(x) \wedge ride_j(y, x)])$

**Challenge (ii): interaction of DP and CP complements.** One could try to avoid the above problem by interpreting clausal depiction complements instead as type- $\langle s, t \rangle$  propositions. (This is indirectly suggested by the treatment in [3, 24].) However, the resulting *different-type* interpretation of nominal complements (type  $\langle s, \langle e, t \rangle \rangle$ ) and clausal complements (type  $\langle s, t \rangle$ ) disables an easy (i.e. type-shift-free) modelling of DP/CP coordinations like (8) and of entailments from non-finite to nominal and finite depiction reports, respectively (e.g. (9), (10)):

- (8) Uli is imagining  $[[_{DP} a unicorn]$  and  $[_{CP} that a girl is riding it]]$

- (9) a.  $\underline{Uli is imagining [_{NON-FIN} a girl riding [_{DP} a unicorn]]}$   
 $\Rightarrow$  b.  $\underline{Uli is imagining [_{DP} a unicorn]}$

- (10) a.  $\underline{Uli is imagining [_{NON-FIN} a girl riding a unicorn]}$   
 $\Rightarrow$  b.  $\underline{Uli is imagining [_{FIN} that a girl is riding a unicorn]}$

**Challenge (iii): restrictions on the selection of CP complements.** The interpretation of clausal depiction complements is further challenged by the observation that – in contrast to *imagine* (see (6)) – many depiction verbs (incl. *paint*) do not license finite complements (see (11b)):

- (11) a.  $\checkmark$  Paul is painting  $[_{NON-FIN} a penguin diving into the sea]$   
 b.  $*$  Paul is painting  $[_{FIN} that a penguin is diving into the sea]$

Uniform accounts, which interpret finite and non-finite depiction complements in the same semantic type, must do some extra work to explain the deviance of (11b) (see [20]; cf. [27]). Non-uniform accounts, which interpret these complements in different types, need to resort to type-shifts or other syntactic/semantic mechanisms to explain the acceptability of (5) and (6) ([7, 15, 22]).

## 2 Strategy: comparison with approaches to the semantics of responsive verbs

To see the landscape of options for the explanation of DP/CP flexibility in depiction complements, it is instructive to consider different accounts of the selectional flexibility of responsive verbs (i.e. verbs like *know* that accept both declarative and interrogative complements). Theiler *et al.* [27] outline four different approaches to this flexibility. These split into approaches which assume that declarative complements have a different (viz. simpler) semantic type from interrogative complements (i.e. type  $\langle s, t \rangle$  *vis-à-vis*  $\langle \langle s, t \rangle, t \rangle$ ) (hereafter, *non-uniform approaches*) and approaches which assume that declarative complements have the same type as interrogative complements (hereafter, *uniform approaches*, ④; e.g. [27]). Non-uniform approaches include approaches ① that semantically reduce interrogative to declarative complements (*reductive approaches*; e.g. [11]), ② that generalize declarative to interrogative complements (*inverse reductive approaches*; e.g. [28]), and ③ that connect the (different) lexical entries for verbs with declarative and interrogative complements (*twin approaches*; e.g. [9, 25]).<sup>3</sup>

property-pairs typically have a different type from properties, this strategy would disable an easy answer to Challenge (ii).

<sup>3</sup>We here deviate from [27] in classifying meaning-postulate approaches (e.g. [25]) with twin approaches.

Analogues of the above are – at least in principle – also available for DP and CP complements of depiction reports (see Figure 1). These include the semantic reduction of DP to CP complements (below, ‘ $\llbracket \text{DP} \rrbracket \mapsto \llbracket \text{CP} \rrbracket$ ’; see ❶), the generalization of CP to DP complements ( $\llbracket \text{CP} \rrbracket \mapsto \llbracket \text{DP} \rrbracket$ ; see ❷), the (type-)distinction between DP and CP complements ( $\llbracket \text{CP} \rrbracket \neq \llbracket \text{DP} \rrbracket$ ; see ❸), and the (type-)identification of DP and CP complements ( $\llbracket \text{CP} \rrbracket = \llbracket \text{DP} \rrbracket$ ; see ❹):

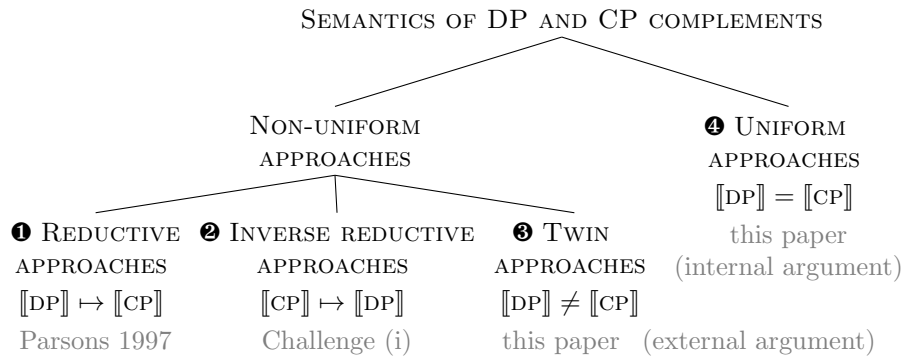


Figure 1: Options for the semantics of DP and CP complements (cf. [27, p. 410]).

The inability of CP complements of depiction verbs to be interpreted in the type of DP complements of such verbs,  $\langle s, \langle e, t \rangle \rangle$  (see Challenge (i)) already rules out an inverse reductive approach to DP/CP flexibility (❷).<sup>4</sup> Given the original motivation for property-type semantics (see esp. [29, 33], *pace* [21]) and the non-existence of an injective function from properties to propositions, similar observations hold for a reductive approach (❶). Since DPs and CPs in non-complement position are traditionally interpreted in different types (i.e.  $\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle$  resp.  $\langle s, t \rangle$ ) – and since the lifting of CPs to the type  $\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle$  would require re-thinking a large part of contemporary formal semantics –, a uniform approach is also not tenable (❹). This seemingly leaves twin approaches (i.e. ❸) as the only option for explaining DP/CP flexibility. However, the different-type interpretation of DP and CP complements in these approaches disables an easy modelling of DP/CP interaction (see Challenge (ii)).

This paper provides an alternative approach to DP/CP flexibility that combines a twin approach with aspects of a uniform approach. Our semantics assumes that DPs and CPs make a different compositional contribution (i.e. of type  $\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle$  resp.  $\langle s, t \rangle$ ) to the interpretation of depiction reports to which they serve as complements (see ❸). As a result, our semantics requires different lexical entries for DP- and CP-taking occurrences of depiction verbs. However, since this semantics interprets these different occurrences through the same non-logical constant, it still enables a (type-)uniform interpretation of DPs and CPs *in the complement* of these verbs. This is achieved by incorporating an  $\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle$ -to- $\langle s, t \rangle$  type-shifter into the lexical entry for DP-taking occurrences of depiction verbs. DP-denotations are then converted into CP-denotations during semantic composition. Since this type-shifter is built into the semantics of the verb (and is not freely available for all DPs), it does not falsely predict DP/CP flexibility for the complements of CP-biased verbs. In this regard, our semantics improves upon reductive approaches.

The rest of the paper is organized as follows: below, we first present our proposed semantics for depiction reports with DP complements (in Sect. 3). To answer Challenges (i) to (iii), this semantics is then extended to depiction reports with different kinds of CP complements (in

<sup>4</sup>This approach is further rejected by the fact that a  $\langle s, t \rangle$ -to- $\langle s, \langle e, t \rangle \rangle$  type-shifter (which would allow CPs to serve as input to the verbs’ DP-entry) would fail to explain the syntactic selection behavior of DP-biased verbs.

Sect. 4.1). We show that the relations between depiction reports with nominal and clausal complements capture the interaction of DP and CP complements in depiction reports (see Sect. 4.2) and that they help explain the selectional restriction on CP complements (see Sect. 4.3). The paper closes with a summary of our results and with pointers to future work.

### 3 Proposal and Background

We have suggested above that our proposed semantics uniformly interprets DP and CP complements of depiction verbs in the type  $\langle s, t \rangle$ . More precisely, this semantics identifies the semantic complements of DP-taking occurrences of depiction verbs as propositionally coded situations. This identification is motivated by the observation that transitive occurrences of depiction verbs (e.g. (2)) select for a semantic situation- (event- or state-)argument and by the possibility of representing (or *coding*) situations by sets of situations (i.e. in the type for propositions).

The ability of *imagine* to take a semantic situation argument is supported by the possibility of replacing the object DP in nominal imagination reports by an explicitly situation-denoting construction (see (12)) and/or by an event-denoting *how*-phrase (see (13)), and by the possibility of modifying the matrix verb in these reports through an ‘experiential’ modifier like *vividly* or in *vivid/lifelike detail* (see (14); cf. [26, p. 156]). These possibilities are corroborated by the observation that physical and mental images typically do not represent isolated items of information (e.g. Kratzer-style *facts*), but informationally richer objects (see [33, p. 433]).

- (12) a. Uli is imagining  $[_{DP} a \text{ unicorn}]$   
        $\equiv$  b. Uli is imagining  $[_{DP} a \text{ situation/scene in which there is a unicorn}]$
- (13) a. Uli is imagining  $[_{DP} a \text{ unicorn cantering}]$   
        $\equiv$  b. Uli is imagining  $[_{DP} \text{how a unicorn is cantering}]$
- (14) a. Uli is *vividly* imagining  $[_{DP} a \text{ unicorn}]$   
       b. Uli is imagining  $[_{DP} a \text{ unicorn}]$  *in vivid/lifelike detail*

The representation of the situation-argument through *sets* of situations is motivated by the fact that – in contrast to visual scenes – depicted situations are often not anchored in a particular world/time, and by the possibility of representing non-anchored situations by sets of isomorphic (= qualitatively identical) situations (see [17, p. 667]; cf. [5, 136]). The latter are situations in which exactly the same propositions are true (resp. false).

Arguably, the identity of the depicted situation depends on the particular manner of depicting (e.g. painting *vis-à-vis* imagining), the depicting agent, and the time of depicting. To capture this dependence, we use a choice function,  $f$ , that selects a subset from a given set of situations  $\lambda j [ \dots ]$  in dependence on a parameter,  $e$ , for the described depicting event (cf. [12, 16]). For  $e$  an imagining event, this subset represents the imagined situation/event (for (2): the situation or imaginary scene to which Uli stands in the imagining relation in the evaluation situation). The obtaining of this relation requires that Uli forms a mental image of this situation that has the same experiential character as perceptual witnessing (see [26, p. 153]).

The interpretation of the *de dicto*-reading of (2) is given in (15).<sup>5</sup> In what follows, *imagine* is a non-logical constant of type  $\langle s, \langle v, \langle \langle s, t \rangle, \langle e, t \rangle \rangle \rangle \rangle$ , where  $v$  is the type for events:

<sup>5</sup>For reasons of simplicity, we hereafter neglect tense and aspect in the interpretation of our example sentences.

$$(15) \quad \llbracket \text{Uli is imagining } [\text{DP a unicorn}] \rrbracket^i = (\exists e)[\mathbf{imagine}_i(e, \text{uli}, \underbrace{f_e(\lambda j \exists x. \text{unicorn}_j(x))}_{\text{Uli's (coded) imagined situation / mental image}})]$$

(15) is obtained from the standard semantics for proper names (type  $e$ ) and object DPs (type  $\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle$ ) through the lexical entry for nominal/transitive *imagine* in (16). In this entry,  $E$  is a situation-relative existence predicate (see [19, p. 117 ff.]):

$$(16) \quad \llbracket \text{imagine-DP} \rrbracket^i = \lambda \mathcal{Q}^{\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle} \lambda z^e (\exists e)[\mathbf{imagine}_i(e, z, f_e(\lambda j. \mathcal{Q}_j(E)))]$$

This entry interprets *imagine* as a relation between an evaluation situation ( $i$ ), an event ( $e$ ), an agent ( $z$ ), and (a propositional representation,  $f_e(\lambda j. \dots)$ , of) the depicted situation. The last argument is obtained by applying the  $e$ -parametrized choice function  $f$  to the proposition,  $\lambda j[\mathcal{Q}_j(E)]$ , that results from applying the standard semantic value of the DP to the property of situation-relative existence. In virtue of this application, the compositional interpretation of (2) initially comes out as (17):

$$(17) \quad \begin{array}{c} \llbracket \text{imagine-DP} \rrbracket^i (\llbracket \text{Uli} \rrbracket, \llbracket \text{a unicorn} \rrbracket) \\ (\exists e)[\mathbf{imagine}_i(e, \text{uli}, f_e(\lambda j \exists x. \text{unicorn}_j(x) \wedge E_j(x)))] \\ \swarrow \quad \searrow \\ \begin{array}{cc} \llbracket \text{Uli} \rrbracket & \llbracket \text{imagine-DP} \rrbracket^i (\llbracket \text{a unicorn} \rrbracket) \\ \text{uli} & \lambda z (\exists e)[\mathbf{imagine}_i(e, z, f_e(\lambda j \exists x. \text{unicorn}_j(x) \wedge E_j(x)))] \end{array} \\ \swarrow \quad \searrow \\ \begin{array}{cc} \llbracket \text{imagine-DP} \rrbracket^i & \llbracket \text{a unicorn} \rrbracket \\ \lambda \mathcal{Q} \lambda z (\exists e)[\mathbf{imagine}_i(e, z, f_e(\lambda j. \mathcal{Q}_j(E)))] & \lambda k \lambda P (\exists x)[\text{unicorn}_k(x) \wedge P_k(x)] \end{array} \end{array}$$

The assumption that every individual that is a unicorn in a situation exists in this situation (formally:  $(\forall j)(\forall x)[\text{unicorn}_j(x) \rightarrow E_j(x)]$ ) then warrants the equivalence of (17) with (15).

## 4 Answering the Challenges

With our semantics for nominal depiction reports in place, we are now ready to address Challenges (i) to (iii):

### 4.1 Answering challenge (i): CP complements

We have seen in Section 1 that property-type semantics for depiction reports resist an extension to clausal reports. In contrast, our semantics from the previous section can be straightforwardly generalized to such reports: in particular, to give a semantics for clausal depiction reports, we use Stephenson's [26] observation that gerund depiction complements are also<sup>6</sup> interpreted as (coded) situations (see our interpretation of (5) in (18)). This interpretation uses the semantics for non-finite occurrences of *imagine* in (19):

$$(18) \quad \llbracket \text{Uli is imagining } [\text{NON-FIN a girl riding a unicorn}] \rrbracket^i \\ = (\exists e)[\mathbf{imagine}_i(e, \text{uli}, \underbrace{f_e(\lambda j \exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(y, x)))}_{\text{Uli's (coded) imagined situation}})]$$

$$(19) \quad \llbracket \text{imagine-NON-FIN} \rrbracket^i = \lambda p^{\langle s, t \rangle} \lambda z^e (\exists e)[\mathbf{imagine}_i(e, z, f_e(p))]$$

<sup>6</sup>In fact, [26] focuses on gerund complements. Nominal depiction reports are only treated in passing.

To interpret depiction reports with a finite **that**-clause complement (e.g. (6)), we use a close variant (in (20)) of the familiar semantics for finite occurrences of **imagine**. This semantics interprets **imagine** as a relation to a proposition (i.e. to an object of the form  $\lambda j [\dots]$ ), rather than to a (propositionally coded) situation (i.e. to an object of the form  $f_e(\lambda j. \dots)$ ). The former differs from the latter in that it does not contain any information that is not encoded in  $p$ .

$$(20) \quad \llbracket \text{imagine-FIN} \rrbracket^i = \lambda p^{\langle s, t \rangle} \lambda z^e (\exists e) [\text{imagine}_i(e, z, p)]$$

Our semantics for finite **imagine** enables a unique interpretation of the doubly non-specific reading of (6), as desired:

$$(21) \quad \begin{aligned} & \llbracket \text{Uli is imagining } [\text{FIN that a girl is riding a unicorn}] \rrbracket^i \\ &= (\exists e) [\text{imagine}_i(e, \text{uli}, \underbrace{\lambda j \exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(y, x))}_{\text{a proposition whose content Uli is imagining}})] \end{aligned}$$

## 4.2 Answering challenge (ii): interaction of DP and CP complements

Above, we have introduced separate lexical entries for differently complemented occurrences of **imagine**. However, in contrast to other non-uniform semantics for selectionally flexible verbs (e.g. [2, 7, 26]), our semantics uses entries with the same non-logical constant (above: **imagine**). This facilitates the modelling of DP/CP coordinations in the complements of depiction reports (I) and enables the obtaining of entailment relations between reports with different kinds of complements (II).

Admittedly, the ability of our semantics to capture DP/CP coordinations in depiction complements (see (I)) is still impeded by the fact that the lexical entries for DP- and finite CP-taking **imagine** (see (16), (20)) require different inputs (of type  $\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle$  and  $\langle s, t \rangle$ , respectively). To compensate for these different requirements, we use the semantics for coordinating **and** in (22). This semantics interprets the conjunction of a DP and a finite CP as an intensional generalized quantifier of the form ‘ $\lambda j \lambda P [\llbracket \text{DP} \rrbracket^j(P) \wedge \llbracket \text{CP-FIN} \rrbracket^j]$ ’:

$$(22) \quad \llbracket \text{DP-and-FIN} \rrbracket = \lambda p^{\langle s, t \rangle} \lambda Q^{\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle} \lambda j^s \lambda P^{\langle s, \langle e, t \rangle \rangle} [\mathcal{Q}_j(P) \wedge p_j]$$

The above enables the interpretation of (8) as (25). This interpretation treats **that** in (8) as the trivial complementizer, **that<sub>T</sub>** (in (23); see [18]), and assumes that coordinations of a DP and a finite CP complement are licensed by **imagine-DP**. Our interpretation uses the intermediate step in (24). In (24) and (25),  $\mathcal{T}_1$  is a higher-order trace that ranges over intensional quantifiers:<sup>7</sup>

$$(23) \quad \llbracket \text{that}_T \rrbracket = \lambda p \lambda j [p_j]$$

$$(24) \quad \begin{aligned} & \llbracket [\text{a unicorn}] [\lambda_1 [\mathcal{T}_1 \text{ and that}_T \text{ a girl is riding } \mathcal{T}_1]] \rrbracket \\ & \equiv (\lambda_1. \llbracket \text{and} \rrbracket (\llbracket \text{that} \rrbracket (\llbracket \text{a girl is riding } \mathcal{T}_1 \rrbracket), \llbracket \mathcal{T}_1 \rrbracket)) (\llbracket \text{a unicorn} \rrbracket) \\ &= (\lambda Q (\lambda p \lambda P \lambda j \lambda P [\mathcal{P}_j(P) \wedge p_j] (\lambda k \exists y. \text{girl}_k(y) \wedge Q_k(\lambda l \lambda z. \text{ride}_l(z, y)), Q))) \\ & \quad (\lambda j' \lambda Q \exists x. \text{unicorn}_{j'}(x) \wedge Q_{j'}(x)) \\ & \equiv \lambda Q \lambda j \lambda P [\mathcal{Q}_j(P) \wedge (\exists y. \text{girl}_j(y) \wedge Q_j(\lambda l \lambda z. \text{ride}_l(z, y)))] (\lambda k \lambda Q \exists x. \text{unicorn}_k(x) \wedge Q_k(x)) \\ & \equiv \lambda j \lambda P [(\exists x. \text{unicorn}_j(x) \wedge P_j(x)) \wedge (\exists y. \text{girl}_j(y) \wedge (\exists z. \text{unicorn}_j(z) \wedge \text{ride}_j(z, y)))] \end{aligned}$$

<sup>7</sup>The assumption of higher-order traces enables us to simultaneously allow for the *de dicto*-reading of (8) and for the anaphoric binding of the pronoun in (8). However, this move is independent of the proposed semantics. Our use of an example with anaphoric binding is motivated by the observation that the most natural embedded DP/CP coordinations are ones where the DP binds a pronoun in the CP.

$$\begin{aligned}
(25) \quad & \llbracket \text{Uli imagines-DP} \llbracket \text{a unicorn} \rrbracket [\lambda_1 [\mathcal{T}_1 \text{ and that a girl is riding } \mathcal{T}_1]] \rrbracket \\
& \equiv \llbracket \text{imagine-DP} \rrbracket^i (\llbracket \text{Uli} \rrbracket, \llbracket \llbracket \text{a unicorn} \rrbracket [\lambda_1 [\mathcal{T}_1 \text{ and that a girl is riding } \mathcal{T}_1]] \rrbracket) \\
& = \lambda \mathcal{Q} \lambda u (\exists e) [\text{imagine}_i(e, u, f_e(\lambda l. \mathcal{Q}_i(E)))](uli, \\
& \quad \lambda j \lambda P [(\exists x. \text{unicorn}_j(x) \wedge P_j(x)) \wedge (\exists y. \text{girl}_j(y) \wedge (\exists z. \text{unicorn}_j(z) \wedge \text{ride}_j(z, y)))] \\
& = (\exists e) [\text{imagine}_i(e, uli, f_e(\lambda j. [(\exists x. \text{unicorn}_j(x) \wedge E_j(x) \wedge \\
& \quad (\exists y. \text{girl}_j(y) \wedge (\exists z. \text{unicorn}_j(z) \wedge \text{ride}_j(z, y)))]))] \\
& \equiv (\exists e) [\text{imagine}_i(e, uli, f_e(\lambda j \exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(x, y)))]
\end{aligned}$$

On (II): since our interpretation of transitive *imagine* (in (16)) applies the parametrized choice function  $f_e$  to the proposition  $\lambda j [\mathcal{Q}_j(E)]$  (where  $\mathcal{Q}$  is the semantic value of the DP complement), it predicts that (2) is equivalent to the result (in (26)) of enriching the object DP in (2) with VP being there (see [21, pp. 375–376]):

$$(26) \quad \text{Uli is imagining } [\text{NON-FIN} [\text{DP a unicorn}] \text{ being there (in his imagined situation)}]$$

Our proof of this equivalence (in (30)) interprets the predicate *be there* through the existence predicate  $E$  (in (27)). This proof uses the intermediate step in (29). It is based on the assumption that every individual that is a unicorn in a situation exists in this situation (see (28)):

$$(27) \quad \llbracket \text{be there} \rrbracket^i \equiv \llbracket \text{exist} \rrbracket^i = \lambda \mathcal{Q} \lambda j [\mathcal{Q}_j(\lambda k \lambda y. E_k(y))]$$

$$(28) \quad (\forall j)(\forall x)[\text{unicorn}_j(x) \rightarrow E_j(x)]$$

$$\begin{aligned}
(29) \quad & \llbracket \text{a unicorn being there} \rrbracket \equiv \llbracket \text{be there} \rrbracket (\llbracket \text{a unicorn} \rrbracket) \\
& = \lambda \mathcal{Q} \lambda j [\mathcal{Q}_j(\lambda k \lambda y. E_k(y))] (\lambda l \lambda P (\exists x)[\text{unicorn}_l(x) \wedge P_l(x)]) \\
& \equiv \lambda j (\exists x)[\text{unicorn}_j(x) \wedge E_j(x)] \\
& \equiv \lambda j (\exists x)[\text{unicorn}_j(x)] \tag{by (28)}
\end{aligned}$$

$$\begin{aligned}
(30) \quad & \llbracket \text{Uli is imagining } [\text{NON-FIN a unicorn being there}] \rrbracket^i \\
& \equiv \llbracket \text{imagine} \rrbracket^i (\llbracket \text{Uli} \rrbracket, \llbracket \text{a unicorn being there} \rrbracket) \\
& = \lambda p \lambda z (\exists e) [\text{imagine}_i(e, z, f_e(p))](uli, \lambda j (\exists x)[\text{unicorn}_j(x)]) \\
& \equiv (\exists e) [\text{imagine}_i(e, uli, f_e(\lambda j \exists x. \text{unicorn}_j(x)))] \\
& = \llbracket \text{Uli is imagining } [\text{DP a unicorn}] \rrbracket^i \tag{see (15)}
\end{aligned}$$

(MP1), below, generalizes the observed equivalence between (2) and (26) to the relation between nominal and gerund occurrences of depiction verbs (below, *imagine*):

$$(MP1) \quad (\forall \mathcal{Q})(\forall z) [\llbracket \text{imagine-DP} \rrbracket^i(z, \mathcal{Q}) \equiv \llbracket \text{imagine-NON-FIN} \rrbracket^i(z, \lambda j. \mathcal{Q}_j(E))]$$

(MP1) enables a straightforward account of the obtaining of entailment relations between clausal and nominal depiction reports. In particular, the entailment in (9) (see (31)) is based on the fact that all situations in which a girl is riding a unicorn are situations in which there is a unicorn (see (31b)). The entailment further relies on the intuitive parthood principle in (31c) and on the upward-monotonicity of the complement of *imagine* (see (31d)):

$$\begin{aligned}
(31) \quad & \text{a. } \llbracket \text{Uli is imagining } [\text{NON-FIN a girl riding } [\text{DP a unicorn}]] \rrbracket^i \\
& \quad = (\exists e) [\text{imagine}_i(e, uli, f_e(\lambda j \exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(y, x)))] \\
& \quad \text{b. } (\forall j)[(\exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(y, x))) \rightarrow (\exists z. \text{unicorn}_j(z))] \\
& \quad \text{c. } (\forall p)(\forall q)[p \subseteq q \rightarrow (\forall e. f_e(p) \subseteq f_e(q))] \\
& \quad \text{d. } (\forall p)(\forall z)(\forall e) [\text{imagine}_i(e, z, p) \rightarrow (\forall q. p \subseteq q \rightarrow \text{imagine}_i(e, z, q))] \\
& \Rightarrow \text{e. } \frac{(\exists e) [\text{imagine}_i(e, uli, f_e(\lambda j \exists x. \text{unicorn}_j(x)))]}{\llbracket \text{Uli is imagining } [\text{DP a unicorn}] \rrbracket^i}
\end{aligned}$$

Analogously to the above, the entailment in (10) is based on the relation between gerund and finite clause occurrences of depiction verbs. This relation is specified in (MP2). It is supported by the lexical entries for *imagine* in (19) and (20):

$$(MP2) \quad (\forall p)(\forall z)[\llbracket \text{imagine-NON-FIN} \rrbracket^i(z, p) \equiv \llbracket \text{imagine-FIN} \rrbracket^i(z, f_e(p))],$$

where  $e$  is bound by the quantifier  $(\exists e)$  in  $\llbracket \text{imagine-FIN} \rrbracket^i$

Given the upward monotonicity of *imagine* (see (32c)), the membership of the semantic value of a girl riding a unicorn in the value of that a girl is riding a unicorn (see (32b)) then supports the entailment in (10) (see (32)):

$$(32) \quad \begin{aligned} \text{a. } & \llbracket \text{Uli is imagining } [_{\text{NON-FIN}} \text{a girl riding a unicorn}] \rrbracket^i \\ & = (\exists e)[\text{imagine}_i(e, \text{uli}, f_e(\lambda j \exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(y, x)))))] \\ \text{b. } & (\forall p)(\forall e)[f_e(p) \subseteq p] \\ \text{c. } & \frac{(\forall p)(\forall z)(\forall e)[\text{imagine}_i(e, z, p) \rightarrow (\forall q. p \subseteq q \rightarrow \text{imagine}_i(e, z, q))]}{\text{(see (31d))}} \\ \Rightarrow \text{d. } & \llbracket \text{Uli is imagining } [_{\text{FIN}} \text{that a girl is riding a unicorn}] \rrbracket^i \\ & = (\exists e)[\text{imagine}_i(e, \text{uli}, \lambda j \exists x. \text{unicorn}_j(x) \wedge (\exists y. \text{girl}_j(y) \wedge \text{ride}_j(y, x)))] \quad \text{(see (21))} \end{aligned}$$

### 4.3 Answering challenge (iii): restrictions on the selection of CP complements

Our semantics interprets finite and non-finite complements of depiction verbs as classical propositions (see (20)) respectively as propositionally coded situations (see (19)). This interpretation opens up new possibilities for the explanation of the distribution of DPs and (finite) CPs: since it predicts that depiction verbs combine with both DPs and CPs, this interpretation suggests that the selectional restrictions of these verbs (e.g. the fact that – unlike *imagine* – *paint* rejects finite CP complements; see (11b)) are explained in terms of independently observable semantic properties of these verbs.<sup>8</sup>

Specifically, to account for the deviance of (11b), we argue that the degree of abstractness of classical propositions – *vis-à-vis* propositionally coded situations – is correlated with the possible degree of abstractness of the content of imagining vs. painting: while it is, in principle,<sup>9</sup> possible to imagine that there is a unicorn without imagining any specific properties of this unicorn, constraints on physical depiction (e.g. painting, drawing, and sculpting) block this possibility. These constraints include the choice of color and stroke (for painting/drawing) or of a particular shape (for sculpting). Since these choices are often influenced by the concrete properties of the depicted object, they allow for (some) inferences to the depicted object having these properties.

To capture the lower degree of abstractness of the content of physical pictures (*vis-à-vis* the content of mental images), we assume that – in contrast to *imagine* – the logical translation, *paint*, of *paint* is undefined for (classical propositions, or sets of situations that code) minimal situations in the sense of [17]. This undefinedness then underlies the deviance of (11b).

## 5 Conclusion

In this paper, we have presented a compositional semantics for DP/CP-neutral depiction verbs that interprets the DP and CP complements of these verbs as propositions or propositionally coded situations. This is achieved combining a twin approach to DP/CP flexibility with aspects of a uniform approach: in line with the former, our semantics assumes that DPs and CPs provide

<sup>8</sup>I thank Floris Roelofsen directing my attention to this point.

<sup>9</sup>We interpret the difficulty of this endeavor as support for the situation argument of depiction reports.



different-type inputs to the denotation of depiction verbs. As a result, our semantics postulates different lexical entries for DP- and CP-taking occurrences of such verbs. However, in contrast to existing semantics, our semantics assumes that these entries incorporate different type-shifters that send the different-type inputs to the type  $\langle s, t \rangle$  (i.e.  $\lambda Q \lambda j (\exists e)[f_e(\lambda k. Q_k(E))_j]$  for DPs,  $\lambda p \lambda j (\exists e)[f_e(p)_j]$  for non-finite CPs, and  $\lambda p \lambda j [p_j]$  for finite CPs). The resulting same-type interpretation of different depiction complements exemplifies a uniform approach to DP/CP flexibility. We have shown that our semantics answers several challenges for semantics of depiction reports, including the difficulty of interpreting CP complements of these reports, of accounting for the semantic interaction of DPs and CP in depiction complements, and of explaining the selectional restrictions of depiction verbs.

This paper has focused on depiction verbs that show all three hallmark features of intentionality, viz. referential opacity, non-specificity, and lack of existential import (see [30, p. 516]). We leave it to future work to apply the presented semantics to depiction verbs (e.g. *portray*) that only show some of these features (here: referential opacity).

## 6 References

- [1] J. Barwise, *Scenes and other situations*, J. of Philosophy **78/7** (1981), 369–397.
- [2] T. Bernard and L. Champollion, *Negative events in compositional semantics*, Proc. of SALT 28, 2018, pp. 512–532.
- [3] A.R. Deal, *Property-type objects and modal embedding*, Proc. of SuB 12, 2008, pp. 92–106.
- [4] J. van der Does, *A generalized quantifier logic for naked infinitives*, Linguistics and Philosophy **14/3** (1991), 241–294.
- [5] K. Fine, *Properties, propositions, and sets*, J. of Philosophical Logic **6** (1977), 135–191.
- [6] G. Forbes, *Attitude Problems: An essay on linguistic intensionality*, Oxford University Press, 2006.
- [7] ———, *Content and theme in attitude ascriptions*, Non-Propositional Intentionality (A. Grzankowski and M. Montague, eds.), Oxford University Press, 2018, pp. 114–133.
- [8] V. van Geenhoven and L. McNally, *On the property analysis of opaque complements*, Lingua **115** (2005), 885–914.
- [9] B. Ross. George, *Question Embedding and the Semantics of Answers*, Doctoral dissertation, University of California, Los Angeles, 2011.
- [10] J. Groenendijk and M. Stokhof, *Studies on the Semantics of Questions and the Pragmatics of Answers: Doctoral dissertation*, University of Amsterdam, 1984.
- [11] I Heim, *Interrogative semantics and Karttunen’s semantics for know*, Proc. of IATL (R. Buchalla and A. Mittwoch, eds.), Academion, 1994, pp. 128–144.
- [12] K. von Stechow, *The salience theory of definiteness*, Perspectives on Linguistic Pragmatics, Springer, 2013, pp. 349–374.
- [13] J. Higginbotham, *Remembering, imagining, and the first person*, manuscript (2003).



- [14] L. Karttunen, *Syntax and semantics of questions*, *Linguistics and Philosophy* **1/3** (1977), 3–44.
- [15] I. Kastner, *Factivity mirrors interpretation: The selectional requirements of presuppositional verbs*, *Lingua* **164, Part A** (2015), 156–188.
- [16] A. Kratzer, *Scope or pseudoscope: Are there wide-scope indefinites?*, *Events and Grammar* (S. Rothstein, ed.), Kluwer, 1998, pp. 163–96.
- [17] ———, *Facts: particulars or information units?*, *Linguistics and Philosophy* **25/5-6** (2002), 655–670.
- [18] ———, *Decomposing attitude verbs*, Honoring Anita Mittwoch on her 80th birthday at The Hebrew University of Jerusalem (2006).
- [19] K. Liefke, *A Single-Type Semantics for Natural Language*, Doctoral dissertation, Center for Logic and Philosophy of Science, Tilburg University, 2014.
- [20] K. Liefke and M. Werning, *Evidence for single-type semantics – an alternative to e/t-based dual-type semantics*, *Journal of Semantics* **35/4** (2018), 639–685.
- [21] T. Parsons, *Meaning sensitivity and grammatical structure*, *Structures and Norms in Science* (M.L. Dalla Chiara et al., ed.), Kluwer, 1997, pp. 369–383.
- [22] C. Potts, *The lexical semantics of parenthetical -as and appositive -which*, *Syntax* **5/1** (2002), 55–88.
- [23] J. Pustejovsky, *Situating events in language*, *Conceptualizations of Time* **52** (2016), 27–46.
- [24] F. Schwarz, *On needing propositions and looking for properties*, *Proceedings of SALT XVI*, 2006, pp. 259–276.
- [25] B. Spector and P. Egré, *A uniform semantics for embedded interrogatives: An answer, not necessarily the answer*, *Synthese* **192/6** (2015), 1729–1784.
- [26] T. Stephenson, *Vivid attitudes: Centered situations in the semantics of ‘remember’ and ‘imagine’*, *Proceedings of SALT XX*, 2010, pp. 147–160.
- [27] N. Theiler, F. Roelofsen, and M. Aloni, *A uniform semantics for declarative and interrogative complements*, *Journal of Semantics* **35/3** (2018), 409–466.
- [28] W. Uegaki, *Content nouns and the semantics of question-embedding*, *Journal of Semantics* **33** (2016), 623–660.
- [29] T.E. Zimmermann, *On the proper treatment of opacity in certain verbs*, *Natural Language Semantics* **1/2** (1993), 149–179.
- [30] ———, *Unspecificity and intensionality*, *Audiatur Vox Sapientiae* (C. Féry and W. Sternefeld, eds.), Akademie Verlag, Berlin, 2001, pp. 524–543.
- [31] ———, *Coercion vs. indeterminacy in opaque verbs*, *Intensionality* (R. Kahle, ed.), AK Peters/CRC Press, 2005, pp. 227–279.
- [32] ———, *Monotonicity in opaque verbs*, *Linguistics and Philosophy* **29/6** (2006), 715–761.
- [33] ———, *Painting and opacity*, *Von Rang und Namen* (W. Freitag, H. Rott, H. Sturm, and A. Zinke, eds.), Mentis, 2016, pp. 427–453.

# Not few but all quantifiers can be negated: Towards a referentially transparent semantics of quantified noun phrases

Andy Lücking<sup>1</sup> and Jonathan Ginzburg<sup>2</sup>

<sup>1</sup> Université de Paris and Goethe University Frankfurt  
luecking@em.uni-frankfurt.de

<sup>2</sup> Université de Paris and Institut Universitaire de France  
yonatan.ginzburg@univ-paris-diderot.fr

## Abstract

The main concern of this paper is to introduce *not* as a noun phrase negation operator. It is used in order to compositionally derive structures such as *not(many)(bicycles)*. However, noun phrase negation is part of a much larger account of plural semantics and quantification, guided by the notion of *referential transparency*. Therefore the paper splits into two parts: the first part gives a minimal background on a referentially transparent NP semantics. A foundation of the semantics of plural count nouns in terms of ordered set bipartitions is given. This includes as a side effect a significant reduction in the number of possible quantifier denotations and a derivation of the conservativity universal. The second part deals with *not* and how it interacts with an negation and anaphoric accessibility.

## 1 Motivation

In generalised quantifier theory (GQT), quantifiers are interpreted according to their *type* [18]. The type of a quantifier is given in terms of tuples of natural numbers,  $\langle n_1, \dots, n_k \rangle$ , indicating number and kind of the arguments the quantifier takes, where ‘1’ denotes a set, ‘2’ denotes a binary relation, and so on. These quantifier types figure as templates in terms of which the truth conditions of quantified sentences are spelled out. Here we take a quantifier to be a determiner and a generalised quantifier (GQ) to be determiner–noun projection. While the denotation of, say, *many* is of type  $\langle 1, 1 \rangle$ , the denotation type of *many glasses* is  $\langle 1 \rangle$ .

GQs are very elegant in terms of quantifier logic. However, they exhibit several features which make it difficult to embed them into compositional and more crucially incremental natural language grammar and semantics:

- In virtually every grammatical theory, the sentential head is the main verb and the verb phrase *predicates* of the subject. In GQT, however, the quantifier is treated as the semantic head of a sentence and expresses a *relation* (the arity of the relation is determined by the quantifier type);
- It is commonly assumed that meaning are derived in a compositional way. However, QGs are typically lexicalised holistically, violating compositionality, e.g., *every \_ . . . a different \_* (though see e.g. [12] for an adjectival analysis of ‘different’);
- An even stronger constraint than compositionality is incrementality: natural language input is not only processed in a systematic manner, but also word for word (and indeed at a higher, sub-lexical latency). Quantified noun phrases (QNPs) are no exception, when used in pragmatically supporting, comprehension-oriented contexts [26], in particular those in spoken conversation [6]. Holistic and relational quantifiers seem to be a serious obstacle to this well-established empirical insight, in particular wrt. quantifier floating.

$$NP_{sem} := \left[ \begin{array}{l} \text{q-params} : \left[ \begin{array}{l} \text{maxset} : \text{Set}(Ind) \\ \text{c1} : \overrightarrow{PType}(\text{maxset}) \\ \text{refset} : \text{Set}(Ind) \\ \text{compset} : \text{Set}(Ind) \\ \text{c2} : \text{union}(\text{maxset}, \text{compset}, \text{refset}) \end{array} \right] \\ \text{q-cond} : \text{Rel}(\text{q-params.refset}, \text{q-params.compset}) \\ \text{q-persp} : \text{refset} = \emptyset \vee \text{refset} \neq \emptyset \vee \text{none} \end{array} \right]$$

Figure 1: Semantic anatomy of a quantified noun phrase ( $NP_{sem}$ )

Here we show how associating a significantly distinct and independently motivated type system with (Q)NPs, formulated in *Type Theory with Records* [4], resolves the compositionality and incrementality issues mentioned above and provides a predicational semantics for quantified subjects. A brief introduction into the methodological guidelines and the type-theoretical set-up are therefore necessary (Sec. 2; they have been discussed in more detail elsewhere<sup>1</sup>). Sec. 3 relates the type-theoretical set-up to plural denotations. Noun phrase negation is introduced in Sec. 4 and a brief note on incrementality is given in Sec. 5.

## 2 Background: Referentially transparent QNPs

To blurt things out, the type we associate with NPs is given in Fig. 1. The motivation for such contents derives from the need to attain *referential transparency*, a dialogical desideratum incorporating (i) anaphoric potential, clarificational potential, co-verbal gesture; and (ii) several key recent psycholinguistic results on processing GQs mentioned below (Sec. 5). So far, we recognise four sources of referential transparency:

- (1) **Referential transparency:** an NP is *referentially transparent* if (a) it provides the semantic type required by a clarification request, (b) it provides antecedents for pronominal anaphora, (c) it provides an attachment site for co-verbal gestures, (d) its content parts can be identified and addressed.

Wrt. (1a), [21] show that clarificational potential provides data against higher order denotations (as postulated by GQT) in that answers to reprise fragment clarifications always provide individuals or sets, but not sets of sets. This finding carries particular weight since reprise fragments query (exactly) the semantic (not the pragmatic) content of the fragments being reprised. For such reasons, [21] argue in favour of a more transparent NP denotation in terms of *witness sets*. We refine and generalise their proposal to all NPs.

As is widely accepted, *anaphoric potential* covers two kinds of antecedents, a so-called maximal set and a reference set. Both are exemplified in (2), where the plural pronoun in (2a) refers back to demonstrators that actually took part in the rally (the reference set, or refset), and the plural pronoun in (2b) picks up the antecedent which denotes the whole of demonstrators that potentially could have come to the rally (the maximal set, or maxset). When the antecedent NP is formed with a quantifier which is downward monotone on the restriction, like *few*, even a further witness set can be picked out, namely the complement set (compset) [17]; in (2c) the potential demonstrators that did not come to the rally.

<sup>1</sup>Lücking, Cooper and Ginzburg, manuscript under review.

- (2) Few demonstrators were allowed in near the Place d’Italie.
- a. They raised their signs defiantly.
  - b. But they had all received an invitation.
  - c. They went to a football game instead.

The eponymous contents within the  $\overrightarrow{q}$ -params field in Fig. 1 correspond to these three antecedent sets. The ‘arrow type’,  $\overrightarrow{PType}$ , is the plural variant of a singular predicate and provides the group-constituting property of the plurality (there are collective and distributive subtypes, but this is yet another story). For instance, a group of students is represented as ‘ $\overrightarrow{\text{student}}(\text{maxset})$ ’, where a single student is represented as ‘ $\text{in}(\text{refind}, \text{refset})$ ’. The relation between maxset, refset and refind will become clear shortly in Sec. 3.

The attribute q-persp encodes a quantifier’s perspective, that is, the expectancy whether the refset may be empty (q-persp =  $\emptyset$ ) or not (q-persp  $\neq \emptyset$ ) (following work on GQ processing by [24]). On our account, a denotational underpinning for q-persp’s value is provided, namely whether an empty refset is part of a QNP’s denotation (Sec. 3). That is, other things being equal (e.g. a non-negative context, to which we turn below in Sec. 4), ‘q-persp =  $\emptyset$ ’ labels downward monotonicity. The q-persp value is also the distinguishing semantic feature between the quantificational minimal pair *few* vs. *a few*, which are alike in terms of their q-conds. If a perspective is not applicable, as with proper names, the feature is ‘none’. Given this set-up, anaphoric accessibility is simply governed by the rule that refset and maxset are, *ceteris paribus*, available as antecedents while the compset is only available when the ‘refset =  $\emptyset$ ’ flag is on.

The *descriptive quantifier condition* (q-cond) provides the relation of a QNP in terms of a cardinality comparison between its refset and the compset (readings which draw on a contextually given standard of comparison [16] can always be derived when a threshold  $\theta$  is instantiated in context). That is, q-cond is only defined in terms of NP-internal denotations and does not make reference to a so-called scope set. A closely related move has also been made in GQT: it is known that the semantics of conservative quantifiers of type  $\langle 1, 1 \rangle$  can be reduced to unary functions which only take the restrictor set as argument [7, 11]. On the account in [11], the denotation of a QNP is a pair  $\langle R, W \rangle$  consisting of a restrictor set  $R$  and a *set of witnesses*  $W$ , where  $W$  is the set of subsets of  $R$ . A quantifier denotes a unary function from subsets of  $R$  onto  $W$ . For a universe of two objects there are 512 possible quantifier denotations [11]; for comparison: there are  $2^{16} = 65,536$  quantifier denotations according to GQT [9, p. 632].

While this is a step in the (right) direction of a referentially transparent QNP semantics,  $W$  still faces the clarification request argument of [21]. Therefore, yet another denotational and conceptual underpinning is required.

### 3 Denotational construal

Our starting point to a denotational approach to QNPs is the simple consideration that quantification should rest on a plural semantics. After all, most quantifiers combine with plural nouns.<sup>2</sup> Plural semantics models the extension of a plural count noun in terms of the power set (or something equivalent such as a join semi-lattice) of the extension of the base noun [13]. We argue that referential transparency offers a more detailed picture on plural denotations, namely the need for sets of *ordered set bipartitions*.

<sup>2</sup>Exceptions being *every* and *each*, whose head nouns are syntactically singular (though semantically plural), and—if one wants to subsume it to quantificational determiners—the indefinite article (the definite is ambiguous in this respect, as is ‘no’). And there is no dearth of cross-linguistic variation.

(3) **Ordered set bipartition**

An ordered set bipartition  $p$  of a set  $S$  is a 2-tuple of pairwise disjoint subsets of  $S$  including the empty set such that the union of these subsets is  $S$ .

(An example follows shortly in (4).) The extensions of count nouns are computed as follows (we give both the type denotations and interpreted language denotations (i.e. ‘ $\llbracket \cdot \rrbracket$ ’)):

- Extension of a type:  $\llbracket \forall T \rrbracket = \{a \mid a : T\}$ .
- P-extension of a predicate:  $\llbracket \downarrow P \rrbracket = \{a \mid \exists s[s : P(a)]\}$  (adopted from the  $\beta$ -reduced property extension of [3, Sec. 3.4]). Some explanation is required here. A predicate type in TTR is a complex type  $P(a)$  which is constructed out of a predicate  $P$  and an argument  $a$  (usually of basic type *Ind*). The extension of a predicate type is a set of situations, namely those situations that make  $a$  is  $P$  true. The P-extension of a predicate thus is the set of objects that figure in situations of  $P$ -ness. This corresponds to  $\llbracket P \rrbracket = \{x \mid P(x) = 1\}$ .
- Q-extension of a plural predicate:  $\llbracket \downarrow \vec{P} \rrbracket = p(\llbracket \downarrow P \rrbracket)$  (where  $p$  is the ordered set bipartition operation from (3)). This corresponds to  $p(\llbracket P \rrbracket)$ .

A simple example should illustrate how Q-extensions look like:

- (4) Let  $\llbracket \downarrow \text{bicycle} \rrbracket = \{\overrightarrow{\text{red}}, \overrightarrow{\text{blue}}, \overrightarrow{\text{green}}\}$  (i.e. there are three differently coloured bicycles in our universe  $U$ ). Then  $\llbracket \downarrow \text{bicycle} \rrbracket = p(\llbracket \downarrow \text{bicycle} \rrbracket) = \{\langle \emptyset, \{\overrightarrow{\text{red}}, \overrightarrow{\text{blue}}, \overrightarrow{\text{green}}\} \rangle, \langle \{\overrightarrow{\text{red}}\}, \{\overrightarrow{\text{blue}}, \overrightarrow{\text{green}}\} \rangle, \langle \{\overrightarrow{\text{blue}}\}, \{\overrightarrow{\text{red}}, \overrightarrow{\text{green}}\} \rangle, \langle \{\overrightarrow{\text{green}}\}, \{\overrightarrow{\text{red}}, \overrightarrow{\text{blue}}\} \rangle, \langle \{\overrightarrow{\text{red}}, \overrightarrow{\text{blue}}\}, \{\overrightarrow{\text{green}}\} \rangle, \langle \{\overrightarrow{\text{red}}, \overrightarrow{\text{green}}\}, \{\overrightarrow{\text{blue}}\} \rangle, \langle \{\overrightarrow{\text{blue}}, \overrightarrow{\text{green}}\}, \{\overrightarrow{\text{red}}\} \rangle, \langle \{\overrightarrow{\text{red}}, \overrightarrow{\text{blue}}, \overrightarrow{\text{green}}\}, \emptyset \rangle\}$ .

Each ordered set bipartition in the set of ordered bipartitions is structured in the form ‘(compset, refset)’. That is, the first ordered set bipartition in (4), the one with an empty compset, is the denotation of *every bicycle* in the sample universe. Note that it is just a structured set of bicycles rather than a set of bicycles which is a subset of all other sets, as assumed in GQT. Ordered set bipartitions therefore provide a considerable simplification of quantificational complexities and are therefore preferable in terms of processing and comprehension. To exemplify: For  $U = \{a, b\}$  there are four ordered set bipartitions:  $p(\{a, b\}) = \{\langle \emptyset, \{a, b\} \rangle, \langle \{a\}, \{b\} \rangle, \langle \{b\}, \{a\} \rangle, \langle \{a, b\}, \emptyset \rangle\}$ . The middle two (those without empty sets) are indistinguishable for a quantifier, since a quantifier just looks at cardinalities. Virtually collapsing the two middle bipartitions, there are seven combinatorially possible quantifier denotations left ( $2^n - 1$ , where  $n = 3$  (due to collapsing) and  $-1$  is due to skipping  $\emptyset$  which is already part of the bipartitions). Thus, for two objects there are seven possible quantifier denotations. This is a significant reduction even compared to the already reduced 512 quantifier denotations of the approach of [11], cf. Sec. 2. Note further that our quantifiers ‘live on’ [1] the extensions of their head noun, a property which is equivalent to conservativity [10], as recently argued by [27]. That is, if quantification rests on a plural semantics as envisaged by referential transparency, then the semantic universal that all natural language quantifiers are conservative [10] will be a sequitur (cf. also [11]).

Now the bipartitions of a Q-extension are mutually exclusive: they cannot be realised simultaneously (nor can just a pair of them). They define the logical space of possible plural extensions. That is, each denotation corresponds to a separate type-theoretical model. Quantificational determiners impose constraints on the logical space of the plural extension of their head noun. The constraining component is the quantifier condition (q-cond) by dint of which subsets of plural denotations are excised:  $\llbracket \downarrow \text{DET } \vec{N} \rrbracket = \{p \in p(\llbracket \downarrow N \rrbracket) \mid \text{Rel}(p.\text{second}, p.\text{first}) = 1\}$  (where ‘Rel’ is DET’s q-cond value). Since an individual is always an individual within a refset,

singular is as much a special case of plural as definite is of indefinite. We model this intuitively appealing pattern in exactly this way, where a singular count noun introduces a refind (‘reference individual’) which is an element of a refset. The maxset, finally, is simply the union of the refset and the compset which constitute an ordered set bipartition.

In order to distinguish between referential and quantificational uses of singular and plural QNPs, we employ the mechanism developed by [21]: extensional argument roles can either be referentially grounded (when coerced into contextual parameters (c-params)) or existentially quantified away (when coerced into quantificational parameters (q-params)). Depending on which element goes where, referential and quantificational uses are distinguished in terms of their witnessing conditions:

- **quantificational:** maxset is part of *q-params*.
 
$$a : \left[ \begin{array}{l} \text{q-params} : \left[ \begin{array}{l} \text{maxset} : \text{Set}(\text{Ind}) \\ \text{c1} : \vec{P}(\text{maxset}) \\ \text{refset} : \text{Set}(\text{Ind}) \\ \text{compset} : \text{Set}(\text{Ind}) \end{array} \right] \\ \text{q-cond} : \text{Rel}(\text{q-params.refset}, \text{q-params.compset}) \end{array} \right],$$
 iff  $a \in p(\downarrow P) \wedge \text{Rel}(a.\text{second}, a.\text{first}) = 1$
- **indefinite:** an individual is part of *q-params*.
 
$$a : \left[ \begin{array}{l} \text{q-params} : \left[ \begin{array}{l} \text{maxset} : \text{Set}(\text{Ind}) \\ \text{c1} : \vec{P}(\text{maxset}) \\ \text{refset} : \text{Set}(\text{Ind}) \\ \text{compset} : \text{Set}(\text{Ind}) \\ \text{refind} : \text{Ind} \\ \text{c2} : \text{in}(\text{refind}, \text{refset}) \end{array} \right] \end{array} \right],$$
 iff  $a \in p(\downarrow P) \wedge \text{Rel}(a.\text{second}, a.\text{first}) = 1 \wedge \exists x[x \in a.\text{second}]$
- **plural reference:** maxset is part of *c-params*.
 
$$a : \left[ \begin{array}{l} \text{c-params} : \left[ \begin{array}{l} \text{maxset} : \text{Set}(\text{Ind}) \\ \text{c1} : \vec{P}(\text{maxset}) \\ \text{refset} : \text{Set}(\text{Ind}) \\ \text{compset} : \text{Set}(\text{Ind}) \end{array} \right] \\ \text{q-cond} : \text{Rel}(\text{q-params.refset}, \text{q-params.compset}) \end{array} \right],$$
 iff  $a = \iota x[x \in p(\downarrow P) \wedge \text{Rel}(x.\text{second}, x.\text{first}) = 1 \wedge x \in \text{common-ground}(\text{spkr}, \text{addr})]$
- **singular reference:** refind is part of *c-params*.
 
$$a : \left[ \begin{array}{l} \text{c-params} : \left[ \begin{array}{l} \text{maxset} : \text{Set}(\text{Ind}) \\ \text{c1} : \vec{P}(\text{maxset}) \\ \text{refset} : \text{Set}(\text{Ind}) \\ \text{compset} : \text{Set}(\text{Ind}) \\ \text{refind} : \text{Ind} \\ \text{c2} : \text{in}(\text{refind}, \text{refset}) \end{array} \right] \end{array} \right],$$
 iff  $a \in p(\downarrow P) \wedge \text{Rel}(a.\text{second}, a.\text{first}) = 1 \wedge \iota x[x \in a.\text{second}] \wedge x \in \text{common-ground}(\text{spkr}, \text{addr})$

An  $\text{NP}_{\text{sem}}$  witness is a triplet of sets—maxset, refset, and compset—such that the maxset is the union of refset and compset and the cardinalities of refset and compset comply with the quantifier condition (q-cond). Singular involves a refind on top. Reference is accounted for in terms of common ground membership, which is compatible with various approaches (e.g. with [19]) but not further developed here. Most importantly, structure and type of  $\text{NP}_{\text{sem}}$  is in accordance with the demands of referential transparency as required in (1).

## 4 Predication, ‘anti-predication’ and *not<sub>Q</sub>*

Visual appearance this time is not deceptive: as the record type representation format suggests, our QNP approach is implemented in a TTR-based variant of a constraint-based grammar (i.e. HPSG [20]). HPSG<sub>TTR</sub> has been developed and motivated in [2, 8] and further refined in [14]. An example for a HPSG<sub>TTR</sub> structure is given in Fig. 2, more details can be found in the just-given sources. The basic idea underlying subject predication is that the verb makes use of the ordered set bipartitions of its argument: A verb phrase, a plural predicate,  $\vec{PType}$ , predicates of the refset of its syntactic subject (feature ‘nucl’) and exerts an ‘anti-predication’ on the compset (‘anti-nucl’). In particular the compset and ‘anti-predication’ provide the prerequisites for integrating noun phrase negation.

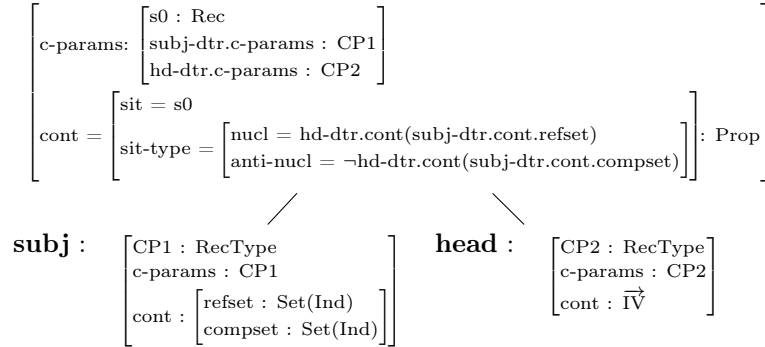


Figure 2: Declarative plural head-subject rule (where  $\vec{\text{IV}}$  labels the type of a plural intransitive verb and  $\text{CP}_x$  the c-params values that get inherited to the mother node)

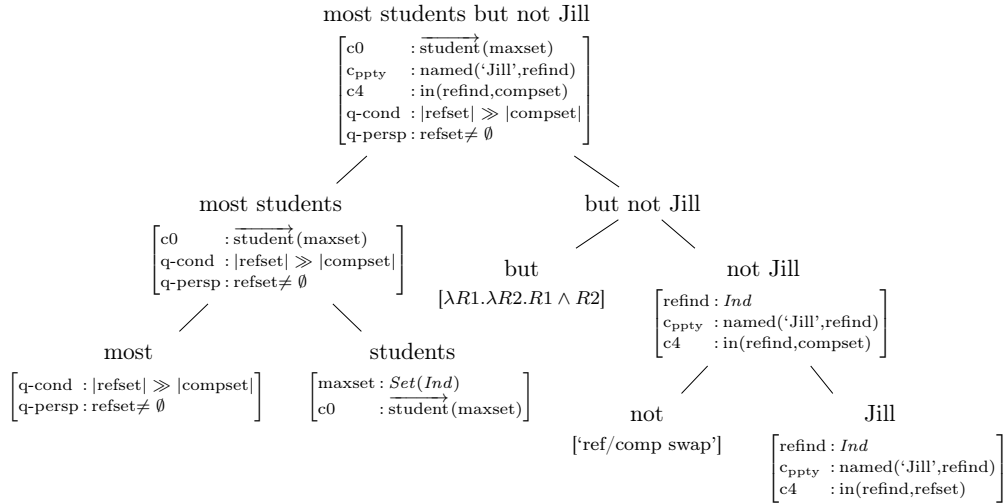


Figure 3: Derivation of the QNP *most students but not Jill*

*Not* when applied to a left decreasing quantifier makes the compset available as an antecedent for anaphora:

- (5) a. Many music lovers admire Mozart. #They [= music lovers not admiring Mozart] prefer Reger instead.
- b. Not many music lovers admire Reger. They [= music lovers not admiring Reger] prefer Mozart instead.

The pattern in (5) might be explained in terms of the QNP’s monotonicity. Compset anaphora is licensed by downward monotone quantifiers [17]. Negation reverses the direction of the upward monotone quantifier *many*, giving rise to (5b). However, compset anaphora under negation is also obtained from non-monotone quantifiers, as shown in (6):



- (6) a. All music lovers admire Mozart. #They [= music lovers not admiring Mozart] prefer Reger instead.
- b. Not all music lovers admire Reger. They [= music lovers not admiring Reger] prefer Mozart instead.

How to model the semantics of the b-sentences in (5) and (6)? Lexical decomposition of the form  $\neg \text{many}.x(\phi)(\psi) \leftrightarrow \text{many}.x(\phi)(\neg\psi)$  [23], for instance, cannot generally be applied, since for the universal quantifier the inequivalence  $\neg \text{all}.x(\phi)(\psi) \not\leftrightarrow \text{all}.x(\phi)(\neg\psi)$  holds. The ‘discourse referent insertion’ account of [25] can derive negated universal quantifiers, but only with a collective scope reading (‘not all  $P$   $Q$ ’ becomes  $\exists X : *P'(X) \wedge \neg Q'(X)$ ). Thus, a full compositional analysis of QNPs including negation is still a semantic desideratum, which we address in terms of a novel negation operation, namely *NP negation*. NP negation, ‘not<sub>Q</sub>’, in essence moves its argument to the ‘anti-witnesses’ contributed by the compset. The basic semantics of ‘not<sub>Q</sub>’ is given by the recursive rule in (7):

(7) **Not<sub>Q</sub>**

- a. ‘refset’ and ‘compset’ labels are swapped.
- b. The relation symbol used in q-cond and q-persp is reversed ( $= \rightarrow \neq$ ,  $\geq \rightarrow <$ ,  $< \rightarrow \geq$ ,  $> \rightarrow \leq$ ,  $\leq \rightarrow >$ ,  $, \gg \rightarrow \not\gg$ ,  $\not\gg \rightarrow \gg$ ,  $\ll \rightarrow \not\ll$ ,  $\not\ll \rightarrow \ll$ ,  $\neq \rightarrow =$ ).

Instead of lexicalising the QNP *most NP<sub>pl</sub> but not  $\alpha$*  ( $\alpha$  being some NP including proper names) as a holistic quantifier of type  $\langle 1, 1, 1 \rangle$  (e.g., *most NP<sub>pl</sub> but not  $\alpha'$*  =  $\{\langle X, Y, Z \rangle : |X \cap Y| > |X \setminus Y| \wedge Z \not\subseteq Y\}$ ), its semantic content is compositionally derived in a way indicated in Fig. 3. Recall that the tree is licensed by grammar rules of the sort exemplified in Fig. 2. The top node in Fig. 3 is the result of two composition methods following the syntactic structure of the NP: firstly, merge (e.g., for combining determiners and head nouns, corresponding to unification), secondly, functional application (*not* acting on its modifiee). The witnessing conditions for the semantic representation of the full QNP follow the pattern given in Sec. 3. Any verb phrase that predicates of this QNP will predicate of the refset but exclude the compset, including Jill. The right witnessing conditions for the sentence are obtained since noun phrase negation is related to the incompatibility semantics of type negation [22, 4]:  $a : \neg T$  iff  $a : T'$  and  $T'$  precludes  $T$ . Now refset and compset are necessarily disjoint. If  $a \in b.\text{second}$  ( $=$  refset) for some  $b \in p(\downarrow P)$ , then  $a \notin b.\text{first}$  ( $=$  compset), for all  $a \in b.\text{second} \cup b.\text{first}$  ( $=$  maxset). It follows that  $[\vee \vec{P}(\text{refset})] \cap [\vee \vec{P}(\text{compset})] = \emptyset$ . For all sets of objects  $A$ , if  $\vec{P}(\text{refset}=A)$  is of a non-empty (now higher-order) type  $\mathcal{T}$ , then it follows that  $\vec{P}(\text{compset}=A)$  is of an empty type  $\mathcal{T}'$ — $\mathcal{T}$  and  $\mathcal{T}'$  preclude each other. From incompatibility we get for any predicate type  $PT$ : if  $PT : \mathcal{T}$  then  $PT : \neg \mathcal{T}'$  and *vice versa*.

The QNP anatomy in addition to noun phrase negation enables the right witnessing conditions for examples like (8a) in a similarly straightforward manner, which eludes a deletion based account in light of (8b) [8]:

- (8) a. Who supports Corbyn? Not Boris.
- b. \*Not Boris supports Corbyn. (Following [15])

Our analysis proceeds as follows: The wh-question introduces a propositional abstract, which abstracts from a record of type [refind : *Ind*] onto the type of supporting Corbyn [8], see (9a). The answer supplies such an individual (which is represented like the *not Jill* node in Fig. 3), and the propositional abstract applies to *not Boris*, resulting in the structure in (9b). Note



that obvious relabelling [3] is employed in order to keep subject (‘-subj’) contributions apart from object (‘-obj’) ones and that structures are shortened to the essential components.

(9) a. *Who supports Corbyn?*  $\mapsto$

$$\left[ \begin{array}{l} \text{sit} = s \\ \text{abstr} = \lambda r : \left[ \begin{array}{l} \text{reset-sbj} : \text{Set}(\text{Ind}) \\ \text{compset-sbj} : \text{Set}(\text{Ind}) \\ \text{c0} : \overrightarrow{\text{person}}(\text{refind-sbj}) \end{array} \right] \left( \left[ \begin{array}{l} \text{c4} : \overrightarrow{\text{support}}(r.\text{refset-sbj},c) \\ \text{c5} : \neg \overrightarrow{\text{support}}(r.\text{compset-sbj},c) \end{array} \right] \right) \end{array} \right]$$

b. Applying *Who supports Corbyn?* to *Not Boris*  $\mapsto$

$$\left[ \begin{array}{l} \text{sit} = s \\ \text{q-params} = \left[ \begin{array}{l} \text{compset-sbj} : \text{Set}(\text{Ind}) \end{array} \right] \\ \text{c-params} = \left[ \begin{array}{l} \text{refind-sbj} : \text{Ind} \\ \text{c1} : \text{in}(\text{refind-sbj}, \text{q-params.compset-sbj}) \\ \text{c6} : \text{named}(\text{refind-sbj}, \text{'Boris'}) \end{array} \right] \\ \text{sit-type} = \left[ \begin{array}{l} \text{c4} : \overrightarrow{\text{support}}(\text{c-params.refset-sbj},c) \\ \text{c5} : \neg \overrightarrow{\text{support}}(\text{q-params.compset-sbj},c) \end{array} \right] \end{array} \right]$$

The question-answer pair in (9b) provides the information that whoever supports Corbyn (i.e. is member of the refset), Boris is part of the people that cannot (the compset).

Since  $\text{NP}_{\text{sem}}$  structures, including singular ones, involve both sets and refinds, anaphora in cases such as (10), which have long been tricky for dynamic semantics [5], can easily be explained:

- (10) A: Go get a bike from the vélib station. B: Oh, but I don’t see any bike that works there.
- a. They are probably rented out. [refset]
- b. It is probably rented out. [refind]

The negative polarity item *any* meaning ‘something’ or ‘a random one out of many’ has two facets. One facet is the existential interpretation which corresponds to a refind analysis. The other facet is the collection of things out of which *any* is picked, which is given by the refset. Both aspects are referred to by the singular and plural pronouns in (10a) and (10b).

The proper name *Peter* refers to an individual from a set of individuals which share the property of being named Peter. This is evidenced in (11a) and (11b):

- (11) A: Where is Peter ... I can’t read the surname? B: There is no Peter in class today.
- a. He is at home with flu. [refind]
- b. They are at home with flu. [compset]

Since *they* refers back to the Peters not in class, (11b) is an instance of compset anaphora. Likewise, *no* shifts the refind in (11a) into the compset. Thus, the compset has at least one Peter in it. In obvious contexts in which (11) could have been uttered some kind of familiarity with the people named Peter can be assumed (cf. the common ground condition in Sec. 3). That is, additional context knowledge can contribute in licensing plural compset anaphora: in case of (11b) there is apparently more than one Peter missing from class.

## 5 A note on processing and incrementality

Our account is additionally motivated by a variety of experimental evidence. We mention two here. First, the extensive work on the reset/compset partition [24] which in particular shows that compset and maxset are not constructed as fallback interpretation options, but, when available, have the same processing status as default antecedents. Second, experimental evidence for incremental interpretation of QNPs in pragmatically supporting contexts [26]—which our account that assigns a QNP an ‘*in situ*’, ‘internal’ (in the sense of not incorporating a projected verbal argument) meaning, is well placed to capture. We achieve this by a mixture of unification (type merge) and functional application as semantic composition techniques. All semantic bits can be addressed by labels, which secures correct semantic composition as well as anaphoric accessibility.

## 6 Conclusion

We presented the main ideas of a ‘referential transparent noun phrase semantics’. Given space considerations, obviously not all complex quantifiers have been discussed here (and the ones discussed deserve a more detailed discussion). However, the motivation and potential of the approach should we hope become apparent. With regards to the range of QNPs which can be analysed in this way, there seems to be only one principled restriction: the NP-internal architecture precludes non-conservative quantifiers. However, one can always resort to a contextually instantiated threshold  $\theta$  in order to account for non-conservative effects.

## 7 Acknowledgments

This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the program ‘Investissements d’Avenir’ (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris – ANR-18-IDEX-0001. This research was also supported by a senior fellowship to the second author by the Institut Universitaire de France.

## References

- [1] Jon Barwise and Robin Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2):159–219, 1981.
- [2] Robin Cooper. Type theory with records and unification-based grammar. In Fritz Hamm and Stephan Kepser, editors, *Logics for Linguistic Structures*, number 201 in Trends in Linguistics: Studies and Monographs, pages 9–33. Mouton de Gruyter, Berlin and New York, 2008.
- [3] Robin Cooper. From perception to communication: An analysis of meaning and action using a theory of types with records (TTR). <https://github.com/robincooper/ttl>, 2019. Unpublished book draft.
- [4] Robin Cooper and Jonathan Ginzburg. Type theory with records for natural language semantics. In Shalom Lappin and Chris Fox, editors, *The Handbook of Contemporary Semantic Theory*, chapter 12, pages 375–407. Wiley-Blackwell, Oxford, UK, 2 edition, 2015.
- [5] Tim Fernando. The donkey strikes back: extending the dynamic interpretation constructively. In *Proceedings of the sixth conference on European chapter of the Association for Computational Linguistics*, pages 130–138. Association for Computational Linguistics, 1993.
- [6] Dominik Freunberger and Mante S. Nieuwland. Incremental comprehension of spoken quantifier sentences: Evidence from brain potentials. *Brain Research*, 1646:475–481, 2016.

- [7] Jean Mark Gawron and Stanley Peters. Some puzzles about pronouns. In Robin Cooper, Kuniaki Mukai, and John Perry, editors, *Situation Theory and Its Applications, I*, CSLI Lecture Notes Number 22, pages 395–431. CSLI Publications, Stanford, 1990.
- [8] Jonathan Ginzburg. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford, UK, 2012.
- [9] Edward L. Keenan. Some properties of natural language quantifiers: Generalized quantifier theory. *Linguistics and Philosophy*, 25(5/6):627–654, 2002.
- [10] Edward L. Keenan and Jonathan Stavi. A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9(3):253–326, 1986.
- [11] Udo Klein. Quantification, witness sets and conservativity. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, and Alex Trueman, editors, *Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 125–133, Somerville, MA, 2012.
- [12] David Lahm. “Different” as a restriction on Skolem functions. In *Proceedings of SALT*, pages 546–565, 2016.
- [13] Godehard Link. The logical analysis of plurals and mass terms: A lattice-theoretic approach. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, Use, and Interpretation of Language*, pages 302–323. de Gruyter, Berlin, 1983.
- [14] Andy Lücking, Jonathan Ginzburg, and Robin Cooper. Grammar in dialogue. In Stefan Müller, Anne Abeillé, Robert D. Borsley, and Jean-Pierre Koenig, editors, *Head-Driven Phrase Structure Grammar: The handbook*. Language Science Press, Berlin, 2019. In production.
- [15] Jerry Morgan. Sentence fragments and the notion ‘sentence’. In B Kachru, editor, *Issues in Linguistics: papers in honour of Henry and Rene Kahane*. UIP, 1973.
- [16] Johan van Benthem. *Essays in Logical Semantics*, chapter 2: Quantifiers. Number 29 in Studies in Linguistics and Philosophy. D. Reidel Publishing Company, Dordrecht, The Netherlands, 1986.
- [17] Rick Nouwen. Complement anaphora and interpretation. *Journal of Semantics*, 20(1):73–113, 2003.
- [18] Stanley Peters and Dag Westerståhl. *Quantifiers in Language and Logic*. Oxford University Press, Oxford, NY, 2006.
- [19] Massimo Poesio. A situation-theoretic formalization of definite description interpretation in plan elaboration dialogues. In Peter Aczel, David Israel, Yasuhiro Katagiri, and Stanley Peters, editors, *Situation Theory and Its Applications*, volume 3, chapter 12, pages 339–374. CSLI Publications, Stanford, CA, 1993.
- [20] Carl Pollard and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. CSLI Publications, Stanford, CA, 1994.
- [21] Matthew Purver and Jonathan Ginzburg. Clarifying noun phrase semantics. *Journal of Semantics*, 21(3):283–339, 2004.
- [22] Aarne Ranta. *Type-Theoretical Grammar*. Clarendon Press, Oxford, 1994.
- [23] Manfred Sailer. Complement anaphora and negative polarity items. In Estela Puig-Waldmüller, editor, *Proceedings of Sinn und Bedeutung 11*, SuB 11, pages 494–508, 2007.
- [24] Anthony J. Sanford, Eugene J. Dawydiak, and Linda M. Moxey. A unified account of quantifier perspective effects in discourse. *Discourse Processes*, 44(1):1–32, 2007.
- [25] Viola Schmitt, Edgar Onea, and Friederike Buch. Restrictions on complement anaphora. In *Proceedings of SALT*, volume 27, pages 212–229, 2017.
- [26] Thomas P. Urbach, Katherine A. DeLong, and Marta Kutas. Quantifiers are incrementally interpreted in context, more than less. *Journal of Memory and Language*, 83:79–96, 2015.
- [27] Kai von Fintel and Edward L. Keenan. Determiners, conservativity, witnesses. *Journal of Semantics*, pages 1–11, 2018.

# Crises of *Identity*\*

Matthew Mandelkern

All Souls College, Oxford, OX1 4AL, United Kingdom  
matthew.mandelkern@all-souls.ox.ac.uk

## Abstract

*Identity* says that any conditional with the form ‘If  $p$ , then  $p$ ’ is a logical truth. I show that a wide range of theories invalidate *Identity*. I argue this is due to a tension between *Identity* and *Import-Export*, and sketch a potential solution.

## 1 Introduction

There is much controversy about the logic of the conditional. One principle that has so far evaded controversy is *Identity*, which says that conditionals with the form ‘If  $p$ , then  $p$ ’ are logical truths. In the first part of this paper, I show that, despite its overwhelming plausibility, a wide variety of theories of the conditional invalidate *Identity*. I then argue that the culprit behind this failure is the *Import-Export* (*IE*) principle, which says that ‘If  $p$ , then if  $q$ , then  $r$ ’ and ‘If  $p$  and  $q$ , then  $r$ ’ are always equivalent. I show that there is a deep and surprising tension between *IE*, on the one hand, and *Identity*, on the other. In light of this tension, and the overwhelming plausibility of *Identity*, I argue we should reject *IE*. In the final part of the paper, I explore how we might reject *IE* while still accounting for the intuitive evidence that supports it.

## 2 Failures of *Identity*

*Identity* says that conditionals with the form ‘If  $p$ , then  $p$ ’, like those in (1), are logical truths:

- (1) a. If it rained, then it rained.
- b. If it had rained, then it would have rained.

*Identity* is extremely natural. Nonetheless, it is invalidated by a wide range of theories of the conditional, as I will show in the rest of this section.

I work with a simple propositional language with atoms  $A, B, C \dots$ , closed under ‘ $\wedge$ ’ (‘and’), ‘ $\neg$ ’ (‘not’), ‘ $\vee$ ’ (‘or’), the material conditional ‘ $\supset$ ’ (‘ $p \supset q$ ’ abbreviates ‘ $\neg p \vee q$ ’), the material biconditional ‘ $\equiv$ ’ (‘ $p \equiv q$ ’ abbreviates ‘ $(p \supset q) \wedge (q \supset p)$ ’), and the natural language conditional connective ‘ $>$ ’ (‘ $p > q$ ’ abbreviates ‘If  $p$ , then  $q$ ’). Lower-case italics range over sentences. Where  $\Gamma$  is a set of sentences, ‘ $\Gamma \models p$ ’ means that  $\Gamma$  logically entails  $p$ , in the standard classical sense, i.e. that  $p$  is true in every intended model where all the elements of  $\Gamma$  are true.

A key player in what follows is the *Import-Export* (*IE*) principle, which says that  $(p > (q > r)) \equiv ((p \wedge q) > r)$  is a logical truth. In other words, *IE* says that what we do with two

---

\*Many thanks to audiences at Oxford, NYU, MCMP, and UCL, three reviewers for the Amsterdam Colloquium, Andrew Bacon, Kyle Blumberg, Susanne Bobzien, David Boylan, Fabrizio Cariani, Sam Carter, Ivano Ciardelli, Nilanjan Das, Kevin Dorst, Kit Fine, Branden Fitelson, Vera Flocke, Simon Goldstein, Jeremy Goodman, Ben Holguín, Justin Khoo, Cameron Domenico Kirk-Giannini, Angelika Kratzer, Vann McGee, Eliot Michaelson, Jonathan Phillips, Milo Phillips-Brown, Daniel Rothschild, Ian Rumfitt, Bernhard Salow, Philippe Schlenker, Ginger Schultheis, Robert Stalnaker, and especially Cian Dorr for very helpful discussion.

successive conditional antecedents is the same as what we do with the corresponding conjunctive antecedent. So, e.g., *IE* says the members of pairs like the following are equivalent:

- (2) a. If the coin is flipped, then if it lands heads, then we'll win.
- b. If the coin is flipped and it lands heads, then we'll win.
- (3) a. If the coin had been flipped, then if it had landed heads, then we would have won.
- b. If the coin had been flipped and it had landed heads, then we would have won.

My central claim in this section is that all extant theories, apart from the material analysis, which validate *IE* also *invalidate Identity*. To see the point informally, consider what it takes to validate *IE*. *IE* says, in essence, that information in successive antecedents is agglomerated: a conditional with two antecedents is evaluated in the same way as a conditional with one corresponding conjunctive antecedent. That means that, to validate *IE*, we need some way of “remembering” successive conditional antecedents. Different *IE*-validating theories have different mechanisms for doing this. For instance, in McGee (1985)’s framework, conditional antecedents are added sequentially to a set of sentences; the consequent is then evaluated at the closest world where all the sentences in that set are true. In Kratzer (1981, 1991)’s framework, conditional antecedents are similarly added to the value of a “modal base” function which takes each world to a set of propositions, which in turn provides the domain of quantification for evaluating the consequent. Relevantly similar approaches are developed in von Stechow (1994); Dekker (1993); Gillies (2004, 2009). These theories differ in important ways, but they all have what I’ll generically call a *domain parameter* of some kind which is in the business of somehow remembering successive conditional antecedents, so that these can be agglomerated when we arrive at the consequent. Intuitively, that is exactly what is needed in order to validate *IE*.

Structurally, this has an important consequence. What proposition a conditional expresses depends on the setting of this domain parameter. And thus, since conditional antecedents change this parameter, what proposition a conditional expresses can change depending on whether it is embedded under a conditional antecedent. In particular, consider a sentence with the form  $p > p$  and suppose that  $p$  itself contains a conditional. Then the first instance of  $p$  will be interpreted relative to a domain parameter with a different value from that used to evaluate the second  $p$ : when we get to the second, the domain parameter will have been updated with the information that  $p$  is true. And that, in turn, means that the two instances of  $p$  can express different propositions, and so the conditional as a whole can be false.

More concretely, think about a conditional of the form  $(\neg(B > A) \wedge A) > (\neg(B > A) \wedge A)$ , where  $A$  and  $B$  are arbitrary atoms. This has the form  $p > p$ . Now consider what happens when we arrive at the consequent of this conditional, if we validate *IE*. At that point, the antecedent will have been added to our domain parameter. So the domain parameter will now entail the antecedent, and so in particular will entail  $A$ . That means that the parameter will only make available  $A$ -worlds for the evaluation of the conditional in the consequent. The consequent,  $(\neg(B > A) \wedge A)$ , entails that a certain conditional,  $B > A$ , is false. The problem is that if our domain parameter—which gives the domain of worlds which matter for evaluating the conditional—includes only  $A$ -worlds, then  $B > A$  can’t be false. That means that  $B > A$ , as it appears in the consequent of our target conditional, must be true; so its negation must be false; so the whole consequent must be false; and so the target conditional as a whole will be false, provided its antecedent is possible.

We can illustrate this in more detail by looking at McGee’s McGee (1985) theory, a variant on Stalnaker (1968)’s theory which validates *IE*. Let  $f$  be a Stalnakerian selection function from any consistent proposition and world to the “closest” world where that proposition is true.  $\Gamma$  is any set of sentences (a *hypothesis set*).  $\mathcal{J}$  is an atomic valuation function. McGee’s

theory says that any sentence is true relative to an absurd hypothesis set, i.e.  $\llbracket p \rrbracket^{f,\Gamma,w} = 1$  if  $\bigcap_{r \in \Gamma} \llbracket r \rrbracket^{f,\emptyset} = \emptyset$ ; an atom is true iff it is true at the closest world where the hypothesis set is true, i.e.  $\llbracket A \rrbracket^{f,\Gamma,w} = 1$  iff  $f(\bigcap_{r \in \Gamma} \llbracket r \rrbracket^{f,\emptyset}, w) \in \mathcal{J}(A)$ ; a negation is true iff the negatum is false, i.e.  $\llbracket \neg p \rrbracket^{f,\Gamma,w} = 1$  iff  $\llbracket p \rrbracket^{f,\Gamma,w} = 0$ ; a conjunction is true iff both conjuncts are, i.e.  $\llbracket p \wedge q \rrbracket^{f,\Gamma,w} = 1$  iff  $\llbracket p \rrbracket^{f,\Gamma,w} = 1$  and  $\llbracket q \rrbracket^{f,\Gamma,w} = 1$ ; finally, a conditional is true iff the consequent is true relative to a hypothesis set updated with the antecedent, i.e.  $\llbracket p > q \rrbracket^{f,\Gamma,w} = \llbracket q \rrbracket^{f,\Gamma \cup \{p\},w}$ . This matches Stalnaker's theory for simple conditionals (e.g.  $A > B$  is true iff  $B$  is true at the closest  $A$ -world); but the hypothesis set lets us agglomerate successive antecedents, so we validate *IE*, unlike Stalnaker. Now consider  $\llbracket (\neg(B > A) \wedge A) > (\neg(B > A) \wedge A) \rrbracket^{f,\emptyset,w}$ . By the clause for conditionals, this is true iff  $\llbracket (\neg(B > A) \wedge A) \rrbracket^{f,\{(\neg(B > A) \wedge A)\},w} = 1$ , which, by the clause for conjunction, holds only if  $\llbracket \neg(B > A) \rrbracket^{f,\{(\neg(B > A) \wedge A)\},w} = 1$ . That in turn holds, by the clause for negation, iff  $\llbracket B > A \rrbracket^{f,\{(\neg(B > A) \wedge A)\},w} = 0$ , which, by the clause for conditionals, holds iff  $\llbracket A \rrbracket^{f,\{(\neg(B > A) \wedge A), B\},w} = 0$ . But  $\llbracket A \rrbracket^{f,\{(\neg(B > A) \wedge A), B\},w} = 1$  by the first clause, since  $\llbracket \neg(B > A) \wedge A \rrbracket^{f,\emptyset} \cap \llbracket B \rrbracket^{f,\emptyset} = \emptyset$  (since  $B \wedge A$  entails  $B > A$  in McGee's system); so  $\llbracket (\neg(B > A) \wedge A) > (\neg(B > A) \wedge A) \rrbracket^{f,\emptyset,w} = 0$ . Similar reasoning leads to the same conclusion for other *IE*-validating theories.

### 3 Locating the tension

Must we invalidate *Identity* if we validate *IE*? No: these principles are jointly consistent. A quick way to see this is that the material conditional ' $\supset$ ' validates both. Recall that the material conditional is the truth-function such that  $p \supset q$  is true whenever  $p$  is false or  $q$  true. This connective validates both *IE* and *Identity*.

But this is not much help, because the material conditional is not an adequate analysis of the natural language 'if... then'. This is the near consensus view, and there are myriad arguments for it. A quick way to see the implausibility of the material analysis is that, since  $p \supset q$  is equivalent to  $\neg p \vee q$ , its negation is equivalent to the conjunction  $p \wedge \neg q$ . But it is clear that the negation of  $p > q$  is not equivalent to  $p \wedge \neg q$ . For instance, 'It's not the case that, if Patch had been a rabbit, she would have been a rodent' and 'It's not the case that, if Patch is a rabbit, she is a rodent' are both clearly true, thanks just to taxonomic facts; neither entails that Patch is a rabbit, *pace* the material view.

The material conditional is the only extant theory I know of which validates both *Identity* and *IE*. It is not the only logically possible one. Nonetheless, I will argue that there is no theory which validates *Identity* and *IE* together in a *plausible* way. I will argue for this by showing that the material conditional is the only connective which validates both *Identity* and *IE* together with two very weak, and very plausible, background principles.

The first principle is a very weak monotonicity principle, which says that  $p > (q \wedge r)$  entails  $p > q$ ; call this *Left Consequent Monotonicity (LCM)*. *LCM* is as far as I know validated by every extant theory, and is of course a very limited corollary of the widely accepted principle that conditionals are monotonic in their consequents. *LCM* is what is required to predict that inferences like the following will be valid: 'If it rains, the picnic will be cancelled and Sue will be upset; therefore, if it rains, the picnic will be cancelled'. The second principle says that  $\neg p$  follows from  $p > q$  together with  $p > \neg q$ ; call this principle *Ad Falsum*. *Ad Falsum*, like *LCM*, is validated by every theory I know of (with the exception of the existential theories in Bassi and Bar-Lev (2018); Herburger (2019)). The most direct evidence for *Ad Falsum* comes from logical and mathematical contexts, where it is very natural to argue that  $p$  is false by showing

that, if  $p$ , then  $q$ , and if  $p$ , then not  $q$ . This reasoning, however, is only valid if *Ad Falsum* is. While this reasoning is most at home in mathematical and logical contexts, it also seems perfectly valid in non-mathematical contexts, as in Gibbard (1981)'s Sly Pete case, where we learn both 'If Pete called, he won' and 'If Pete called, he lost' and can conclude with certainty that he didn't call. More generally, it seems unimaginable that two conditionals with this form could be true, while their antecedent was also true.

But if we validate *Identity*, *IE*, *LCM*, and *Ad Falsum*, then ' $>$ ' must be logically equivalent to ' $\supset$ '. We assume classical logics for ' $\wedge$ ', ' $\vee$ ', and ' $\neg$ ', uniform substitutability for sentence letters, and substitutability of logical equivalents. For arbitrary  $p$  and  $q$ , given *Identity*, we have  $\models (\neg(p > q) \wedge q) > (\neg(p > q) \wedge q)$ ; given *LCM*, we then have  $\models (\neg(p > q) \wedge q) > \neg(p > q)$ . *Identity* also gives us  $\models ((\neg(p > q) \wedge q) \wedge p) > ((\neg(p > q) \wedge q) \wedge p)$ ; substitution of logical equivalents and *LCM* then give us  $\models ((\neg(p > q) \wedge q) \wedge p) > q$ ; thus by *IE*,  $\models (\neg(p > q) \wedge q) > (p > q)$ . So by *Ad Falsum* we have  $\models \neg(\neg(p > q) \wedge q)$ . By classical logic we have  $q \models p > q$ . By classical logic and uniform substitution we have  $((p \wedge (p > q)) \wedge \neg q) \models ((p > q) \wedge (p > \neg q))$ ; *Ad Falsum* thus tells us that  $((p \wedge (p > q)) \wedge \neg q) \models \neg p$ ; by *reductio*, we have  $p \wedge (p > q) \models q$ , i.e. *Modus Ponens (MP)* for ' $>$ '. By *Identity* we have  $\models ((p \supset q) \wedge p) > ((p \supset q) \wedge p)$ ; by substitutability of logical equivalents and *LCM* we have  $\models ((p \supset q) \wedge p) > q$ ; by *IE*  $\models (p \supset q) > (p > q)$ ; by *MP*,  $p \supset q \models p > q$ ; *MP* also gives us  $p > q \models p \supset q$ ; and so  $p > q$  and  $p \supset q$  are logically equivalent. In sum: the only conditional that validates *Identity*, *IE*, *LCM*, and *Ad Falsum* is the material conditional. Since we know that the conditional is not material, one of these principles must be invalid.

Our result strengthens a result of Gibbard (1981), which showed that only the material conditional validates all three of *IE*, *MP*, and *Logical Implication (LI)*, which says that, if  $p \models q$ , then  $\models p > q$ . *LI* follows from *Identity* and *LCM* given substitutability of logical equivalents, so Gibbard's result shows that only the material conditional validates all of *IE*, *MP*, *Identity*, and *LCM*. Our result replaces *MP* with *Ad Falsum*. *Ad Falsum* follows from *MP*, but is weaker than *MP*, so our result strengthens Gibbard's. This is dialectically important, because, while a strong case has been mounted against *MP* by McGee (1985), and many theories have been advanced which invalidate *MP* (including all the *IE*-validating theories cited above), I know of no case against *Ad Falsum*, and every theory I know of, even those that invalidate *MP*, still validates *Ad Falsum*.<sup>1</sup> Moreover, I do not see any prospects for a case against *Ad Falsum* on the basis of examples like McGee's or indeed on any other. The reaction to Gibbard's result has mainly focused on the choice between *MP* and *IE*. The present result shows that that reaction misses an important tension which already exists between *IE* and *Identity*.

## 4 Rejecting *Identity*?

If we accept *Ad Falsum* and *LCM*, then the present result shows we must choose between rejecting *Identity* or rejecting *IE*, given that the conditional is not material. Let us consider first whether there could be a case against *Identity*. It is hard to see *a priori* how there could be. If  $p$  holds, then it seems certain that, whatever else holds,  $p$  does. But we should not be too quick to dismiss a potential case against *Identity*; it is at least conceivable that *Identity* could fail in the case of complex conditionals, exactly where theories like McGee's predict it will.

Indeed, the foregoing discussion gives us a precise way to explore this possibility. Consider sentences with the form  $(B \wedge \neg(A > B)) > \neg(A > B)$ , a slightly simpler variant on the sentences considered above. Sentences with this form are predicted by *Identity* to be logical truths

<sup>1</sup>With the exception, again, of the existential theories mentioned above.



(assuming *LCM*). By contrast, theories that validate *IE* predict that the internal negation of such conditionals are instead logical truths. So, to assess whether a case can be made against *Identity* (and thus in favor of *IE*) on the basis of complex conditionals, we can look at sentences with forms  $(B \wedge \neg(A > B)) > \neg(A > B)$  and  $(B \wedge \neg(A > B)) > (A > B)$ , respectively, as in:

- (4) a. If the match had lit, and it's not the case that the match would have lit if it had been wet, then it's not the case that the match would have lit if it had been wet.  
 b. If the match had lit, and it's not the case that the match would have lit if it had been wet, then the match would have lit if it had been wet.
- (5) a. If the vase had broken, and it's not the case that the vase would have broken if it had been wrapped in plastic, then it's not the case that the vase would have broken if it had been wrapped in plastic.  
 b. If the vase had broken, and it's not the case that the vase would have broken if it had been wrapped in plastic, then the vase would have broken if it had been wrapped in plastic.

Although these are complicated, it seems clear that the first variants in each pair are logical truths, while the second variants are logical falsehoods (assuming their antecedents are possible). Thus it seems that *Identity*, not *IE*, has the correct verdict here. This makes me pessimistic that any case can be constructed against *Identity* on the basis of complex conditionals.

## 5 Rejecting *Import-Export*

We should thus reject *IE*. But the motivation for this so far is relatively indirect. It would be nice to find more direct evidence against *IE*—that is, pairs with the form  $p > (q > r)$  and  $(p \wedge q) > r$  which are intuitively inequivalent. Strikingly, subjunctive conditionals seem to yield such pairs, while indicatives appear not to. Consider first subjunctive pairs like Etlin (2008)'s (6), and Stephen Yablo's (7) (p.c.):

- (6) a. If the match had lit, then it would have lit if it had been wet.  
 b. If the match had lit and it had been wet, then it would have lit.
- (7) a. If I had been exactly 6' tall, then if I had been a bit taller than 6', I would have been 6'1".  
 b. If I had been exactly 6' tall and a bit taller than 6', I would have been 6'1".

These pairs instantiate *IE* but are intuitively inequivalent. (6-a) can be false in some circumstances, whereas (6-b) cannot. Conversely, (7-a) is plausibly true in some circumstances; whereas (7-b) feels like nonsense (if it is ever true, then it is surely only true trivially, unlike (7-a)). These felt inequivalences target the two directions of *IE*. Pairs like this suggest that neither direction is valid for subjunctives.

When we turn to indicatives, however, things look different. Consider the indicative versions of the pairs we have just looked at:

- (8) a. If the match lit, then it lit if it was wet.  
 b. If the match lit and it was wet, then it lit.
- (9) a. If I am exactly 6' tall, then if I am a bit taller than 6', then I am 6'1".  
 b. If I am exactly 6' tall and a bit taller than 6', then I am 6'1".

Unlike the corresponding subjunctive pairs, these appear to be pairwise equivalent: both of the



sentences in (8) feel equally like logical truths; and both of the sentences in (9) feel equally incoherent. More generally, I have not been able to find concrete counterexamples to *IE* for indicatives: the very same pairs which seem to witness the *invalidity* of *IE* for subjunctives seem to provide further support for its *validity* when it comes to indicatives.

## 6 Strawson validity

We thus need a theory which (i) accounts for the contrast just observed between indicatives and subjunctives; and (ii) accounts for the *felt* validity of *IE* in the case of indicatives without also invalidating *Identity*. Let us focus for the moment on indicatives. Summing up our evidence, we have strong indirect evidence that *IE* is invalid, but we do not seem to find concrete counterinstances to it. One response to this kind of situation is to say that the inference pattern in question is not logically valid, but still preserves truth in an important subset of cases. This, in turn, makes it hard to find concrete counterinstances to it. Following Strawson (1952); von Fintel (1999), we can distinguish *logical entailment* (preservation of truth in all intended models) from *Strawson entailment*: preservation of truth in all intended models *where we might actually use the sentences in question*. More precisely, we can associate sentences with presuppositions—conditions on their felicitous use—and then define Strawson entailment as follows:

**Strawson entailment:**  $\Gamma$  Strawson entails  $p$  iff for any intended model  $\mathcal{M}$ , context  $c$  and world  $w \in c$ , if the presuppositions of all the elements of  $\Gamma$  and of  $p$  are satisfied in  $\langle c, w \rangle$  in  $\mathcal{M}$  and all the elements of  $\Gamma$  are true at  $\langle c, w \rangle$  in  $\mathcal{M}$ , so is  $p$ .

If an inference is Strawson valid, it doesn't necessarily preserve truth in all worlds in all intended models. But it does preserve truth in any context-world pair where all the premises and the conclusion have their presuppositions satisfied—which includes all contexts where the sentences in question can be naturally used. So, if an inference is Strawson valid, there won't be *natural* concrete counterexamples to it, even if the inference is not *logically* valid.

There are various ways we can model presuppositions, and thus Strawson validity. Here I follow multi-dimensional approaches (Herzberger (1973); Karttunen and Peters (1979)) which distinguish truth and presupposition as two dimensions of content.<sup>2</sup> The first dimension, which I underline for mnemonic reasons, records presupposition satisfaction; the second records truth. So, where  $*$  is shorthand for either 1 or 0, the inference from  $p$  to  $q$  is *Strawson* valid iff, for any intended model and any  $c$  and  $w \in c$ ,  $\llbracket q \rrbracket^{c,w}$  is  $\langle \underline{1}, 0 \rangle$  if  $\llbracket p \rrbracket^{c,w}$  is  $\langle \underline{1}, 1 \rangle$  in that model; whereas the inference is *logically* valid iff, for any intended model and any  $c$  and  $w \in c$ ,  $\llbracket q \rrbracket^{c,w}$  is  $\langle *, 1 \rangle$  if  $\llbracket p \rrbracket^{c,w}$  is in that model.

## 7 The local indicative constraint

With this discussion in hand, I can state my proposal: to find an account of the presuppositions of indicative conditionals which predicts that *IE* is Strawson valid for indicatives, but not subjunctives. There is, in fact, an existing proposal which associates indicatives with a presupposition which subjunctives lack. The idea, proposed in Stalnaker (1975), slightly strengthened in von Fintel (1998), and widely accepted, is that indicatives presuppose that their antecedent is evaluated at a contextually possible world. One motivation for this is contrasts like (10):

<sup>2</sup>The kind of presupposition I focus on here is *expressive*, not semantic, insofar as it projects universally, and cannot be filtered; I leave the universal projection rules implicit.

- (10) Beau Balou won the race.
- a. #If he didn't win, we lost a lot of money.
  - b. If he hadn't won, we would have lost a lot of money.

In general, using ' $>_i$ ' now for the indicative conditional and ' $>_s$ ' for the subjunctive,  $p >_i q$  is infelicitous when  $p$  has been ruled out in a context, while  $p >_s q$  remains acceptable. Another motivation is the felt validity of the 'or'-to-'if' inference for indicatives, but not subjunctives:

- (11) a. It was the gardener or the butler, and it might have been either.  
 b.  $\rightsquigarrow$  If it wasn't the gardener, it was the butler.  
 c.  $\not\rightarrow$  If it hadn't been the gardener, it would have been the butler.

In general, the inference from  $p \vee q$  to  $\neg p >_i q$  feels legitimate (provided  $p$  is contextually possible), while the inference to  $\neg p >_s q$  does not.

To capture these two patterns, start with the semantics for the conditional from [Stalnaker \(1968\)](#), on which a conditional says that the closest antecedent-world (according to the contextual selection function) is a consequent-world. Then we say that  $p >_i q$  presupposes that, for any context world  $w$ , the closest  $p$ -world to  $w$  is in the context; while  $p >_s q$  does not have a parallel presupposition. This accounts for our two generalizations.

This presupposition, which I will call the *indicative constraint*, does not on its own help with *IE*. But a close extension does. The motivation for the extension comes from the observation that the compatibility requirement which motivates the indicative constraint resurfaces at a *local* level. Consider (12):

- (12) I don't know whether Bob came to the party.
- a. #But suppose that Bob came, and that if he didn't, he went to work.
  - b. But suppose that Bob came, and that if he hadn't, he would have gone to work.

The embedded indicative in (12-a) is infelicitous, in contrast to the subjunctive variant in (12-b). This is surprising because, relative to the global context in (12), it is possible that Bob didn't go to the party, and so the indicative antecedent is compatible with the global context. To account for the contrast in (12), it looks like we need to compute our compatibility requirement for indicatives relative to a *local* context which takes into account the information in the left conjunct in (12-a)—that Bob came to the party. Similar contrasts in other environments support this point. For instance, consider nested conditionals. Suppose we have a die which is either weighted towards evens or odds; we don't know which, and we don't know whether the die was thrown. Compare (13-a) and (13-b):

- (13) a. #If the die was thrown and landed four, then if it didn't land four, it landed two or six.  
 b. If the die had been thrown and landed four, then if it hadn't landed four it would have landed two or six.

The antecedent of the embedded conditional in (13-a) and (13-b)—that the die didn't land four—is compatible with the global context. But, embedded under a conditional antecedent that entails that the die landed four, only the subjunctive variant in (13-b) seems acceptable, while the indicative variant is not. Once more, it looks like the indicative conditional's compatibility constraint in the consequent of a conditional is calculated relative to a local context: in this case, one which entails the information in the conditional's antecedent.<sup>3</sup>

<sup>3</sup>[Boylan and Schultheis \(2019\)](#) provide yet another motivation for a local version of the indicative constraint,

That the indicative constraint is calculated locally in fact looks unsurprising from the point of view of the recent literature on epistemic modality. That literature has suggested that epistemic accessibility is calculated in a local manner in general, and so it is not surprising that indicative’s epistemic compatibility constraint is also local.<sup>4</sup> There are different ways we could capture the locality of the indicative constraint. For each “local” theory of epistemic modality, we could build a corresponding local indicative constraint in roughly similar fashion. Here I will build loosely on the bounded theory of Mandelkern (2019). That theory borrows Schlenker (2009)’s account of local contexts from his theory of presupposition. A local context, on Schlenker’s account, is a set of worlds which represents the information locally available relative to a given syntactic environment and global context: in other words, the information that could be added to that environment without changing the contextual meaning of the sentence as a whole. The local context for a conditional’s consequent, for instance, entails its antecedent; the local context for a right conjunct entails the left conjunct. The bounded theory posits that epistemic modals presuppose that their accessibility relation is *local* in the sense that only local context worlds can be accessed from local context worlds.

I propose to localize the indicative constraint in parallel to this. Recall that the indicative constraint says that the indicative selection function must take any context world and indicative antecedent to a context world. We need only change ‘context’ for ‘local context’ to get an appropriately local version. In other words, where  $\kappa$  represents the local context for the conditional, our *local indicative constraint* says that  $p >_i q$  presupposes that, for any world  $w$  in  $\kappa$ , the closest  $p$ -world to  $w$  is also in  $\kappa$ . By contrast, the subjunctive has no similar constraint.

For unembedded conditionals, the local indicative constraint is equivalent to the global indicative constraint. But things are different for embedded conditionals. The local context for a right conjunct will be the intersection of the global context and the left conjunct.<sup>5</sup> So, in a global context  $c$ , the local context for the conditional in  $\neg p \wedge (p >_i q)$  will be  $c \cap \neg p$ . That means that, whenever the local indicative constraint is satisfied, for any world  $w'$  in  $c \cap \neg p$ ,  $f_i(p, w') \in (c \cap \neg p)$ . But this is incoherent, since  $f_i(p, w')$  must be a  $p$ -world (if  $p$  is consistent), and thus cannot be in  $c \cap \neg p$ .<sup>6</sup> So there will be no way to satisfy the local indicative constraint for an indicative conditional in a conjunction like this. Crucially, this reasoning goes through whether  $\neg p \wedge (p >_i q)$  is embedded or unembedded, accounting for the infelicity of sentences like (12-a). Likewise, the local context for the consequent of an indicative conditional is the global context intersected with the antecedent. So, in a conditional with the form  $p >_i (\neg p >_i q)$  in global context  $c$ , the local context for the consequent is  $c \cap p$ ; so the local indicative constraint for the embedded conditional will require that the closest  $\neg p$ -world to the closest  $p$ -world be in  $c \cap p$ , and once more will be unsatisfiable. Since subjunctives do not have a local indicative constraint, none of this reasoning will go through for them, accounting for the contrasts above.

Summing up, with  $f_i$  the indicative selection function and  $f_s$  the subjunctive selection function (for readability I sometimes write ‘ $p$ ’ for ‘ $\llbracket p \rrbracket^{\kappa, f_i, f_s}$ ’ when there is no danger of ambiguity):

- $\llbracket p >_s q \rrbracket^{\kappa, f_i, f_s, w} = \langle \underline{1}, 1 \rangle$  iff  $\llbracket q \rrbracket^{\kappa, f_s(p, w') : w' \in \kappa, f_i, f_s, f_s(p, w)} = \langle \underline{1}, 1 \rangle$

from embedded cases of ‘or’-to-‘if’.

<sup>4</sup>E.g. Groenendijk et al. (1996); Aloni (2001); Yalcin (2007); Dorr and Hawthorne (2013); Mandelkern (2019).

<sup>5</sup>In general, incorporating local contexts for the Boolean connectives yields:  $\llbracket p \wedge q \rrbracket^{\kappa, f_i, f_s} = \llbracket p \rrbracket^{\kappa \cap [q]}^{\kappa, f_i, f_s, f_i, f_s} \cap \llbracket q \rrbracket^{\kappa \cap [p]}^{\kappa, f_i, f_s, f_i, f_s}$ ;  $\llbracket p \vee q \rrbracket^{\kappa, f_i, f_s} = \llbracket p \rrbracket^{\kappa \cap [\neg q]}^{\kappa, f_i, f_s, f_i, f_s} \cup \llbracket q \rrbracket^{\kappa \cap [\neg p]}^{\kappa, f_i, f_s, f_i, f_s}$ ; and  $\llbracket \neg p \rrbracket^{\kappa, f_i, f_s} = \mathcal{W} \setminus \llbracket p \rrbracket^{\kappa, f_i, f_s}$ . I omit the world parameter to denote the set of worlds where the sentence in question is true, so  $\llbracket p \rrbracket^{\kappa, f_i, f_s} = \{w : \llbracket p \rrbracket^{\kappa, f_i, f_s, w} = \langle \ast, 1 \rangle\}$ . I leave implicit universal projection rules for presupposition satisfaction.

<sup>6</sup>Selection functions must meet four constraints. Where  $\varphi$  and  $\psi$  are propositions: *Strong Centering*:  $f(\varphi, w) = w$  iff  $w \in \varphi$ ; *Success*:  $f(\varphi, w) \in \varphi$  provided  $\varphi \neq \emptyset$ ; *CSO*: if  $f(\varphi, w) \in \psi$  and  $f(\psi, w) \in \varphi$ , then  $f(\varphi, w) = f(\psi, w)$ ; and *Absurdity*:  $f(\emptyset, w) = \lambda$ , where  $\lambda$  is an absurd world that makes all sentences true.

[Stalnaker truth conditions]

- $\llbracket p >_i q \rrbracket^{\kappa, f_i, f_s, w}$ 
  - =  $\langle \underline{1}, * \rangle$  iff  $\forall w' \in \kappa : f_i(p, w') \in \kappa$  [local indicative constraint]
  - =  $\langle \underline{*}, 1 \rangle$  iff  $\llbracket q \rrbracket^{\kappa \cap p, f_i, f_s, f_i(p, w)} = \langle \underline{1}, 1 \rangle$  [Stalnaker truth conditions]

Now back to *IE*: strikingly, the local indicative constraint guarantees that *IE* is Strawson valid for indicatives, i.e. that the conjunction  $((p >_i (q >_i r)) \supset ((p \wedge q) >_i r)) \wedge (((p \wedge q) >_i r) \supset (p >_i (q >_i r)))$  is never  $\langle \underline{1}, 0 \rangle$  for any  $c$  and  $w \in c$  in any intended model. Take each conjunct in turn, considering first  $(p >_i (q >_i r)) \supset ((p \wedge q) >_i r)$ . Suppose there is an index  $\mathbf{i} = \langle c, f_i, f_s, w \rangle$  with  $w \in c$  such that the presuppositions of all the indicative conditionals are satisfied at  $\mathbf{i}$  but this material conditional has a true antecedent and false consequent at  $\mathbf{i}$ . The local context for the consequent of a material conditional is the global context together with its antecedent. Thus  $(p \wedge q) >_i r$  is false at  $\langle c \cap (p >_i (q >_i r)), f_i, f_s, w \rangle$ . By the local indicative constraint, since  $w \in c \cap (p >_i (q >_i r))$ ,  $f_i(p \wedge q, w) \in c \cap (p >_i (q >_i r))$ . But any index that makes both  $p \wedge q$  true and makes  $p >_i (q >_i r)$  true makes  $r$  true; so  $(p \wedge q) >_i r$  is true at  $\langle c \cap (p >_i (q >_i r)), f_i, f_s, w \rangle$ , contrary to assumption. Next consider  $((p \wedge q) >_i r) \supset (p >_i (q >_i r))$ . Suppose the presuppositions of all the indicative conditionals are satisfied at some index  $\mathbf{i} = \langle c, f_i, f_s, w \rangle$  with  $w \in c$  but the material conditional is false at  $\mathbf{i}$ . Then  $p >_i (q >_i r)$  is false at  $\langle c \cap ((p \wedge q) >_i r), f_i, f_s, w \rangle$ , and  $w \in c \cap ((p \wedge q) >_i r)$ . By the local indicative constraint,  $f_i(p, w) \in c \cap ((p \wedge q) >_i r)$ , and by *Success*  $f_i(p, w) \in c \cap ((p \wedge q) >_i r) \cap p$ . So again by the local indicative constraint,  $f_i(q, f_i(p, w)) \in c \cap ((p \wedge q) >_i r) \cap p$ , and so will be a  $p \wedge q$ -world and a  $(p \wedge q) >_i r$ -world and hence an  $r$ -world, so  $p >_i (q >_i r)$  is true at  $\langle c \cap ((p \wedge q) >_i r), f_i, f_s, w \rangle$ , contrary to assumption. This reasoning turns crucially on the local indicative constraint, so nothing similar follows for subjunctives.

## 8 Conclusion

The only way to validate *Identity*, *IE*, *LCM* and *Ad Falsum* together is with the material conditional. This helps explain why every extant theory of the conditional which validates *IE*, other than the material analysis, invalidates *Identity*. *Identity*, however, appears to be valid. And I cannot see any case against *LCM* or *Ad Falsum*. So we must reject *IE*. This fits well with the empirical evidence in the case of subjunctives, where the existing literature contains concrete counterinstances to *IE*. But the same pairs of conditionals that constitute counterexamples in the subjunctive mood feel pairwise equivalent in the indicative mood. This makes it difficult to reject *IE* for indicatives. I have explored one way we might account for this situation: ascribe to indicative conditionals a presupposition which predicts that *IE* is Strawson valid, but not logically valid. That presupposition, the *local indicative constraint*, is independently motivated on the basis of unrelated contrasts between indicatives and subjunctives, and naturally accounts for the apparent differences in their logics which I have emphasized here.

## References

- Maria D. Aloni. *Quantification Under Conceptual Covers*. PhD thesis, University of Amsterdam, 2001.
- Itai Bassi and Moshe E. Bar-Lev. A unified existential semantics for bare conditionals. In *Sinn und Bedeutung*, volume 21, pages 125–142, 2018.

- David Boylan and Ginger Schultheis. Attitudes, conditionals, and the qualitative thesis. Manuscript, NYU and Rutgers, April 2019.
- Paul Dekker. *Transsentential Meditations: Ups and Downs in Dynamic Semantics*. PhD thesis, University of Amsterdam, 1993.
- Cian Dorr and John Hawthorne. Embedding epistemic modals. *Mind*, 122(488):867–913, 2013.
- David Etlin. Modus ponens revisited. Manuscript, MIT, February 2008.
- Kai von Fintel. *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts at Amherst, 1994.
- Kai von Fintel. The presupposition of subjunctive conditionals. In *The Interpretive Tract*, volume 25, pages 29–44. MITWPL, Cambridge, Massachusetts, 1998.
- Kai von Fintel. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, 16(2):97–148, 1999.
- Allan Gibbard. Two recent theories of conditionals. In *Ifs: Conditionals, Beliefs, Decision, Chance, and Time*, pages 211–247. Reidel, Dordrecht, 1981.
- Anthony Gillies. Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616, 2004.
- Anthony Gillies. On truth conditions for *If* (but not quite only *If*). *Philosophical Review*, 118(3):325–349, 2009.
- Jeroen Groenendijk, Martin Stokhof, and Frank Veltman. Coreference and modality. In *Handbook of Contemporary Semantic Theory*, pages 179–216. Oxford: Blackwell, 1996.
- Elena Herburger. Bare conditionals in the red. *Linguistics and Philosophy*, 42(2):131–175, 2019.
- Hans G. Herzberger. Dimensions of truth. *Journal of Philosophical Logic*, 2(4):535–556, 1973.
- Lauri Karttunen and Stanley Peters. Conventional implicatures in Montague grammar. In *Syntax and Semantics 11: Presupposition*. Academic Press, New York, 1979.
- Angelika Kratzer. The notional category of modality. In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter, 1981.
- Angelika Kratzer. Modality. In *Semantics: An International Handbook of Contemporary Research*, pages 639–650. de Gruyter, Berlin, 1991.
- Matthew Mandelkern. Bounded modality. *The Philosophical Review*, 128(1):1–61, 2019.
- Vann McGee. A counterexample to modus ponens. *The Journal of Philosophy*, 82(9):462–471, 1985.
- Philippe Schlenker. Local contexts. *Semantics and Pragmatics*, 2(3):1–78, 2009.
- Robert Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Oxford: Blackwell, 1968.
- Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–86, 1975.
- Peter Strawson. *Introduction to Logical Theory*. Methuen, London, 1952.
- Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, 2007.

# Negation and Alternatives in Conditional Antecedents

Dean McHugh<sup>1</sup> and Alexandre Cremers

<sup>1</sup> ILLC – University of Amsterdam

## Abstract

A number of authors, beginning with [Alonso-Ovalle \(2006\)](#), have used conditional antecedents to argue for the presence of alternatives in semantics. In this tradition, recent experimental work from [Ciardelli et al. \(2018b\)](#) and [Schulz \(2018\)](#) uses data from conditional antecedents to investigate the interaction between negation and alternatives. We contribute to this line of inquiry with an experiment to test a number of semantics of conditionals through their predictions on the relationship between alternatives and negation (namely, [Fine, 2012](#); [Alonso-Ovalle, 2006](#); [Ciardelli et al., 2018b](#); [Willer, 2018](#); [Schulz, 2018](#)). We find experimental support for a variant of [Schulz's](#) theory and against all other accounts we consider.

## 1 Introduction

Ever since [Hamblin \(1973\)](#), many semanticists have felt the need to enrich their semantic framework so that expressions do not have their ‘traditional’ denotation, but are represented as sets of such denotations. The traditional denotations are then known as an expression’s alternatives. Alternatives have been applied in the semantics of questions (e.g. [Hamblin, 1973](#); [Ciardelli et al., 2018a](#)), indefinites (e.g. [Kratzer and Shimoyama, 2002](#)), quantified expressions (e.g. [Hagstrom, 1998](#); [Shimoyama, 2001](#)), modals (e.g. [Simons, 2005](#); [Aloni, 2007](#)), conditionals (e.g. [Alonso-Ovalle, 2006](#); [Santorio, 2018](#)), unconditionals ([Rawlins, 2008](#)), and exclusive strengthening operators ([Menéndez-Benito, 2005](#); [Alonso-Ovalle, 2006](#); [Roelofsen and van Gool, 2010](#)), to name a few.

While the importance of alternatives in semantics is widely acknowledged, different authors use different frameworks to represent them. As [Ciardelli et al. \(2017\)](#) and [Ciardelli and Roelofsen \(2017\)](#) point out, different frameworks for alternatives make different predictions, making the choice of which framework to adopt an empirical issue. This is especially true in the semantics of conditionals. Conditional antecedents offer an interesting test case for broader issues in the semantics of alternatives. This is due to the idea—going back to [Alonso-Ovalle \(2006\)](#) and variously implemented since—that the truth conditions of conditionals involve direct quantification over the antecedent’s alternatives. [Ciardelli \(2016\)](#) represents this idea schematically in (1) (we have rephrased [Ciardelli's](#) scheme to apply generally, since the scheme is written to apply to inquisitive semantics).

- (1) **Conditional semantics scheme** ([Ciardelli, 2016](#)): A conditional *if*  $A, C$  is true just in case for every alternative  $p$  of  $\llbracket A \rrbracket$ , there is an alternative  $q$  of  $\llbracket C \rrbracket$  such that  $p \Rightarrow q$  holds, where  $\Rightarrow$  is given by one’s favorite semantics of conditionals, defined over propositions.

Beginning with the work of [Alonso-Ovalle \(2006\)](#), there are now a number of semantics of conditionals that appeal to alternatives. [Table 1](#) provides an overview of recent contributions.

Somewhat bewilderingly, each theory listed in [Table 1](#) proposes a different entry for negation. One particular difference between them concerns whether negated expressions can have multiple alternatives. To illustrate, while recent semantics of conditionals that use alternatives all agree that (2a)’s antecedent raises two distinct alternatives, they disagree as to whether a negated antecedent, as in (2b), raises one or many alternatives.

Semantics of conditionals	Semantic framework
Alonso-Ovalle (2006)	Alternative semantics (e.g. Hamblin, 1973; Kratzer and Shimoyama, 2002)
Fine (2012)	Truthmaker semantics (Fine, 2017)
Willer (2018)	Dynamic semantics (Veltman, 1996), enriched with alternatives
Ciardelli et al. (2018b)	Inquisitive semantics (Ciardelli et al., 2018a)
Schulz (2018)	Inquisitive semantics with a modified clause for negation

Table 1: Recent work in the semantics of conditionals that uses alternatives

- (2) a. If you had taken the train or the metro, you would have arrived on time.  
 b. If Mary and her ex husband had not both come to the party, we would've had more fun.

The interaction between negation and alternatives offers a useful dimension by which to compare frameworks for alternatives, and thereby attain a greater understanding of the empirical adequacy of a number of frameworks for alternatives simultaneously.

We therefore ran an experiment to test the predictions of each of the theories listed in Table 1, using a design similar to that of (Ciardelli et al., 2018b, henceforth: CZC). Specifically, we investigated the interaction between negation and alternatives in conditional antecedents.

## 2 The role of alternatives in conditional antecedents

A central reason why Alonso-Ovalle (2006) introduced alternatives into the semantics of conditionals was to prohibit an unwanted interaction observed between the disjuncts of a disjunctive conditional antecedent. His example, based on Nute (1975), is (3).

- (3) If we had had good weather this summer or the sun had grown cold, we would have had a bumper crop.

In the semantics of Stalnaker (1968) and Lewis (1973), a counterfactual is true just in case its consequent holds in all the most similar worlds to the actual world that make the antecedent true. Presumably, since the sun growing cold is such a bizarre event, in all the most similar worlds to the actual world where there is good weather this summer or the sun grows cold, there is good weather this summer, in which case we have a bumper crop. However, unlike the semantics of similarity, on hearing (3) speakers do not ignore the scenario where the sun grows cold, in which case there is not a bumper crop.

Since the theory of Stalnaker (1968); Lewis (1973) only takes the truth conditions – or, *informative content* – of the antecedent into account, it allows the two disjuncts of a disjunctive antecedent to interact when they are fed into the mechanism for making counterfactual assumptions. In this case, it is the striking difference in similarity between the two disjuncts in (3) that results in the semantics ignoring the scenario where the sun grows cold.

Alternatives in semantics offer one way to avoid such interaction. This is because they allow one to feed the alternatives of conditional antecedent separately into the mechanism for making counterfactual assumptions, as shown in (1).

## 2.1 Negation as an alternative-flattener: a problem

Schulz (2018) has shown that a version of the problem that Alonso-Ovalle pointed out for Lewis and Stalnaker also arises for alternative and inquisitive semantics. This is because negation in alternative and inquisitive flattens alternatives, in the sense that negated expressions always have a single alternative. In these frameworks, then, the advantages brought by alternatives into the semantics of conditionals cannot be transferred to negated antecedents. In alternative and inquisitive semantics, the contribution of a negated antecedent is based on its informative content alone.

Schulz (2018) presents experimental evidence that alternatives can survive negation. Schulz adapts the scenario from CZC, where a light is on just in case the electricity is working and both switches are in the same position. Currently, the electricity is working, switch A is up and switch B is down, so the light is off, as shown in Figure 1.

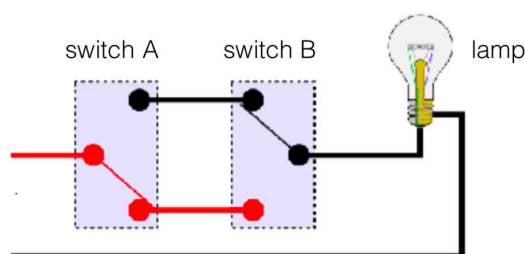


Figure 1: Scenario tested by Schulz (2018)

- (4)
- a. If the electricity was working, then the light would be on.  $E > \text{ON}$
  - b. If the electricity was working and switch A was up, then the light would be on.  $(E \wedge A) > \text{ON}$
  - c. If the electricity was working and switch A and switch B were not both up, then the light would (still) be off.  $[E \wedge \neg(A \wedge B)] > \text{ON}$

Sentences	True	%	False	%	Indet.	%
$E > \text{ON}$	8	16%	42	82%	1	2%
$(E \wedge A) > \text{ON}$	43	84%	5	10%	2	4%
$[E \wedge \neg(A \wedge B)] > \text{OFF}$	14	27%	27	53%	8	16%
$[E \wedge \neg(A \wedge B)] > \text{OFF}^*$ (filtered results)	9	9%	20	59%	5	15%

Table 2: Results from Schulz (2018)

As Table 2 shows, a majority of participants in Schulz’s experiment judged (4c) to be false. This is not predicted by alternative and inquisitive semantics, either under the similarity approach or the background semantics proposed by CZC (§4). Regarding the similarity approach, first note that in the actual scenario the switches are not both up. Then intuitively, the most similar world to the actual world where the electricity is on and the switches are both up is one where the electricity is on and the switches are untouched. In this case the light is off, so the similarity approach predicts the truth of (4c). We achieve the same result under the



background semantics proposed by CZC.<sup>1</sup>

## 2.2 Negation without alternative-flattening

One might conclude from Schulz (2018)'s results that, unlike in alternative and inquisitive semantics, negated conditional antecedents should be capable of having multiple alternatives. Under this strategy, the solution Alonso-Ovalle proposed for disjunctive conditional antecedents in response to (3) can be extended to negated conditional antecedents.

Fine (2012), Willer (2018) and Schulz (2018) offer semantics for conditionals where negations can have multiple alternatives. Fine (2012) and Willer (2018) validate De Morgan's law, and therefore face the challenge posed by CZC, who offer evidence De Morgan's law fails for conditional antecedents. However, Willer suggests that one explanation of their data is due to the particular construction of the sentences they tested, in particular their use of 'both' (Willer, 2018, p. 390). To replicate CZC's results, we sought to test a different equivalence, that of  $\neg\neg(A \vee B)$  and  $A \vee B$ , without using 'both' as a scope marker. The equivalence of  $\neg\neg(A \vee B)$  and  $A \vee B$  is validated by Fine and Willer.<sup>2</sup>

In contrast, Schulz (2018) does not validate De Morgan's law, nor the equivalence of  $\neg\neg(A \vee B)$  and  $A \vee B$ . Schulz adopts inquisitive semantics but proposes that negation introduces an extra requirement on alternatives, one not found in unnegated antecedents: the meaning of a negated sentence  $\neg S$  is given by the set of states that (i) are truth-conditionally incompatible with  $S$  and (ii) specify all and only the values of each atomic sentence in  $S$ .

For reasons that we discuss in section 4.2, we introduce both a binary and  $n$ -ary version of Schulz's theory.

- (5) *Schulz negation*
- a.  $\mathcal{L}(\varphi) = \{a : a \text{ is an atomic sentence appearing in } \varphi\}$
  - b.  $w \sim_{\varphi} v$  iff  $w(a) = v(a)$  for every  $a \in \mathcal{L}(\varphi)$ 
    - (i) Binary version:  $w(a) \in \{0, 1\}$  for every world  $w$  and atomic sentence  $a$ .
    - (ii)  $n$ -ary version:  $w(a)$  can be outside  $\{0, 1\}$ .
  - c. For any information state  $p \subseteq W$ ,
    - (i)  $p \models Q(\varphi)$  iff  $w \sim_{\varphi} v$  for every  $w, v \in p$  ( $p$  'answers the question raised by  $\varphi$ ')
    - (ii)  $p \perp \varphi$  iff  $p \cap |\varphi|$  is empty ( $p$  and  $\varphi$  are mutually exclusive)
  - d. For any proposition  $P \subseteq \wp(W)$ ,  $P \models \neg\varphi$  iff  $p \models Q(\varphi)$  and  $p \perp \varphi$  for every  $p \in P$
  - e.  $\llbracket \text{not } \varphi \rrbracket = \{p \subseteq W : Q(\varphi) \text{ and } p \perp \varphi\}$

<sup>1</sup> Since the background semantics of CZC is rather new, we show this fact here. CZC account for their data by assuming what they call a 'maximal background', according to which, given an alternative  $a$  of the antecedent, we fix the value of every fact that is not called into question by  $a$ . According to CZC, since the antecedent of (4c) contains a single alternative, to assess (4c) we fix the value of every fact that is not called into question by the proposition  $[E \wedge \neg(A \wedge B)]$ , where 'calling a fact into question' is a technical term defined in their paper. Since the scenario of Figure 1 has a particularly simple causal structure (the causal graph representing it consists of just three arrows from the electricity, switch A and switch B, respectively, to the light), calling a fact into question amounts to contributing to its falsity (cf. definition 4 in CZC). To check whether the fact that switch A is down contributes to the falsity of  $E \wedge \neg(A \wedge B)$ , we have to find a set of actual facts  $F \subseteq \{\neg E, \neg A, B, \neg L\}$  such that  $F$  is consistent with  $E \wedge \neg(A \wedge B)$  but  $F \cup \{\neg A\}$  is not. It is easily checked that no such set of facts exists. Similarly, one can show that the fact  $B$  is up does not contribute to the falsity of  $E \wedge \neg(A \wedge B)$ .

<sup>2</sup>To see this, note that all semantic frameworks agree that  $\neg\neg(A \vee B)$  is equivalent to  $\neg(\neg A \wedge \neg B)$ . Then by De Morgan's law  $\neg(\neg A \wedge \neg B)$  is equivalent to  $\neg\neg A \vee \neg\neg B$ , which is equivalent to  $A \vee B$  when  $A$  and  $B$  are atomic.

### 3 Our experiment

We presented 192 Mechanical Turk participants with the wiring diagram in Fig. 2, illustrating how lighting in public buildings such as hospitals is often controlled in such a way that a caretaker can lock the lights OFF or ON (by moving switch A in position bottom or top resp.), or leave it under the control of the normal circuit switch B by leaving A in the middle (letting a patient turn the light ON or OFF as they wish). Participants were then told that switch A is in the middle and switch B is down, and instructed to rate a few sentences on a scale from 1 (clearly false) to 7 (clearly true). Each participant only saw one of T1 and T2, in random order with the True and False filler and the Control item. T3 was presented last, as it had a slightly different structure. 74 participants who responded 4 or less on the True filler were excluded from analyses, as well as 3 participants who didn't report English as their native language. Participants were at chance on the False filler (presumably because of an ambiguity regarding the antecedent of 'if that wasn't the case'), so this item was not used as an exclusion criterion.

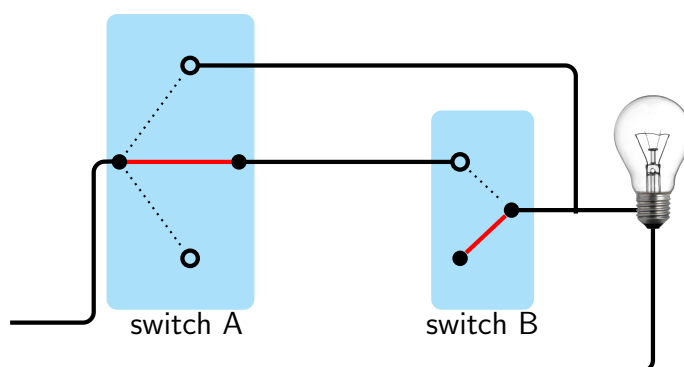


Figure 2: Scenario used in the experiment

Let us use  $A\uparrow$ ,  $A\bullet$  and  $A\downarrow$  to denote that switch A is up, in the middle and down, respectively, and  $B\uparrow$  and  $B\downarrow$  to denote that switch B is up and down, respectively. Of course, the logical formulas representing the sentences in (6) did not feature in the experiment.

- (6) **False:** Currently, switch A is in the middle and switch B is down. If that wasn't the case, the light would be on.  $\neg(A\bullet \wedge B\downarrow) > \text{ON}$
- T1:** Currently, neither switch is up. If that wasn't the case, the light would be on.  $\neg\neg(A\uparrow \vee B\uparrow) > \text{ON}$
- T2:** Currently, switch A is in the middle and switch B is down. If switch A was up or switch B was up, the light would be on.  $A\uparrow \vee B\uparrow > \text{ON}$
- T3:** If switch B was up but not switch A, the light would be on.  $B\uparrow \wedge \neg A\uparrow > \text{ON}$
- Control:** Currently, switch B is down. If that wasn't the case, the light would be on.  $\neg B\downarrow > \text{ON}$
- True:** Currently, switch A is not up. If that was the case, the light would be on.  $A\uparrow > \text{ON}$

## 4 Predictions

### 4.1 Theories where negation flattens alternatives

Alonso-Ovalle (2006) and CZC predict T1 to be false. To see this for the similarity approach, note that the world where switch A is down and B is up is among the most similar worlds to the actual one where  $\neg\neg(A\uparrow \vee B\uparrow)$  is true.<sup>3</sup> Alonso-Ovalle (2006) and CZC predict T2 to be true, since the antecedent  $A\uparrow \vee B\uparrow$  raises two alternatives. These alternatives are each hypothetically assumed separately, making  $(A\uparrow \vee B\uparrow) > \text{ON}$  equivalent to the conjunction  $(A\uparrow > \text{ON}) \wedge (B\uparrow > \text{ON})$ . Finally, they both predict T3 to be false. This is for the same reason they predict (4c) to be false, as we saw in section 2.1. For instance, on the similarity approach, the world where A is kept in the middle and B is up is more similar to the actual one where A is allowed to vary. (For the background semantics, the same reasoning as in footnote 1 applies.)

### 4.2 Theories where negation does not flatten alternatives

As mentioned in section 2.2, Fine and Willer predict the equivalence of T1 and T2. Furthermore, they predict both T1 and T2 to be true. But since these theories allow negated antecedents to have many alternatives, they predict T3 to be false. To see this, note that Fine and Willer (2018) can predict that (7a) and (7b) have the same alternatives.

- (7) a. Switch A is not up.  
b. Switch A is in the middle or down.

Then as these two systems validate the distribution of conjunction of disjunction, and simplification of disjunctive antecedents, they predict the falsity of T3, as shown in (8).

- (8) a. If switch B was up but not A, the light would be on.  $(B\uparrow \wedge \neg A\uparrow) > \text{ON}$   
b.  $\Leftrightarrow$  If switch B was up but A was in the middle or down, the light would be on.  $B\uparrow \wedge (A\bullet \vee A\downarrow) > \text{ON}$   
c.  $\Leftrightarrow$  (by distribution of  $\wedge$  over  $\vee$ ): If switch B was up and A was in the middle, or B was up and A was down, the light would be on.  $(B\uparrow \wedge A\bullet) \vee (B\uparrow \wedge A\downarrow) > \text{ON}$   
d.  $\Rightarrow$  (by simplification): If B was up and A was down, the light would be on.  $(B\uparrow \wedge A\downarrow) > \text{ON}$

<sup>3</sup>CZC's background semantics also predicts T1 to be true. To see this, note that the facts  $A\bullet$  and  $B\downarrow$  are both individually consistent with  $\neg\neg(A\uparrow \vee B\uparrow)$ ; after all,  $A\bullet \wedge B\uparrow$  and  $(A\uparrow \wedge B\downarrow)$  are both possible. However, their conjunction  $(A\bullet \wedge B\downarrow)$  is not consistent with  $\neg\neg(A\uparrow \vee B\uparrow)$ . Thus the facts  $A\bullet$  and  $B\downarrow$  both contribute to the falsity of  $\neg\neg(A\uparrow \vee B\uparrow)$  in the actual world.  $B\downarrow$  contributes to the falsity of  $\neg\neg(A\uparrow \vee B\uparrow)$ , since e.g.  $\{A\bullet\}$  is consistent with  $|\neg\neg(A\uparrow \vee B\uparrow)|$ , but  $\{A\bullet, B\downarrow\}$  is not. And  $A\bullet$  contributes to the falsity of  $\neg\neg(A\uparrow \vee B\uparrow)$ , since e.g.  $\{B\downarrow\}$  is consistent with  $|\neg\neg(A\uparrow \vee B\uparrow)|$ , but  $\{A\bullet, B\downarrow\}$  is not.

So according to CZC these two facts are not backgrounded on the default option of a maximal background; in other words, we do not fix at its actual value the position of switch A or B. So they predict that when listeners counterfactually imagine  $\neg\neg(A\uparrow \vee B\uparrow)$ , they consider all positions of switches A and B where that sentence is true. These are, with their corresponding outcomes for the light:

$A\uparrow$ and $B\uparrow$	$A\uparrow$ and $B\downarrow$	$A\bullet$ and $B\uparrow$	$A\downarrow$ and $B\uparrow$
On	On	On	Off

Note the final scenario: when switch A is down and switch B is up, the light is off. CZC therefore predict the counterfactual  $\neg\neg(A\text{ up } \vee B\text{ up}) > \text{On}$  to be false. Note that they make this prediction regardless of the background parameter. We have just shown the prediction to hold for a maximal background, but since the scenario where switch A is down and B is up is considered under a maximal background, and any non-maximal background only brings more scenarios into consideration, the scenario where switch A is down and B is up is considered under every background.

The last sentence is quite clearly false, since the light is never on when switch A is down. [Willer \(2018\)](#) therefore predicts the falsity of T3 (8a).

[Schulz \(2018\)](#) presents her clause for negation in a bivalent system: a proposition  $p$  “answers the question” raised by an atomic sentence  $a$  just in case  $p$  determines whether  $a$  is true or false. In our experiment, switch A could take three values: up, middle or down. [Schulz’s](#) prediction for T1 depends on whether one uses her bivalent formulation or switches to one where variables can take arbitrary arity. In the bivalent formulation, there are four maximal states that answer the question raised by the set of atomics  $\{A\uparrow, B\uparrow\}$ :

$$|A\uparrow \wedge B\uparrow| \quad |A\uparrow \wedge \neg B\uparrow| \quad |\neg A\uparrow \wedge B\uparrow| \quad |\neg A\uparrow \wedge \neg B\uparrow|$$

However, in the  $n$ -ary formulation, there are six states: each one determining not only whether A is up and whether B is up, but the exact position of each switch.

$$|A\uparrow \wedge B\uparrow| \quad |A\uparrow \wedge B\downarrow| \quad |A\bullet \wedge B\uparrow| \quad |A\downarrow \wedge B\uparrow| \quad |A\bullet \wedge B\downarrow| \quad |A\downarrow \wedge B\downarrow|$$



Figure 3: T1,  $\neg\neg(A\uparrow \vee B\uparrow)$ , in [Schulz’s](#) framework

The binary version predicts T1 to be true. For the two alternatives where switch A is up, the light is on, so the only alternative that could make the consequence false is  $|(A\bullet \vee A\downarrow) \wedge B\uparrow|$ . Now, under the background semantics of CZC, the fact that switch A is in the middle does not contribute to the falsity of this alternative,  $|(A\bullet \vee A\downarrow) \wedge B\uparrow|$ . So the binary version predicts that we keep the position of switch A fixed when hypothetically entertaining that alternative. Since the fact that B is down does contribute to the falsity of that alternative (indeed it is inconsistent with that alternative), we do not keep the position of B fixed. And in that case, the only scenario to check is the one where switch A is in the middle and switch B is up, in which case the light is on. Hence the binary version of [Schulz’s](#) theory predicts T1 to be true. In contrast, in the  $n$ -ary version, one alternative is  $|(A\downarrow \wedge B\uparrow)|$ . In that scenario, the light is off, making T1 false.

As T2 does not involve negation, and [Schulz \(2018\)](#)’s theory only differs from CZC in respect of negation, [Schulz](#) makes the same prediction regarding T2; namely, that it is true.

[Schulz \(2018\)](#) predicts T3 to be false, both on the binary and  $n$ -ary formulation. In the binary formulation,  $|(A\uparrow \vee A\downarrow) \wedge B\uparrow|$  is one alternative raised by the antecedent  $\neg(A\bullet \wedge B\downarrow)$ . Since  $A\bullet$  and  $B\downarrow$  are each consistent on their own (i.e. consistent with the empty set of facts), but are inconsistent with  $|(A\uparrow \vee A\downarrow) \wedge B\uparrow|$ , according to the mechanism of making counterfactual assumptions proposed by CZC, the facts  $A\bullet$  and  $B\downarrow$  both contribute to the falsity of  $\neg(A\bullet \wedge B\downarrow)$ , and hence are not fixed to their actual values. Then when we hypothetically assume the alternative  $|(A\uparrow \vee A\downarrow) \wedge B\uparrow|$ , we consider two scenarios: both switches being up, and switch A

being up and switch B being down. In this latter case, the light is off, making the counterfactual as a whole false. Similarly, in the  $n$ -ary formulation,  $|A\downarrow \wedge B\uparrow|$  is one alternative raised by the antecedent  $\neg(A\bullet \wedge B\downarrow)$ . In this scenario, the light is off and so the counterfactual  $F: \neg(A\bullet \wedge B\downarrow) > On$  is false.

## 5 Results

The results of the experiment are depicted in Figure 4. A cumulative link mixed model on data from the control and test sentences showed that T1 and T3 were rated significantly lower than the control (both  $z < -2.5, p < .01$ ), while T2 was rated significantly higher than control ( $z = 2.1, p = .039$ ). A posthoc comparison of targets T1 and T3 revealed no difference between the two ( $z = -0.5, p = .62$ ).

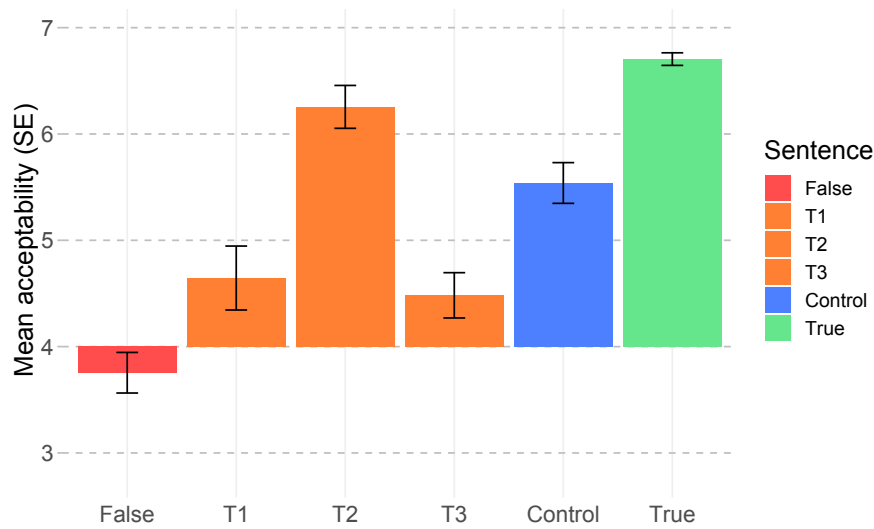


Figure 4: Mean and SE acceptability of each experimental item

We interpret our data as indicating that T1 and T3 are false, and that T2 is true. Table 3 provides an overview of the theories we have considered here and their predictions.

Theory / Antecedent	T1 $\neg\neg(A\uparrow \vee B\uparrow)$	T2 $A\uparrow \vee B\uparrow$	T3 $B\uparrow \wedge \neg A\uparrow$
Our data (interpreted)	✗	✓	✗
Alonso-Ovalle (2006)	✗	✓	✓
Ciardelli et al. (2018b)	✗	✓	✓
Fine (2012)	✓	✓	✗
Willer (2018)	✓	✓	✗
Schulz (2018) binary	✓	✓	✓
Schulz (2018) $n$ -ary	✗	✓	✗

Table 3: Overview of predictions and results

## 6 Discussion

The results of our experiment pose a challenge to all contemporary semantics of conditionals based on alternatives we have considered, besides the  $n$ -ary version of Schulz (2018). Most surprising of all is our experimental support for Schulz’s unique and entirely novel requirement on the semantics of negation.

Our paper also makes a new contribution to the debate surrounding the interpretation of the experimental results from Ciardelli et al. (2018b). Their data previously suggested against semantic frameworks for conditionals that validate De Morgan’s law (e.g. Lewis, 1973; Fine, 2012; Santorio, 2018; Willer, 2018). However, Willer (2018, 390) suggests an alternative explanation for the results of Ciardelli et al.; in particular, by pointing out the undue influence of the word “both” in their experiment. Since any semantic framework validating De Morgan’s law also predicts the equivalence of T1 and T2, our experiment contributes independent evidence—using previously untested sentences without “both”—against De Morgan’s law in conditional antecedents.

Our study has its own limitations of course. To reduce the complexity of some counterfactual antecedents and avoid overt double negations, we resorted to sentential anaphora in T1, the control, and the false filler: We first introduced a sentence to be used as an antecedent, followed by a counterfactual of the form “if that wasn’t the case”, where ‘it’ was to be understood as the previous sentence. There is of course a risk that participants may have interpreted these sentences differently, and that seems to have been the case for the false filler. Unlike the two others, the antecedent sentence contained a conjunction, and it may very well be that some participants resolved ‘it’ as referring to the second conjunct rather than the whole conjunction. This would explain while many accepted this sentence. Crucially however, in the case of T1 and the control, we do not see any salient alternative to the resolution we intended. A more general issue is the sheer complexity of the scenario we tested, reflected in the number of participants we had to exclude (similar to Ciardelli et al.’s original study). Unfortunately, this seems to be an inherent limitation of this experimental design, but to the extent that we find clear contrasts between the different sentences we tested, these differences still call for an explanation.

Finally, we would like to come back to the mismatch between our results and Ciardelli et al.’s proposal. While Schulz’s results calls for a modification of the role of negation in inquisitive semantics, this is not an issue here: inquisitive semantics correctly predicts the difference between T1 and T2. Instead, the problem comes from Ciardelli et al.’s assumptions about backgrounding, which fail to capture T3 (because  $A\bullet$  does not contribute to the falsity of  $A\uparrow$ ). A possible fix then, would be to assume that any fact that is mentioned cannot be backgrounded. At this point, one may question how much we want to encode in the semantics, and what should be left to pragmatics. Ciardelli et al. propose a background *semantics*, but the departures needed to account for the results presented here and in Schulz (2018) suggest that it may be a matter of pragmatics more than semantics, as it must be sensitive to the form of the antecedent and not just its semantic content.

## References

- Aloni, M. (2007). Free choice, modals, and imperatives. *Natural Language Semantics*, 15(1):65–94.
- Alonso-Ovalle, L. (2006). *Disjunction in alternative semantics*. PhD thesis, University of Massachusetts Amherst.

- Ciardelli, I. (2016). Lifting conditionals to inquisitive semantics. In *Semantics and Linguistic Theory*, volume 26, pages 732–752.
- Ciardelli, I., Groenendijk, J., and Roelofsen, F. (2018a). *Inquisitive semantics*. Oxford University Press.
- Ciardelli, I. and Roelofsen, F. (2017). Hurford’s constraint, the semantics of disjunction, and the nature of alternatives. *Natural Language Semantics*, 25(3):199–222.
- Ciardelli, I., Roelofsen, F., and Theiler, N. (2017). Composing alternatives. *Linguistics and Philosophy*, 40(1):1–36.
- Ciardelli, I., Zhang, L., and Champollion, L. (2018b). Two switches in the theory of counterfactuals. *Linguistics and Philosophy*.
- Fine, K. (2012). Counterfactuals without possible worlds. *Journal of Philosophy*, 109(3):221–246.
- Fine, K. (2017). Truthmaker semantics. In *A Companion to the Philosophy of Language*, chapter 22, pages 556–577. Wiley-Blackwell.
- Hagstrom, P. (1998). *Decomposing questions*. PhD thesis, MIT.
- Hamblin, C. L. (1973). Questions in montague english. In *Montague grammar*, pages 247–259. Elsevier.
- Kratzer, A. and Shimoyama, J. (2002). Indeterminate pronouns: The view from japanese. In Otsu, Y., editor, *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, pages 1–25.
- Lewis, D. (1973). *Counterfactuals*. John Wiley & Sons.
- Menéndez-Benito, P. (2005). *The grammar of choice*. PhD thesis, University of Massachusetts, Amherst.
- Nute, D. (1975). Counterfactuals and the similarity of words. *The Journal of Philosophy*, 72(21):773–778.
- Rawlins, K. (2008). *(Un)conditionals: An investigation in the syntax and semantics of conditional structures*. PhD thesis, University of California, Santa Cruz.
- Roelofsen, F. and van Gool, S. (2010). Disjunctive questions, intonation, and highlighting. In Aloni, M., Bastiaanse, H., de Jager, T., and Schulz, K., editors, *Logic, Language and Meaning*, pages 384–394, Berlin, Heidelberg. Springer.
- Santorio, P. (2018). Alternatives and truthmakers in conditional semantics. *The Journal of Philosophy*.
- Schulz, K. (2018). The similarity approach strikes back: Negation in counterfactuals. In Sauerland, U. and Solt, S., editors, *Proceedings of Sinn und Bedeutung 22*, volume 2 of *ZASPiL 61*, pages 343–360. Leibniz-Centre General Linguistics, Berlin.
- Shimoyama, J. (2001). *Wh-constructions in Japanese*. PhD thesis, University of Massachusetts at Amherst.
- Simons, M. (2005). Dividing things up: The semantics of or and the modal/or interaction. *Natural Language Semantics*, 13(3):271–316.
- Stalnaker, R. (1968). A theory of conditionals. In *Ifs*, pages 41–55. Springer.
- Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261.
- Willer, M. (2018). Simplifying with free choice. *Topoi*, 37(3):379–392.

# How prosody disambiguates between Alternative and Polar Questions \*

Erlinde Meertens

University of Konstanz  
erlinde.meertens@uni-konstanz.de

## Abstract

This paper is concerned with the function of prosody in disambiguating Polar Questions and Alternative Questions from English. Based on data from Basque and Finnish, and existing ingredients in the literature, I propose that the underlying disambiguating principle is that the syntax-semantics mapping and the prosody-discourse mapping have to be coherent with each other. In evaluating coherence, I crucially take *Maximize Informativity* to apply to both QUD and utterance level and I propose that two criteria for coherence have to be met: there has to be a proper motherQ-daughter Q relation and a daughterQ cannot be presuppositionally heavier than its motherQ.

## 1 Introduction

Across languages, several strategies are used to disambiguate between an Alternative Question (AltQ - (1-a)) and a Polar Question (PolQ - (1-b)). Firstly, there are languages that rely on prosody. In for example English, AltQs are characterized by an accent on each disjunct and a final fall, whereas PolQs typically have a block accent and a final rise, as illustrated in (1)<sup>1</sup> [Bartels, 1999]. Secondly, there are languages that make use of a special disjunction form that forces an AltQ interpretation, as, for example, Finnish [Haspelmath, 2007] In (2-a), the usage of *vai* results in an AltQ interpretation, whereas the declarative disjunction *tai* in (2-b) leads to a PolQ reading. Third, there are languages, like Basque that make use of a combination of prosody and a special AltQ disjunction form [Saltarelli, 1988]. The combination of an accent on each disjunct and a final fall, and the disjunction *ala* results in an AltQ reading. Disjunctive questions with a block accent, final rise and the disjunction *edo* are interpreted as PolQs<sup>2</sup>.

- (1) a. Did Ana see [Boris]↑ or [Osip]↓? [AltQ]  
Which one of the following did Ana see: Boris or Osip?  
b. Did Ana see [Boris or Osip]↑? [PolQ]  
Is it true that Ana saw Boris or Osip?
- (2) a. Haluatko sokeiria **vai** kermaa? [Finnish- AltQ]  
Want.you sugar **or**<sub>AltQ</sub> cream  
b. Haluatko sokeiria **tai** kermaa? [Finnish- PolQ]  
Want.you sugar **or**<sub>PolQ/decl</sub> cream
- (3) a. [Te-a]↑ **ala** [kafe-a]↓ nahi duzu↓? [Basque-AltQ]  
tea-ART **or**<sub>AltQ</sub> coffee-ART want you-it

\*Thanks to Kajsa Djärv, Felix Frühauf, David Krassnig, Sergio Monforte, Floris Roelofsen, Maribel Romero, and the audience of the Workshop 'Exhaustivity in Questions', Tübingen for discussion and feedback. All errors are my own.

<sup>1</sup>Throughout the paper, ↑ indicates a rising pitch accent L\*H- or , when used sentence final, the rising boundary tone L\*H-H% and ↓ indicates the falling boundary tone H\*L-L%.

<sup>2</sup>Many thanks to Sergio Monforte for helping me with the Basque data.



- b. [Te-a      edo              kafe-a]      nahi duzu↑?                      [Basque-PolQ]  
 tea-ART **or**<sub>PolQ/AltQ</sub> coffee-ART want you-it

In this paper, I propose that the interpretations in (1)-(3) are the result of the same underlying, independently motivated mechanisms. The key idea is that the syntax-semantics mapping and the prosody-discourse mapping have to be coherent with each other. The structure of the paper is as follows. In §2, I lay out my assumptions about the syntax and semantics of disjunctive questions. In §3, I discuss the relationship between prosody and discourse structure. In §4, I propose that the contents of §2 and §3 have to be coherent with each other. §5 serves to present the analytical tools that are used to evaluate coherence between the syntax-semantics and the prosody-discourse mapping. In §6, I apply the formalisms to the relevant data and show that the proposal makes the right predictions. In §7, I briefly discuss the advantages of the current account in comparison to prominent accounts in the literature. I conclude in §8.

## 2 The Syntax-Semantics Backbone

Let us start with the ingredients of the compositional analysis. In what follows, I make use of a Hamblin semantics for questions and take the meaning of a question to be a set of propositions that describe the possible answers to the question. The relevant ingredients for disjunctive questions are (i) disjunction, (ii) the question operator  $Q$ , and (iii) the existential operator  $\exists$ . I follow [Alonso-Ovalle \[2006\]](#) and assume a Hamblin analysis of disjunction, as defined in (4).

$$(4) \quad \llbracket \alpha \text{ or } \beta \rrbracket = \{p \in D_{\langle e, t \rangle} \mid p \in \llbracket \alpha \rrbracket^{w, g} \vee p \in \llbracket \beta \rrbracket^{w, g}\}$$

Disjunction collects the (usually) singleton propositions that are denoted by the disjuncts, resulting in an a set consisting of propositional alternatives. This set can be combined with a number of propositional operators. In the case of disjunctive questions, the relevant operators are  $Q$  and  $\exists$ . I assume a simple definition of  $Q$ , as proposed by [Biezma and Rawlins \[2012\]](#), which has as its only job to leave the alternatives generated by disjunction in tact, see (5). I assume the definition of  $\exists$  in (6) following [Shimoyama \[2006\]](#).

$$(5) \quad \llbracket Q\alpha \rrbracket = \llbracket \alpha \rrbracket \quad \text{[Biezma and Rawlins, 2012]}$$

$$(6) \quad \llbracket \exists\alpha \rrbracket = \lambda w. \exists p \in \llbracket \alpha \rrbracket : p(w) = 1 \quad \text{[Shimoyama, 2006]}$$

The alternatives that are collected by disjunction can either associate with  $Q$ , leading to an AltQ, or with  $\exists$ , resulting in a PolQ. Hence, applying the definitions for disjunction,  $Q$  and  $\exists$  to the string *Did Ana see Boris or Osip* results in two possible denotations, see (7-a) and (7-b).

- (7) a. Disjunction associates with  $Q$  : [ $Q_i$  Ana saw Boris or<sub>i</sub> Osip].  
 Hamblin set: { $\lambda w. \text{Ana saw}_w \text{ Boris}, \lambda w. \text{Ana saw}_w \text{ Osip}$ }  
 Interpretation: **Alternative Question**
- b. Disjunction associates with  $\exists$ : [ $Q \exists_i$  [<sub>IP</sub> Ana saw Boris or<sub>i</sub> Osip]]  
 Hamblin set: { $\lambda w. \exists p \in \{A \text{ saw } B, A \text{ saw } O\} : p(w) = 1$ }  
 = { $\lambda w'. A \text{ saw}_{w'} B \vee A \text{ saw}_{w'} O$ }  
 Interpretation: **Polar Question**

In short, on the syntax-semantics level, English disjunctive questions are ambiguous. For Basque and Finnish, I propose that the association properties of disjunction are encoded in the lexical

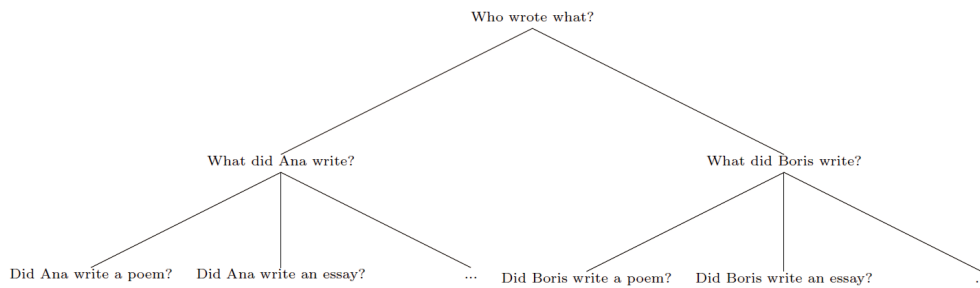
entry for disjunction<sup>3</sup>. In the case of disjunctive questions,  $or_{\text{PolQ/Decl}}$  (*tai* in Finnish/*edo* in Basque) associates with  $\exists$ .  $Or_{\text{AltQ}}$  (*vai* in Finnish/*ala* in Basque) only occurs in disjunctive questions and always forces association with Q, leading to an AltQ interpretation.

### 3 Prosody and Discourse Structure

This section serves to lay out the relationship between the relevant prosodic cues and discourse structure. I make use of the Question under Discussion (QUD) Framework [Roberts, 1996]. For this part of the proposal, I combine existing ingredients from the literature (cf. Meertens et al. [2019]) and take (i) the role of the accents on the disjuncts to be *shaping* the QUD [Roberts, 1996, Biezma, 2009] and (ii) adopt Westera [2017]’s account and argue that the final boundary tone serves to restrict/not restrict the QUD.

To begin with, I follow Roberts [1996], Büring [2003] and take discourse structure to include a stack of (often implicit) QUDs. This produces a set of hierarchically ordered, (mother and daughter) questions, as in (8).

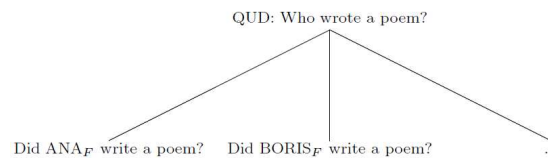
(8)



Furthermore, I adopt Rooth [1992]’s analysis of focus. An utterance  $\phi$  has an ordinary value  $\llbracket\phi\rrbracket$  and a focus semantic value  $\llbracket\phi\rrbracket^f$ . The focus semantic value consists of alternative denotations that are of the same semantic type as  $\phi$ . See, for example, the declarative in (9).

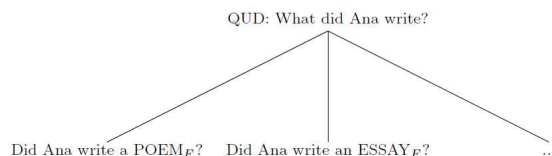
(9)  $\llbracket\text{Ana}_F \text{ wrote a poem}\rrbracket^f = \{\text{Ana wrote a poem, Boris wrote a poem, Osip wrote a poem,}\dots\}$

I follow Roberts [1996]’s and Biezma [2009]’s application of Roothian focus-marking to QUDs. They argue that the effect of focus-marking is to constrain the shape of the immediately higher question in the QUD stack. This is illustrated in (10) and (11). The immediate higher question is dependent on what element is focus marked. In other words, focal accents in the daughter question (daughterQ) signal the position of the wh-element in its mother question (motherQ).



<sup>3</sup>This is reminiscent of the Japanese particle *ka* and its association properties with indeterminate phrases [Shimoyama, 2006].

(10)



(11)

WhQ-PolQ-sequences can be used as a diagnostic for testing the relation between questions. If a WhQ can be followed by a PolQ, this PolQ is an appropriate daughterQ of the WhQ. In the case of the focal accent in PolQs, the intuitions in (10) and (11) are reflected by the (in)felicity of the sequences in (12) and (13) [Roberts, 1996].

- (12) a. What did Ana write? Did Ana write a POEM?  
 b. Who wrote a poem? Did ANA write a poem?

- (13) a. #What did Ana write? Did ANA write a poem?  
 b. #Who wrote a poem? Did Ana write a POEM?

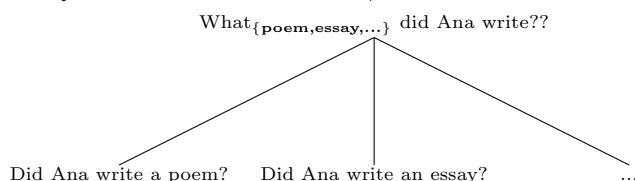
Let me now turn to the role of the falling boundary tone characterizing AltQs like (1-a). Following Westera [2017], I take the final falling boundary tone to signal that the speaker believes only the pronounced are relevant and epistemically live answers to the QUD. A final rise, in contrast, indicates that the speaker considers the possibility that there are other alternatives than the mentioned ones that are relevant and epistemically possible. This is illustrated in (14).

- (14) a. Are you from Denmark↑? → other alternatives are relevant and live  
 b. Are you from Denmark↓? → no other alternatives are relevant and live

The question with the final rise in (14-a) signals that the speaker considers the possibility that there are relevant epistemically live alternatives other than the addressee being from Denmark. Such a question is felicitous, for example, in a context in which the speaker is interested in where the addressee is from, also when it is not Denmark. In contrast, the form with the final fall in (14-b) signals that the speaker believes that the only relevant epistemically live possibility is that the addressee is from Denmark or not. Such a sentence is used, for example<sup>4</sup>, in a context in which there is a special service for people from Denmark at an airport or an insurance company. In that case, the speaker is only after a bare yes/no answer and the addressee being from Sweden or Norway is irrelevant [Westera, 2017].

The combination of the role of the focal accent in questions and the final boundary tone, as described above, results in the following trees for plain PolQs (see (15)), disjunctive PolQs (see (16)), and AltQs (see (17)) respectively [Meertens et al., 2019].

- (15) PolQ: Did Ana write a POEM↑?



<sup>4</sup>See Bartels [1999] for other uses of falling PolQs, for example Quiz Questions.

- (16) PolQ: Did Ana write [a poem or an essay]<sup>↑</sup>?  
 What<sub>{poem,essay,...}</sub> did Ana write??  
 Did Ana write a poem or an essay? ...
- (17) AltQ: Did Ana write a POEM<sup>↑</sup> or an ESSAY<sup>↓</sup>?  
 What<sub>{poem,essay}</sub> did Ana write??  
 Did Ana write a poem? Did Ana write an essay?

In the PolQ in (15), the object *poem* is focus-marked, meaning that the QUD consists of propositions that share the same VP property and differ only in the value of the object. No further constraints are imposed on the QUD, since the final boundary tone is a rise. The analysis for the PolQ in (16) is identical. The difference lies only in the size of the object and the set of alternatives that is the result of focus marking is therefore different. There are, in fact, two candidate sets of alternatives for {Boris or Osip}: the set containing {b∨o,c∨d,...} and the one à la Sauerland [2004] containing {b∨o,b∧o,b,o}. I remain neutral as to what the set exactly looks like, as it has no effect on my analysis. For the AltQ in (17), each disjunct is focus-marked, meaning that each disjunct separately contributes to the QUD. The final boundary tone is a fall, signalling that there are no further relevant epistemically live alternatives. This leads to the QUD in (17).

## 4 Syntax-Semantics and Prosody-Discourse Coherence

I propose that the prosody-discourse mapping always has to be coherent with the syntax-semantics mapping. Recall that I considered three language types. Finnish is the ‘simple’ case. The syntactic-semantic mechanism determines the final interpretation of the utterance and because prosody does not play an important role in disjunctive questions, there are no further coherence restrictions. Concerning Basque, I predict infelicity if the syntax-semantics as generated by disjunction is not coherent with the discourse structure as indicated by prosody. In the case of English, I take prosody to be the guide towards the syntax-semantics. The prosody associated with PolQs in English is only compatible with the denotation in which *or* is associating with  $\exists$  (PolQ LF) and AltQ prosody is only compatible with the semantics in which *or* associates with Q (AltQ LF). Because each available prosodic structure is only compatible with one LF, prosody on its own can get us to the right interpretation. Note that this proposal entails that the disambiguating effects of prosody in English are not encoded semantically or syntactically, but rather the result of coherence requirements.

## 5 Analytical Tools

In this section, I propose that to assess coherence, we can make use of two analytical tools: (i) *Maximize Informativity* and (ii) criteria that define a proper motherQ-daughterQ relation.

### 5.1 Maximize Informativity

Dayal [1996] proposed the principle in (18) to capture the contrast between (19-a) and (19-b).

- (18) *Maximize Informativity*: The Hamblin set of a question must contain a maximally informative true answer.
- (19) a. A: Which poet did Ana see? (= {Ana saw Boris, Ana saw Osip})  
 B: # Ana saw Boris and Osip.
- b. A: Which poets did Ana see? = ({Ana saw Boris, Ana saw Osip, Ana saw Boris  $\oplus$  Osip})  
 B:  $\checkmark$  Ana saw Boris and Osip

Applying (18) to (19-a) leads to the unavailability of the *both* answer. The conjunction of Boris and Osip is not part of the question denotation. As it would be more informative than the propositions in the set, the conjunction is ruled out as an answer. This is different for the plural, because the conjunction is part of the denotation and can thus be taken as the most informative answer, in other words, the *both* answer is available.

I follow Spector [2010] and adopt his idea that *Maximize Informativity* not only applies to constituent questions, but also to AltQs. From there it's an uncontroversial step to assume it also applies to PolQs. On top of that, I propose that (18) applies to both utterance level and QUD level. As will become clear in §6, this is a crucial component of the current proposal.

## 5.2 MotherQ-DaughterQ Relation

There are two criteria that have to be satisfied for an utterance to be a proper daughterQ to a motherQ. First, I adopt Roberts [1996]'s notion, given in (20).

- (20) A complete answer to a daughterQ  $q$  entails an at least partial answer to its motherQ  $Q$ .

This criterion is illustrated in the sequences in (21). An answer to *does she want a sandwich?* in (21-a) provides an (at least) partial answer to the question *What does Ana want to eat?* In contrast, an answer to *does she want a coke?* does not.

- (21) a. What does Ana want to eat? Does she want a sandwich?  
 b. # What does Ana want to eat? Does she want a coke?

Furthermore, I propose the novel restriction in (22) which says that a daughterQ cannot be presuppositionally heavier than its motherQ.

- (22) Any presupposition carried by a daughterQ  $q$  has to be carried by its motherQ  $Q$  as well.

Again, this is reflected by the (in)felicity of sequences. In (23-b), the first criterium is fulfilled, i.e. an answer to the question *which female professor killed the victim?* provides an (at least) partial answer to the motherQ. The sequence is still infelicitous. This infelicity follows from the definition in (22): the daughterQ carries a presupposition that is not carried by the motherQ.

- (23) a. Which professor killed the victim? Did prof. Woland kill the victim?  
 b. # Which professor killed the victim? Which female professor killed the victim?

This restriction is a crucial component of the current proposal, as will become clear in the next section. To summarize, I propose that to settle whether a daughterQ (as provided by the syntax-semantics), is coherent with the motherQ (as shaped by the prosody), the criteria in (20) [Roberts, 1996] and (22) have to be met.

## 6 Back to the Data

### 6.1 The Mechanism

In this section, I apply the described mechanisms to the relevant data. Let us start with the disjunctive PolQ LF with PolQ prosody (hence a successful PolQ) in (24). To assess coherence, let us first see what prosody gives us. The block accent on [saw Boris or Osip] generates propositions that are alternatives to *Boris or Osip*. The final rise signals that other alternatives than Ana having seen Boris or Osip could be relevant or epistemically live.

- (24) *or* associates with  $\exists$  (*edo* in Basque) [PolQ LF]  
 [Q  $\exists_i$  [<sub>IP</sub>Ana saw Boris or<sub>i</sub> Osip $\uparrow$ ]]  
 QUD-coherence:  $\checkmark$  Who<sub>{boris or osip,...}</sub> did Ana see?  
 Hamblin set:  $\{\lambda w'.A \text{ saw}_{w'} B \vee A \text{ saw}_{w'} O\}$

Now let us turn to evaluating the coherence between the syntax-semantics and prosody-discourse. As a first step, *Maximize Informativity* is applied to both the QUD and the Hamblin set. This is semantically vacuous, because both sets contain the disjunction and thus no alternatives are ruled out by the principle. The second step is to see whether the criteria in (20) and (22) are met. An answer to the question that is denoted by the Hamblin set indeed provides an at least partial answer to the QUD, and there are also no presuppositions in the Hamblin set that are not carried by the QUD. Thus, the syntax-semantics and prosody-discourse are coherent with each other.

This is different for the same PolQ LF with AltQ prosody in (25). In this case, prosody gives us a final fall that signals that the only two relevant and epistemically live alternatives are {Ana saw Boris} and {Ana saw Osip}.

- (25) *or* associates with  $\exists$  (*edo* in Basque) [PolQ LF]  
 [Q  $\exists_i$  [<sub>IP</sub>Ana saw Boris $\uparrow$  or<sub>i</sub> Osip $\downarrow$ ]]  
 QUD-coherence:  $\#$  Who<sub>{boris,osip}</sub> did Ana see?  
 Hamblin set:  $\{\lambda w'.A \text{ saw}_{w'} B \vee A \text{ saw}_{w'} O\}$

Again, to assess whether daughterQ and motherQ are coherent with each other, let us start with applying *Maximize Informativity* to both the Hamblin set and the QUD. Just as in the previous case, the net semantic effect on the Hamblin set is zero. However, when applied to the QUD, which is restricted by the final fall, it gives rise to the presupposition that exactly one out of {Ana saw Boris} and {Ana saw Osip} is true. We now run into problems if we want to check (20): The daughterQ asks whether the proposition  $\{\lambda w'.A \text{ saw}_{w'} B \vee A \text{ saw}_{w'} O\}$  is true or false. Neither the negative answer nor the positive answer to the daughterQ provides an (at least) partial answer to the motherQ that asks whether {Ana saw Boris} or {Ana saw Osip} is true. DaughterQ and motherQ are not coherent with each other, meaning that the PolQ LF is not available when the utterance is pronounced with AltQ prosody. Note that the application of *Maximize Informativity* to the QUD is partially motivated by, and necessary for, cases like (25). If it would only apply at the utterance level, cases like this could go through: The QUD would only ask about {Ana saw Boris, Ana saw Osip} and the daughterQ, that asks whether {Ana saw Boris or Osip} would provide an (at least partial) answer to that QUD.

The next case is the AltQ LF with AltQ prosody in (26). To evaluate coherence for the LF in which *or* associates with Q, consider the prosody. The accent determines the wh-word of the QUD and the final fall signals that the only two relevant and epistemically live alternatives are *Ana saw Boris* and *Ana saw Osip*.

- (26) *or* associates with Q (*ala* in Basque) [AltQ LF]  
 [Q<sub>i</sub> Ana saw Boris↑ or<sub>i</sub> Osip↓].  
 QUD-coherence: ✓ Who<sub>{b,o}</sub> did Ana see?  
 Hamblin set: {λ*w*.Ana saw<sub>w</sub> Boris, λ*w*.Ana saw<sub>w</sub> Osip}

The application of *Maximize Informativity* results in the presupposition that exactly one of the alternatives is true on both QUD and Hamblin-set level. An answer to the daughterQ provides an answer to the motherQ, hence (20) is satisfied. The presupposition is carried by both the motherQ and the daughterQ. Thus, the criterion in (22) is also satisfied.

I'm now at the final case in (27) namely the AltQ LF with PolQ prosody. The Hamblin set consists of the two propositions *Ana saw Boris* and *Ana saw Osip*. Because of the block intonation, the focus set consists of propositions that are alternatives to *Boris or Osip* and the final rise signals that there could be other relevant, epistemically live possibilities.

- (27) *or* associates with Q (*ala* in Basque) [AltQ LF]  
 [Q<sub>i</sub> [Ana saw Boris or<sub>i</sub> Osip↑]].  
 QUD-coherence: # Who<sub>{boris or osip, ...}</sub> did Ana see?  
 Hamblin set: {λ*w*.Ana saw<sub>w</sub> Boris, λ*w*.Ana saw<sub>w</sub> Osip}

Like for the other cases, the first step is to apply *Maximize Informativity* to both levels. Its application to the QUD is semantically vacuous, because the conjunction is not excluded due to a final fall. Application of *maximize informativity* to the Hamblin set gives rise to the presupposition that either {Ana saw Boris} or {Ana saw Osip} is true, and not both. Now let me turn to the criteria in (20) and (22). An answer to the question that is denoted by the Hamblin set provides an (at least) partial answer to the QUD, thus (20) is met. Coherence crashes because of the criterium in (22). The daughterQ carries a presupposition that is the result of applying *Maximize Informativity*, namely that Ana saw exactly one out of the set {Ana saw Boris, Ana saw Osip}. This is not presupposed by the motherQ, which means there is a violation of (22). Hence, the AltQ interpretation is not available with PolQ prosody.

## 6.2 Predictions

The current proposal makes two predictions. First, the application of *Maximize Informativity* accounts for the observation why AltQs, but not PolQs, give rise to exclusivity effects.

- (28) Did Ana see [Boris]↑ or [Osip]↓?  
 → Ana didn't see both Boris and Osip
- (29) Did Ana see [Boris or Osip]↑?  
 ↗ Ana didn't see both Boris and Osip

If the set {Ana saw Boris, Ana saw Osip} entails one maximally informative true answer, the conjunction is automatically ruled out, by means of it being more informative than the individual alternatives in the set<sup>5</sup> Second, we predict infelicity for utterances that are overtly

<sup>5</sup>Note that exclusivity effects are 'at issue', in the QUD sense, for declaratives, but not for questions [Farkas and Roelofsen, 2017].

- (i) a. A: Did Ana see [Boris]↑ or [Osip]↓?  
 B: # No, she saw both.  
 b. A: Ana saw [Boris]↑ or [Osip]↓.  
 B: ✓ No, she saw both



incoherent, i.e. in languages like Basque, which mark both syntax-semantics and prosody-discourse on the surface, they are not coherent with each other.<sup>6</sup>

- (30) a. # [Te-a **ala** kafe-a] nahi duzu↑?  
           tea-ART **or**<sub>AltQ</sub> coffee-ART want you-it  
       b. # [Te-a]↑ **edo** [kafe-a]↓ nahi duzu↓?  
           tea-ART **or**<sub>PolQ/AltQ</sub> coffee-ART want you-it

This prediction is borne out and the data serve as nice support for the current account.

## 7 Comparison to existing accounts

In this section I compare the current proposal to the prominent accounts in the literature from [Biezma and Rawlins \[2012\]](#) and [Roelofsen and van Gool \[2010\]](#). A detailed review of the accounts outscopes the goals of this paper, but the key points are as follows. Both [Biezma and Rawlins \[2012\]](#) and [Roelofsen and van Gool \[2010\]](#) argue that the final falling boundary tone is the crucial ingredient for AltQ composition. Furthermore, in both accounts the final fall is given a purely semantic function. Simplified, this function is to signal the presence of an exclusive strengthening operator (in [[Roelofsen and van Gool, 2010](#)]) or an operator that exhausts the set of possible answers to the question (in [[Biezma and Rawlins, 2012](#)]). Finally, [Biezma and Rawlins \[2012\]](#) do not model the focal accent on the disjuncts and [Roelofsen and van Gool \[2010\]](#) make its semantic net contribution null. There is a number of crucial advantages of the current proposal, as compared to the described accounts. First, the present proposal accounts for the data in Basque and Finnish, which rely on other surface cues than prosody to compose an AltQ. At this point, I do not see how an account that only models the final falling intonation could account for the crosslinguistic data. Basque and Finnish encode AltQ meaning in the lexical item for disjunction, which cannot be explained if the final fall is the crucial ingredient for AltQ composition. Secondly, the current account models both the accents on the disjuncts and the final fall. Modelling both prosodic surface cues is a fundamental requirement for a satisfying account of AltQs, even if one is only concerned with English [[Meertens et al., 2019](#)]. Thirdly, [Biezma and Rawlins \[2012\]](#) and [Roelofsen and van Gool \[2010\]](#) take the exclusivity effects of AltQs to be encoded in an operator as a further stipulation. Within the current proposal, exclusivity effects follow naturally from pragmatic principles that are independently motivated.

## 8 Conclusion

I have proposed that the disambiguating effects of prosody in English are the result of coherence requirements between syntax-semantics mapping and prosody-discourse mapping. This proposal makes use of coherence requirements that are independently motivated. I have also showed that the current proposal is able to deal with crosslinguistic variation in surface cues for AltQ composition and naturally explains the exclusivity effects of AltQs in English.

---

*Maximize Informativity* says it is presupposed that there is one maximally informative answer in the Hamblin set. We thus expect the exclusivity effects to be not at issue and therefore the data in (i) are accounted for.

<sup>6</sup>Thanks to Sergio Monforte for these specific data.



## References

- Luis Alonso-Ovalle. *Disjunction in alternative semantics*. University of Massachusetts Amherst Amherst, 2006.
- Christine Bartels. *The intonation of English statements and questions*. Garland, New York, 1999.
- María Biezma. Alternative vs polar questions: The cornering effect. In E. Cormany, S. Ito, and D. Lutz, editors, *Proceedings of Semantics and Linguistic Theory (SALT) 19*, pages 37–54. DOI: 10.3765/salt.v19i0.2519, 2009.
- María Biezma and Kyle Rawlins. Responding to alternative and polar questions. *Linguistics and Philosophy*, 35(5):361–406, 2012.
- Daniel Büring. On D-trees, beans and B-accents. *Linguistics and Philosophy*, 26:511–545, 2003.
- Veneeta Dayal. *Locality in WH quantification: Questions and relative clauses in Hindi*. Kluwer, Dordrecht, 1996.
- Donka Farkas and Floris Roelofsen. Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics*, 34(2):237–289, 2017.
- Martin Haspelmath. Coordination. *Language typology and syntactic description*, 2:1–51, 2007.
- Erlinde Meertens, Sophie Egger, and Maribel Romero. Multiple accent in alternative questions. In *Proceedings of Sinn und Bedeutung*, volume 23, pages 179–196, 2019.
- Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. In J. H. Yoon and A. Kathol, editors, *Ohio State University Working Papers in Linguistics*, Vol. 49: *Papers in Semantics*, pages 91–136, Columbus, OH, 1996. Ohio State University.
- Floris Roelofsen and Sam van Gool. Disjunctive questions, intonation, and highlighting. In H. Bastiaanse M. Aloni and T. de Jager, editors, *Logic, Language and Meaning: Selected Papers from the 17th Amsterdam Colloquium*, pages 384–394. Springer, Heidelberg, 2010.
- Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116, 1992.
- Mario Saltarelli. *Basque*. Croom Helm Ltd, 1988.
- Uli Sauerland. Scalar implicatures in complex sentences. *Linguistics and philosophy*, 27(3): 367–391, 2004.
- Junko Shimoyama. Indeterminate phrase quantification in Japanese. *Natural Language Semantics*, 14(2):139–173, 2006.
- Benjamin Spector. Topics in the semantics of interrogative clauses. Minicourse at the University of Vienna. <http://lumiere.ens.fr/~bspector/Questions2010/>, 2010.
- Matthijs Westera. *Exhaustivity and intonation: A unified theory*. PhD thesis, University of Amsterdam, 2017.

# Combining neurophysiology and formal semantics and pragmatics: the case of the N400

Ralf Naumann<sup>1</sup> and Wiebke Petersen<sup>2</sup>

<sup>1</sup> Heinrich-Heine-Universität Düsseldorf  
naumann@phil.hhu.de

<sup>2</sup> Heinrich-Heine-Universität Düsseldorf  
petersen@phil.hhu.de

## Abstract

We present an outline of how experimental data from neurolinguistics related to one particular ERP-component, the N400, can be analyzed in a probabilistic extension of Incremental Dynamics with frames and situation models. We show that none of the semantic and/or pragmatic properties proposed in the neurolinguistic literature alone can explain the whole range of data. Our own approach is similar to that of Werning et al. in taking the pragmatic dimension seriously by incorporating both the perspective of the speaker and the listener using RSA using a Bayesian model. Probabilities are calculated by using both semantic information which is based on an information ordering on situation models and discourse information which is based on a linking relation between discourse referents.

## 1 Introduction

An ERP-component is the summation of the post-synaptic potentials of large ensembles (in the order of thousands or millions) of neurons synchronized to an event. When measured from the scalp, continuous ERP-components manifest themselves as voltage fluctuations that can be divided into components. A component is taken to reflect the neural activity underlying a specific computational activity carried out in a given neuroanatomical module. The N400 component is a negative deflection in the ERP signal that starts around 200 - 300 ms post-word onset and peaks around 400ms.

The N400 amplitude on a word  $w$  in a context  $c = w_1 \dots w_t$  is typically inversely related to its conditional probability given this context:  $P(w|c)$ , [11]. Underlying this relation is a model of online processing according to which at every step during this processing a prediction about the upcoming word is made guided by the probability distribution  $P(w|c)$  (see [11] for an overview). The N400 amplitude shows a gradient effect. It is smallest for the most predicted words, intermediate for the words with moderate predictability and larger for words with the lowest predictability (see e.g. [18] and [13]). This conditional probability can be measured either by the human judged cloze probability or by the information-theoretic notion of surprisal. Given an initial sequence of words  $w_1 \dots w_{t-1}$ ,  $w_t$  can be viewed as a random variable. Its surprisal (or self-information) is defined as follows:<sup>1</sup>  $surprisal(w_t) := -\log P(w_t | w_1 \dots w_{t-1})$ . If defined in this way surprisal of a word is typically directly related to the word's N400 amplitude, [9].

The context  $w_1 \dots w_{t-1}$  must not be restricted to a single clause or sentence because the wider context can have an influence on the modulation of the N400 amplitude. For example, the probabilities for the target words in the sentences “The peanut was salted/in love” are inverted, if the sentence is not presented in isolation but in the context of a comic or a fiction story in which the peanut is ascribed typical human properties like being able to sing and dance. As an

---

<sup>1</sup>The base of the logarithm is an arbitrary scaling factor.

effect, a property like ‘being salted’ is now highly unlikely if not even impossible. By contrast, the property of being in love now receives a high probability. This inversion of probabilities is reflected in the N400 amplitude: for ‘salted’ this amplitude was enhanced compared to that for ‘in love’, [16].

The two most prominent interpretations of the underlying neuro-cognitive function of the N400 are the integration and the retrieval view. On the integration account, the N400 amplitude ‘indexes the effort involved in integrating the word meaning of the eliciting word form with the preceding context, to produce an updated utterance interpretation’, [7]. On the retrieval/access account ‘the N400 amplitude reflects the effort involved in retrieving from long-term memory conceptual knowledge associated with the eliciting word which is influenced by the extent to which this knowledge is cued (or primed) by the preceding context, [7]. What is left open by the above characterization is which properties of words and the context underly the N400 amplitude. Three prominent properties that have been suggested are (i) semantic features, (ii) plausibility, and (iii) semantic similarity.

Evidence for semantic features as being correlated with the N400 amplitude comes from the fact that the correlation between cloze probability and the N400 amplitude is not monotone.

- (1) They wanted to make the hotel look more like a tropical resort. So along the driveway they planted rows of palms/pines/tulips. [8]

In (1) ‘pine’ but not ‘tulips’ comes from the same semantic category ‘tree’ as the best completion ‘palms’. Though ‘pines’ and ‘tulips’ have the same low cloze probability ( $< 0.05$ ), their N400 amplitudes differ. Within category violations (pines) elicit smaller N400 amplitudes than between category violations (tulips). Federmeier & Kutas argue that this result suggests that it is feature overlap like being tall or having a similar form that affords within category violations a processing benefit relative to between category violations, [8, p.485].

However, feature overlap with the best completion is not without exceptions, as shown by the following examples.

- (2) a. A huge blizzard swept through town last night. My kids ended up getting the day off from school. They spent the whole day outside building a big jacket in the front yard, [14].  
 b. The wreckage of the sunken ship was salvaged by the victims ... [17].

Though the critical words share few semantic features with the best completions (snowman, divers), either no, (2-b), or only a small N400 effect, (2-a), is observed.

A second candidate is plausibility which can be quantized by offline rating tasks using, e.g., a Likert scale. Plausibility is often related to the integration view of the N400. The less plausible a resulting interpretation is the more difficult must it have been to integrate the critical word in the preceding context. Evidence for the role of plausibility comes from the fact that in the Federmeier & Kutas study best completions elicited the smallest N400 amplitude and the highest plausibility ratings. Between category violations elicited the largest N400 amplitudes and got the lowest plausibility ratings. Within category violations were intermediate on both variables, [8, p.486]. However, this monotone relation no longer holds if contextual constraint is taken into account. Most importantly, in low-constraint contexts the plausibility for within category violations is significantly higher compared to high-constraint contexts. By contrast, the N400 amplitudes are significantly different in the opposite direction. The more plausible within category violation in low-constraint contexts have a higher N400 amplitude than the less plausible within category violations in a high-constraint context. Furthermore, in semantic illusion data as given in (3) no N400 effect is observed although the sentence has an implausible

interpretation.

(3) The fox that on the poacher hunted . . . .

A third candidate is semantic similarity. On this account the N400 amplitude is modulated by the degree to which a critical word in a target sentence is semantically related to the words preceding it in the context. For example, in the peanut example one has the ‘being in love’ has a higher semantic similarity to words like ‘dancing’ and ‘singing’ than to ‘being salted’. One way of quantifying semantic similarity is to use Latent Semantic Analysis. On this account pairwise term-to-document semantic similarity values (SSVs) are extracted from corpora (see [13] for an application). Semantic similarity underlies the Retrieval-Integration model of [20]. One of its strengths is that it can explain semantic illusion data as given in (3). As there is a semantic relation between the arguments preceding the verb (‘fox’, ‘poacher’) and the verb itself (‘hunted’) no N400 effect is expected for the verb.

However, similar to both the notions of semantic feature overlap and plausibility, there are counterexamples to the thesis that the N400 amplitude is (monotonically) related to the corresponding LSA value. Kuperberg et al., [12] showed that the degree of causal relationship in three-sentence scenarios with matched SSVs influences the N400 amplitude: highly related < intermediately related < causally unrelated. The authors conclude that it is the situation model constructed from the context (message-level meaning) that influences semantic processing of the critical word and not semantic relatedness. Similarly, [13] could show an influence of high-versus low-constraint contexts on the N400 amplitude for controlled SSVs.

Let us summarize the findings of this section: The modulation of the N400 amplitude is sensitive to (i) semantic feature overlap and not simply to words (for a similar argument, see [13]). Exceptions are cases in which the sort of the word actually found does not share semantic features with the best completion, but with the preceding context. (ii) Plausibility judgements accounts for the preceding context, but fail to explain semantic illusion data. Finally, (iii) semantic similarity as defined by LSA abstracts away from thematic roles and thus captures semantic illusion data. However, it is not restrictive enough, as it does not account for causal relatedness and degree of context constraint.

An alternative is proposed by Werning and colleagues, [21], [22]. They start from the assumption, already discussed in section 1, that at any moment in a communicative situation a comprehender generates a probabilistic prediction about how a sentence or a discourse uttered by a speaker will most likely be continued, [21, p. 3504]. This communicative act is goal-directed, i.e. the speaker wants to describe a particular situation. Hence, he will choose a context  $c$  which makes the referent denoted by  $w$  (highly) relevant. They define  $P(w|c)$  not in terms of a single property. They rather follow the rational speech act model (RSA) and assume a Bayesian model, see e.g. [6]. Such a model allows comprehenders to update their priors regarding a word  $w$  following a context  $c$  with pragmatic considerations on speakers’ intentions thereby arriving at a probabilistic prediction of  $w$ . The conditional probability  $P(w|c)$  is defined as the product of a probabilistic semantic factor, given by the prior, and a pragmatic (discourse) component that is represented by the likelihood term.

(4)  $P(w|c) \propto P(w) \cdot P(c|w)$ .

The prior is a function of the semantic similarity between  $w$  and  $c$  and/or another word  $w'$  in  $c$  and is defined using LSA. This reflects overall statistical co-occurrence patterns and hence statistical regularities. The likelihood term models the pragmatic dimension. Given that a speaker wants to convey information about the referent of a word  $w$  he will choose a context  $c$  that makes the occurrence of this referent relevant (plausible). Hence  $P(c|w)$  strictly increases

with the relevance (plausibility) of  $c$  for  $w$ . In their empirical studies they showed that relating the modulation of the N400 amplitude to (4) yields empirical better results than relating this modulation to either only semantic similarity (prior) or relevance (plausibility) (likelihood). Problematic for this approach are semantic illusion data and examples like (2) for which there is no or only an attenuated N400 effect despite the fact that there is no semantic feature overlap between the critical word and the best completion. In the case of semantic illusion data the semantic similarity between the verb and the two preceding arguments is the same in the expected and the switched thematic role variants because switching thematic roles does not affect the SSV value. However, they differ w.r.t. relevance (plausibility) due to the difference in thematic role assignments. As a result, an N400 effect for the switched roles variants compared to the expected role assignments should be observed. Similarly, for the data in (2) the semantic similarity between the critical word and the expected best completion is nearly identical (e.g. in the case of (2-b) one has  $SSV(\text{divers, context})=0.22$  and  $SSV(\text{victims, context}) = 0.18$ ). By contrast, the difference in plausibility is significant: 6.3 vs 2.9 on a 7-point Likert scale. Hence, an N400 effect for ‘victims’ relative to ‘divers’ is expected, contrary to what was observed. Given the results of the first section, these problems reflect the failure of plausibility being an underlying factor of the modulation of the N400 amplitude. Relevance/plausibility applies at the propositional/discourse level and hence at the level of event structures. This raises the question whether there is another property, possibly related to a different semantic object, that underlies this modulation.

## 2 Towards an RSA account on language processing that respects neurophysiological findings

If quantized by LSA, semantic similarity is based on semantic relationships between words and concepts, including (inferential) schema-based relationships. It is insensitive to word order and both syntax and thematic relation. An example of such inferential relationships is the qualia structure in the Generative Lexicon. It links a sort of objects, say cream, to a particular action (or a set of actions) that specifies the function or purpose of objects of this sort. For a comprehender who processes the verb denoting this action the interpretative task is to relate the corresponding discourse referent to the discourse referent of the noun to whose qualia structure the action belongs. This suggests that the N400 is sensitive to establishing such relations between discourse referents. This hypothesis raises two questions: (i) is there direct evidence that the N400 is sensitive to such relations between objects?, and (ii) How can these relations be defined? Consider the examples in (5).

- (5) a. Peter hatte einen langen Tag und wollte ein Bier. Die Kneipe war bis Mitternacht geöffnet/Das Essen war bereits auf dem Tisch, [7].  
Peter had a long day and wanted a beer. The bar was open till midnight./ The meal was already on the table.
- b. Tobias besuchte einen Dirigenten/ein Konzert in Berlin/unterhielt sich mit Nina. Er erzählte, daß der Dirigent sehr beeindruckend war, [3].  
Tobias visited a concert/a conductor in Berlin/talked to Nina. He said that the conductor was very impressive.

In contrast to factive verbs, the existence of the direct object of non-factive verbs like ‘want’ is neither presupposed nor asserted. ‘Want’ raises a particular question under discussion: ‘Did the actor get what he wanted?’ This question triggers the expectation that this question will be

answered. A bar is a paradigmatic place where one gets beer (or not if, e.g. it is already closed) and where one can go if one wants a beer. By contrast, a meal can be served without any beverages and, in addition, a beer being served is only one among many possibilities. Delogu et al. found an N400 effect for ‘Essen’ compared to ‘Kneipe’. The examples in (5-b) involve bridging inferences. Burkhardt found an attenuated N400 effect for bridged DPs (Konzert - Dirigent) and an enhanced effect for new DPs (Nina - Dirigent) compared to the given DP (Dirigent - Dirigent). She calls the general phenomenon ‘discourse linking’.

Examples like those in (2) show that discourse linking cannot be restricted to (single) event structures. Rather, it is related to situation models. Such models go beyond the propositional content conveyed by the words in a context and which essentially involve world knowledge, [23] see [1] on learning such models from texts]). For example, in (2-a) a wintery scene and in (2-b) a ship wreckage scenario is described. Situation models basically are complex events. Since situation models comprise sequences of events, predictions are not restricted to single events. For example, given a context that specifies a situation model whose prototypical realization consists of the action sequence  $e_1 \dots e_r$  and in which the initial sequence  $e_1 \dots e_k$  has been introduced, predictions are possibly related to any of the events  $e_{k+1} \dots e_r$  and objects participating in them. In the wintery scenery in (2-a) the jackets are expected because they are related to the children in a (yet to be introduced) state of wearing which is a background state that constantly holds while the children were playing outside. As a second example, consider (6), which is an example of semantic illusion data.

(6) The restaurant owner forgot which waitress the customer had ...

Given a restaurant scenario, in which a waitress and a customer have already been introduced, actions like ‘serve’, ‘ask’, ‘order’ or ‘pay’ are expected. In this case these actions are not only related to the restaurant scenario but also to the current event whose sort is still unknown. Hence, situation models possibly set up predictions to objects in event structures that have not yet been introduced.

We hypothesize that the modulation of the N400 amplitude is sensitive to (i) semantic similarity between situation models and (ii) the way situation models are related by discourse linking. Let us relate this to the RSA approach (see e.g. [6]). The main insight underlying the RSA model is that a listener not only uses the literal meaning possibly enriched by world and script knowledge but also takes into account that a speaker chooses an expression in such a way that he is able to infer the intended referent of the expression. Hence, speaker and listener recursively reason about each others’ goals to arrive at pragmatically enriched meanings. Formally, one has that a speaker  $S_1$  chooses a term  $t$  to (soft-max) optimize expected utility given a meaning (referent)  $r$ :  $S_1(t|r) \propto e^{\lambda(U_1(t|r))}$  ( $\lambda$  is the gain on the speaker’s softmax decision rule). A literal listener interprets utterances literally without reasoning about the speaker. He has a prior distribution over referents and uses Bayesian inference to (eliminatively) update her belief about the intended referent given the utterance’s literal meaning. A pragmatic listener  $L_1$  then reasons about  $S_1$  by inverting  $S_1$ ’s model using Bayes’ rule in order to infer the referent  $r$  given utterance  $u$ , where  $P(r)$  is the prior probability over referents:  $L_1(r|t) \propto P(r)S_1(t|r)$ . In our application  $t$  and  $r$  are situation models  $sm$  and  $sm'$ . The task, therefore, is to define  $P(sm)$  and the utility function  $U_1(sm, sm')$ . In order to solve this task, one has to define the meaning of common nouns and verbs in terms of semantic features. This will be done by a decompositional analysis. Second, probabilities have to be defined in terms of such decompositional structures. To this end, frames and their properties will be introduced.

In order to account for the modulation of the N400 amplitude at the semantic feature level, common nouns and verbs are not interpreted as sets of objects, either individuals or events.

Rather, they are interpreted as sets of pairs  $\langle o, f \rangle$  consisting of an object and a frame. Frames are elements of a separate domain  $D_f$  of frames. Each frame is related to a particular object (an individual or a (complex) event) as its root and is a partial description of that object in a particular world. Being a partial description of an object, a frame is linked to a relational structure that is built by (finite) chains of attributes. This link is captured by a function  $\theta$  which maps a frame to a set of relations. For a given object, its associated frame stores information got during a discourse so far as well as world knowledge. Besides the domain  $D_f$ , there are the domains  $D_i$  of individuals, the domain  $D_e$  of events and the domain  $D_w$  of possible worlds.

At the discourse level we use Incremental Dynamics [19] enriched with frames [15]. We extend our approach in [15] by set-valued frames for the current situation. Situation models or possibilities are triples  $sm = \langle c_{sm}, f_{sm}, w_{sm} \rangle$  consisting of a stack  $c_{sm}$ , a current situation frame  $f_{sm}$  and a world  $w_{sm}$ . A particular stack position is a pair  $\langle o, f \rangle$  with  $o \in D_i \cup D_e$  and  $f \in D_f$ . An information state is a set of possibilities. Every situation frame  $f_{sm}$  has an attribute ACTIONS whose value is the set of actions (events) occurring in this scenario (denoted by  $a(f_{sm})$ ). A second attribute is PARTICIPANTS whose value is a set of individuals  $p(f_{sm})$ . Each element of this set is related to at least one action or one other participant, the set of these pairs  $pr(f_{sm})$  is the value of the attribute PARTICIPANCY\_RELATION. The value of the attribute ORDER is a set  $o(f_{sm})$  of pairs of events that preorders the value of ACTIONS attribute. Situation frames are sorted by  $SM$  which are sorts of complex events like ‘wintery scenario’ or ‘restaurant scheme’.

In order to capture the sensitivity of the modulation of the N400 amplitude to both semantic features and world knowledge probabilities need to be defined not at the level of words but at the frame level and particularly at the level of situation frames. Whenever a new word is processed, the current situation model is updated. In the case of a newly introduced discourse referent, the set values in the situation frame are extended and the stack is prolonged by one position (see [15] for details). In case of an anaphora there exists already a corresponding object frame pair in the stack of which the frame is updated, i.e. refined, by the newly gained information (see again [15] for details). On the level of ordinary frames the refinement can be defined in terms of the following information ordering  $\sqsubseteq$  on frames. Let  $\theta(f) = \{R_1, \dots, R_n\}$  be the set of chains associated with  $f$ . One has  $f \sqsubseteq f'$  if (i)  $f$  and  $f'$  have the same root and (iii)  $\theta(f) \subseteq \theta(f')$ . When taken together, this means that  $f'$  possibly contains information about more attributes or  $f'$  possibly contains more specific information about the values of attributes. For set-valued situation frames and full situation models the information order is as follows:

- (7) a.  $f_{sm} \sqsubseteq f_{sm'}$  iff  $a(f_{sm}) \subseteq a(f_{sm'})$  and  $p(f_{sm}) \subseteq p(f_{sm'})$  and  $pr(f_{sm}) \subseteq pr(f_{sm'})$  and  $o(f_{sm}) \subseteq o(f_{sm'})$ .  
 b.  $sm \sqsubseteq sm'$  iff (i)  $w_{sm} = w_{sm'}$ , (ii)  $f_{sm} \sqsubseteq f_{sm'}$  and (iii)  $\forall o : \text{if } \langle o, f \rangle \in c_{sm} \text{ and } \langle o, f' \rangle \in c_{sm'} \text{ then } f \sqsubseteq f'$ .

Thus a situation model  $sm'$  extends or refines a situation model  $sm$  if both belong to the same world and the attribute values of  $f_{sm'}$  are supersets of the corresponding values of  $f_{sm}$  and for all objects belonging to  $sm$  there corresponding frames stored in the stack  $c_{sm}$  are possibly refined in  $sm'$ . An update is a move along the information hierarchy: Let the context be given by the words  $w_1 \dots w_{t-1}$  with corresponding frames  $f_1 \dots f_{t-1}$  resulting in a situation model  $sm_{w_1^{t-1}}$ . The meaning of the word  $w_t$  is a context change potential. In the present context this is the change it brings about with respect to  $sm_{w_1^{t-1}}$  yielding the updated situation model  $sm_{w_1^t}$ . The contribution of  $w_t$  is an object frame pair  $\langle o, f_t \rangle$ . The operation of updating  $sm_{w_1^{t-1}}$  with  $\langle o, f_t \rangle$  will be denoted by  $\oplus$ :  $sm_{w_1^t} = sm_{w_1^{t-1}} \oplus \langle o, f_t \rangle$ . Let's assume that  $w_t$  introduces a new



discourse referent  $o \notin p(sm)$ . Then it extends the stack,  $c_{sm'} = c_{sm} \frown \langle o, f_t \rangle$ , and introduces a new participant,  $p(f_{sm'}) = p(f_{sm}) \cup \{o\}$ , that is linked to at least one other participant or action,  $\exists o' \in p(f_{sm'}) \cup a(f_{sm'}) : pr(f_{sm}) \cup \{o, o'\} \subseteq pr(f_{sm'})$ . The case for a new action discourse referent  $o \notin a(f_{sm})$  is similar. If  $w_t$  denotes an object  $o$  that is already on the stack,  $\langle o, f \rangle c_{sm}$  then the corresponding frame is updated by the information  $f_t$  via unification:  $\langle o, f \sqcup f_t \rangle \in c_{sm'}$ .

Next we define the prior for a situation model  $sm$  given a situation model  $sm'$ . In order to do so, we need to abstract away from the concrete worlds and the concrete participants. Two situation models  $sm$  and  $sm'$  are alphabetic variants of each other,  $sm \approx sm'$ , iff there are bijections  $\phi : p(sm) \rightarrow p(sm')$  and  $\phi : a(sm) \rightarrow a(sm')$  such that  $f_{sm'} = \phi(f_{sm})$  and  $c_{sm'} = \phi(c_{sm})$ . Thus alphabetic variants differ only in their worlds, their participants and their actions, but not in their structure and thus not in their general properties. Given a situation model  $sm$  resulting from the processing of a context  $w_1 \dots w_{t-1}$ , the prior for a possible extension of  $sm$  by the frame  $f_t$  is

$$P(sm \oplus f_t) = \frac{|\bigcup_{sm' \approx sm \oplus f_t} \uparrow sm'|}{|\bigcup_{f \in D_f \wedge sm' \approx sm \oplus f} \uparrow sm'|}$$

with  $\uparrow sm$  being the filter of  $sm$ , i.e.  $\uparrow sm = \{sm' | sm \sqsubseteq sm'\}$ .

$P(sm \oplus f_t)$  will be used as the prior. Note that the prior is calculated without taking thematic role information into account. The reason why predictions are based on semantic features and not on information based on thematic roles has to do with ambiguity triggered by literal meaning which considerably increases processing load with the effect that it is not possible to arrive at a stable representation without making use of predictions, [5]. Semantic processing in the brain is done in a left-to-right fashion. This makes it necessary to rethink the way arguments are related to verbs (see [2] for evidence and discussion). Following [4] and [2], we assume the following structure of a DP:  $[[DetN]_{DP_1} [TR]]_{DP_2}$  in which the contribution of sortal and thematic role information are separated. On this interpretation the assignment of a thematic role can be taken as a non-deterministic operation which introduces branching (at least for case-less languages like English or Dutch). The crucial difference between the two kinds of information is that for a listener only the sortal information is directly given whereas thematic role information has to be inferred.

Recall that on a pragmatic-discourse oriented perspective texts are goal-oriented. In our approach the global aim of the speaker is to describe a particular situation. To this end, he will locally introduce objects that are part of this situation and attribute properties to them. He will therefore choose a (prior) context that makes the mention of these objects (most) likely. Such a mention is highly probable if the context already contains this object as an element. Next come objects that are (directly) related by an attribute to objects already introduced in the context since they, at least indirectly, extend information about objects already introduced. This has the effect of making the text coherent by maximizing anaphoricity. We hypothesize that these probabilities are related to particular accessibility relations between situation models that are defined in (8).

- (8) a.  $sm \underline{\text{Id}} sm'$  iff (i)  $f_{sm} \sqsubseteq f_{sm'}$  and (ii)  $a(sm) \cup p(sm) = a(sm') \cup p(sm')$ .  
 b.  $sm \underline{\text{BI}} sm'$  iff (i)  $f_{sm} \sqsubseteq f_{sm'}$  and (ii)  $\forall o' \in (a(sm') \cup p(sm')) \setminus (a(sm) \cup p(sm)) : \exists \langle o, f \rangle \in c_{sm} \exists f' \exists R$  with  $f \sqsubseteq f'$  and  $R(f')(o)(o')$ .  
 c.  $\underline{\text{Id}} \cup \underline{\text{BI}} \subseteq \underline{\text{DL}}; \underline{\text{DL}} \subseteq \underline{\text{DL}}$ .

$\underline{\text{Id}}$  requires the two situation models to have the same participants and events. Hence, this relation captures the case of given DPs.  $\underline{\text{BI}}$  requires that for each object that belongs to  $sm'$



but not to  $sm$  there is an object belonging to  $sm$  and a relation in a minimal extension of the frame of this object that links the two objects. Thus, this relation captures the case of bridged DPs. Finally, discourse linking,  $\underline{DL}$ , comprises these two relations and is itself a subrelation of  $\underline{\square}$ . This can be tested by using a procedure based on the next-mention bias used by [10] for anaphoric relations involving pronouns. This bias will be largest for objects already introduced, in particular if they are related to the topic. Similarly, bridged DPs answering a QuD as in the example (5-a), will get a high probability.

How is discourse linking based on the accessibility relations in (8) related to the information conveyed by a word? Recall that the aim of a speaker is to describe a particular situation. At each point in processing there is some degree of uncertainty for the listener about which situation model is described. This uncertainty decreases by getting to know the frame of the next word and the extent of this decrease corresponds to the information contained in that frame. According to discourse linking, this decrease should be related to making the discourse coherent by maximizing anaphoricity. This will be the case if the interpretation of a DP does decrease uncertainty only minimally. Let us make this formally precise. Recall that in RSA the speaker's utility function must be related to the information conveyed by a word in an utterance. This information is often quantized by surprisal discussed in the first section. However, as also shown there, this metric can neither account for semantic illusion data nor for the examples in (2) where there is no feature overlap between the critical word and the expected best completion. As alternative to surprisal, we use entropy reduction which is based on the notion of (n-step) entropy and which is directly related to uncertainty about a referent. For  $S_{f_{comp}}$  the set of all completions, the listener wants to know which situation model  $sm$  is described by determining a frame  $f_{sm} \in S_{f_{comp}}$ . The uncertainty about  $f_{sm}$  is defined as the entropy of the probability distribution over  $S_{f_{comp}}$ :  $H(S_{f_{comp}}) = -\sum_{f_{sm} \in S_{f_{comp}}} P(f_{sm}) \log(P(f_{sm}))$ . When the first  $t$  words of a sentence have been processed, the probability distribution over  $S_{f_{comp}}$  has changed from  $P(S_{f_{comp}})$  to  $P(S_{f_{comp}} | f_{w_1^t})$  with  $f_{w_1^t}$  being the frame got after the first  $t$  words. The corresponding entropy equals  $H(S_{f_{comp}}; f_{w_1^t}) = -\sum_{f_{sm} \in S_{f_{comp}}} P(f_{sm} | f_{w_1^t}) \log P(f_{sm} | f_{w_1^t})$ . The amount of information that the next word  $w_{t+1}$  gives about the random variable  $S_{f_{comp}}$  is defined as the reduction in entropy due to that word:  $\Delta H(S_{f_{comp}}; f_{w_{t+1}}) = H(S_{f_{comp}}; f_{w_1^t}) - H(S_{f_{comp}}; f_{w_1^{t+1}})$ . Let's apply this to some examples from discourse linking. Consider first the case of a given DP like 'the conductor' in (5-b). This kind of DP does not exclude any extensions that were possible before this DP was encountered because the information related to this DP was already known in the input information state. For bridged DPs, this will in general not be the case because some extensions are excluded by establishing a linking relation that was not known before. Take, for example, the case of the jackets in (2-a). This excludes situations in which the children were wearing coats or ski suits. A bridged DP need not always exclude extensions. In the case of the ship wreckage, e.g., it is not excluded that there were other material casualties besides the victims. This example shows that entropy reduction is more fine-grained than the subrelations of  $\underline{DL}$  because it allows for distinctions in the  $\underline{BI}$  relation. By contrast, a new DP leads to an increase in entropy because in general it cannot be directly linked to an object already introduced. For example, in the case of the conductor about whom Tobias talks to Nina there is no direct link relating the former to the latter two objects. What is missing, e.g., is a discourse referent that establishes such a relation. This can be a concert that Tobias attended. When taken together, one gets that a speaker chooses a context in such a way that the reduction in entropy is minimal for the object he wants to talk about next. Entropy reduction is calculated relative to the ordering  $\underline{DL}$ . Hence, the contribution of  $f_t$  in determining an element of  $S_{f_{comp}}$  is always related to the transition from  $sm$  to  $sm'$  triggered by  $f_t$  relative to

this ordering. This contribution is denoted by  $\oplus_{\text{DL}}(sm, sm', f_{w_{t+1}})$ . When taken together, the utility function  $U_1$  is defined as follows:  $U_1(sm | sm') := -\Delta H(S_{f_{comp}}; \oplus_{\text{DL}}(sm, sm', f_{w_{t+1}}))$ .

Let us next discuss some examples used above. We will begin with the two examples in (2). Both ‘jackets’ and ‘victims’ can be linked to an object that is an element of the current situation model. The jackets are the value of the CLOTHES attribute and the victims are the value of the CASUALTIES attribute. As already stated above, they differ w.r.t. the information they provide. Whereas ‘victims’ does not exclude other casualties, ‘jackets’ excludes other sorts of outer clothing like coats for example. Similarly, the prior for ‘victims’ will in general be greater than that for ‘jackets’ due to the fact that the children can wear other clothes. Comparing ‘victims’ with the most expected ‘divers’, one gets: in both cases a new object is introduced into the situation model. Similar to ‘victims’, ‘divers’ does not exclude any sort because salvaging a shipwreck requires objects of this sort. The priors will also be equal so that no N400 effect is expected. For the example (6) one has: The prior for an event of serving is the same for a restaurant scenario in which a waitress and a customer have been introduced. Note that this only holds if thematic role assignments are not taken into account. Similarly, the action of serving can equally be linked to objects of sort ‘waitress’ and ‘customer’ as actions in which objects of this sort are involved. Since the action is identical for both assignments of thematic roles, entropy reduction does not differ. Consider next example (1). The critical words differ with respect to their prior probabilities. Given the preceding context, ‘palm’ are more expected because they satisfy more of the (soft) constraints imposed on the object planted, e.g. that its typical origin are the tropics. For linking at the discourse level, one gets: in the input information state an event of planting has already been introduced which expects objects of sort ‘plant’. This constraint is satisfied by all three sorts which occur in the position of the critical word. They do not differ w.r.t. entropy reduction because they all paradigmatically exclude the other possibilities.

### 3 Acknowledgments

We would like to thank Harm Brouwer for illuminating discussions on the neurolinguistic data and the relation to the Retrieval-Integration model. Markus Werning provided helpful discussions of his own approach as well as of possible ways to combine neurolinguistic data with formal semantic theories. The research was supported by the German Science Foundation (DFG) funding the Collaborative Research Center 991.

### References

- [1] Simon Ahrendt and Vera Demberg. Improving event prediction by representing script participants. In *Proceedings NAACL 2016*, pages 546–551. ACL, 2016.
- [2] Oliver Bott and Wolfgang Sternefeld. An event semantics with continuations for incremental interpretation. *Journal of Semantics*, 34(2):201–236, 2017.
- [3] Petra Burkhardt. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98(2):159 – 168, 2006.
- [4] Lucas Champollion. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy*, 38(1):31–66, Feb 2015.
- [5] Herbert Clark. *Using language*. Cambridge University Press, 1996.
- [6] Judith Degen, Michael Henry Tessler, and Noah D. Goodman. Wonky worlds: Listeners revise world knowledge when utterances are odd. In *Proc. of CogSci 2015*, page 548–553.

- [7] Francesca Delogu, Harm Brouwer, and Matthew W. Crocker. Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension. *Brain and Cognition*, 135:103569, 2019.
- [8] Kara D. Federmeier and Marta Kutas. A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41:469, 1999.
- [9] Stefan L. Frank, Leun J. Otten, Giulia Galli, and Gabriella Vigliocco. The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, 140:1 – 11, 2015.
- [10] Andrew Kehler, Laura Kertz, Hannah Rohde, and Jeffrey Elman. Evaluating an expectation-driven question-under-discussion model of discourse interpretation. *Discourse Processes*, 53:1–20, 2016.
- [11] Gina Kuperberg and T. Florian Jaeger. What do we mean by prediction in language comprehension? *Language, cognition and neuroscience*, 31(1):32–59, 2016.
- [12] Gina Kuperberg, Martin Paczynski, and Tali Ditman. Establishing causal coherence across sentences: an ERP study. *Journal of Cognitive Neuroscience*, 23(5):1230–1246, 2011.
- [13] Gina R. Kuperberg, Trevor Brothers, and Edward W. Wlotko. A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, 0(0):1–24, 0.
- [14] Ross Metusalem, Marta Kutas, Thomas P. Urbach, Mary Hare, Ken McRae, and Jeffrey L. Elman. Generalized event knowledge activation during online sentence comprehension. *Journal of memory and language*, 66 4:545–567, 2012.
- [15] Ralf Naumann and Wiebke Petersen. Bridging inferences in a dynamic frame theory. In Alexandra Silva, Sam Staton, Peter Sutton, and Carla Umbach, editors, *Language, Logic, and Computation*, pages 228–252, Berlin, Heidelberg, 2019. Springer Berlin Heidelberg.
- [16] Mante S. Nieuwland and Jos J. A. Van Berkum. When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18(7):1098–1111, 2006.
- [17] Martin Paczynski and Gina Kuperberg. Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language*, 67(4):426–448, 2012.
- [18] Dianne E. Thornhill and Cyma Van Petten. Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, 83(3):382 – 392, 2012.
- [19] Jan van Eijck. On the proper treatment of context in nl. In Paola Monachesi, editor, *Computational Linguistics in the Netherlands 1999. Selected Papers from the Tenth CLIN Meeting*. UILOTS Utrecht, 1999.
- [20] Noortje J. Venhuizen, Matthew W. Crocker, and Harm Brouwer. Expectation-based comprehension: Modeling the interaction of world knowledge and linguistic experience. *Discourse Processes*, 0(0):1–27, 2018.
- [21] Markus Werning and Erica Cosentino. The interaction of bayesian pragmatics and lexical semantics in linguistic interpretation: Using event-related potentials to investigate hearers’ probabilistic predictions. In *Proceedings of CogSci 39*, pages 3504–09, 2017.
- [22] Markus Werning, Matthias Unterhuber, and Gregor Wiedemann. Bayesian pragmatics provides the best quantitative model of context effects on word meaning in EEG and cloze data. In *Proceedings of CogSci 41*, pages 3085–91, 2019.
- [23] Rolf Zwaan and Gabriel A. Radvansky. Situation models in language comprehension and memory. *Psychological Bulletin*, 123:162 – 185, 1998.

# On Supererogation: One Should Go When Going Is Good Enough and Not Going Is Not\*

David Y. Oshima

Nagoya University, Nagoya, Japan  
davidyo@nagoya-u.jp

## Abstract

Major existing approaches to prioritizing modality (which subsumes deontic and bouletic modality) suffer from the problem of supererogation ('giving too much'), validating the inference from the premises (i) that  $p$  should be the case and (ii) that ' $p$  and  $q$ ' is better than ' $p$  (and not  $q$ )' to the conclusion that  $q$  should be the case. I propose a novel account of *should* and related prioritizing-modal expressions, under which *should*( $p$ ) essentially means that making  $p$  happen is, or constitutes part of, the sole way (among the contextually prominent alternative options) to make the situation 'good enough'.

## 1 Introduction

The two major existing approaches to *prioritizing modality* (which subsumes deontic and bouletic modality), summarized in (1), suffer from the problem of *supererogation* ('giving too much'; Chisholm 1963).

- (1) i. *should*( $p$ ) is true if and only if  $p$  holds true in all highest-ranked (best) worlds within the modal base, where the ranking of worlds is based on how well they conform to the relevant set of rules, goals, etc. (e.g., Kratzer 2012).
- ii. *should*( $p$ ) is true if and only if the (expected) goodness of  $p$  significantly exceeds that of  $\bigcup(\mathbf{ALT}(p))$ , where goodness of propositions is determined by the relevant set of rules, goals, etc. and  $\mathbf{ALT}(p)$  represents the set of the contextually prominent alternatives of  $p$  (e.g., Lassiter 2011).

These accounts support the invalid inference shown in (2), and fail to predict the consistency of (3), adapted from Lassiter (2017:197), and that of (4).

- (2) (premise 1)  $p$  should be the case.  
(premise 2) ' $p$  and  $q$ ' is better than ' $p$  (and  $\neg q$ )'.  
(conclusion) Therefore, ' $p$  and  $q$ ' should be the case.
- (3) a. (On moral grounds,) Adam should visit his friend Bob, who has been ill.  
b. Visiting *and* cooking dinner is morally better than just visiting.  
c. However, cooking dinner is optional; it is not the case that (on moral grounds) Adam should visit and cook dinner.
- (4) a. According to my doctor, I should not have more than one cup of coffee in the morning.  
b. He says it is best for my health if I avoid having any coffee.  
c. However, he also says that having just one cup won't hurt too much; it is not the case that I should not have a cup of coffee in the morning.

---

\*Thanks to Daniel Lassiter, Christopher Tancredi, and Teruyuki Mizuno for helpful discussion and comments.

One might argue that the inference in (2) is actually valid (and thus supererogation is not a real issue), and the apparent consistency of (3)/(4) is to be attributed to implicit shifting of the ordering source; in the situation of (3), for example, one might suppose that the interpreter is inclined to accept the truth of (3c) with an unspoken qualifier along the lines of ‘If we take Adam’s personal desires into consideration in addition to what morality demands’. However, if such shifting is readily possible, then the interpreter would be at least as strongly inclined to refute the truth of (3b) as to accept the truth of (3c). If the interpreter adopts the same ordering source in (3a) and (3b), it is natural to expect him to maintain it when he evaluates (3c).

## 2 Proposal

I propose that *should*( $p$ ) essentially means that making  $p$  hold true is, or constitutes part of, the sole way (among the contextually prominent alternative options) to make the situation ‘good enough’. In other words, *should*( $p$ ) means that if  $p$  is (were) not the case, the situation cannot (could not) be good enough, and if  $p$  is (were) the case, the situation can (could) be good enough. Formally:

- (5) *should*( $p$ ) holds true with respect to modal base  $f$  and ordering source  $g$  if and only if:
- a. for any  $f'$  such that  $f' \in \mathbf{F}_{\cup^*}(f, \{p\})$ , (i)  $\theta_g \prec \mathbf{V}_g(f')$ , and (ii) for any  $q$  such that  $q \in \mathbf{ALT}(p)$  and  $q$  does not entail  $p$ , there is no  $f''$  such that  $f'' \in \mathbf{F}_{\cup^*}(f, \{q\})$  and  $\theta_g \prec \mathbf{V}_g(f'')$ , or;
  - b. there is some  $r$  such that (i)  $r$  contextually entails  $p$  and (ii) *should*( $r$ ) holds true.

where

- (6) a.  $\mathbf{F}_{\cup^*}(P, Q)$  is a set of sets of propositions, such that each member of  $\mathbf{F}_{\cup^*}(P, Q)$  is the union of (i)  $Q$  and (ii) some  $P'$  such that  $P'$  is maximally similar to (and potentially equal to)  $P$  but is compatible with  $Q$ .
- b.  $\mathbf{V}_g$  is a function that assigns goodness values in view of ordering source  $g$  to situations (i.e. sets of propositions).
- c.  $\theta_g$  represents the context-sensitive acceptability threshold value relative to  $g$ , such that ‘ $\theta_g \prec \mathbf{V}_g(P)$ ’ means that situation  $P$  is ‘good enough’ in view of ordering source  $g$ .

In intuitive terms,  $\mathbf{F}_{\cup^*}(P, Q)$  can be understood to be a (possibly singleton) set of realities alternative to  $P$  that are compatible with  $Q$ ; for example:

- (7) a.  $\mathbf{F}_{\cup^*}(\{p\}, \{q\}) = \mathbf{F}_{\cup^*}(\{p, q\}, \{q\}) = \mathbf{F}_{\cup^*}(\{p, \neg q\}, \{q\}) = \{\{p, q\}\}$   
 b.  $\mathbf{F}_{\cup^*}(\{p, q\}, \{\neg[p \wedge q]\}) = \{\{p, \neg q\}, \{\neg p, q\}\}$

The modal base is typically the common ground (CG), and the ordering source is the relevant set of rules, goals, etc.

### 3 Illustration

#### 3.1 The ‘sick friend’ scenario

I will now illustrate how the proposed analysis circumvents the supererogation problem with the scenario in (3). Let  $\mathbf{v}$  and  $\mathbf{d}$  represent ‘Adam visits Bob’ and ‘Adam cooks dinner for Bob’ respectively ( $\mathbf{v}$  is contextually entailed by  $\mathbf{d}$ , but this is not a crucial factor here). I take the contextually prominent alternatives of  $\mathbf{v}$ ,  $\mathbf{ALT}(\mathbf{v})$ , to be  $\{\mathbf{v}, \mathbf{d}, \mathbf{v} \wedge \mathbf{d}\}$ , and further that  $\mathbf{ALT}(\mathbf{v}) = \mathbf{ALT}(\mathbf{d}) = \mathbf{ALT}(\mathbf{v} \wedge \mathbf{d}) (= \{\mathbf{v}, \mathbf{d}, \mathbf{v} \wedge \mathbf{d}\})$ . Intuitively, (i) situations (8a,b) are good enough, while (8c) is not, and (ii) (8a) is better than (8b) and can be said to be ideal and ‘supererogatory’.

- (8) a. {‘Adam and Bob are friends’, ‘Bob has been sick’,  $\mathbf{v}, \mathbf{v} \wedge \mathbf{d}, \dots$ } (ideal/supererogatory)  
 b. {‘Adam and Bob are friends’, ‘Bob has been sick’,  $\mathbf{v}, \mathbf{v} \wedge \mathbf{d}, \dots$ } (good enough)  
 c. {‘Adam and Bob are friends’, ‘Bob has been sick’,  $\mathbf{v}, \mathbf{v} \wedge \mathbf{d}, \dots$ } (not good enough)

The addition of  $\mathbf{v}$  to the CG guarantees that the situation will be good enough, and  $\mathbf{v}$  does not have any alternative that (i) does not entail it and (ii) makes the situation good enough.  $\mathbf{v} \wedge \mathbf{d}$  too will make the situation good enough, but it has an alternative, namely  $\mathbf{v}$ , that (i) does not entail it and (ii) makes the situation good enough. Thus, the proposed analysis rightly predicts the truth of *should*( $\mathbf{v}$ ) and the falsity of *should*( $\mathbf{v} \wedge \mathbf{d}$ ). Furthermore, by (5b), it is guaranteed that ‘Adam should do something for Bob’, ‘Adam should go out’, etc., whose prejacent-propositions are contextually entailed by  $\mathbf{v}$ , will be true.

#### 3.2 The ‘coffee’ scenario

Consider next the scenario in (4). (10) illustrates situations that are ‘more than good’, ‘just good enough’, and ‘not good enough’ in relation to how many cups of coffee the speaker of (4) has, with  $\neg\mathbf{1}$  standing for ‘I do not drink one (or more) cup(s) of coffee’ and  $\neg\mathbf{2}$  standing for ‘I do not drink two (or more) cups of coffee’.

- (9) a. {‘I have such-and-such health conditions’,  $\neg\mathbf{2}, \neg\mathbf{1}, \dots$ } (ideal/supererogatory)  
 b. {‘I have such-and-such health conditions’,  $\neg\mathbf{2}, \neg\mathbf{1}, \dots$ } (good enough)  
 c. {‘I have such-and-such health conditions’,  $\neg\mathbf{2}, \neg\mathbf{1}, \dots$ } (not good enough)

The proposed analysis rightly predicts that *should*( $\neg\mathbf{1}$ ) does not hold true,  $\neg\mathbf{2}$  being an alternative that does not entail  $\neg\mathbf{1}$  and makes the situation good enough.

## 4 Varieties of prioritizing modals

This section addresses some modal expressions other than *should*.

### 4.1 Strong vs. weak necessity

Expressions of obligation come in different strengths. For example, *must* is said to convey a stronger obligation than *should* and *ought* do (e.g. Portner and Rubinstein 2016); to illustrate, (10a–c) could all be true in the same discourse context where the interlocutors are interested in the academic standing of the addressee.

- (10) a. You must turn in at least one paper. (Otherwise, you will fail the class.)

- b. It is not the case that you must turn in two papers.
- c. You should turn in two papers. (Otherwise, your grade will be a B– or worse.)

Under the proposed analysis, the difference in strength between *must* and *should* can be attributed to the difference in threshold values associated with them. I suggest that the threshold value for *should*,  $\theta$ , is set higher than that for *must*,  $\theta'$  (i.e.  $\theta' \prec \theta$ ), in a way similar to how the threshold values for *small* and *tiny* (or *very small*) differ. The goodness of a situation where the addressee receives, say, a C+, may exceed  $\theta'$  while not exceeding  $\theta$ .

## 4.2 Permission

The proposed analysis of prioritizing necessity (obligation) is fully compatible with the received wisdom that obligation and permission are duals of each other, so that the following approximation holds:

$$(11) \quad \textit{may}(p) \approx \neg \textit{must}(\neg p).$$

One caveat here is that it is not clear whether the threshold value associated with *may* is equivalent to, and has exactly the same degree and range of context-sensitive variability as, the value associated with *must* or any other particular expression of obligation. This is an empirical issue that calls for investigations.

## 5 Manifestation of conditional reasoning in Japanese and Korean modal constructions

In Japanese and Korean, prioritizing-modal statements are often and systematically expressed with idiomatic constructions—referred to as ‘conditional evaluative constructions’ by Kaufmann (2018)—along the lines of ‘If  $p$  is not the case, it will not be good’ (for obligations) or ‘Even if  $p$  is the case, it will be good’ (for permissions).<sup>1</sup>

- (12) (Japanese)
- a. Ken wa tomato o tabenakereba ikenai.  
K. Top tomato Acc eat.Neg.Prov go.Pot.Neg.Prs  
‘Ken should eat tomato.’ (lit. ‘One cannot go if Ken doesn’t eat tomato.’)
  - b. Mari wa kaette mo yoi.  
M. Top return.Ger also good.Prs  
‘Mari may go home.’ (lit. ‘It will be good even if Mari goes home.’)

It is a matter of debate whether these constructions count as genuine, semantically transparent conditionals (Kaufmann 2018; Chung 2019); it seems fair to say, at any event, that they provide fairly strong evidence that conditional reasoning is a key component of prioritizing modality.

## 6 Priority under uncertainty

An adequate theory of prioritizing modality needs to account for the readings available to (14) under the premises described in (13) (Kolodny and MacFarlane 2010, Cariani *et al.* 2013).

<sup>1</sup>The abbreviations in glosses are: Acc = accusative, Ger = gerund, Top = topic-marker, Neg = negation, Pot = potential, Prov = provisional, Prs = present.



- (13) Ten miners are trapped either in shaft A or in shaft B, but we do not know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block shaft A or shaft B, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed.
- (14) We should block neither shaft.

Many judge (14) as TRUE on its prominent reading, called the *subjective* reading; note that this does not readily follow from the Kratzerian analysis described in (1-i), as in the ideal worlds the right shaft is blocked and all miners will be saved. (14) has a second, less prominent reading (the *objective* reading), on which it is judged by many as FALSE ('No, we should block the right shaft—whichever it is.'). We take the inferences leading to these judgements to involve calculation of the *expected utility*—the expected number of miners saved. We suggest that, we, more or less rational agents, tend to adopt an ordering source that dictates the maximization of the expected utility, so that 'One maximizes the expected number of people saved' is, or is entailed by, a top-priority condition (call it  $c_1$ ) in  $g$ . Note that, unlike Lassiter (2011) and Chung (*forthcoming*), I do *not* take expected utility to be a hardwired component of the semantics of *should*. Rather, it is an extra-linguistic factor affecting what kind of goals we try to fulfill and what kind norms we try to abide by.

When the modal base is equivalent to the CG (as typically is the case), only the option of blocking neither shaft leads to the satisfaction of  $c_1$ , saving nine miners (and losing one) instead of incurring a 50% chance of losing 10. We suggest that the modal base is allowed to be indeterminate to some extent, and could contain some relevant facts unknown to the interlocutors. When the modal base is taken to contain either 'the miners are in Shaft A' or 'the miners are in Shaft B' (although the interlocutors do not know which), the objective reading comes about, on which the only option to satisfy  $c_1$  is blocking A or blocking B (although, again, the interlocutors do not know which).

Cariani *et al.* (2013) remark that the conditional statement in (15) too has two readings, called *reflecting* and *non-reflecting*.

- (15) If the miners are in shaft A, we ought to block shaft A.

In the described situation, (15) is TRUE on its typical, reflecting reading, while it is judged (by many) as FALSE on its non-reflecting reading, which is marginal but could be elicited as in the following.

- (16) If the miners are in shaft A, we (still) ought to block neither shaft, for their being in shaft A doesn't mean that we know where they are. Indeed, no matter where the miners are, we ought to block neither shaft. (Cariani *et al.* 2013: 227)

I suggest that the reflecting reading is brought about by (temporarily) adding 'the miners are in Shaft A' to the modal base as well as to the CG, whereas the non-reflecting reading is brought about by (temporarily) adding 'the miners are in Shaft A' only to the CG and not to the modal base.

## 7 *Or* under prioritizing modality

When embedded under *should*, *may*, etc. (on their prioritizing interpretation), the conjunction *or* allows and typically receives a 'non-disjunctive' interpretation.



## 7.1 *Or* under *should*

Generally, *should*(*p*) entails *should*(*q*) if *p* entails *q*.

- (17) a. Adam made chicken soup for Bob.  
       *entails*: ‘Adam made something for Bob.’  
 b. Adam should have made chicken soup for Bob.  
       *entails*: ‘Adam should have made something for Bob.’

This pattern does not hold, however, when *q* involves *or*. (18b) allows two readings, and on the prominent one, it does not entail that Adam should have made chicken soup or a burrito for Bob.

- (18) a. Adam made chicken soup for Bob.  
       *entails*: ‘Adam made chicken soup or a burrito for Bob.’  
 b. Adam should have made chicken soup for Bob.

I propose that *should*(*[p or q]*) is typically interpreted as: ‘If *p* does not hold, then *should*(*q*) holds, and if *q* does not hold, then *should*(*p*) holds’ ( $[\neg p \rightarrow \textit{should}(q)] \wedge [\neg q \rightarrow \textit{should}(p)]$ ), although it allows the straightforward disjunctive interpretation as well. The two readings of (19) can be approximated as in (20).

- (19) John should take the subway or the bus.  
 (20) **Reading #1** (typical): ‘If John does not take the subway, then he should take the bus, and if he does not take the bus, then he should take the subway.’  
**Reading #2** (less common): ‘Either John should take the subway, or he should take the bus.’

It is not clear to me how the more prominent ‘conditional obligation’ interpretation comes about. It is interesting to note, however, that similar patterns are found in two other places that involve *or*. One is the ‘Imperative *or* Declarative’ construction, exemplified in (21).

- (21) Take the subway, or you will be late.  
       *entails*: ‘If you do not take the subway, then you will be late; if you are not late, then you will have taken the subway.’

This construction shares the semantic scheme: ‘If the prejacent of one disjunct does not hold, then some component of the other disjunct holds’ with the *or*-under-*should* configuration.

Another is the projection of presuppositions in statements with *or*. According to Karttunen (1974), ‘either *p* or *q*’ presupposes that if *p* does not hold then the presupposition of *q* has to be common ground, and if *q* does not hold then the presupposition of *p* has to be common ground.

- (22) Either *p* or *q*.  
       *presupposes*:  $[\neg p \rightarrow \mathbf{ps}(q)] \wedge [\neg q \rightarrow \mathbf{ps}(p)]$   
 (23) Either Mary doesn’t come, or her husband will come too.  
       *presupposes*: ‘If Mary comes, then somebody other than Mary’s husband will come.’  
       *does not presuppose*: ‘Somebody other than Mary’s husband will come.’

## 7.2 *Or* under *may*

A statement of the form:  $may([p \text{ or } q])$  too allows two readings, again the straightforward disjunctive reading being less typical. In many cases, the typical reading appears to allow paraphrasing with *and*.

- (24) John may take the bus or (take) the subway.  
 $\approx$  John may take the bus, and he may take the subway.

I suggest that the two interpretations of  $may([p \text{ or } q])$  are entirely parallel to those of  $should([p \text{ or } q])$ , the typical one being: ‘If  $p$  does not hold, then  $may(q)$  holds, and if  $q$  does not hold, then  $may(p)$  holds’ ( $[\neg p \rightarrow may(q)] \wedge [\neg q \rightarrow may(p)]$ ). In the case of (25), the two readings are as in (26).

- (25) Mary may play video games or read comic books.  
 (26) **Reading #1** (typical): ‘If Mary does not play video games, then she may read comic books, and if she does not read comic books, then she may play video games.’  
**Reading #2** (less common): ‘Either Mary may play video games, or she may read comic books.’

The data set in (27) endorses the supposition that *or* under *may* does not simply mean logical conjunction; rather, the typical interpretation of  $may([p \text{ or } q])$  involves ‘conditional permission’.

- (27) a. Mary may play video games and read comic books. #But she is not allowed to do both.  
 b. Mary may play video games or read comic books. But she is not allowed to do both.

Note that the ‘conditional permission’ interpretation leaves unspecified whether  $may([p \wedge q])$  holds true or not, so that (25) is also compatible with the continuation: ‘And she is allowed to do both (too)’.

## 8 Summary

I proposed a novel analysis of *should*, where  $should(p)$  means that if  $p$  is (were) not the case, the situation cannot (could not) be good enough, and if  $p$  is (were) the case, the situation can (could) be good enough. The proposed account circumvents the issue of supererogation, as well as that of the miners’ paradox, which pose challenges for the major alternative accounts. The relation between *should* and *must*, *may*, etc. can be accounted for under the proposed account at least as straightforwardly as under the alternative accounts. The proposed account is also resonant with the fact that in some languages prioritizing modality is systematically expressed with constructions that look much like conditionals. I also discussed that the conjunction *or* embedded under a prioritizing modal induces a special interpretation, which may be dubbed as the conditional obligation/permission interpretation.

## References

- [1] Fabrizio Cariani, Magdalena Kaufmann, and Stefan Kaufmann. Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy*, 36:225–259, 2013.

- [2] Roderick Chisholm. Supererogation and offence: A conceptual scheme for ethics. *Ratio*, 5:1–14, 1963.
- [3] WooJin Chung. Decomposing deontic modality: Evidence from Korean. *Journal of Semantics*, 36, 2019. doi:10.1083/jos/ffz016.
- [4] WooJin Chung. Expected value of conditionals and expected utility: A probabilistic account of conditional evaluative constructions. In *Proceedings of Sinn und Bedeutung 24*. forthcoming.
- [5] Lauri Karttunen. Presupposition and linguistic context. *Theoretical Linguistics*, 1:181–194, 1974.
- [6] Magnalena Kaufmann. What ‘may’ and ‘must’ may be in Japanese. In Kenshi Funakoshi, Shigeto Kawahara, and Chris Tancredi, editors, *Japanese/Korean Linguistics*, volume 24, pages 103–126. CSLI Publications, 2018.
- [7] Nico Kolodny and John MacFarlane. Ifs and oughts. *Journal of Philosophy*, 107:115–143, 2010.
- [8] Angelika Kratzer. *Modality and conditionals: New and revised perspectives*. Oxford University Press, 2012.
- [9] Daniel Lassiter. *Measurement and modality: The scalar basis of modal semantics*. PhD thesis, New York University, 2011.
- [10] Daniel Lassiter. *Graded modality: Qualitative and quantitative perspectives*. Oxford University Press, 2017.
- [11] Paul Portner and Aynat Rubinstein. Extreme and non-extreme deontic modals. In Nate Charlow and Matthew Chrisman, editors, *Deontic Modals*, pages 256–282. Oxford University Press, 2016.

# Speaker discourse roles and the discourse profile of reportative evidentials\*

Roumyana Pancheva and Deniz Rudin

University of Southern California  
{pancheva, drudin}@usc.edu

## Abstract

Reportative evidentials introduce an asymmetry between what their utterance makes at issue and what their utterance commits the speaker to. We propose that this is due to reportative evidentials enforcing a contextual requirement of nonidentity between the speaker and the agent whose commitment is being described. A general mechanism for assertive update that is sensitive to this distinction delivers the exceptionality of reportative updates while preserving a uniform characterization of assertion. We then discuss ramifications for the typology of evidentials, and sketch an extension to the use of reportative evidentials in questions.

## 1 Introduction

A speaker who utters a declarative sentence with an evidential marker, EVID- $p$ , expresses commitment to the prejacent proposition  $p$  in case the evidential is direct, (1a), or inferential (1b), but not when the evidential is reportative (1c).

(1) Cuzco Quechua (Quechua), Faller (2002): ex. (2))

a. <i>Para-sha-n-mi</i>	b. <i>Para-sha-n-chá</i>	c. <i>Para-sha-n-si</i>
rain-PRG-3-DIRECT	rain-PRG-3-INFERENTIAL	rain-PRG-3-REPORTATIVE
‘It is raining (I see).’	‘It is raining (I infer).’	‘It is raining (I was told).’

The emerging consensus among dynamic accounts of reportative evidentials (Murray 2014, AnderBois 2014, 2019a, Faller 2019) is that they introduce an asymmetry between what their utterance makes at issue (the prejacent  $p$ ) and what their utterance commits the speaker to (the existence of a report with content  $p$ ). These theories allow evidentials to rewrite basic update operations like speaker commitment. A worry is that this undercuts recent advances in our understanding of the relationship between clause type and context update (e.g. Gunlogson 2001, Farkas & Roelofsen 2017), which pursue an explanatory link between the semantic type and illocutionary potential of declarative sentences. This paper explores a way to connect the discourse profile of reportative evidentials to the morphosemantics of reportative evidentials, while maintaining the explanatory power of a general account of the context update carried out by assertive utterances of declarative sentences.

Building on the insight that reportative evidentials involve perspective-shifting (AnderBois 2014) that relies on a fine-grained decomposition of discourse roles (Faller 2019), we propose that reportative evidentials introduce a presupposition that enforces distinctions between the referents of discourse roles in context. We formalize a basic update operation for assertive uses of declarative sentences that delivers standard assertive updates for non-reportative sentences,

---

\*For helpful comments, thanks to the anonymous reviewers for this and another conference.

but updates similar to those proposed on the dynamic accounts above for reportative sentences, via sensitivity to the discourse roles involved in the presupposition of reportative evidentials. The resulting account links the discourse behavior of reportative assertions to the meaning of reportative evidentials without taking reportative evidentials to be special update operators, resulting in a more constrained, predictive account of the space of possible assertive updates.

## 2 The morphosemantics of reportative evidentials

Pronouns introduce presuppositions about discourse participants, evaluated relative to the context parameter.<sup>1</sup> For example, first person pronouns presuppose that their referent is the speaker of the context. But ‘speaker’ is a cover term for distinct discourse roles which are often coextensional (Goffman 1979, Sells 1987, a.o.), e.g. the animator (*a*, the person who is actually talking) and the principal (*pr*, the person whose commitment is being described, or, in Goffman’s terms, ‘the party to whose position the words attest’).<sup>2</sup> We propose that these two roles are represented independently in the context parameter, though coextensional by default (in addition to the addressee *ad*; see 2).

(2) Example context (preliminary): [*a* = MARY, *pr* = MARY, *ad* = BETH]<sup>c</sup>

These presuppositions are associated with person features—i.e., a pronoun bearing the feature [+ADDRESSEE] (or, rather, the combination of features [−SPEAKER, +PARTICIPANT]—q.v. Halle 1997, Nevins 2007 or alternative person feature representations as in Harley & Ritter 2002) presupposes that the context maps the variable it introduces to the addressee. We assume, with Pancheva & Zubizarreta (2019), that evidentiality is partially encoded with person features, and that different types of evidentiality (e.g. direct, inferential, reportative) can be captured with different combinations of person features.

We propose here that reportative evidentials are associated with particular values of the person features representing the discourse roles of *animator* and *principal*, thus decomposing [+SPEAKER] so it is no longer a unitary person feature. Specifically, we propose that reportative evidentials are associated with the person features [+ANIMATOR, −PRINCIPAL], serving to introduce a presupposition that the context contains an entity that is equivalent to *a* and not equivalent to *pr* (3).

(3) PRELIMINARY SEMANTICS<sup>3</sup>  
 a.  $\llbracket \text{REP} \rrbracket^c$  is defined iff  $a_c \neq pr_c$       b. If defined,  $\llbracket \text{REP} \rrbracket^c$  is an identity function

A reportative simply presupposes that the animator is not describing their own commitment with the prejacent. Note that, on our proposal, this is the entire contribution of the reportative evidential. The distinctive illocutionary profile of reportative assertions is not directly encoded

<sup>1</sup>We follow convention in describing this aspect of the meaning of pronouns as presuppositional. What matters for our purposes is only that pronouns enforce such requirements on the context one way or another. Throughout, what we describe as presuppositions could just as easily be implemented as some other contextual requirement.

<sup>2</sup>Note that though we borrow terminology from Goffman, his work is focused on sociologically complex contexts, such as speeches given by proxy. We redeploy these terms to describe simpler cases in which a speaker is attributing a commitment to another agent. We suspect that a significant amount of pretense is involved in the more convoluted cases, and we do not intend our proposal to extend to them. In particular, we do not make the prediction that a reportative evidential would be used in the special contexts that Goffman considers. This will become even clearer when we update the presupposition of the reportative evidential in §6 to not directly invoke the animator role.

<sup>3</sup>We will revise this semantics in §6 in a way that is irrelevant to its predictions for declarative sentences.

in the meaning of the reportative evidential, but rather falls out of how the contextual requirement enforced by the evidential interacts with a general mechanism for assertive context update, which we turn to now.

### 3 The theory of assertion

#### 3.1 A general mechanism for assertion

We assume a central role for discourse commitments ( $DC$ ) in context update (Hamblin 1971, Gunlogson 2001). We assume that all agents  $x$  are associated with a set of public discourse commitments  $DC_x$ , the set of all propositions that  $x$  has made a public discourse commitment to. Following Faller (2019) but diverging in the implementational details, we propose that when  $a$  utters a declarative sentence (modulo intonation; Pierrehumbert & Hirschberg 1990, Bartels 1999, Gunlogson 2001, Westera 2017, Rudin 2018 a.o.), she commits to the proposition that the principal is (discourse-)committed to the prejacent  $p$  (a proposition we notate  $\text{COMM}(pr, p)$ ), and pushes  $p$  onto the Table stack, making it at issue (Farkas & Bruce 2010). An assertion by  $a$  of a sentence denoting  $p$  is a function from an input context  $k_i$  to an output context  $k_o$ :

- (4)  $\text{ASSERT}(a, p, k_i) \rightarrow k_o$  such that
- $DC_{a,o} = DC_{a,i} + \text{COMM}(pr, p)$
  - $\text{TABLE}_o = \text{TABLE}_i +_{\text{push}} p$
  - In all other respects,  $k_o = k_i$

This account predicts an asymmetry in the assertive profile of reportative and non-reportative sentences. In the default case, in which  $a = pr$ , this simply entails that  $a$  is committed to  $p$  itself ( $\text{COMM}(a, \text{COMM}(a, p)) \leftrightarrow \text{COMM}(a, p)$ , q.v. Condoravdi & Lauer 2017 ex. 20), as in standard accounts of assertion. However, in the case of reportative assertions,  $a \neq pr$  via accommodation of the presupposition of the reportative, so the speaker’s commitment to  $\text{COMM}(pr, p)$  does not entail a commitment to  $p$  itself, unlike in standard assertions.

- (5) Output context following a reportative assertion of  $p$ :

$DC_a$	Table	$DC_{ad}$
$\text{COMM}(pr, p)$	$\{p\}$	

$pr \neq a$

The speaker makes no indication of who  $pr$  is, other than that it’s not them. In some contexts, it is clear who  $pr$  is, as when the reportative assertion is prefaced by *I was talking to Susie on the phone, and...*, or when the speaker adds *Susie told me so* after the reportative assertion, a pattern Aikhenvald (2004) calls ‘lexical reinforcement’. But out of the blue, the identity of  $pr$  remains vague. If the context suggests that  $a$  trusts  $pr$ , we can conclude that  $a$  is suggesting that the issue they’ve raised should be resolved by making  $p$  Common Ground. If the context suggests otherwise, no such suggestion will be inferred. It may also be that appeal is sometimes made to a generic  $pr$  (see §5).

#### 3.2 Three Clarifications

##### 3.2.1 The nature of discourse commitments

What an agent is discourse-committed to and what an agent is taken to believe are not the same thing (a distinction made quite clearly by Grice 1975 and Stalnaker 1978). We highlight two

differences. First, a discourse commitment is incurred only by a conversational move, whereas we can by many non-conversational means infer what an agent believes. This means that if somebody is committed to  $p$ , it must be that they made an utterance. Second, we can watch somebody lie in defiance of manifest evidence to the contrary, and they have committed to the lie in spite of our firm evidence that it contravenes their own beliefs. So somebody being committed to  $p$  does not mean that they are taken to actually believe it.

### 3.2.2 Committing to the existence of a commitment

We do not claim that a reportative assertion adds anything to any commitment set other than the animator's. We assume that no agent can add propositions to any commitment set other than their own. In the case of a reportative assertion, the speaker commits to the proposition that the principal is committed to  $p$ ; if the principal is not in fact committed to  $p$ , then the animator has simply spoken falsely, not added anything to the principal's commitments.

### 3.2.3 Commitments outside of the current conversation

We assume that agents keep track of representations of the discourse commitments of other agents outside of the scope of individual conversations. This is common-sensical: we keep track of what somebody said to us yesterday, and are surprised if they contradict it today; likewise, we notice when a description of a situation by one agent differs from how a different agent described it in a different conversation. So if an agent hears an assertion that  $p$  from a party they are not currently engaged in conversation with—say, in a radio interview they are listening to—they nonetheless add  $p$  to their representation of the commitment state of the speaker.

If an agent's commitment to some principal being committed to  $p$  is sincere, that means that in their current representation of the principal's commitments, the principal is committed to  $p$ . As of the last time they had reason to update their record of the principal's commitments,  $p$  was a member of that set—i.e., last they heard the principal speak (or heard a credible report of what they'd said), the principal was committed to  $p$ —and they have no particular reason to believe that the principal has rescinded that commitment since.

## 4 Predictions

### 4.1 The existence of a prior communicative event

In making a reportative assertion,  $a$  makes a discourse commitment to  $\text{COMM}(pr, p)$ . As discourse commitments can only be incurred by communicative acts (§3.2.1), if  $pr \neq a$ , it can only be true that  $pr$  is committed to  $p$  if there is a prior communicative event,  $u_0$ , whose animator  $a_0$  was  $pr$ . Thus, it is predicted that reportative assertions  $\text{REP-}p$  are felicitously used only if there is a prior communicative event whose speaker made a discourse commitment to  $p$ .

### 4.2 The principal is not the addressee

Interpretations of reportatives in which  $ad = pr$  are not available, but our account permits them in principle. We suggest that such interpretations are indeed compatible with the semantics of reportative evidentials, but are ruled out by the same pragmatic pressures that rule out interpretations of attitude reports like 'I hear that  $p$ ' where  $p$  is a discourse commitment of the addressee. Under normal circumstances, it is uncooperative to make utterances that simply communicate what the addressee is committed to, as this information is redundant.

### 4.3 Speaker bias toward $p$

In making a reportative assertion,  $a$  places  $p$  on the Table while committing to  $\text{COMM}(pr, p)$ . This may or may not suggest her epistemic bias toward  $p$ , depending on the context. If the report is taken to be credible, then positive bias toward  $p$  is predicted. If the report is not taken to be credible, then negative bias is predicted. We assume, contra Faller (2019), that Walker’s (1996) Collaborative Principle, which obligates interlocutors to voice disagreement with others’ commitments immediately if they do not share them, only applies to commitments made by the current interlocutors. We predict that neither  $a$  nor  $ad$  is automatically taken to share  $pr$ ’s commitment to  $p$  following the utterance of a reportative sentence.

## 5 Typology of reportative evidentials

Evidential systems with a single, general reportative evidential are well represented in existing studies of evidentiality. Our proposal concerns the meaning of such general reportative markers. Beyond the claim that the (current) animator is not the principal, nothing more is said about the identity of the principal (and original animator): it could be a particular individual or it could be generic. Suitable paraphrases of this distinction are ‘he/she says’, ‘I am told’ vs. ‘one says’, ‘it is said’. An example of a general reportative evidential, explicitly acknowledging the vagueness with respect to the source of the report, is given in (6).

- (6) Tagalog (Austronesian) AnderBois (2017): ex. (5)

Scenario: We are discussing our opinions about a local politician.

*Matapat daw siya.*  
honest REPORTATIVE DIR.3SG

‘He’s honest, I heard.’ or ‘He’s honest, they say.’

Less commonly, some languages have more than one reportative evidential. Several subtypes of reportatives have been proposed in the typological literature, aiming to capture the more specialized meanings: *second-hand*, *third-hand*, and *folklore* (Willett 1988, Aikhenvald 2004).<sup>4</sup> Unfortunately, there is much terminological confusion, with *second-hand* sometimes used as a synonym for a general reportative. But when meaning distinctions are made, *second-hand* is used for reports ‘from someone who said it’, while *third-hand* is used for reports ‘from someone else who in their turn acquired the information through another [...] report’ (Aikhenvald 2004: 395). This is the same distinction that direct evidentials and reportatives make with respect to the described event, see (1a) vs. (1c), but here the distinction is applied to the communicative event itself: the speaker of an utterance REP- $p$  has directly witnessed an utterance with the content of  $p$  (second-hand) vs. the content of REP- $p$  (third-hand). Tsafiki (Barbacoan) is said to distinguish third-hand from second-hand reportative evidence through the repetition of the reportative marker (Aikhenvald 2004: 54, 179). *Folklore* evidentials are used in the genre of culturally-shared narratives (e.g., folk tales, myths) (Aikhenvald 2004).

<sup>4</sup>Quotative markers used for direct quotation are sometimes grouped with reportative evidentials. Putting aside cases where the term *quotative* is simply used as an alternative for *reported* (also called *hearsay*) or *second-hand* (see Aikhenvald 2004 for examples and discussion of both the use of terminology and the need to differentiate among the terms), quotative markers proper have different properties from reportative evidentials (see AnderBois 2019a,b for a theoretically-based comparison of reportative and quotative markers in Yucatec Maya and several other Mayan languages).



Clearly, more research is needed to establish the robustness of the reportative sub-categories cross-linguistically.<sup>5</sup> Nevertheless, of relevance here is the fact that our account is sensitive to these distinctions. A communicative event  $u_n$ , in which an animator  $a_n$  utters REP- $p$ , commits  $a_n$  to the proposition that COMM( $pr, p$ ), and thus to the existence of a prior communicative event,  $u_0$ , whose animator  $a_0$  was  $pr$ , i.e., COMM( $a_0, p$ ).  $a_n$  (the reportative evidence holder for  $p$  in  $u_n$ ) may have directly witnessed  $u_0$ , as an addressee (i.e., a participant in  $u_0$ ), or just a recipient (i.e., an observing bystander), in which case the reportative evidence is *second-hand*. Alternatively,  $a_n$  may have been a participant/recipient in a different communicative event,  $u_i$ , whose animator  $a_i$  used a reportative evidential herself, i.e., she committed to COMM( $pr, p$ ) but not to COMM( $a_i, p$ ). In this case  $a_n$ 's reportative evidence is *third-hand*. This distinction concerns the potential distinctness of *animator* and *principal*, not only in the current communicative event  $u_n$  but also in the reported one, and thus could be naturally specified in the presuppositions of individual reportative evidentials. In the case of *folklore* evidentials, an appeal could be made to a generic principal.

In addition to the various subtypes of reportatives, a very common typological pattern is underspecification of evidential morphemes, such that indirect evidentials can give rise to both inferential and reportative meanings. Because we do not make specific reference to reportativity in our semantics for reportative evidentials, our proposal allows for this kind of underspecification to be modeled without resorting to simply listing evidence types in the lexical semantics of evidentials.

## 6 Extension to questions

### 6.1 Interrogative flip

In languages where evidentials may appear in questions, the perspectival orientation typically undergoes '*interrogative flip*': the evidence holder switches from speaker to addressee and the evidence concerns the expected answer to the question (see 10a). The phenomenon of interrogative flip suggests that the presupposition of the reportative evidential should be formulated more broadly than we did in §2, and it should be *evidence holder (eh) ≠ principal*, rather than the more specific *animator ≠ principal*, given that the animator role only concerns the speaker.

- (7) Example context (final): [ $a = \text{MARY}, pr = \text{MARY}, eh = \text{MARY}, ad = \text{BETH}$ ]<sup>c</sup>  
 (8) REVISED SEMANTICS  
 a.  $\llbracket \text{REP} \rrbracket^c$  is defined iff  $eh_c \neq pr_c$       b. If defined,  $\llbracket \text{REP} \rrbracket^c$  is an identity function

The discussion in §2 and §3 is not affected because in matrix declaratives *evidence holder = animator* by default, and so it follows that *animator ≠ principal*.<sup>6</sup> We assume, however, that utterances of interrogatives shift the value of *eh* to the addressee.

- (9) Output context following a reportative polar question  $\{p, \neg p\}$ :

$DC_a$	Table	$DC_{ad}$
$\{p, \neg p\}$		
$pr \neq eh; eh = ad$		

<sup>5</sup>In particular, there seem to be no clear examples of *second-hand* and *third-hand* reportatives marked with different morphemes within one and the same language (Aikhenvald 2004: 59).

<sup>6</sup>The more general formulation of the presupposition of reportative evidentials, *evidence holder ≠ principal*, is also needed for evidentials in attitude reports, for the cases where the evidence holder shifts from the speaker to the attitude holder (the matrix subject of the attitude verb).



## 7 Significance

The decomposition of speaker roles within grammatically-accessible representations of discourse participants is motivated independently, as these roles can pull apart in logophora (Sells 1987, Culy 1997, Sundaresan 2012) and free indirect discourse (Doron 1991, Schlenker 2004, Sharvit 2008, Eckardt 2014). The proposal fits reportative evidentials into this broader typology of perspectival constructions.

The update account builds on previous dynamic proposals without sacrificing a uniform conception of declarative updates. We show that asymmetric behavior of reportative and non-reportative sentences can be derived from such a uniform conception, resulting in a more constrained, explanatory theory of the space of possible discourse moves.

Because the proposal does not hard-code the illocution of reportative assertions into the meaning of the reportative evidential, it opens up avenues to explain the behavior of reportative evidentials in questions in terms of the same decoupling of participant roles that is at play in evidential assertions. The account of reportative assertions extends straightforwardly to interrogative flip interpretations of reportative questions.

## References

- Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford University Press.
- AnderBois, Scott. 2014. On the exceptional status of reportative evidentials. In Todd Snider, Sarah D'Antonio & Mia Weigand (eds.), *Proceedings of SALT 24*, 234–254. CLC Publications.
- AnderBois, Scott. 2017. An illocutionary account of reportative evidentials in imperatives. In Dan Burgdorf, Jacob Collard, Sireemas Maspong & Brynhildur Stefánsdóttir (eds.), *Proceedings of SALT 27*, 459–479. CLC Publications.
- AnderBois, Scott. 2019a. At-issueness in direct quotation: The case of Mayan quotatives. SALT 29 abstract.
- AnderBois, Scott. 2019b. Reportatives and quotatives in Mayan languages. In Ryan Bennett, Robert Henderson, Pedro Mateo Pedro & Megan Harvey (eds.), *Proceedings of form and analysis in Mayan linguistics (FAMLi) 5*, 1–16.
- Bartels, Christine. 1999. *The intonation of English statements and questions: A compositional interpretation*. Routledge.
- Bhadra, Diti. 2018. Evidentials are syntax-sensitive: The view from Bangla. *Glossa* 3. 1–39.
- Condoravdi, Cleo & Sven Lauer. 2017. Conditional imperatives and endorsement. In A. Lamont & K. Tetzloff (eds.), *Proceedings of NELS 47*, vol. 1, 185–204.
- Culy, Christopher. 1997. Aspects of logophoric marking. *Linguistics* 32(6). 1055–1094.
- Doron, Edit. 1991. Point of view as a factor of content. In Steven K. Moore & Adam Zachary Wyner (eds.), *Proceedings of SALT 1*, 51–64.
- Eckardt, Regine. 2014. *The semantics of free indirect discourse*. Brill.
- Faller, Martina. 2002. *Semantics and pragmatics of evidentials in Cuzco Quechua*: Stanford dissertation.

- Faller, Martina. 2019. The discourse commitments of illocutionary reportatives. *Semantics and Pragmatics* 12(8).
- Farkas, Donka & Kim Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27(1). 81–118.
- Farkas, Donka & Floris Roelofsen. 2017. Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics* 34(2). 237–289.
- Goffman, Erving. 1979. Footing. *Semiotica* 25(1-2). 1–30.
- Grice, Paul. 1975. Logic and conversation. In P. Cole & J. Morgan (eds.), *Syntax and semantics 3: Speech acts*, 41–58. Cambridge, MA: Academic Press.
- Gunlogson, Christine. 2001. *True to form: Rising and falling declaratives as questions in English*: University of California, Santa Cruz dissertation.
- Halle, Moris. 1997. Impoverishment and fission. In Benjamin Bruening et al. (ed.), *Pf: Papers at the interface*, vol. 30, 425–450. MITWPL.
- Hamblin, C.L. 1971. Mathematical models of dialogue. *Theoria* 37. 130–155.
- Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3). 482–526.
- Murray, Sarah. 2014. Varieties of update. *Semantics and Pragmatics* 7.
- Murray, Sarah. 2017. *Semantics of Evidentials*. Oxford University Press.
- Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory* 25. 273–313.
- Pancheva, Roumyana & Maria Luisa Zubizarreta. 2019. On the role of person features in the evidential-temporal connection. *Canadian Journal of Linguistics* 65. A special volume on Person.
- Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan & M. Pollack (eds.), *Intentions in communication*, 271–311. MIT Press.
- Rudin, Deniz. 2018. *Rising above commitment*: University of California, Santa Cruz dissertation.
- Schlenker, Philippe. 2004. Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind and Language* 19(3). 279–304.
- Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18(3). 445–479.
- Sharvit, Yael. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31. 353–395.
- Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics 9*, 315–32. Academic Press.
- Sundaresan, Sandhya. 2012. *Context and (co-)reference in the syntax and its interfaces*: Tromsø and Stuttgart dissertation.

- Walker, Marilyn. 1996. Inferring acceptance and rejection in dialog by default rules of inference. *Language and Speech* 39(2-3). 265–304.
- Westera, Matthijs. 2017. *Exhaustivity and intonation: A unified theory*: Institute for Logic, Language and Computation, Universiteit van Amsterdam dissertation.
- Willett, Thomas. 1988. A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language* 12. 51–97.

# *Just* exhaustification. A ‘two stage’ theory of exclusives\*

Daniele Panizza<sup>1</sup> and Gennaro Chierchia<sup>2</sup>

<sup>1</sup> University of Goettingen, Goettingen, Germany  
daniele.panizza@gmail.com

<sup>2</sup> Harvard University, Cambridge, MA, U.S.A.  
chierch@fas.harvard.edu

## Abstract

We propose a novel unified semantics for *only/just*, which addresses the following problems: the problem of *only* taking scope over a disjunction (cf. [6]), the problem of rank-order readings ([8], [4]) and the problem of what *only* presupposes (cf. [9], [1]). Our proposal is based on the following assumptions: a) *only* and *just* are subject to a non-vacuity presupposition, b) they only ‘see’ entailment-based alternatives, c) they employ Innocent Exclusion [6], and d) they are not presuppositional but they have however a factive implicature, as proposed by Romoli [12] for soft triggers, i.e. they exhaustify in two stages.

## 1 Three issues in the semantics of *only*

The semantics of exclusives such as *just* or *only* has been the object of intense scrutiny over many years but some recalcitrant problems remain open. According to the standard analysis in 1 (cf. [7], [13], among many others) *only* operates on a proposition (i.e. the prejacent) and a set of contextually defined alternatives. Its contribution to the interpretation of a sentence is such that the prejacent is presupposed, and every alternative that is not entailed by the proposition (i.e. non-weaker) must be false.

$$(1) \text{ Only}_{\text{ALT}}(p) = \lambda w: p(w) = 1. \forall q \in \text{ALT} [ p \not\leq q \rightarrow q(w) = 0 ]$$

The semantics in 1 applied to an affirmative sentences such as 2 yields the following result: the proposition ‘Mary talked to John’ is presupposed and every non-weaker alternative (i.e. propositions of the form ‘Mary talked to x’ where x is a relevant individual in the domain) is false.

(2) Mary *only/just* talked to John.

Under these assumptions a sentence like 2 winds up meaning, correctly, that Mary talked to John and she did not talk to any other salient individual. The reason why the prejacent of *only*<sup>1</sup> is taken to be presupposed and not asserted becomes evident when *only* applies to a negative sentence such as 3.

(3) Mary did not *only* talk to John.

---

\*The authors would like to thank Jacopo Romoli, Clemens Steiner-Mayr, Yasutada Sudo, Jovana Gaji, Samer Al Khatib and three anonymous reviewers for their help and insightful comments on this work. All the mistakes are our own.

<sup>1</sup>Henceforth we will just write *only* as referring to both *just* and *only*

Here, the exclusive component is negated, hence the truth condition of **1** are reversed: at least one of the alternatives must be true. Yet, the truth of the prejacent (“Mary talked to John”) is preserved, in that **3** means that Mary talked to John, and she talked to someone else too.

In the present work, we focus on three problems that this standard analysis faces. First, the semantics in **1** generates contradictions when only takes scope over a disjunction as in ‘John only spoke to Mary or Sue’ [6]. Second, the presupposition of *only* displays unusual projective properties. For instance, it seems to disappear when *only* occurs in the antecedent of a conditional (cf. [9], [1]), as in **4**, or in the consequent of a conditional in **5** (cf. [5], [2]).

- (4) If John spoke only to Mary at the party, he will be depressed.  
 (5) If you want good cheese, you only have to go to North End.

Clearly, **4** is felicitous in a context in which John did not talk to Mary, or, for that matter, to anyone at all. Similarly, sentence **5** does not presuppose that if you want good cheese you have to go to North End (and can’t go anywhere else). Third, in many languages exclusives like *only* and just give rise to so-called rank-order readings, as in **6a** [8], [4]. In this case, the relevant alternatives, e.g., {first year student, second year student,...}, are mutually exclusive, which makes the contribution of *only* as spelled out in **1** very unclear, since the prejacent by itself already excludes the alternatives. Moreover, these readings add an extra twist to the unruly presuppositional behavior of *only*, as the presupposition of the prejacent disappears under negation, as in **6b**.

- (6) a. John is only a first-year student.  
 b. John is not only a first-year student, he is a second-year.

In this paper we provide an analysis of *only* which addresses these three issues. Our main point of departure from the classic account is that we propose that the prejacent of *only* is not presupposed and its behaviour in negative sentences such as **3** is captured through a second-order exhaustification, in the spirit of Romoli’s [12] proposal on soft triggers. In Sections 2-4 we expand on the three issues just sketched elaborating on why they all need to be taken into account together. In section 5 we present our proposal and explore some of its consequences.

## 2 *Only* and Innocent Exclusion

The problem with *only* taking scope over disjunctive sentences was addressed by Fox [6]. Consider:

- (7) a. Who of Bill, Paul, Mary and Sue will you talk to at the party?  
 b. I will only talk to Sue or Mary.

Sentence **7b** in answer to **7a** is naturally interpreted as in **8**:

- (8) I will talk to Sue or Mary, and I will not talk to Bill and I will not talk to Paul.

Notice that the question in **7a** explicitly introduces Bill, Paul, Mary and Sue as the set of relevant individuals, which means that the set of alternatives for *only* will have to include all such individuals:

- (9) ALT-set of **7b**  $\supseteq$  {**I will talk to Bill**, **I will talk to Paul**, I will talk to Sue, I will talk to Mary}

The alternatives in boldface are those who need to be negated in order to obtain the right result in 8. The problem is that the standard characterization of *only* in 1 cannot distinguish between the alternatives in boldface and those not in boldface, namely the individual disjuncts. If the set of alternatives is as in 9, the classical definition of *only* immediately and straightforwardly leads to a contradiction. One might try to get around this problem by excluding the disjuncts from the set of relevant alternatives. But this move is problematic on two counts. First, it is unprincipled: if disjuncts are not relevant when uttering a disjunction, then what is? Second, including disjuncts among the set of relevant alternatives is the only way to get Free Choice effects, as argued by Fox and others.

It is perhaps worth pointing out that the same point can be made for rank-order readings of *only*:

- (10) John is *only* second or third year.

Sentence 10 simply says that John is second or third year, which entails he is not first or fourth, and suggests that being second or third constitutes some ‘less than maximal’ stage of the program (i.e. 10 would be infelicitous if applied to a three year study program). Note that including the single disjuncts among the relevant alternatives would lead to contradiction here too. These considerations point to the conclusion that the definition of *only* has to be modified in a more radical way than tinkering with what set of alternatives are relevant. Fox [6] proposes modifying exclusives in terms of the notion of ‘Innocent Exclusion’. Given a set of alternatives ALT to a proposition p, the set of Innocently Excludable (IE) alternatives is defined as follows:

- (11) a.  $I-E(p,ALT) = \cap\{A' \subseteq ALT: A' \text{ is a maximal set in ALT, s.t., } A' \neg \cup \{p\} \text{ is consistent}\}$   
 b. For any A,  $A \neg = \{\neg a: a \in A\}$

An alternative is IE if and only if it belongs to all maximal subsets of alternatives that can be negated consistently with the assertion. Exclusives can only ‘eliminate’ alternatives that are IE. In other words, what examples like 7b show is that the exclusions that (overt) *only* brings about must be systematically relativized to IE subsets of the Focal alternatives associated with it. Our proposal in Section 5 below acknowledges and incorporates this view. Our contribution in this connection will be simply to show that this shift of perspective has interesting consequences in connection with how weak Negative Polarity Items (NPIs) are licensed by exclusives like *only*.

### 3 Rank-order readings and the non-vacuity of *only*

The main issues to be born in mind in connection with rank-order readings are the following. First, as noted with example 6a, in rank-order sentences, we seem to be dealing with a set of alternatives that are already mutually exclusive, and hence the role of *only* appears to be unclear. This is especially dramatic in view of the observation that usually *only* is deviant, when vacuous:

- (12) a. Q: Who of John, Bill and Sue did you invite?  
 b. I invited *only* John and Bill.  
 c. \* I invited *only* John, Bill and Sue.



Second, the deviance of 12c, i.e. the 'vacuous' case, seems to be parallel to the observation that in rank-order readings, the associate of *only* cannot be 'highest' on the relevant scale:

- (13) a. Context: John is in a 5 year program.  
 b. Q: Which year is John in? How close is he to finishing?  
 c. He is very far from finishing. He is *only* first year.  
 d. He is not quite so close to finishing. He is *only* fourth year.  
 e. \* He is *only* fifth year.

The parallelism between 12c and 13e is evident: in both cases the 'top' element of the relevant scale cannot be felicitously be operated on by *only*.

The third problem can be best appreciated by looking at the different behavior of negation vis-à-vis rank ordered vs. non rank ordered *only*:

- (14) a. John isn't *only* a linguist, he is a poet.  
 b. John isn't *only* a 1st year, he is a 2nd year.

The most natural understanding of 14a is additive, that John is a linguist and a poet; the most natural interpretation of 14b is of denial, that he isn't 1st but 2nd year; the presuppositional approach to *only* predicts that 14a and 14b should be parallel. But this is not so; notice, by the way, how 14b is rather seamless, and doesn't seem to involve the reinterpretation that goes with a presupposition cancellation. This contrast is mysterious from the traditional presuppositional point of view on *only*.

One very reasonable way of dealing with all this is to maintain that rank-order *only* calls upon an 'at least' operator: *John is only first year* means something like *John is only at least first year* (= first year or more). The latter, in turn, according to the classical theory, would presuppose that John is at least first year; this suffices to account for the pattern of *only* under negation (see, e.g., [4] for an account along these very lines). The problem is that the alleged presupposition of *only* just does not behave as a presupposition in conditionals. To this we now turn.

## 4 *Only* in conditionals: problems for presuppositional theories of exclusives

Let us consider again the example in 4, repeated below.

- (15) If John spoke *only* to Mary at the party, he will be depressed.

According to the semantics in 1, the example above presupposes that John spoke to Mary at the party. Yet, the sentence in 15 is perfectly compatible with the falsity of the prejacent of *only*. Namely, this sentence is felicitous if John spoke with no-one at the party. Indeed, in this case 15 suggests that John will still be depressed. This problem was first noticed by Ippolito [9], who proposed another lexical entry for *only*, which does not run into the problem pointed out above. Ippolito's proposal involves a conditional presupposition, namely, that if any member of the set of alternatives is true then the prejacent must be true as well. In negative sentences such as 3, which overtly asserts that one of the alternatives is true, such presupposition immediately derives the truth of the prejacent. In positive sentences like 2, instead, the truth of the prejacent comes from a scalar implicature that arises from the competition between the

assertion and some stronger alternative. In the case of 2, e.g., this alternative would be ‘Mary did not talk to anyone’. Ippolito [9] points out that even local accommodation of the prejacent (cf. [16]) does not solve the puzzle in 15, but rather it derives the wrong results, as shown in 16.

- (16) If John spoke to someone and he spoke to Mary and no-one else, he will be depressed.

Sentence 16 is perfectly compatible with the situations in which the antecedent is false (‘John spoke to no-one or to other people than Mary’), and if anything it suggests that it’s possible that he spoke to no-one and he will be happy, which is not what 15 intuitively means. Ippolito’s conditional presupposition, instead, only states that if John spoke to anyone, than he must have spoken to Mary, hence it correctly predicts that 15 is felicitous in a context where the prejacent is false, where nothing is presupposed.

The problem with Ippolito’s proposal, as pointed out by [4], is that this account does not work for rank-order readings under negation, such as 6b. In this case it would incorrectly predict that the prejacent is true (i.e. ‘if John is enrolled in any year, he is a first-year’). This contradicts the meaning of the proposition ‘John is a second-year or more’, in that it would predict that John is at the same time a first-year and a second-year or more.

In the literature other proposals (cf. [5], [2], [4]) have been put forward in order to account for other cases in which *only* does not seem to presuppose its prejacent such as, for instance, in sentences like 5. The idea shared by these accounts is that the presupposition introduced by *only* is, in fact, something logically weaker than its prejacent, such as that at least one of the alternatives is true. While we cannot discuss the details of these proposals here, let us remark that all these accounts run into troubles with examples like 15, in which *only* does not seem to presuppose anything at all.

## 5 The proposal: a two-stage theory of *only*

Before diving into our proposal, it is worth discussing whether it is plausible to consider the prejacent of *only* as part of the assertoric content, contrary to what has been claimed in the literature since the original proposal by Horn [7]. First of all, insofar as our intuitions are concerned, *only*-statements seem to assert the truth of the prejacent, just like focus does. Consider the next examples:

- (17) a. Who came to the party yesterday?  
 b. Only [John]<sub>F</sub> came.  
 c. [John]<sub>F</sub> came.
- (18) a. How many books have you read this year?  
 b. I only read [two books]<sub>F</sub>.  
 c. I read [two books]<sub>F</sub>.

According to the standard account, 17b and 18b are truth-conditionally different from 17c and 18c. While the latter statements directly address the questions in 17a and 18a via their assertoric component, the sentences including *only* address the questions via presuppositional accommodation. Yet, both structures, *only* and Focus, exclude relevant alternatives from the conveyed interpretation via some sort of exclusive inference. In fact, the difference between the interpretations generated by *only* vs. exclusive Focus does not seem to exist. Sentences 18b and 17b, just like their counterpart without overt *only*, do not seem to involve any sort

of pragmatic accommodation. Rather, unlike other run-of-the-mill presuppositional triggers, their alleged presuppositions offer a completely natural way to address the questions at stake. Another observation suggesting that *only* does not presuppose but asserts the prejacent is the following. Normally, presuppositions are conditions of assertability of an utterance. Sentences 19b and 20b uttered in contexts 19a and 20a are deviant, and standard accounts (cf., e.g., [14]) maintain that their truth values are undefined in such contexts.

- (19) a. Context: France does not have a king  
 b. This morning I had breakfast with the King of France
- (20) a. Context: John does not have a brother  
 b. John realized his brother is the president of U.S.A.

In contrast, sentence 17b is intuitively false in a context where John did not come to the party, as it is false if uttered in a context where John did come to the party along with other salient individuals. These considerations suggest that the idea, discussed in the following paragraphs, that the prejacent of *only* is part of the assertion of an *only*-sentence is worth being pursued.

Our proposal is based on the following assumptions:

- (21) a. *Only* and *just* are subject to a non-vacuity presupposition (cf. [1])  
 b. They only ‘see’ entailment-based sets of alternatives  
 c. They employ Innocent Exclusion (cf. [6])  
 d. They are non presuppositional, but trigger a factive implicature along the lines proposed by Romoli [12] for soft presuppositional triggers

Assumptions 21a, 21b and 21c are fairly straightforward; we will first show how to implement them and then turn to 21d. More specifically, assumptions 21a, 21b and 21c can be made formally explicit along the following lines:

- (22) a.  $\text{Only}_{\text{ALT}}(p) = \lambda w [ p(w) = 1 \wedge \forall q \in \text{I-E}(p, \text{ALT}) \rightarrow q(w) = 0 ]$ ,  
 if  $\text{Only}_{\text{ALT}}(p) \subset p$ , else undefined  
 Where ALT is entailment based (i.e.  $\forall p \in \text{ALT} [ \exists q \in \text{ALT} (p \subset q \vee q \subset p) ]$  )  
 b.  $\text{I-E}(p, A) = \cap \{ A' \subseteq A : A' \text{ is a maximal subset of } A, \text{ s.t., } A' \cup \{p\} \text{ is consistent} \}$   
 where  $A' \neg = \{ \neg q : q \in A \}$

Note first that the prejacent is now taken to be part of the assertion, rather than presupposed. Second, the non-vacuity presupposition built in 22a can simply be viewed as an economy condition: Do not use *only* if its effects are vacuous. And third, limiting the ALTs visible to *only* to entailment based scales has as a consequence that a rank-order scale (like *being first year, being second year,...* etc.) cannot be directly operated on by *only*. It first has to be mapped onto an entailment based one in order for *only* to see it. This can be done in a general way using an at least-like operator:

- (23) a. Rank-order scale: 1st year < 2nd year < ... < 5th year  
 Assume that each of the elements of this scale are of type  $\langle e, t \rangle$  and mutually exclusive

- b. Corresponding Entailment based scale:

$$[1ys \vee 2ys \vee 3ys \vee 4ys \vee 5ys] \supseteq [2ys \vee 3ys \vee 4ys \vee 5ys] \supseteq \dots \supseteq 5ys$$

Let  $A$  be a rank order scale ordered by ‘<’, then the corresponding entailment based scale  $EB(A)$  is:  $\{AT\ LEAST(a_i): a_i \in A\}$ , where for each  $a_i \in A$ ,

$$AT\ LEAST(a_i) = \lambda x. a_i(x) \vee a_j(x) \dots \vee a_n(x), \text{ for any } a_j, \text{ an such that } a_i \leq a_j, a_n$$

While many formal details would need to be spelled out further, the idea here is pretty clear, indeed straightforward. We assume that rank-order readings result from covert uses of an ‘at least’ operator, prompted by the fact that *only* cannot deal with non-entailment based scales, in the spirit of [4] (cf. also [2]). All the peculiarities of rank-order scales immediately fall into place under this view. Consider for example 24a:

- (24) a. John is only first year.  
 b. Prejacent: John is AT LEAST (1st year) = John is  $(1y \vee 2y \vee 3y \vee 4y \vee 5y)$   
 c. ALTs: John is  $(2y \vee 3y \vee 4y \vee 5y)$ , John is  $(3y \vee 4y \vee 5y)$ , etc.  
 d.  $Only_{ALT}(\text{John is AT LEAST (1st year)}) =$   
 $\text{John is } (1y \vee 2y \vee 3y \vee 4y \vee 5y) \wedge \neg \text{John is } (2y \vee 3y \vee 4y \vee 5y) = \text{John is 1st year.}$

Note that all of the ALTs in 24c are IE. Hence use of *only* as in 24d straightforwardly delivers the desired result. The non vacuity presupposition is met with respect to the modified scale. Finally, the fact that *only* cannot associate with the highest degree in a rank-order scale is just a consequence of non-vacuity. Further effects of ‘low rank’ are, we think, just pragmatics. So far so good. Our proposal so far is, perhaps, just a variant of [4]. Where we depart from all available proposals is in the treatment of ‘projection’. We suggest that *only* enters in paradigmatic alternative with its prejacent, in the manner proposed by Romoli [12] for factives. I.e. as part of the lexical specification of the semantic contributions of *only*, we include the following (as a way of implementing 21d above):

$$(25) \quad ALT(Only_{ALT}(p)) = \{p, Only_{ALT}(p)\}$$

Formally activated set of alternatives need to be factored into meaning. This is generally done via a covert counterpart of *only*, namely *Exh*, along the lines independently argued for by [6] and much subsequent literature (cf. [3]), in connection with Free Choice and polarity sensitive phenomena. Let us see what consequences this has.

In positive environments, there are no consequences. To see this, consider, e.g., 26a, in an environment where the relevant set of alternatives are (the closure under meet and join of)  $\{\text{linguist, philosopher}\}$ :

- (26) a. John is only a linguist.  
 b.  $ONLY_{\{\text{LINGUIST, PHILOSOPHER}\}}(\text{john is a linguist}) =$   
 $\text{John is a linguist} \wedge \neg \text{John is a philosopher}$   
 c.  $ALT(b) = \{\text{John is a linguist, John is a linguist} \wedge \neg \text{John is a philosopher}\}$   
 d.  $Exh_{ALT(b)} ONLY_{\{\text{LINGUIST, PHILOSOPHER}\}}(\text{john is a linguist}) =$   
 $\text{John is a linguist} \wedge \neg \text{John is a philosopher}$

Sentence 26a is interpreted as 26b. The set of alternatives to this interpretation are as in 26c. They must undergo a second round of (this time, covert) exhaustification as in 26d. However, the prejacent to this covert exhaustification is logically stronger than its alternative (the sentence without *only*); hence nothing happens and we get back simply 26b. In negative environments, however, this is not so:

- (27) a. John is not only a linguist.  
 b.  $\neg \text{ONLY}_{\{\text{LINGUIST}, \text{PHILOSOPHER}\}}(\text{John is a linguist}) =$   
 $\neg (\text{John is a linguist} \wedge \neg \text{John is a philosopher}) =$   
 $\text{John is a linguist} \rightarrow \text{John is a philosopher}$   
 c.  $\text{ALT}(b) = \{\neg \text{John is a linguist}, \neg(\text{John is a linguist} \wedge \neg \text{John is a philosopher})\}$   
 d.  $\text{Exh}_{\text{ALT}(b)} \neg \text{ONLY}_{\{\text{LINGUIST}, \text{PHILOSOPHER}\}}(\text{John is a linguist}) =$   
 $\neg(\text{John is a linguist} \wedge \neg \text{John is a philosopher}) \wedge \neg \neg \text{John is a linguist} =$   
 $(\text{John is a linguist} \rightarrow \text{John is a philosopher}) \wedge \text{John is a linguist} =$   
 $\text{John is a linguist} \wedge \text{John is a philosopher}$

Under negation the sentence with *only* becomes weaker than its alternative (the negated prejacent of *only*); hence in this case, the second round of (covert) exhaustification is not vacuous: it brings in a strengthening which is tantamount to the additive inference.

Taking stock, we notice that by replacing the familiar stipulation that  $\text{ONLY}(p)$  presupposes  $p$ , with the stipulation that  $\text{ONLY}(p)$  and  $p$  are alternatives to each other, given the existence of an independently motivated mechanism for covert exhaustification a la Fox, we obtain results which are very similar to those of the traditional presuppositional approach when it comes to negation: the prejacent of *only* is passed up across negation. Notice, furthermore, that this extends without further ado to the case of negated rank-order reading under the version of the use of the AT LEAST operator sketched in the present section (we have to leave the relevant computations to the patience of our readers). Where the present proposal and the traditional theory clearly come apart is when *only* is embedded in conditionals. We begin by noticing that *only* in antecedents of conditionals behaves as a soft trigger in the sense of Romoli:

- (28) I don't know whether John is a linguist or a philosopher or both. But if he is only a linguist, he won't be of any use to us.

Sentences like 28 show that *only* just does not project across antecedents of conditionals. Romoli's approach, which we adapted here to the case of *only* predicts this behavior fully:

- (29) a. If John is only a linguist, he won't be of any use to us.  
 b. If John is a linguist, he won't be of any use to us.  
 c. It is not the case that if John is a linguist, he won't be of any use to us.  
 d. i.  $\text{ONLY}_{\{\text{LING}, \text{PHIL}\}} \text{John is a linguist} \rightarrow \text{John is of no use}$   
 ii.  $\text{ALT: John is a linguist} \rightarrow \text{John is of no use}$   
 iii.  $\text{ONLY}_{\{\text{LING}, \text{PHIL}\}} \text{John is a linguist} \rightarrow \text{John is of no use} \wedge \neg$   
 $\text{John is a linguist} \rightarrow \text{John is of no use}$

The alternative to 29a is 29b; the latter is non-weaker than the assertion 29b; therefore it gets denied, as in 29c, and the assertion gets strengthened to the conjunction of 29a and 29b. All this is expressed semi-formally in 29d. The result is compatible with a variety of outcomes, depending on the specific of the contexts. But the point is that the so-called presuppositions of *only* is predicted **not** to project in this case. Similar considerations apply, mutatis mutandis, to the consequent of conditionals, as in 5.

We finally note that the present proposal on *only* also predicts without any specific assumption (and without resorting to Strawson entailment, cf. [15]) that exclusives will be good licensors of weak NPIs. In order to show why, let us adopt an exhaustification based approach to NPIs, along the lines of [10] and [11], or [3]. Let us assume in particular that weak NPIs like *anyone*

are existentials which obligatorily activate subdomain alternatives, which must be factored into meaning via Exh. Thus, for example, a sentence like **30a** has the interpretation in **30b**, with alternatives in **30c**.

- (30) a. John saw anyone.  
 b.  $\text{Exh}_{\text{ALT}}(\exists x \in D[\text{John saw } x])$   
 c.  $\text{D-ALT} = \{\exists x \in D'[\text{John saw } x]: D' \subseteq D\}$

On the assumption that Exh in **30b** is not relativized to IE subsets, **30b** is contradictory, for all of the alternatives in **30c** are stronger than the assertion and all would have to be excluded<sup>2</sup>. Accordingly, **30a** is ruled out, as desired. This contrasts with what happens in cases like (23):

- (31) a. Only John saw anyone.  
 b.  $\text{ONLY}_{\text{ALT}}(\exists x \in D[\text{John saw } x])$   
 c.  $\text{D-ALT} = \{\exists x \in D'[\text{John saw } x]: D' \subseteq D\}$   
 d.  $\text{F-ALT} = \{\exists x \in D'[\text{a saw } x]: \text{a} \in \text{REL}\}$ ,  
     where REL is the set of contextually salient individuals  
 e.  $\text{ALT} = \text{D-ALT} \cup \text{F-ALT}$   
 f.  $\exists x \in D[\text{John saw } x] \wedge \neg \exists x \in D[\text{a saw } x]$  (for any a in U)

The switch from Exh to ONLY brings in the set of focal alternatives in **31e**, which plausibly gets added to the lexically triggered subdomain alternatives associated with any as in **31c**. However, ONLY excludes IE alternatives. All of the F-ALTs are, while none of the D-ALTs is. So the result is **31f**, which is exactly what we want. And the subdomain alternatives are properly factored into meaning (vacuously in this case)<sup>3</sup>. Strictly speaking, use of *only* brings along its own set of alternatives as in **32a**, which in turns must be factored into meaning via Exh.

- (32) a.  $\text{ALT}((23b)) = \{\text{ONLY}_{\text{ALT}}(\exists x \in D[\text{John saw } x]), \exists x \in D[\text{John saw } x]\}$   
 b.  $\text{Exh ALT}((23b))(\text{ONLY}_{\text{ALT}}(\exists x \in D[\text{John saw } x])) = \text{ONLY}_{\text{ALT}}(\exists x \in D[\text{John saw } x])$

But the exhaustification in **32a** turns out to be vacuous, as is generally the case for *only* in positive environments.

In conclusion, we have proposed a set of arguably minor modifications of the traditional theory of *only*:

- (33) a.  $\text{ONLY}_{\text{ALT}}(p)$  asserts p and excludes all IE members of ALT  
 b. It is subject to non-vacuity  
 c. ALT must be partially or totally ordered by entailment  
 d.  $\text{ONLY}_{\text{ALT}}(p)$  forms a (formal) alternative with p

<sup>2</sup>This in turns entails that the version of Exh employed in Free Choice phenomena must be relativized to IE alternative. This is precisely the parameter that differentiates ‘pure NPIs’ from FCI. See [3] for one way of developing this idea, within a parametric approach to polarity sensitivity.

<sup>3</sup>This is generally so for weak NPIs. The semantic effects of D-alternatives becomes visible only in situations of contrastive stress like:

Speaker A: who ate potatoes?

Speaker B: Only John

Speaker A: Are you sure that nobody else ate maybe some leftovers?

Speaker B: Yes. Only John ate ANY potatoes.

Cf., again [3] for relevant discussion.

Of these, the most substantive departure from tradition is 33d, which replaces the stipulation that  $\text{ONLY}_{\text{ALT}}(p)$  presupposes  $p$ . We argued that these changes account for the behavior of *only*, including the coming about of its rank-order readings, in a way that explains at least as well as the traditional approach its core properties, such as *only*’s capacity for NPI licensing. The present proposal, moreover, seems to begin to make sense of how and why the preja cent ‘projects’ across negation, but not in conditionals, which from the point of view of the traditional theory keeps being elusive, in spite of several valuable attempts. There is, we think, a rich set of potential consequences that the present approach opens for future research. For example, the availability of rank-order readings for exclusives depends on the appeal to an AT LEAST operator, necessary to shift rank-order scales into entailment based ones. The availability of such an operator might be item-specific, with some exclusives allowing for it, while others disallowing it. We believe this expectation to be born out, but we must leave further exploration of this and other consequences to future research.

## References

- [1] Sam Samer Al Khatib. *‘Only’and association with negative antonyms*. PhD thesis, Massachusetts Institute of Technology, 2013.
- [2] Luis Alonso-Ovalle and Aron Hirsch. Keep’only’strong. In *Semantics and Linguistic Theory*, volume 28, pages 251–270, 2018.
- [3] Gennaro Chierchia. *Logic in Grammar. Polarity, Free Choice, and Intervention*. Oxford University Press, Oxford, 2013.
- [4] Elizabeth Coppock and David I Beaver. Principles of the exclusive muddle. *Journal of Semantics*, 31(3):371–432, 2013.
- [5] Kai von Stechow and Sabine Iatridou. Anatomy of a modal construction. *Linguistic Inquiry*, 38(3):445–483, 2007.
- [6] Danny Fox. Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, pages 71–120. Springer, 2007.
- [7] Larry Horn. A presuppositional theory of onlyand even. *CLS 5, Chicago Linguistics Society*, 1969.
- [8] Laurence R Horn. Pick a theory, not just any theory. *Negation and Polarity. Syntactic and semantic perspectives*, pages 147–192, 2000.
- [9] Michela Ippolito. On the meaning of only. *Journal of Semantics*, 25(1):45–91, 2007.
- [10] Manfred Krifka. The semantics and pragmatics of polarity items. *Linguistic analysis*, 25(3-4):209–257, 1995.
- [11] Utpal Lahiri. Focus and negative polarity in hindi. *Natural language semantics*, 6(1):57–123, 1998.
- [12] Jacopo Romoli. *Soft but strong. Neg-raising, soft triggers, and exhaustification*. PhD thesis, 2012.
- [13] Mats Rooth. A Theory of Focus Interpretation. *Natural Language Semantics*, 1:75–116, 1992.
- [14] Robert Stalnaker. Pragmatic presuppositions. In *Proceedings of the Texas conference on per~formatives, presuppositions, and implicatures. Arlington, VA: Center for Applied Linguistics*, pages 135–148. ERIC, 1977.
- [15] Kai Von Stechow. Npi licensing, strawson entailment, and context dependency. *Journal of semantics*, 16(2):97–148, 1999.
- [16] Kai Von Stechow. What is presupposition accommodation, again? *Philosophical perspectives*, 22:137–170, 2008.



# Fixing de Morgan’s Laws in Counterfactuals\*

Jacopo Romoli<sup>1</sup>, Paolo Santorio<sup>2</sup>, and Eva Wittenberg<sup>3</sup>

<sup>1</sup> Ulster University, Belfast, U.K.  
j.romoli@ulster.ac.uk

<sup>2</sup> University of Maryland, College Park, College Park MD, U.S.A.  
santorio@umd.edu

<sup>3</sup> UC San Diego, La Jolla CA, U.S.A.  
ewittenberg@ucsd.edu

## Abstract

Classical semantics for counterfactuals are based on a notion of comparative similarity. These semantics are intensional, hence they predict that logically equivalent clauses can be substituted in counterfactuals *salva veritate*. A recent truth-value judgment study by Ciardelli, Zhang, and Champollion ([6]; CZC) appears to challenge both the idea that comparative similarity plays a role in counterfactual semantics and the prediction that logical equivalents are substitutable. CZC account for their data via an inquisitive semantics for disjunction and a semantics for counterfactuals that does not exploit the standard similarity algorithm. We report on a study consisting of two experiments that start from CZC’s general idea, but use a simpler scenario, manipulate negation more systematically, and add an extra task based on the selection of pictures. Our results replicate the differences found by CZC, but they also suggest that the effect is linked to the presence of overt negation rather than disjunction. We conclude that (i) inquisitive disjunction is neither necessary nor sufficient to account for the problem in full generality, and (ii) the evidence does not encourage rejecting a similarity semantics.

## 1 Introduction

### 1.1 Classical semantics and disjunctive antecedents

Classical theories of counterfactuals start from a simple idea, which is pithily put by Stalnaker: “Consider a possible world in which  $A$  is true, and which otherwise differs minimally from the actual world. “If  $A$ , then  $B$ ” is true (false) just in case  $B$  is true (false) in that possible world.” ([19]) Theories in the tradition of Stalnaker and Lewis ([19], [14], [15], a.o.) model this idea formally via a relation of comparative closeness, represented as ‘ $\preceq_w$ ’.  $\preceq_w$  compares worlds with respect to their closeness to a benchmark world  $w$ : ‘ $w' \preceq_w w''$ ’ says that  $w'$  is closer to  $w$  than  $w''$ . Comparative closeness singles out a set of ‘maximally close’ antecedent worlds, which are then used to evaluate the consequent. The schematic truth conditions of a counterfactual are:<sup>1</sup>

- (1)  $A \Box \rightarrow C$  is true at  $w$  iff all  $A$ -worlds that are maximally  $\preceq_w$  close to  $w$  are  $C$ -worlds

This paper discusses a challenge raised to classical accounts of counterfactuals by Ciardelli, Zhang, and Champollion ([6]; henceforth, CZC). This challenge has its roots in a classical

---

\*We would like to thank Moysh Bar-Lev, Matthew Mandelkern and Paul Marty for very helpful discussion, and audiences at NELS 50 at MIT, Ulster University, University of Maryland, University of California San Diego, and University of Chicago. Work on this project was partially supported by the Leverhulme trust grant RPG-2018-425 to Jacopo Romoli.

<sup>1</sup>Here we strike a compromise between Stalnaker’s and Lewis’s actual theories, using universal quantification like Lewis but making the so-called limit assumption like Stalnaker. For discussion of the latter, see [20], [10].



problem for semantics of counterfactuals, first raised by Kit Fine ([7]). Fine observes that counterfactuals with disjunctive antecedents seem to entail the 'simplified' counterfactuals with the individual disjuncts as antecedents.

- (2) If John had taken Syntax or Semantics, he would solve this exercise.  
 $\rightsquigarrow$  If John had taken Syntax, he would solve this exercise.  
 $\rightsquigarrow$  If John had taken Semantics, he would solve this exercise.

Examples like (2) seem to suggest that (3) is a valid principle of counterfactual logic.

- (3) **Simplification:**  $(A \vee B) \Box \rightarrow C \models A \Box \rightarrow C, B \Box \rightarrow C$

Unfortunately, comparative closeness semantics does not validate Simplification, nor can it be tweaked to validate it without substantial consequences. The reason is that Simplification is inconsistent with two logical principles that standard semantics validates.

- (4) **Substitution:**  $A \Box \rightarrow C \models A' \Box \rightarrow C$  (with  $A$  and  $A'$  logically equivalent)  
(5) **Failure of Antecedent Strengthening:**  $A \Box \rightarrow C \not\models A^+ \Box \rightarrow C$  (with  $A^+ \models A$ )

In the face of this problem, theorists have split into two camps. The first camp tries to accommodate Simplification as a broadly pragmatic effect (see e.g. [11], [12], [8], [2], [3]). The second camp tries to account for Simplification by rejecting Substitution (see e.g. [1], [16], [21]). CZC's proposal falls in this second camp.

## 1.2 CZC's challenge

CZC claim that Substitution fails in a specific and striking way: In counterfactual antecedents, classical de Morgan equivalences between disjunctions of negations and negations of conjunctions fail. In particular, CZC argue that counterfactuals where antecedents are disjunctions of negations don't entail counterfactuals where antecedents are negated conjunctions:

- (6) **de Morgan Failure in Counterfactuals:**  $(\neg A \vee \neg B) \Box \rightarrow C \not\models \neg(A \wedge B) \Box \rightarrow C$

CZC claim is based on experimental evidence. They run one main experiment, preceded by two pre-tests and followed by three post-hoc tests; here we report their main experiment.

Participants were presented with a brief description of a scenario involving a lightbulb and two switches, together with a picture (Figure 1). The description makes clear that the light is on whenever the two switches, A and B, are in the same position (both up or both down). Subject are asked to provide a truth value judgment (TVJ), choosing between 'true', 'false', or 'indeterminate' on a sentence from one of five conditions:<sup>2</sup>

- |     |    |  |   |
|-----|----|--|---|
| (7) | a. | If A was down, the light would be off.               | $\bar{A} \Box \rightarrow C$                |
|     | b. | If B was down the light would be off.                | $\bar{B} \Box \rightarrow C$                |
|     | c. | If A or B was down, the light would be off.          | $(\bar{A} \vee \bar{B}) \Box \rightarrow C$ |
|     | d. | If A and B were not both up, the light would be off. | $\neg(A \wedge B) \Box \rightarrow C$       |
|     | e. | If A and B were not both up, the light would be on.  | $\neg(A \wedge B) \Box \rightarrow \bar{C}$ |

Their result raise two main challenges for the classical theories. First, the counterfactuals with de Morgan-equivalent antecedents in (7-c) and (7-d) are endorsed at very different rates

<sup>2</sup>Following CZC, we use  $\ulcorner \bar{A} \urcorner$  to denote a clause that is equivalent to the negation of A but does not include overt negation, and  $\ulcorner \neg A \urcorner$  to denote clauses that involve overt negation.

(69.33% vs 22.04%). This appears to be an obvious violation of **Substitution**, which predicts that (7-c) and (7-d) are equivalent. Second, (7-d) is endorsed at a much lower rate than the corresponding counterfactuals with simple negated antecedents in (7-a) and (7-b) (66.02% and 65.11%, respectively). This seems to show that the following principle is invalid:

(8) **Negated Conjunction:**  $\neg A \Box \rightarrow C, \neg B \Box \rightarrow C \vDash \neg(A \wedge B) \Box \rightarrow C$

But, CZC argue, **Negated Conjunction** is a key principle for any similarity-based semantics.<sup>3</sup>

(9) **Negated Conjunction:**  $\neg A \Box \rightarrow C, \neg B \Box \rightarrow C \vDash \neg(A \wedge B) \Box \rightarrow C$

We now move to discuss two general approaches to accounting for these results.

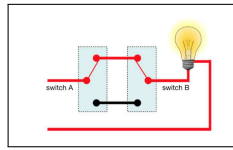


Figure 1: Stimulus of CZC’s main experiment.

## 2 Two approaches

In this section, we review two solutions to CZC’s challenge. One is due to CZC themselves and exploits inquisitive semantics; the other locates the source of the problem in negation.

### 2.1 The disjunction approach

The first approach is couched within an inquisitive semantics framework ([5]). Inquisitive semantics aims at unifying clause types across declarative and interrogative sentences: every clause denotes a set of propositions, i.e. a set of sets of worlds. Declarative and interrogative clauses differ in that the denotations of declarative sentences include a single weakest proposition, while the denotations of interrogative clauses don’t. Following [5], we say that declarative clauses include exactly *one alternative*, while interrogative clauses include *multiple alternatives*. Clauses that denote sets of multiple alternatives are said to have *inquisitive meanings*.

Crucially, in the inquisitive framework disjunctive clauses have an inquisitive meaning:  $A \vee B$  denotes a set with two alternatives, namely  $\{a, b\}$ . Conversely, negated sentences have a declarative meaning and denote a set with a single alternative. This is at the basis of the different treatment of  $(\neg A \vee \neg B) \Box \rightarrow C$  and  $\neg(A \wedge B) \Box \rightarrow C$ . CZC show how to import counterfactuals from classical truth-conditional semantics to an inquisitive framework. Let ‘ $\Rightarrow$ ’ be a binary connective between propositions encoding any truth-conditional meaning for counterfactuals. We can ‘lift’ the basic meaning of  $\Rightarrow$  to an inquisitive framework in the following way:

<sup>3</sup>Standard axiomatizations of counterfactuals (see e.g. [4]) involve or entail the variant of Negated Conjunction that we get by swapping  $\neg(A \wedge B) \Box \rightarrow C$  for the DeMorgan-equivalent  $(\neg A \vee \neg B) \Box \rightarrow C$ . Even if we reject **Negated Conjunction**, however, this doesn’t necessarily mean that comparative similarity should simply drop out of consideration. In fact, there are a number of semantics for counterfactuals that are still based on similarity and closeness and that invalidate principles logically related to Negated Conjunction: see e.g. [17], [9]; see also [18] for discussion. What the failure of **Negated Conjunction** establishes is that we cannot evaluate counterfactuals against a *fixed* closeness ordering, i.e. an ordering that is unaffected by counterfactual suppositions.

(10)  $A \Box \rightarrow C$  is true at  $w$  iff all  $p$  in  $Alt(A)$  there is some  $q$  in  $Alt(B)$  s.t.  $p \Rightarrow q$  is true in  $w$

When an antecedent  $A$  has an inquisitive meaning, each of the alternatives denoted by  $A$  combines separately with the rest of the counterfactual. Hence (11-a) is equivalent to (11-b).

- (11) a. If A or B was down, the light would be off  
 b. If A was down, the light would be off and if B was down, the light would be off

On the other hand, the antecedent of (12) only denotes one alternative, hence it differs from (11-a). This accounts for the asymmetry in endorsement between the two.

(12) If A and B were not both up, the light would be off

CZC also define a meaning for  $\Rightarrow$  that allows for failures of **Negated conjunction**. Surveying the details of their account would take us too far afield. Suffice it to say that they employ a nonstandard premise semantics (see [13]) where antecedents can selectively remove propositions from the premise set (not dissimilarly from [17]). This allows for the weakening of the logic required to invalidate **Negated conjunction**.

## 2.2 The negation approach

The second approach explains CZC's results by appealing to negation. Two accounts following this approach have been proposed. The first is due to Bar-Lev and Fox ([2, 3]; henceforth BLF) and is based on an implicature account of Simplification. The second is due to Kathrin Schulz ([18]) and also exploits an inquisitive framework, though in crucially different ways from CZC. For reasons of space, here we limit ourselves to considering the implicature account. So far as we can see, the two accounts make the same predictions for the cases we're interested in.

BLF's account exploits an exhaustification-based approach to implicatures. This approach postulates a syntactically realized exhaustivity operator EXH, which is appended to a sentence and returns the meaning of that sentence together with its implicatures. For example:

(13) EXH[some of the kids are playing] = some and not all of the kids are playing

BLF show how, given certain assumptions about EXH, we can derive Simplification as an implicature.<sup>4</sup> (14-a) is predicted to be equivalent to the conjunction of (14-b)-(14-c).

- (14) a. EXH[If A or B was down, the light would be off]  
 b. if A was down, the light would be off  
 c. if B was down the light would be off

By itself, this mechanism doesn't explain the asymmetry between disjunction and negated conjunction. But BLF suggest a key extra assumption, i.e. that EXH is an alternative of negation. This predicts that antecedents with negated conjunctions give rise to a different (and stronger) Simplification implicature: (15-a) is equivalent to the conjunction of (15-b)-(15-d). Given that (15-d) is false in the relevant context, this accounts for CZC's results.

- (15) EXH[If A and B not both up, the light would be off] =  
 a. If A was not up, the light would be off  
 b. If B was not up, the light would be off  
 c. If A was not up and B was not up, the light would be off

<sup>4</sup>Though see [16] for some arguments against an implicature approach to simplification.



- (17) a. If A or B were on the right, the see-saw would be balanced. DISJ
- b. If A and B were not both on the left, the see-saw would be balanced. NEG CONJ
- c. If A and B were both on the right, the see-saw would be balanced. CONJ
- d. If A or B were not on the left, the see-saw would be balanced. NEG DISJ

(17-a) and (17-b) are equivalent to CZC's positive disjunction and negated conjunction conditions. (17-c) is a sanity check predicted to be *false* by both approaches, corresponding to CZC's filler sentence. The crucial novel condition is (17-d), which involves both disjunction and overt negation. As we discuss below, comparing the latter to (17-a) and (17-b) is what allows us to test the potential role of connective type vs. the presence of overt negation. Each participant was randomly assigned to one condition and saw one item only.

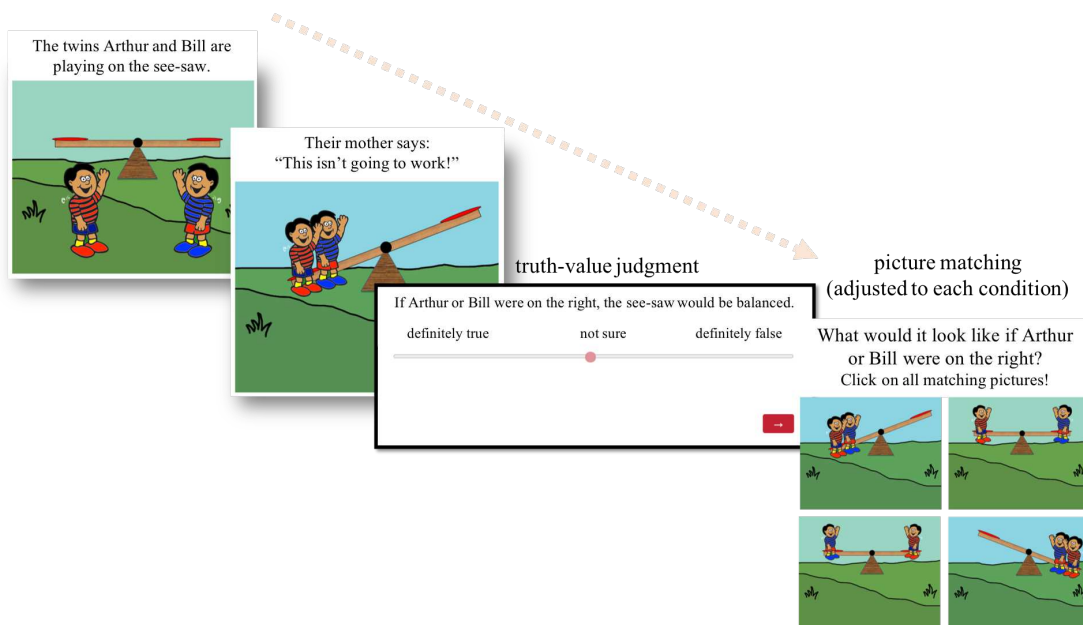


Figure 2: Trial structure in both our experiments.

Participants were presented with two tasks. The first was a TVJ task, as in CZC's study. The only difference was that, rather than giving participants the three options true/false/indeterminate, we presented them with a slider ranging from *definitely false* to *definitely true* (see Fig. 2) and asked them to adjust it according to their judgments. After the TVJ, participants were presented with a picture selection task. They were asked to select pictures that matched a linguistic description, via a question that involved the same antecedent they had seen for the TVJ task (e.g.: *What would it look like if Arthur and Bill were not both on the left?*). Subjects were explicitly told that they could select more than one sentence.

The crucial aspect of this task was in what conditions participants would select the bottom right picture in Fig. 2 ('BOTH RIGHT' picture) corresponding to both twins being on the right and the see-saw not being balanced. We take the selection of the BOTH RIGHT picture to suggest that participants imagined a scenario in which the two twins are on the right (and the see-saw is thus still imbalanced). This would, in turn, correlate with lower endorsement rates of the relevant sentences in the TVJ task.

Let us review the predictions of the two approaches. As for the TVJ task, the disjunction approach predicts higher endorsement rate whenever disjunction is present, regardless of whether negation is also involved. Therefore DISJUNCTION and NEGATED DISJUNCTION should receive higher endorsement rates than NEGATED CONJUNCTION. On the other hand, the negation approach predicts that the conditions involving negation will receive lower endorsement rates, so DISJUNCTION will score higher than both NEGATED DISJUNCTION and NEGATED CONJUNCTION. Both approaches predicts that the sanity check with positive conjunction will be judged definitely false. The predictions are summarized in (18) and (19) ( $x > y$  indicates that  $x$  is predicted to score higher than  $y$ ).

(18) **Predictions of the Disjunction approach:**  
 $\{\text{DISJUNCTION, NEGATED DISJUNCTION}\} > \text{NEGATED CONJUNCTION}$

(19) **Predictions of the Negation approach:**  
 $\text{DISJUNCTION} > \{\text{NEGATED DISJUNCTION, NEGATED CONJUNCTION}\}$

As for the picture selection task, the disjunction approach predicts that the BOTH RIGHT picture will be selected for the conditions not involving disjunction, while the negation approach predicts that it will be selected whenever negation is present. The key case is again the NEGATED DISJUNCTION condition. The disjunction approach predicts that the BOTH RIGHT picture *will not* be selected in this case, while the negation approach predicts that it *will* be selected.

### 3.1.2 Results

The results of the TVJ are plotted in the left half of Fig. 3. Bonferroni-corrected Mann-Whitney tests on contrast-coded conditions revealed significant differences between all comparisons (all  $ps < .006$ ), except between the conditions that contained negation, regardless of whether they involved disjunction or conjunction ( $p > .8$ ). The sanity check POSITIVE CONJUNCTION worked as expected, with participants overwhelmingly judging it as “definitely false”. We also replicated CZC’s difference between the POSITIVE DISJUNCTION and the NEGATED CONJUNCTION condition, with the former being endorsed at a lower rate than the latter. Crucially, however, we found equally low level of endorsement for the NEGATED DISJUNCTION condition.

The results of the picture-selection task are in Fig. 3. The y-axis corresponds to the TVJs, and participants are divided in three groups, depending on which picture(s) they selected. Most importantly for us is the green group, which includes subjects who selected the BOTH RIGHT picture in addition to the pictures where the twins are on different sides. This picture was selected most often when overt negation was present, regardless of the connective involved. That is, subjects selected all three pictures in the NEGATED CONJUNCTION and the NEGATED DISJUNCTION conditions. In the positive disjunction condition, on the other hand, they only considered the pictures where the twins are on different sides (red group). In sum, the results from both tasks reveal an effect of overt negation and are therefore in line with the negation approach, while they are challenging for the disjunction approach.

## 3.2 Experiment 2

In Experiment 2, we consider antecedents without binary connectives and investigate whether negation plays a role also in this case. As we will show, the results also suggest that **Negated conjunction** holds after all, once we control for the presence of negation.

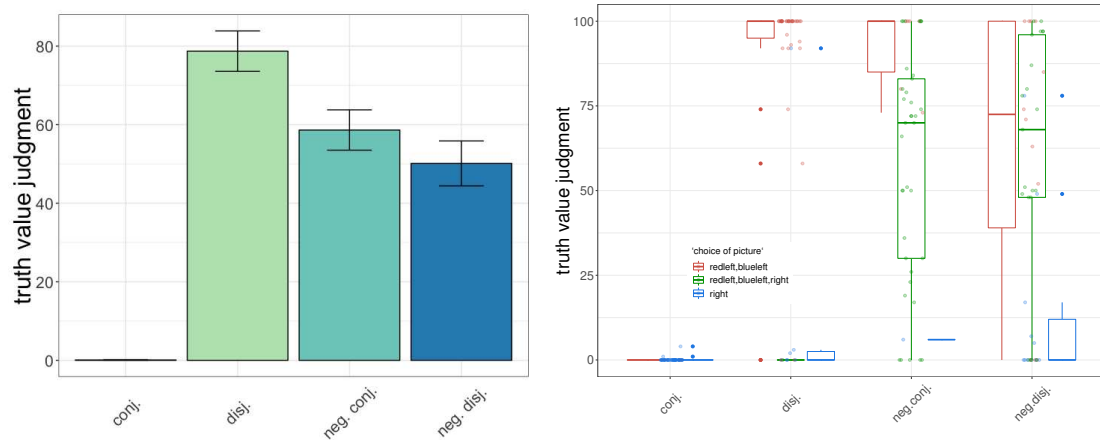


Figure 3: Results from the TVJ task (left) and picture choice task (right) of Experiment 1. Both plots show the TVJs across the four conditions. In the left plot we also divided participants in three groups, depending on their selection choices: participants who only selected the scenario in which both twins are on the right (blue group), participants who only selected the two scenarios in which each twin is one one side (red group), and participant who selected both the latter two scenarios, and in the addition, the one in which both twins were on the right (green group).

### 3.2.1 Methods

**Participants** 200 participants, recruited through AMT, participated in our experiment and were randomly assigned to one of the four conditions described below. They were all self-declared native speakers of English. No participant was excluded from the analysis.

**Procedure and Material** The procedure and tasks were the same as Experiment 1. The conditions involved again the NEGATED CONJUNCTION, which was now compared to counterfactuals with simple antecedents, either involving negation or not, all summarised in (20-a)-(20-e).

- (20)
- |    |  |            |
|----|--|------------|
| a. | If A was on the right, the see-saw would be balanced.                | POSITIVE A |
| b. | If B was on the right, the see-saw would be balanced.                | POSITIVE B |
| c. | If A was not on the left, the see-saw would be balanced.             | NEGATIVE A |
| d. | If B was not on the left, the see-saw would be balanced.             | NEGATIVE B |
| e. | If A and B were not both on the left, the see-saw would be balanced. | NEG CONJ   |

The predictions of the two approaches are as follows. In the TVJ task, the disjunction approach predicts that each of the simple counterfactuals should be endorsed at a higher rate than the NEGATED CONJUNCTION condition, while the negation approach predicts that all conditions involving negation, thus also NEGATIVE A and NEGATIVE B, will be endorsed at lower rate than the positive ones (at least, on the assumption that *A is on the right* and *B is on the right* work as alternative to each other in the relevant context). The predictions are summarized in (21) and (22).

- (21) **Predictions of the Disjunction approach:**  
 $\{\text{POSITIVE A, POSITIVE B, NEGATIVE A, NEGATIVE B}\} > \text{NEGATED CONJUNCTION}$

- (22) **Predictions of the Negation approach:**  
 $\{\text{POSITIVE A, POSITIVE B}\} > \{\text{NEGATIVE A, NEGATIVE B, NEGATED CONJUNCTION}\}$

In the picture selection task, both approaches expects participants to select the BOTH RIGHT picture in the NEGATED CONJUNCTION condition but not in the positive conditions.

### 3.2.2 Results

The results of the TVJ are plotted in Fig. 4. A Mann-Whitney tests with Bonferroni correction on contrast-coded conditions revealed significant differences between the positive and negative conditions ( $p < .0001$ ), but no difference among the negative conditions ( $p = .2$ ). We found lower endorsement of the NEGATED CONJUNCTION condition than in the POSITIVE simple antecedents conditions. Crucially, however, the NEGATIVE simple conditions were as low as the NEGATED CONJUNCTION one, showing that the key factor is the presence of overt negation.

The results of the picture-selection task are in Fig. 4 on the right. As before, the y-axis corresponds to the TVJs, and participants are divided in three groups, depending on the selection they made. As in Experiment 1, participants selected the BOTH RIGHT picture in addition to the pictures where the twins are on different sides whenever overt negation was present.

These results again lend support to the negation approach, and are challenging for the disjunction approach. In addition, they suggest that the **Negated conjunction** inference holds after all, once we control for the presence of overt negation. This is sufficient to fend off the second challenge raised by CZC. (This said, no approach predicts the results of the picture selection task, so more work is in order for a full account of the data.)

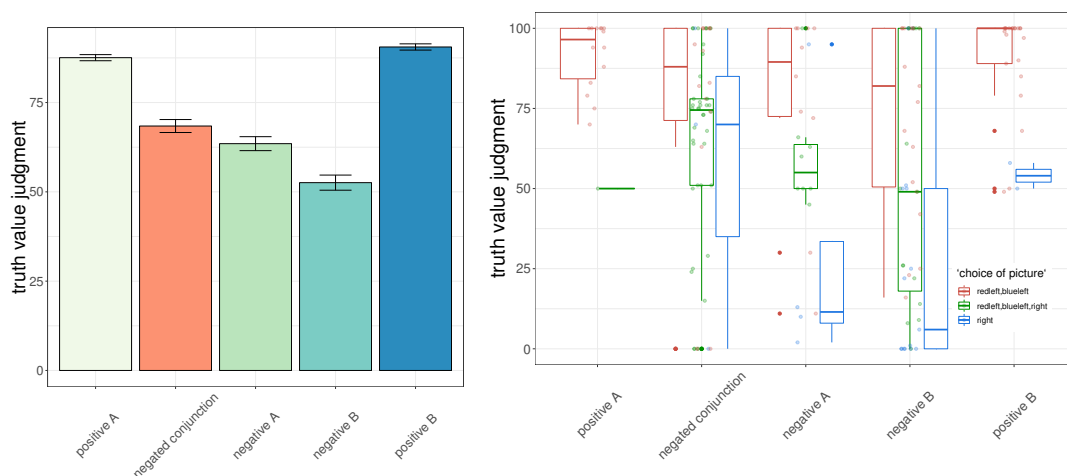


Figure 4: Results from the TVJ task (left) and picture selection task (right) of Experiment 2.

## 4 Discussion and conclusion

We tested counterfactuals with complex antecedents in a simple scenario across two experiments. Our subjects were asked to perform a TVJ and a picture selection task. Overall, we found a main effect of overt negation. In the TVJ task, subjects endorse a counterfactual with a



complex antecedent at a lower rate iff overt negation is present. In the picture selection task, participants consider the scenario in which both twins are on the right if and only if the sentence contains negation. Overall, our results lend support to an account that blames the phenomena observed by CZC on negation rather than disjunction. To be sure, we confirmed CZC's main result: sometimes disjunctions of negations and negations of conjunctions produce different TVJ. But this happens only when overt negation is involved, challenging the idea that the explanation lies with the inquisitiveness of disjunction. Moreover, our result in Experiment 2 suggests that, once we control for the presence of overt negation, **Negated conjunction** holds, addressing the challenge that CZC raise for comparative similarity semantics.

## References

- [1] Luis Alonso-Ovalle. Counterfactuals, correlatives, and disjunction. *Linguistics and Philosophy*, 32(2):207–244, 2009.
- [2] Moshe Bar-Lev. *Free choice, homogeneity and innocent inclusion*. PhD thesis, The Hebrew University of Jerusalem, 2018.
- [3] Moshe Bar-Lev and Danny Fox. Free choice, simplification, and innocent inclusion. Ms Ecole Normal Supérieure and MIT, 2019.
- [4] John P Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, 22(1):76–84, 1981.
- [5] Ivano Ciardelli, Jeroen Groenendijk, and Floris Roelofsen. *Inquisitive Semantics*. Oxford University Press, 2018.
- [6] Ivano Ciardelli, Zhang Linmin, and Lucas Champollion. Two switches in the theory of counterfactuals: A study of truth conditionality and minimal change. *Linguistic and Philosophy*, 2018.
- [7] Kit Fine. Truth, vagueness and logic. *Synthese*, 1975.
- [8] Michael Franke. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics*, 4(1):1–82, June 2011.
- [9] Thomas F. Icard. From programs to causal models. In *Proceedings of the 21st Amsterdam Colloquium*, pages 35–44, 2017.
- [10] Stefan Kaufmann. The limit assumption. *Semantics & Pragmatics*, 18(10), 2017.
- [11] Nathan Klinedinst. *Plurality and possibility*. PhD thesis, UCLA, 2007.
- [12] Nathan Klinedinst. (simplification of)<sup>2</sup> disjunctive antecedents. *MIT Working Papers in Linguistics*, 60, 2009.
- [13] Angelika Kratzer. Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic*, 10(2):201–216, 1981.
- [14] David K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, MA, 1973.
- [15] David K. Lewis. Counterfactuals and comparative possibility. *Journal of Philosophical Logic*, 2(4):418–446, 1973.
- [16] Paolo Santorio. Alternatives and truthmakers in conditional semantics. *Journal of Philosophy*, 115(10):513–549, 2018.
- [17] Paolo Santorio. Interventions in premise semantics. *Philosophers' Imprint*, 19, 2019.
- [18] Katrin Schulz. The similarity approach strikes back: Negation in counterfactuals. *Proceedings of Sinn und Bedeutung*, 22(2):343–360, May 2019.
- [19] Robert Stalnaker. A theory of conditional. In N. Rescher, editor, *Studies in Logical Theory*. Oxford, 1968.
- [20] Robert C. Stalnaker. A defense of conditional excluded middle. In William Harper, Robert C. Stalnaker, and Glenn Pearce, editors, *Ifs*, pages 87–104. Reidel, 1981.
- [21] Malte Willer. Simplifying with free choice. *Topoi*, 37(3):379–392, 2018.

# Common ground: *In sensu composito* or *in sensu diviso* \*

Merel Semeijn

University of Groningen  
m.semeijn@rug.nl

## Abstract

Traditional definitions of common ground in terms of iterative de re attitudes do not apply to conversations where at least one conversational participant is not acquainted with the other(s). I propose and compare two potential refinements of traditional definitions based on Abelard’s distinction between generality *in sensu composito* and *in sensu diviso*.

## 1 Introduction: de re common ground

Stalnaker’s [12] widely adopted notion of the ‘common ground’ of a conversation (i.e. the set of presuppositions mutually shared by conversational participants) has a dual function: First, assertions are analysed as proposals to update the common ground. Second, the notion of common ground is used to explain how a shared background can guide the production and interpretation of speech acts. For instance, I can say (and you can interpret) “Bob is coming back next week” because it is common ground between us who Bob is. Standard Stalnakerian common ground definitions are based on ‘face-to-face’<sup>1</sup> conversations where speaker  $a$  and addressee  $b$  have iterative attitudes towards one another. For instance, common ground can be defined as common belief:  $p$  is common ground between  $a$  and  $b$  iff (where  $\mathbf{B}_x\phi$  means  $x$  believes that  $\phi$ ):

$$\begin{array}{c} \frac{\mathbf{B}_ap \quad \mathbf{B}_bp}{\mathbf{B}_b\mathbf{B}_ap \quad \mathbf{B}_a\mathbf{B}_bp} \\ \frac{\mathbf{B}_a\mathbf{B}_b\mathbf{B}_ap \quad \mathbf{B}_b\mathbf{B}_a\mathbf{B}_bp}{\vdots \quad \vdots} \end{array}$$

In words, speaker and addressee both believe  $p$ ; both believe that the other believes  $p$ ; etc.<sup>2</sup> Alternatively, common ground can be defined in terms of common acceptance (i.e.  $\mathbf{A}_ap$ ,  $\mathbf{A}_bp$ ,  $\mathbf{B}_b\mathbf{A}_ap$ ,  $\mathbf{B}_a\mathbf{A}_bp$ , etc. where  $\mathbf{A}_x\phi$  means  $x$  accepts that  $\phi$  in a doxastically neutral sense. e.g. Stalnaker [13], Stokke [14]) or in terms of common commitments (i.e.  $\mathbf{C}_{a,b}p$ ,  $\mathbf{C}_{b,a}p$ ,  $\mathbf{C}_{b,a}\mathbf{C}_{a,b}p$ ,  $\mathbf{C}_{a,b}\mathbf{C}_{b,a}p$ , etc. where  $\mathbf{C}_{x,y}\phi$  means  $x$  is committed to  $y$  to act on  $\phi$ . e.g. Geurts [3]). Moreover, these iterative structures can be extended to common grounds between more than two participants (e.g. Stalnaker [13], Lewis [6], Schiffer [11]):  $p$  is common ground between all conversational participants in some community iff (where  $Cx$  means  $x$  is a conversational participant):

---

\*This research is supported by NWO Vidi grant 276-80-004 (Maier). Many thanks to Emar Maier, Bart Geurts, Christopher Badura and three anonymous reviewers.

<sup>1</sup>Here ‘face-to-face’ conversations do not require conversational participants to actually be in front of each other. It is sufficient if they know who they are talking to (e.g. an online chat conversation with a friend).

<sup>2</sup>Of course actual people don’t form infinitely many beliefs about each other’s mental state and will usually not come further than third or fourth order beliefs. This does not entail that there never exists common ground. Rather, these infinite structures represent “a chain of implications [that follow from our beliefs], not of steps in anyone’s actual reasoning. Therefore there is nothing improper about its infinite length.” (Lewis, [6, p.53]).

$$\frac{\frac{\forall x(Cx \rightarrow \mathbf{B}_x p)}{\forall x \forall y (Cx \wedge Cy \rightarrow \mathbf{B}_x \mathbf{B}_y p)}}{\forall x \forall y \forall z (Cx \wedge Cy \wedge Cz \rightarrow \mathbf{B}_x \mathbf{B}_y \mathbf{B}_z p)}$$

$$\vdots$$

In words, everyone in the community believes that  $p$ ; everyone in the community believes that everyone in the community believes that  $p$ ; etc.

Note that if  $a$  believes about  $b$  that he has a certain belief, this means that  $a$  has a *de re* belief about  $b$ . Closer inspection reveals that the iterative structures of traditional common ground definitions are all comprised of *de re* attitudes. This is problematic because the concept of common ground has – without much hesitation – been extended to non-face-to-face communication (e.g. Stokke [14]; Schiffer [11]) in which the speaker is known to the addressee but the addressee is not known to the speaker, such as books, broadcasted speeches or blogposts. For instance, it is common ground between biographer Ray Monk and myself that Wittgenstein was Austrian. However, definitions of common ground in terms of *de re* attitudes do not apply to this type of communication; I may *de re* believe about Monk but obviously Monk does not have *de re* beliefs about me. Monk merely has beliefs about the mental state of ‘the reader(s)’, whoever that may be. Hence there can actually never be any common ground between Monk and his readership. This is unsatisfactory since Monk and myself *do* seem to respectively produce and interpret the biographical text against a background of shared assumptions.

In this paper, I spell out the challenge to traditional common ground definitions posed by non-face-to-face communication using a relational analysis of *de re* belief (Section 2) that makes explicit the acquaintance relations involved in *de re* beliefs. In this paper I focus on belief based common ground definitions but the discussed issues and solutions extend to common ground definitions in terms of other *de re* attitudes. Next, I introduce Abelard’s distinction between generality *in sensu composito* and *in sensu diviso* (Section 3) and, in line with this distinction, propose two novel refinements of common ground definitions (Section 3.1 and 3.2). I show how from both definitions we can derive the original iteration of *de re* beliefs of face-to-face communication as a limit case (Section 4). Finally, I show how the case of an acquaintance that hasn’t revealed themselves as a conversational participant may aid us in deciding between the two definitions (Section 5).

## 2 Non-face-to-face communication

In the relational analysis of *de re* belief (See Kaplan [5], Lewis [7]) if  $a$  believes *de re* of  $b$  that he is  $Q$  this means that there is an acquaintance relation between  $a$  and  $b$ , and that  $a$  believes that the person he knows through this acquaintance relation is  $Q$ . This analysis implies that  $a$  has a *de se* belief (e.g.  $a$  has a belief about “the person that I saw on the beach”). I follow Lewis [7] in analysing all attitudes (including *de re* attitudes) as essentially *de se* attitudes, i.e. as self-ascription of a property. So if  $a$  believes *de re* of  $b$  that he is  $Q$ , then  $a$  is acquainted with  $b$ , and  $a$  self-ascribes the property of being such that “the person that I am acquainted with is  $Q$ ”. This is represented as  $\exists R_1[R_1(a, b) \wedge \mathbf{B}_a^* \lambda i [Q(\iota R_1(i, v))]]$ , where  $\exists R_n[R_n(x, y)]$  means there is an acquaintance relation  $R_n$  such that  $x$  is acquainted with  $y$  through  $R_n$  and  $\mathbf{B}_x^* \phi$  means  $x$  self-ascribes the property  $\phi$ . Here and henceforth I assume that everyone is acquainted with themselves (i.e.  $\forall x \exists R_1[R_1(x, x)]$ ) and that all relevant acquaintance relations with oneself are of identity (i.e. thoughts about oneself relevant for common ground are *de se*). *De se* belief is denoted as being about  $i, i', i''$  etc. In the rest of this paper I will abbreviate  $\iota R_1(i, v)$  (i.e. “the person that I am acquainted with”) as  $\iota R_1^i$ .

Iterative de re beliefs further complicate this picture. I adopt Maier’s [8] analysis of iterative *de re* attitudes where if  $a$  believes de re of  $b$  that  $b$  believes de re of  $c$  that she is  $Q$  (i.e. in the earlier notation  $\mathbf{B}_a\mathbf{B}_bQc$ ) then this entails that  $a$  is acquainted with both  $b$  and  $c$  and that  $a$  self-ascribes the property of being such that  $b$  is acquainted with  $c$  and that the former self-ascribes the property of being such that “the person I am acquainted with is  $Q$ ” (i.e.  $\exists R_1[R_1(a, b) \wedge \exists R_2[R_2(a, c) \wedge \mathbf{B}_a^*\lambda i[\exists R_3[R_3(\gamma R_1^i, \gamma R_2^i)] \wedge \mathbf{B}_{\gamma R_1^i}^*\lambda i'[Q(\gamma R_3^i)]]]]]$ ).

We can thus rewrite the common ground definition in terms of de re beliefs as follows (making the required acquaintance relations explicit):  $p$  is common ground between speaker  $a$  and addressee  $b$  iff:

$$\frac{\frac{\mathbf{B}_a^*\lambda i[p] \quad \mathbf{B}_b^*\lambda i[p]}{\exists R_1[R_1(b, a) \wedge \mathbf{B}_b^*\lambda i[\mathbf{B}_{\gamma R_1^i}^*\lambda i'[p]]]} \quad \exists R_1[R_1(a, b) \wedge \mathbf{B}_a^*\lambda i[\mathbf{B}_{\gamma R_1^i}^*\lambda i'[p]]]}{\exists R_1[R_1(a, b) \wedge \mathbf{B}_a^*\lambda i[\exists R_2[R_2(\gamma R_1^i, i) \wedge \mathbf{B}_{\gamma R_1^i}^*\lambda i'[\mathbf{B}_{R_2^i}^*\lambda i''[p]]]]] \quad \exists R_1[R_1(b, a) \wedge \mathbf{B}_b^*\lambda i[\exists R_2[R_2(\gamma R_1^i, i) \wedge \mathbf{B}_{\gamma R_1^i}^*\lambda i'[\mathbf{B}_{R_2^i}^*\lambda i''[p]]]]]}{\vdots \quad \vdots}$$

So, both speaker and addressee self-ascribe the property of being such that  $p$ ; both are acquainted with the other and self-ascribe the property of being such that the person they are acquainted with self-ascribes the property of being such that  $p$ , etc.

The above reformulation of the traditional definition makes explicit why de re common ground definitions do not apply to non-face-to-face communication. I suggest that there are three distinct types of non-face-to-face communication that we do intuitively describe in common ground terminology but that traditional definitions do not apply to. Namely, non-face-to-face conversations where [1] the speaker is known to the addressee but the addressee is not known to the speaker (e.g. Intuitively, it is common ground between Monk and myself that Wittgenstein was Austrian). Here the speaker  $a$  is not acquainted with the addressee  $b$  (i.e.  $\neg\exists R_1[R_1(a, b)]$ ). In a similar vein, traditional definitions do not apply to communication where [2] the addressee is known to the speaker but the speaker is not known to the addressee (e.g. Intuitively, it is common ground between you and the writer of the anonymous love letter that your voice is like the morning sun). Here the addressee is not acquainted with the speaker (i.e.  $\neg\exists R_1[R_1(b, a)]$ ). Nor do de re common ground definitions apply to communication where [3] neither conversational participant is known to the other (e.g. Intuitively, it is common ground between reviewer and author in a double blind peer review process that the submitted paper should not exceed 10,000 words).<sup>3</sup> In these cases neither speaker nor addressee is acquainted with the other (i.e.  $\neg\exists R_1[R_1(a, b)] \wedge \neg\exists R_2[R_2(b, a)]$ ). Hence the above iteration of de re beliefs cannot materialize in these cases and so there can exist no common ground between speaker and addressee (i.e. between Monk and his readership, between you and your admirer or between anonymous author and reviewer). This is unsatisfactory since these conversations *do* seem to involve producing and interpreting the relevant texts against a shared background.

The same problem arises in a generalized definition of common ground in terms of de re beliefs. If we make all acquaintance relations explicit such definitions would be rewritten as follows:  $p$  is common ground between all conversational participants in some community iff:

<sup>3</sup>Actually, these distinctions raise questions about what constitutes an acquaintance relation. Intuitively, I am acquainted with Monk, but not with the writer of the anonymous love letter because – even though I’ve never met either – I know Monk through his book, reading about him on Wikipedia, someone referring to him etc. and this is not true for the anonymous admirer. However, am I not in (an impoverished sense but in) essentially the same way also acquainted with the anonymous admirer (or with the anonymous reviewer) *through* the love-letter (or the review)? (See e.g. Jeshion [4], Recanati [10]) Maybe conversations of type [2] and [3] (where the addressee is not acquainted with the speaker) are in fact not possible. However, such a concession would not dissolve the problem with de re common ground; Discourse where the addressee is unknown (type [1]) is still possible.

$$\frac{\forall x(Cx \rightarrow \mathbf{B}_x^* \lambda i[p])}{\frac{\forall x \forall y(Cx \wedge Cy \rightarrow (\exists R_1[R_1(x, y) \wedge \mathbf{B}_x^* \lambda i[\mathbf{B}_{\gamma R_1^i}^* \lambda i'[p]]])}{\forall x \forall y \forall z(Cx \wedge Cy \wedge Cz \rightarrow (\exists R_1[R_1(x, y) \wedge \exists R_2[R_2(x, z) \wedge \mathbf{B}_x^* \lambda i[\exists R_3[R_3(\gamma R_1^i, \gamma R_2^i) \wedge \mathbf{B}_{\gamma R_1^i}^* \lambda i'[\mathbf{B}_{R_3^i}^* \lambda i''[p]]]]])])})}}{\vdots}$$

In words, everyone in the community self-ascribes the property of being such that  $p$ ; everyone in the community is acquainted with everyone in the community and self-ascribes the property of being such that the person they are acquainted with self-ascribes the property of being such that  $p$ ; etc. Here in conversations of type [1], [2] or [3] at least one of the conversational participants (speaker or one of the addressees) is not acquainted with at least one of the other conversational participants (i.e.  $Cx \wedge Cy \wedge \neg \exists R_1[R_1(x, y)]$ ) and hence there can be no common ground in the relevant community.

### 3 Redefining common ground

In this section I propose two potential fixes to make traditional common ground definitions applicable to both face-to-face conversations *and* conversations of types [1], [2] and [3]. They fall in line with Abelard's distinction between two types of generality (as discussed by Lewis [6]): *in sensu composito* or 'collective' (Section 3.1) and *in sensu diviso* or 'distributive' (Section 3.2).<sup>4</sup> If I believe a general rule *in sensu composito* then I have a general *de dicto* belief. For instance, if  $a$  considers all flowers to be pretty *in sensu composito*, then she believes 'that all flowers are pretty' (i.e.  $\mathbf{B}_a \forall x(Fx \rightarrow Px)$  where  $Fx$  and  $Px$  respectively mean  $x$  is a flower and  $x$  is pretty). This means that there may be flowers that  $a$  does not believe to be pretty (for instance because she fails to realize that they are flowers). Conversely, if I believe a general rule *in sensu diviso* then I have a general disposition to form singular *de re* beliefs in every relevant situation. For instance, if  $a$  considers all flowers to be pretty *in sensu diviso*, then she believes of every flower, if she sees it, that it is pretty (i.e.  $\forall x((Fx \wedge Sax) \rightarrow \mathbf{B}_a Px)$  where  $Sxy$  means  $x$  sees  $y$ )<sup>5</sup>. This means that  $a$  might not recognize every flower as a flower (might even lack the concept of 'flower' altogether) but still believe of every flower that she comes across that it is pretty.

#### 3.1 *In sensu composito* common ground

First, I will present the *in sensu composito* definition of common ground. An *in sensu composito* understanding of general thought by  $a$  about the mental states of conversational partners would be as follows:  $a$  believes (or self-ascribes the property of being such) that 'all conversational partners in the community believe that  $p$ ' (i.e. in our earlier notation:  $\mathbf{B}_a \forall x(Cx \rightarrow \mathbf{B}_x p)$ , in the present notation:  $\mathbf{B}_a^* \lambda i[\forall x(Cx \rightarrow \mathbf{B}_x^* \lambda i'[p])]$ ). This leads to the following definition of generalized common ground in terms of general *de dicto* belief:  $p$  is common ground between all conversational participants in some community iff:

<sup>4</sup>Bermúdez [1] construes from Braithwaite's [2] account of generality the following intermediate concept of general belief: if  $a$  believes that flowers are beautiful then  $a$  believes of every flower that she sees and considers to be a flower, that it is pretty (i.e.  $\forall x((Fx \wedge Sax \wedge \mathbf{B}_a Fx) \rightarrow \mathbf{B}_a Px)$ ). I leave exploration into the merits of extending this notion to a common ground definition for future research.

<sup>5</sup>Contrary to Bermúdez [1] and Meggle [9]), I represent the fact that  $a$  has a *disposition* to form *de re* beliefs by a conditional: if  $a$  is in the relevant situation (e.g. sees a flower), then  $a$  forms the appropriate beliefs.

$$\frac{\frac{\forall x(Cx \rightarrow \mathbf{B}_x^* \lambda i[p])}{\forall y(Cy \rightarrow \mathbf{B}_y^* \lambda i[\forall x(Cx \rightarrow \mathbf{B}_x^* \lambda i'[p])])}}{\forall z(Cz \rightarrow \mathbf{B}_z^* \lambda i[\forall y(Cy \rightarrow \mathbf{B}_y^* \lambda i'[\forall x(Cx \rightarrow \mathbf{B}_x^* \lambda i''[p])])])}} \\ \vdots$$

In words, everyone in the community believes that  $p$ ; everyone in the community believes that everyone in the community believes that  $p$ ; etc. Hence  $p$  can be common ground in a community even though nobody has any de re beliefs about anyone. All that is required is that people have appropriate beliefs about what ‘everyone in the community’ believes.

Likewise, an *in sensu composito* understanding of general thought about conversational participants would lead to the following definition of common ground between one speaker and one addressee in terms of singular de dicto belief.<sup>6</sup>  $p$  is common ground between speaker  $a$  and addressee  $b$  iff (where  $Sx$  means  $x$  is a speaker and  $Ax$  means  $x$  is an addressee):

$$\frac{\frac{\frac{\mathbf{B}_a^* \lambda i[p] \quad \mathbf{B}_b^* \lambda i[p]}{\mathbf{B}_b^* \lambda i[\mathbf{B}_{\lambda x[Sx]}^* \lambda i'[p]] \quad \mathbf{B}_a^* \lambda i[\mathbf{B}_{\lambda x[Ax]}^* \lambda i'[p]]}}{\mathbf{B}_a^* \lambda i[\mathbf{B}_{\lambda x[Ax]}^* \lambda i'[\mathbf{B}_{\lambda x[Sx]}^* \lambda i''[p]]] \quad \mathbf{B}_b^* \lambda i[\mathbf{B}_{\lambda x[Sx]}^* \lambda i[\mathbf{B}_{\lambda x[Ax]}^* \lambda i''[p]]]}}{\vdots \quad \vdots}}$$

So, both speaker and addressee believe that  $p$ ; the addressee believes that ‘the speaker’ believes that  $p$ ; the speaker believes that ‘the addressee’ believes that  $p$ ; etc. Again,  $p$  can be common ground without addressee or speaker forming any de re beliefs. All that is required is that they have the appropriate beliefs about what ‘the speaker’ or ‘the addressee’ believes.

### 3.2 *In sensu diviso* common ground

Next, I turn to the *in sensu diviso* version of a common ground definition. To formulate the *in sensu diviso* definition we need to rewrite in conditional (or *in sensu diviso*) form the relational analysis of de re attitudes. In words, if  $a$  believes *in sensu diviso* of  $b$  that he is  $Q$ , then *if* there is an acquaintance relation from  $a$  to  $b$ , then  $a$  believes (or self-ascribes the property of being such) that the person he knows through this acquaintance relation is  $Q$ . As a type of donkey sentence this is translated as  $\forall R_1[R_1(a, b) \rightarrow \mathbf{B}_a^* \lambda i[Q_{\lambda R_1^i}]]$ . An *in sensu diviso* understanding of a general thought by  $a$  about the mental states of conversational partners would be as follows: It is true of all conversational partners in some community that if  $a$  is in a relevant situation with a conversational participant, then  $a$  believes of this person that he believes that  $p$  (i.e. in our earlier notation:  $\forall x(Cx \wedge Rax \rightarrow \mathbf{B}_a \mathbf{B}_x p)$  where  $Rxy$  means  $x$  is in a relevant situation with  $y$ ). I assume that ‘the relevant situations’ are situations where an acquaintance relation obtains<sup>7</sup> so that it is true of all conversational partners in some community that if there is an acquaintance relation from  $a$  to the conversational participant, then  $a$  believes that the person she is acquainted with believes that  $p$ . As a type of donkey sentence this gets translated as  $\forall x(Cx \rightarrow \forall R_1[R_1(a, x) \rightarrow \mathbf{B}_a^* \lambda i[\mathbf{B}_{\lambda R_1^i}^* \lambda i'[p]]])$ . This leads to the following definition of generalized common ground in terms of conditional de re belief:<sup>8</sup>  $p$  is common ground between all conversational participants in some community iff:

<sup>6</sup>We derive these iterations from the general common ground definition because  $a$  and  $b$  are conversational participants (i.e.  $Ca$  and  $Cb$ ),  $a$  believes that ‘the addressee’ is a conversational participant (i.e.  $\mathbf{B}_a^* \lambda i[C_{\lambda x[Ax]]}$ ),  $b$  believes that ‘the speaker’ is a conversational participant (i.e.  $\mathbf{B}_b^* \lambda i[C_{\lambda x[Sx]]}$ ), etc.

<sup>7</sup>Arguably, one can put constraints on the relevant types of acquaintance relations (e.g.  $a$  is acquainted with a conversational participant ‘in the context of conversation’). I leave this to further research.

<sup>8</sup>Meggle [9] alludes to an alternative ‘*in sensu diviso*’ version of a generalized common ground definition that boils down to the original definition in terms of de re beliefs. This version does not seem to do justice to the

$$\frac{\forall x(Cx \rightarrow \mathbf{B}_x^* \lambda i[p])}{\frac{\forall x, y((Cx \wedge Cy) \rightarrow \forall R_1[R_1(x, y) \rightarrow \mathbf{B}_x^* \lambda i[\mathbf{B}_{\gamma R_1^i}^* \lambda i'[p]])]}{\forall x, y, z((Cx \wedge Cy \wedge Cz) \rightarrow \forall R_1, R_2[(R_1(x, y) \wedge R_2(x, z)) \rightarrow \mathbf{B}_x^* \lambda i[\forall R_3[R_3(\gamma R_1^i, \gamma R_2^i) \rightarrow \mathbf{B}_{\gamma R_1^i}^* \lambda i'[\mathbf{B}_{\gamma R_3^i}^* \lambda i''[p]]]])]} \\ \vdots}$$

So, everyone in the community believes that  $p$ ; for everyone in the community it is true that for everyone in the community, if the one were acquainted with the other, then the one would believe that the person they are acquainted with believes that  $p$ ; etc. So, again, no de re beliefs are required for  $p$  to be common ground in a community, merely that people would form the appropriate de re beliefs about one another if they were acquainted (and would believe that the others would as well).

Alternatively, common ground between speaker  $s$  and addressee  $a$  is defined as follows:<sup>9</sup>  $p$  is common ground between speaker  $a$  and addressee  $b$  iff:

$$\frac{\frac{\mathbf{B}_a^* \lambda i[p] \quad \mathbf{B}_b^* \lambda i[p]}{\forall R_1[R_1(b, a) \rightarrow \mathbf{B}_b^* \lambda i[\mathbf{B}_{\gamma R_1^i}^* \lambda i'[p]]] \quad \forall R_1[R_1(a, b) \rightarrow \mathbf{B}_a^* \lambda i[\mathbf{B}_{\gamma R_1^i}^* \lambda i'[p]]]}}{\forall R_1[R_1(a, b) \rightarrow \mathbf{B}_a^* \lambda i[\forall R_2[R_2(\gamma R_1^i, i) \rightarrow \mathbf{B}_b^* \lambda i[\forall R_3[R_3(\gamma R_1^i, i) \rightarrow \mathbf{B}_{\gamma R_1^i}^* \lambda i'[\mathbf{B}_{\gamma R_2^i}^* \lambda i''[p]]]])] \quad \forall R_1[R_1(b, a) \rightarrow \mathbf{B}_b^* \lambda i[\forall R_2[R_2(\gamma R_1^i, i) \rightarrow \mathbf{B}_{\gamma R_1^i}^* \lambda i'[\mathbf{B}_{\gamma R_2^i}^* \lambda i''[p]]]])]} \\ \vdots \quad \vdots}$$

So, both speaker and addressee believe that  $p$ ; if the addressee were acquainted with the speaker, then the addressee would believe that the person he is acquainted with believes that  $p$ ; if the speaker were acquainted with the addressee, then the speaker would believe that the person he is acquainted with believes that  $p$ ; etc. So  $p$  can be common ground between speaker and addressee even though neither has any de re beliefs about the other. All that is required is that they would form the right de re beliefs about each other if they were acquainted (and would believe the other would do so as well).

## 4 Deriving de re beliefs

The above definitions are supposed to be general definitions of common ground that apply to all four types of communication (i.e. [1], [2], [3] and face-to-face communication) where one or more conversational participants may form iterative de re beliefs about the other. In the following two subsections I explore whether we can derive the appropriate de re beliefs in all four types of communication from the *in sensu composito* and from the *in sensu diviso* definition. I show that we *can* (albeit in slightly different ways for the two definitions) if we assume that in case a conversational participants is ‘known’, this party’s identity is common ground (e.g. It is common ground between speaker  $a$  and addressee  $b$  that  $a = \iota x[Sx]$ ). Since this is a de re belief about  $a$  that requires an acquaintance relation (i.e.  $\exists R_1[R_1(b, a)]$ ), we can derive the appropriate iterative de re beliefs. The iteration of de re beliefs of face-to-face communication (See p. 3) turns out to be a limit case where the identities of both conversational participants is known. For reasons of space I only show how this works for definitions of common ground between an addressee and a speaker.

fact that a general *in sensu diviso* belief involves a *disposition* to form de re beliefs. Moreover, it suffers from the problems described in section 2 and is hence not the most useful version of an *in sensu diviso* definition.

<sup>9</sup>We arrive at these iterations from the general common ground definition because  $a$  and  $b$  are conversational participants (i.e.  $Ca$  and  $Cb$ )



#### 4.1 *In sensu composito*

First, I will show how the different types of communication fit in the *in sensu composito* definition of common ground. In communication of type [3], where neither conversational participant is known (e.g. double blind peer review), I take the beliefs of the participants to be properly described by the ‘bare’ iteration of de dicto beliefs of the *in sensu composito* definition on p. 5, i.e. neither addressee nor speaker has any de re beliefs about the other, nor believes the other to have these. Addressee and speaker merely have beliefs about the mental state of ‘the speaker’ or ‘the addressee’.

Conversations of type [1] and [2] are different because the identity of either the speaker or the addressee is known. Below I focus on conversations of type [1] (where the speaker is known) but since conversations of type [2] are the exact mirror image of these conversations, the discussion applies to both. For instance, in the case of me (Semeijn) reading Monk’s biography it is common ground between Monk  $m$  and Semeijn  $s$  that Monk is the speaker (i.e.  $m = \iota x[Sx]$ ). For this de re belief to be *in sensu composito* common ground there must be an appropriate acquaintance relation between addressee and speaker (i.e. Semeijn must be acquainted with Monk to have a de re belief about him:  $\exists R_1[R_1(s, m) \wedge \mathbf{B}_s^* \lambda i[\iota R_1^i = \iota x[Sx]]]$ ) and conversational participants must believe that ‘the speaker’ and ‘the addressee’ have the appropriate acquaintance relations in order to have this de re belief about the speaker (e.g. Monk believes that ‘the addressee’ is acquainted with him and that ‘the addressee’ believes that the person she is acquainted with is ‘the speaker’:  $\mathbf{B}_m^* \lambda i[\exists R_1[R_1(\iota x[Ax], i) \wedge \mathbf{B}_{\iota x[Ax]}^* \lambda i'[\iota R_1^{i'} = \iota x[Sx]]]]$ ). See appendix A for details. If we assume that we believe the logical consequences of our beliefs we can from this, and the information that some proposition  $p$  is *in sensu composito* common ground, derive the following iteration of beliefs that incorporates de re beliefs about the speaker Monk (but not about the addressee Semeijn):

$$\frac{\frac{\mathbf{B}_m^* \lambda i[p] \quad \mathbf{B}_s^* \lambda i[p]}{\exists R_1[R_1(s, m) \wedge \mathbf{B}_s^* \lambda i[\mathbf{B}_{\iota R_1^i} \lambda i'[p]]]} \quad \mathbf{B}_m^* \lambda i[\mathbf{B}_{\iota x[Ax]}^* \lambda i'[p]]}{\mathbf{B}_m^* \lambda i[\exists R_1[R_1(\iota x[Ax], i) \wedge \mathbf{B}_{\iota x[Ax]}^* \lambda i'[\mathbf{B}_{\iota R_1^i}^* \lambda i''[p]]]]} \quad \frac{\mathbf{B}_s^* \lambda i[p] \quad \mathbf{B}_m^* \lambda i[\mathbf{B}_{\iota x[Ax]}^* \lambda i'[p]]}{\exists R_1[R_1(s, m) \wedge \mathbf{B}_s^* \lambda i[\mathbf{B}_{\iota R_1^i}^* \lambda i'[\mathbf{B}_{\iota x[Ax]}^* \lambda i''[p]]]]}}{\vdots \quad \vdots}$$

The above iterations represent the mental states of Monk and myself, i.e. speaker and addressee in conversations of type [1]; both Monk and Semeijn believe that  $p$ ; Semeijn is acquainted with Monk and believes that the person she is acquainted with believes that  $p$ ; Monk believes that ‘the addressee’ believes that  $p$ , etc. So, I do have de re beliefs about Monk’s beliefs (and Monk believes that ‘the addressee’ has these) but Monk only has beliefs about ‘the addressee’s’ beliefs.

Lastly, to arrive from this at the original iteration of de re beliefs that properly describes the mental states of people in face-to-face communication (See p. 3), we only need to assume that it is also *in sensu composito* common ground that I am the addressee (i.e.  $s = \iota x[Ax]$ ).

#### 4.2 *In sensu diviso*

Now that I have shown how conversations of type [1], [2], [3] and face-to-face conversations fit into the *in sensu composito* definition of common ground, I will show how they fit into the *in sensu diviso* common ground definition. As in the *in sensu composito* definition, I take the mental states of people engaging in communication of type [3] (where neither conversational participant is known) to be properly described by the ‘bare’ iteration of dispositions to form de re beliefs of the *in sensu diviso* definition on p. 6, i.e. neither addressee nor speaker has any de re



beliefs about the other, nor believes the other to have these. Contrary to the *in sensu composito* definition, the *in sensu diviso* definition also doesn't require the conversational participants to have any de dicto beliefs about 'the speaker's' or 'the addressee's' mental state but instead requires them to have the proper dispositions to form de re beliefs about each other.

As for the other types of communication, in the case of me reading Monk's biography (type [1]) it is again common ground that Monk is the speaker (i.e.  $m = \iota x[Sx]$ ). In order for this de re belief to be *in sensu diviso* common ground there must be an appropriate acquaintance relation between Semeijn and Monk (i.e.  $\exists R_1[R_1(s, m) \wedge \mathbf{B}_s^* \lambda i[\iota R_1^i = \iota x[Sx]]]$ ) and Semeijn and Monk must have the appropriate dispositions to form de re beliefs about the other being appropriately acquainted (e.g. If Monk were acquainted with Semeijn, then Monk would believe that this person is acquainted with him and that this person believes that the person that she is acquainted with is 'the speaker':  $\forall R_1[R_1(m, s) \rightarrow \mathbf{B}_m^* \lambda i[\exists R_2[R_2(\iota R_1^i, i) \wedge \mathbf{B}_{\iota R_1^i}^* \lambda i'[\iota R_2^{i'} = \iota x[Sx]]]]]$ ). See appendix A for details. From this, and the information that some proposition  $p$  is *in sensu diviso* common ground, we can derive the following iteration of beliefs that incorporates de re beliefs about the speaker Monk (but not about the addressee Semeijn):

$$\frac{\frac{\mathbf{B}_m^* \lambda i[p] \quad \mathbf{B}_s^* \lambda i[p]}{\exists R_1[R_1(s, m) \wedge \mathbf{B}_s^* \lambda i[\mathbf{B}_{\iota R_1^i}^* \lambda i'[p]]]} \quad \forall R_1[R_1(m, s) \rightarrow \mathbf{B}_m^* \lambda i[\mathbf{B}_{\iota R_1^i}^* \lambda i''[p]]]}{\forall R_1[R_1(m, s) \rightarrow \mathbf{B}_m^* \lambda i[\forall R_2[R_2(\iota R_1^i, i) \rightarrow \mathbf{B}_{\iota R_1^i}^* \lambda i'[\mathbf{B}_{\iota R_2^{i'}}^* \lambda i''[p]]]]]} \quad \frac{\exists R_1[R_1(s, m) \wedge \mathbf{B}_s^* \lambda i[\forall R_2[R_2(\iota R_1^i, i) \rightarrow \mathbf{B}_{\iota R_1^i}^* \lambda i'[\mathbf{B}_{\iota R_2^{i'}}^* \lambda i''[p]]]]]}{\mathbf{B}_{\iota R_1^i}^* \lambda i'[\mathbf{B}_{\iota R_2^{i'}}^* \lambda i''[p]]]} \quad \mathbf{B}_{\iota R_1^i}^* \lambda i'[\mathbf{B}_{\iota R_2^{i'}}^* \lambda i''[p]]]}{\vdots \quad \vdots}$$

In words, both Monk and Semeijn believe that  $p$ ; Semeijn is acquainted with Monk and believes of the person she is acquainted with that he believes that  $p$ ; if Monk were acquainted with Semeijn, then Monk would believe that the person he is acquainted with believes that  $p$ ; etc. So, I do have de re beliefs about Monk's beliefs (and Monk would believe that I have these if he knew me) but Monk only has a disposition to form de re beliefs about me.

Again, to arrive from this at the original iteration of de re beliefs (See p. 3) we only need to assume that it is also *in sensu diviso* common ground that I am the addressee (i.e.  $s = \iota x[Ax]$ ). In fact – unlike in the *in sensu composito* definition where it needs to be common ground who 'the addressee' and 'the speaker' are – it is enough to assume that  $a$  de re belief about speaker and addressee is common ground to arrive at the original iteration of de re beliefs; This ensures that speaker and addressee are acquainted with each other and have the appropriate dispositions to form beliefs about conversational participants being acquainted with each other. This contrast in derivations reflects the central difference between the two notions of common ground relevant for deciding between the definitions in the next section.

## 5 The shy acquaintance

Now that I have presented two possible strategies to improve upon traditional definitions of common ground, we can try to decide between them. Most importantly, we can compare how well they fit our intuitive understanding of common ground. Lewis [6] has argued that a general rule is convention only if people believe it *in sensu diviso*, i.e. only if people respond in accordance with the rule in all relevant instances (whether they are aware of this or not). Similarly, one could argue that  $q$  is only truly common ground when people form the appropriate de re beliefs about conversational participants whenever they are acquainted with them. However, it is not obvious that Lewis' reasoning extends to the concept of common ground.

To clearly see the difference between the two definitions suppose Thea has a weekly blog on architecture that her favourite nephew Nick is an avid reader of. However, afraid of being asked for his opinion, Thea's nephew has not yet revealed to his aunt that he is a reader of her blog, i.e. a conversational participant. Suppose Thea's latest blogpost contained the information that *The Sims* was originally designed as an architecture simulator ( $q$ ). According to the *in sensu diviso* definition of common ground it is *not* common ground between Thea  $t$  and her nephew  $n$  that  $q$  – Thea is acquainted with the addressee but does not believe of him that he believes that  $q$  (i.e.  $\exists R_1[R_1(t, n) \wedge \neg \mathbf{B}_t^* \lambda i[\mathbf{B}_{\gamma R_1}^* \lambda i'[q]]]$ ). According to the *in sensu composito* definition, it *is* common ground between Thea and her nephew that  $q$  – Thea has the appropriate beliefs about what ‘the addressee’ believes (i.e.  $\mathbf{B}_t^* \lambda i[\mathbf{B}_{\gamma x[Ax]}^* \lambda i'[q]]$ ) and in this situation the addressee is her nephew (even though Thea does not recognize him as such).

An anonymous reviewer has judged a comparable case to form an argument in favour of the *in sensu diviso* definition; If Thea would meet her nephew at her husband's birthday party, surely she would not base the production of her speech acts on the assumption that  $q$  is shared background knowledge. For instance, she would not say something like “I gave my husband that architecture simulation game” since she would not expect her nephew to understand what game she is talking about. Similarly, the linguistic behaviour of the nephew will be as if he has no idea what game Thea is talking about (since he does not want to expose himself as a reader).

I agree with the reviewer's judgement of the above example conversation but argue that in this conversation it would also not be common ground that  $q$  on an *in sensu composito* understanding, i.e. during the birthday conversation Thea does not believe that ‘the addressee’ (of that conversation) believes that  $q$ . The difference between the *in sensu composito* and the *in sensu diviso* definition in fact only comes out in conversations where Thea believes that ‘the addressee’ (of that conversation) believes that  $q$ . So the relevant situation to consider is one where Thea is writing another blogpost; Here Thea believes that ‘the addressee’ believes that  $q$  (because the previous post contained this information) and she is acquainted with one of her readers (i.e. her nephew) but does not believe of him that he believes that  $q$ . Although intuitions may vary concerning this and related cases, I take these considerations to form a prima facie argument for the *in sensu composito* definition of common ground, i.e. in this situation it *is* common ground between Thea and her nephew that  $q$ . For instance, Thea could write “I gave my husband that architecture simulation game” in such a situation since she expects ‘the reader’ to understand what she is talking about. Intuitively, people's linguistic behaviour will depend on their beliefs about ‘the speaker's’ or ‘the addressee's’ beliefs, not on what they *would* believe about conversational participants if they were acquainted with them.

## 6 Conclusions

This paper proposes two novel refinements of traditional definitions of common ground in terms of de re attitudes in order to make them applicable to non-face-to-face communication. These fall in line with Abelard's distinction between generality *in sensu composito* (i.e. speaker and addressee have iterative beliefs about the mental states of ‘the speaker’ and ‘the addressee’) and *in sensu diviso* (i.e. speaker and addressee would form iterative de re beliefs about the mental states of speaker and addressee if they were acquainted with them). I have shown how the four distinguished types of communication fit in either definition taking face-to-face communication as a limit case. Lastly, I have argued that the case of the shy acquaintance forms a prima facie argument in favour of an *in sensu composito* definition.

## References

- [1] Jose Luis Bermudez. *Thinking without words*. OUP USA, 2003.
- [2] Richard Bevan Braithwaite. The nature of believing. *Proceedings of the Aristotelian society*, 33(1):129–146, 1932.
- [3] Bart Geurts. Communication as commitment sharing: speech acts, implicatures, common ground. *Theoretical linguistics*, 45(1-2):1–30, 1997.
- [4] Robin Jeshion. Singular thought: Acquaintance, semantic instrumentalism, and cognitivism. In Robin Jeshion, editor, *New essays on singular thought*, pages 105–141. OUP, 2010.
- [5] David Kaplan. Quantifying in. *Synthese*, 19(1-2):178–214, 1968.
- [6] David Lewis. *Convention: A philosophical study*. Blackwell, 1969.
- [7] David Lewis. Attitudes de dicto and de se. *The philosophical review*, 88(4):513–543, 1979.
- [8] Emar Maier. Iterated de re: A new puzzle for the relational report semantics. In Arndt Rieger and Torgrim Solstad, editors, *Proceedings of Sinn und Bedeutung 13*, pages 347–355, 2009.
- [9] Georg Meggle. Common belief and common knowledge. In Matti Sintonen, Petri Ylikoski, and Kaarlo Miller, editors, *Realism in action: Essays in the philosophy of the social sciences*, pages 251–258. Springer, Dordrecht, 2003.
- [10] François Recanati. Singular thought: In defense of acquaintance. In Robin Jeshion, editor, *New essays on singular thought*, pages 141–188. OUP, 2009.
- [11] Stephen Schiffer. *Meaning*. Clarendon Press, 1972.
- [12] Robert Stalnaker. Pragmatics. *Synthese*, 22(1-2):272–289, 1970.
- [13] Robert Stalnaker. Common ground. *Linguistics and philosophy*, 25(5-6):701–721, December 2002.
- [14] Andreas Stokke. Lying and Asserting. *Journal of Philosophy*, 110(1):33–60, 2013.

## A Appendix

It is *in sensu composito* common ground between  $m$  and  $s$  that  $m = \iota x[Sx]$  iff

$$\begin{array}{c}
 \frac{\mathbf{B}_m^* \lambda i [i = \iota x[Sx]] \quad \exists R_1 [R_1(s, m) \wedge \mathbf{B}_s^* \lambda i [\iota R_1^i = \iota x[Sx]]]}{\exists R_1 [R_1(s, m) \wedge \mathbf{B}_s^* \lambda i [\exists R_2 [R_2(\iota x[Sx], \iota R_1^i) \wedge \mathbf{B}_{\iota x[Sx]}^* \lambda i' [\iota R_2^i = \iota x[Sx]]]]]} \quad \frac{\mathbf{B}_m^* \lambda i [\exists R_1 [R_1(\iota x[Ax], i) \wedge \mathbf{B}_{\iota x[Ax]}^* \lambda i' [\iota R_1^i = \iota x[Sx]]]]}{\mathbf{B}_m^* \lambda i [\exists R_1 [R_1(\iota x[Ax], i) \wedge \mathbf{B}_s^* \lambda i [\exists R_2 [R_2(\iota x[Sx], \iota R_1^i) \wedge \mathbf{B}_{\iota x[Ax]}^* \lambda i' [\exists R_3 [R_3(\iota x[Ax], \iota R_2^i) \wedge \mathbf{B}_{\iota x[Sx]}^* \lambda i'' [\iota R_3^{i''} = \iota x[Sx]]]]]]]} \\
 \vdots \quad \vdots
 \end{array}$$

It is *in sensu diviso* common ground between  $m$  and  $s$  that  $m = \iota x[Sx]$  iff

$$\begin{array}{c}
 \frac{\mathbf{B}_m^* \lambda i [i = \iota x[Sx]] \quad \exists R_1 [R_1(s, m) \wedge \mathbf{B}_s^* \lambda i [\iota R_1^i = \iota x[Sx]]]}{\forall R_1 [R_1(s, m) \rightarrow \mathbf{B}_s^* \lambda i [\mathbf{B}_{\iota R_1^i}^* \lambda i' [i' = \iota x[Sx]]]]} \quad \frac{\exists R_1 [R_1(m, s) \rightarrow \mathbf{B}_m^* \lambda i [\exists R_2 [R_2(\iota R_1^i, i) \wedge \mathbf{B}_{\iota R_1^i}^* \lambda i' [\iota R_2^i = \iota x[Sx]]]]]}{\forall R_1 [R_1(m, s) \wedge \mathbf{B}_m^* \lambda i [\forall R_2 [R_2(\iota R_1^i, i) \rightarrow \mathbf{B}_{\iota R_1^i}^* \lambda i' [\exists R_3 [R_3(\iota R_2^i, i') \wedge \mathbf{B}_{\iota R_2^i}^* \lambda i'' [\iota R_3^{i''} = \iota x[Sx]]]]]]]} \\
 \vdots \quad \vdots
 \end{array}$$

# Singular/Plural contrasts: The case of Informational Object Nouns\*

Peter R. Sutton and Hana Filip

Heinrich-Heine-Universität, Düsseldorf  
peter.sutton@hhu.de  
hana.filip@hhu.de

## Abstract

This paper brings together a number of different topics in semantics: the count/mass distinction, telicity, plurality, and cumulativity and distributivity. We focus on sentences containing Informational Object Nouns (IONs) such as *statement* and *belief* in constructions such as *Alex's/Alex and Billie's statement(s)/belief(s) that p/that p and q*. We observe that the interpretation of the ION in such constructions *vis-à-vis* referring to a proposition/propositions or to an eventuality/eventualities is sensitive to whether the subjective genitive and the ION are singular or plural, and to whether the complement clause is simplex ( $p$ ) or complex ( $p$  and  $q$ ). We derive these patterns based on a theory of the mass/count distinction for IONs developed in Sutton and Filip (2019).

## 1 Introduction.

This paper contributes to the small but growing literature on the countability of abstract nouns (Grimm, 2014; Nicholas, 2010; Tovena, 2001; Zamparelli, 2018, amongst others). Abstract nouns have largely been set aside in semantic theories of the mass/count distinction due to their highly puzzling nature and heterogeneity (Grimm, 2014; Zamparelli, 2018). Therefore, our modest focus is on one lexical subclass, which we label *Informational Object Nouns* (IONs) (following Sutton and Filip, 2019). IONs take propositional complements and accept *is true/false* predications. Many IONs are derived from Psych verbs (*knowledge, belief, thought, opinion*), and from verbs that denote speech act events (*statement, assertion, utterance*). The former also have another sense/use that refers to a stative *relation* to a proposition (e.g., that of believing it, knowing it etc.), which we model as denoting a set of STATES. The latter have an additional sense/use that denotes a set of dynamic EVENTS of stating/asserting etc.

We argue that count IONs can be individuated in different ways, depending on whether the eventuality type specified in their lexical entries is a STATE or an EVENT. In particular, we discuss interactions between count IONs, the grammatical number of a head NP (the ION) and of its modifier, and the complexity of its propositional complement, namely, cases in which count IONs such as *belief* and *statement* are singular or plural, have singular or plural subjective genitives, and can take simplex or complex propositional complements:

	Subj. genitive (SG or PL)	ION (SG or PL)	Prop. complement (simplex or complex)
(1)	Alex's	belief/statement	that $\phi$
	Alex and Billie's	beliefs/statements	that $\phi$ and $\psi$

---

\*We would like to thank our anonymous reviewers and the audience at Universität Wien who heard and gave helpful comments on some of the early ideas presented here. We would especially like to thank Viola Schmitt and Nina Haslinger for their invaluable input. Errors that remain are our own. This research was funded by the DFG project *CRC 991: The Structure of Representations in Language, Cognition, and Science*.

These cases are of particular interest, since, as we will show, these alternations dictate whether a proposition or an eventuality is referred to, whether there is a plurality thereof or not, and how this distributes to agents/experiencers. Such alternations, we will argue, can be predicted by a theory of the mass/count distinction based upon individuation relative to a context as proposed by Sutton and Filip (2019) and Filip and Sutton (2017), amongst others.

## 2 Diagnostics for Informational Object Nouns (IONs)

The class of IONs, in English, include the following nouns:

*assertion, belief, contention, evidence, fact, idea, information, intelligence<sup>1</sup>, judgement, knowledge, opinion, proclamation, pronouncement, proposal, proposition, statement, story, thought, truth, utterance.*

We propose that the class of IONs consists of nouns that pass *both* the tests T1 and T2 below. No other nouns do. So, if N is an ION, then:

(T1) *N that is true/false* is felicitous (truth-evaluability);

(T2) *N that p* is felicitous (propositional complementisers)

For example, *belief* and *statement* take propositional complements: *Alex's belief/statement that it's raining*, and admit of truth/falsity predications: *Alex's belief/statement was true/false*. In contrast, nouns like *feeling* pass the complementiser test (T2): *the feeling that I have forgotten something*. However, they fail test (T1): *that feeling was true* is odd, if *true* is intended in its truth-value sense, and not in the *genuine, real* sense. Although concrete Ns like *book, article* are attested in collocations such as *this article is true* (understood as meaning that its content is true at a given world/time), they fail test (T2).

## 3 Observations

First, as shown by the examples in (2) and (3) IONs do not uniquely determine what counts as one proposition across contexts. In the original corpus examples in (2-a) and (3-a), the singular ION denotes a proposition expressed by a complement clause which is a conjunction of two clauses, but nonetheless they constitute what counts as 'one' proposition in the denotation of the ION. In the minimally modified examples in (2-b) and (3-b), the use of plural IONs individuates each proposition expressed by the two separate conjuncts as *two opinions/statements*

- (2) a. ... the **opinion** that these two German countries belonged together and that the German people should solve their own internal affairs and difficulties. [UKwaC]  
 b. ... the **opinions** that these two German countries belonged together and that the German people should solve their own internal affairs and difficulties.
- (3) a. The Panel is pleased to note the company's **statement** that the product is no longer available and that it would not form part of its Christmas 2001 gift range. [UKwaC]  
 b. The Panel is pleased to note the company's **statements** that the product is no longer available and that it would not form part of its Christmas 2001 gift range.

<sup>1</sup>*Intelligence* in the sense of pieces of military intelligence

In (3-a) and (3-b), the sentences can also denote the stating events: one stating event in (3-a) and possibly multiple stating events in (3-b). However, in (2-a) and (2-b) no parallel reference to opinion-states (i.e. to multiple opinion-states by diverse ‘opinion-holders’) is possible.

Second, we get a meaning contrast for plural subjective genitives when the complement clause is simplex (without a conjunction). Take the minimal pair in (4). In (4-a), *belief* refers to, and is individuated in terms of, a single proposition. But, in the same context, the use of the plural *beliefs* in (4-b) forces a reading in which relations to the same proposition are individuated in terms of the Experiencers (the different cousins).

- (4) a. it certainly fueled my cousins’ **belief** that my family were “snobs”.  
 b. it certainly fueled my cousins’ **beliefs** that my family were “snobs”. [UKWaC]

For IONs such as *statement*, we get a different alternation. In (5-a), unlike in (4), reference can be made either to the proposition or to the stating-event. In (5-b), similarly to (4-b) it is the stating events that are referred to (reference to the proposition expressed is blocked). However, in (5-a), Franks and Vershbow are joint agents in making the statement, whereas in (5-b), the only reading is that across differentiable events, Franks and Vershbow both made statements that conveyed the same contents (either individually or together).

- (5) a. in the wake of US Gen Tommy Franks and US ambassador Alexander Vershbow’s **statement** that the US would produce the evidence of Iraqi WMD.  
 b. in the wake of US Gen Tommy Franks and US ambassador Alexander Vershbow’s **statements** that the US would produce the evidence of Iraqi WMD. [BNC]

Third, we find contrasts such as those in (6) and (7). In (6-b), the plural *statements* allows for different officials each to have stated both or either one of the propositions denoted by the (complex) complement clause, whereas the singular *statement* in (6-a) suggests that the officials jointly stated the same (complex) proposition.

- (6) a. there was no credible evidence to support the Bush administration officials’ **statement** that Iraq had stockpiles of biological and chemical weapons and was close to having a nuclear weapon.  
 b. there was no credible evidence to support the Bush administration officials’ **statements** that Iraq had stockpiles of biological and chemical weapons and was close to having a nuclear weapon.

In (7-b), some complainants might believe that the ads were irresponsible, some that the ads could encourage emulation by children, and others both of these things. In contrast, in (7-a), where we have *belief* as a singular noun, all of the complainants believe the same thing, namely, that the ads were irresponsible and could encourage emulation by children.

- (7) a. We did not agree with the complainants’ **belief** that the ads were irresponsible and could encourage emulation by children  
 b. We did not agree with the complainants’ **beliefs** that the ads were irresponsible and could encourage emulation by children [UKWaC]

The contrast between (6-a) and (6-b), on the one hand, and (7-a) and (7-b), on the other, is that, while all can refer to the proposition(s) expressed by the complement clause, (6-a) and (6-b) can also refer to a single stating-event or multiple stating events respectively. In contrast, the IONs in (7-a) and (7-b) cannot refer to the belief-states of the experiencers.

All of these observations are summarised in Table 1. The general pattern is that STATE-IONs such as *belief* can only refer to the proposition(s) that are, e.g., believed, unless the ION is plural with a plural subjective genitive and a simplex complement (4-b) in which case reference is to STATES individuated relative to experiencers. In contrast, EVENT-IONs generally allow for readings that denote the relevant EVENT(S) or denote the relevant proposition(s). Again, if the

Table 1: Summary of observations

Type	Subj. Gen.	ION	Complement	Denotation of ION	Ex. No.
EVENT	SG	SG	complex	1 proposition / 1 event	(3-a)
STATE	SG	SG	complex	1 proposition (no state reading)	(2-a)
EVENT	SG	PL	complex	$n$ propositions / $n$ events	(3-b)
STATE	SG	PL	complex	$n$ propositions (no state reading)	(2-b)
EVENT	PL	SG	simple	1 proposition / 1 event	(5-a)
STATE	PL	SG	simple	1 proposition (no state reading)	(4-a)
EVENT	PL	PL	simple	$n$ events (no proposition reading)	(5-b)
STATE	PL	PL	simple	$n$ states (no proposition reading)	(4-b)
EVENT	PL	SG	complex	1 joint agency event / 1 proposition	(6-a)
STATE	PL	SG	complex	1 proposition (no state reading)	(7-a)
EVENT	PL	PL	complex	1 joint agency event or $n$ events / 1 or $n$ propositions	(6-b)
STATE	PL	PL	complex	1 or $n$ propositions (no state reading)	(7-b)

EVENT-ION is plural with a plural subjective genitive and a simplex complement, reference to propositions is blocked (5-b), in which case reference is only to EVENTS.

## 4 Analysis

### 4.1 Plural sets and individuation schemas across semantic types

We assume Schmitt’s (2013; 2017) theory for a type of generalised mereological sum operation that applies to e.g., propositions and eventualities. Schmitt defines a unified mereological sum ( $\sqcup$ ) operation over different semantic types. For each domain of type  $a$ ,  $\mathcal{D}_a$ , we have a bijection function  $pl_a$  on the powerset of  $\mathcal{D}_a$  to a plural structure, namely, the set of singularities and pluralities for that domain,  $\mathbf{PL}_a$  (the inverse of  $pl_a$  is  $pl_a^{-1}$ ):

$$pl_a : (\mathcal{P}(\mathcal{D}_a) \setminus \{\emptyset\}) \rightarrow \mathbf{PL}_a \quad (8)$$

Predicates with domains on  $\mathbf{PL}_a$  are members of the power set of this domain:  $\mathbf{S}_a = \mathcal{P}(\mathbf{PL}_a)$ , namely sets of sets of singularities and pluralities for that domain.

To demonstrate why this is helpful, suppose we have three (atomic) possible worlds in our domain:  $w_1, w_2, w_3$ , and so three *atomic functions* of type  $\langle s, t \rangle$  characterised by the singleton sets  $\{w_1\}$ ,  $\{w_2\}$ , and  $\{w_3\}$ . The set of possible propositions is  $\mathcal{P}(\{\{w_1\}, \{w_2\}, \text{and } \{w_3\}\}) \setminus \{\emptyset\}$  which is isomorphic to  $\mathbf{PL}_{\langle s, t \rangle}$ . For example, for some  $\mathbf{p}, \mathbf{q}, \mathbf{r}$ ,  $pl_{\langle s, t \rangle}(\{w_1\}) = \mathbf{p}$ ,  $pl_{\langle s, t \rangle}(\{w_2\}) = \mathbf{q}$ , and  $pl_{\langle s, t \rangle}(\{w_3\}) = \mathbf{r}$ . Count and mass predicates can then be distinguished in terms of quantization by supposing that countable sets of propositions are quantized sets relative to the plural structure formed by  $pl_{\langle s, t \rangle}$ . The definition of a quantized set (Krifka, 1989) is:

$$QUA(X) \leftrightarrow \forall x, y [x \in X \wedge y \in X \rightarrow \neg x \sqsubset y] \quad (9)$$

For example, suppose that there are two predicates of propositions of type  $\langle \langle s, t \rangle, t \rangle$ ,  $P$  and  $Q$  such that  $pl_{\langle s, t \rangle}(P) = \{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$  and  $pl_{\langle s, t \rangle}(Q) = \{\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p} \sqcup \mathbf{q}\}$ . The former differs from the latter insofar as only the denotation of the former is quantized. Hence  $P$  is, by hypothesis, countable, whereas  $Q$  is not ( $Q$  is mass).

In order to model how what counts as one proposition can vary across contexts, we use Sutton and Filip’s (2019) notion of a context indexed individuation schema,  $\mathcal{Q}_c$  that applies to sets and returns maximally quantized subsets of entities that count as ‘one’ in that context and so can be arguments of cardinality functions. (See Krifka (1989) for *quantized*. See, Landman

(2011; 2016) for a related idea in terms of disjointness (maximally disjoint subsets). Cf. Filip’s (2008) use of maximalization.) *Maximally quantized subsets* are defined as follows:

$$Y \subseteq_{max.QUA} X \leftrightarrow Y \subseteq X \wedge QUA(Y) \wedge \forall Z[Z \supseteq Y \wedge Z \subseteq X \wedge QUA(Z) \rightarrow Z = Y] \quad (10)$$

Whereas quantized sets have no two members such that one is the proper part of the other, maximally quantized subsets of a set  $X$  are the largest subsets of  $X$  that are quantized.

Generalising Sutton and Filip’s (2019) definition of  $\mathcal{Q}_c$ ,  $\mathcal{Q}_c$  applies to an expression of any type and returns a maximally quantized subset thereof (a possibly different set depending on the value of the context,  $c$ ).

For all  $c$ , for all  $\tau \in \mathbf{type}$ ,  $\mathcal{Q}_c$  is a polymorphic function of type  $\langle \tau, \tau \rangle$

If  $X$  is a set, the members of which are of type  $a$ , then  $\mathcal{Q}_c(X) = Y$ , such that  $\{\mathbf{y} : \mathbf{y} = pl_a(y), y \in Y\} \subseteq_{max.QUA} \{\mathbf{x} : \mathbf{x} = pl_a(x), x \in X\}$

To build upon the example from above, for the set  $X = \{\{w_1\}, \{w_2\}, \{w_1, w_2\}\}$ , applying  $pl_s$  to the members of this set yields  $\{\mathbf{p}, \mathbf{q}, \mathbf{p} \sqcup \mathbf{q}\}$ . Since there are two maximally quantized subsets of this set,  $\{\mathbf{p}, \mathbf{q}\}$  and  $\{\mathbf{p} \sqcup \mathbf{q}\}$ , there are two distinguishable contexts  $c$  and  $c'$  such that:

$$\mathcal{Q}_c(X) = \{\{w_1\}, \{w_2\}\}; \quad \mathcal{Q}_{c'}(Y) = \{\{w_1, w_2\}\}$$

## 4.2 Lexical entries

Here we take our two main examples of the two types of count IONs, *belief* and *statement*. In parallel to proposals for the verbal predicates they are derived from, we assume that the lexical entries for these nouns includes a predicate of eventualities:  $\lambda s.belief(s)$  and  $\lambda e.statement(e)$ , respectively. (Where  $s$  is a variable over STATES and  $e$  is a variable over EVENTS). Nouns like *belief* also have an Experiencer thematic role ( $\lambda x.\lambda s.exp(s)(x)$ ), and nouns like *statement* have an Agent thematic role ( $\lambda x.\lambda s.agent(e)(x)$ ). Both *belief*-like and *statement*-like nouns have a Contents thematic role, ( $\lambda s.\lambda p.contents(s)(p)$  and  $\lambda e.\lambda p.contents(e)(p)$ , respectively). Following Sutton and Filip’s (2019) proposal for concrete count nouns, we assume that the lexical entries of count IONs include a context indexed individuation schema. For IONs, this applies to the Contents-specifying conjunct in the lexical entry such that it ensures the set of propositions denoted by the ION is quantized. So if  $\sigma$  is a constant that denotes either an EVENT, a sum of EVENTS, a STATE, or a sum of STATES,  $\lambda p.\mathcal{Q}_c(contents)(\sigma)(p)$  is a quantized set of propositions (or, more correctly, the set of members of  $\mathbf{PL}_{(s,t)}$  that are related to members of that set via  $pl_{(s,t)}$  is quantized).

The lexical entry for *belief* is given in (11). It is a function from experiencers to a function from propositions to a set of states. The plural *beliefs* is formed via upward closure under sum (encoded by the  $*$  operator). Assuming that semantic plural formation applies only to quantized predicates, and given that the set of belief states is not quantized (see section 4.3 for justification), the only place for the  $*$  operator to apply is on  $\lambda p.\mathcal{Q}_c(contents_c(s))(p)$  (i.e. allowing reference to pluralities of propositions).

$$\llbracket \text{belief} \rrbracket^c = \lambda x \lambda p \lambda s [belief(s) \wedge exp(s)(x) \wedge \mathcal{Q}_c(contents_c(s))(p)] \quad (11)$$

$$\llbracket \text{beliefs} \rrbracket^c = \lambda x \lambda p \lambda s [belief(s) \wedge exp(s)(x) \wedge * \mathcal{Q}_c(contents_c(s))(p)] \quad (12)$$

Given that the interpretations of *belief(s)* and *belief(s) that S* are the same type (both can denote propositions, both can combine with determiners), we assume the following semantics



for *that S* complements.  $f$  is of type  $\langle e, \langle \langle s, t \rangle, \langle v, t \rangle \rangle \rangle$ , where  $v$  is the type of eventualities and  $\sigma$  is a variable over eventualities (e.g., STATES or EVENTS).

$$\llbracket \text{that } S \rrbracket^c = \lambda f \lambda x \lambda p \lambda \sigma [f(x)(p)(\sigma) \wedge p = \llbracket S \rrbracket^c] \quad (13)$$

$$\llbracket \text{belief that } S \rrbracket^c = \lambda x \lambda p \lambda s [belief(s) \wedge exp(s)(x) \wedge \mathcal{Q}_c(contents_c(s))(p) \wedge p = \llbracket S \rrbracket^c] \quad (14)$$

In subjective genitive constructions, assuming an  $\iota$ -closure licensed by the genitive and  $\exists$ -closure of the remaining argument, this yields an ambiguity between a DP denoting a proposition or a belief-state:

$$\llbracket \text{Jo's belief that } S \rrbracket^c = \begin{cases} \iota p \exists s [belief(s) \wedge exp(s)(j) \wedge \mathcal{Q}_c(contents_c(s))(p) \wedge p = \llbracket S \rrbracket^c] \\ \iota s \exists p [belief(s) \wedge exp(s)(j) \wedge \mathcal{Q}_c(contents_c(s))(p) \wedge p = \llbracket S \rrbracket^c] \end{cases} \quad (15)$$

EVENT-IONs such as *statement* have similar lexical entries, but there is a key differences between the lexical entries of EVENT-IONs and STATE-IONs on top of those detailed above: plural formation for EVENT-IONs can apply with respect to propositions *or* EVENTS:

$$\llbracket \text{statement} \rrbracket^c = \lambda x \lambda p \lambda e [statement(e) \wedge agent(e)(x) \wedge \mathcal{Q}_c(contents_c(e))(p)] \quad (16)$$

$$\llbracket \text{statements} \rrbracket^c = \lambda x \lambda p \lambda e [*statement(e) \wedge agent(e)(x) \wedge *\mathcal{Q}_c(contents_c(e))(p)] \quad (17)$$

Expressions such as *Jo's statement/statements that S* are derived as above *mutatis mutandis*.

### 4.3 Further assumptions

We require eight further assumptions in order to derive the distributional patterns in Table 1. These are divided into three common knowledge assumptions (i)-(iii), two theoretical quantization assumptions (iv)-(v), two theoretical assumptions relating to individuation of propositions (vi)-(vii), and, finally, one theoretical assumption relating to cumulation (viii).

**(i) experience of a token mental state cannot be shared.** Aside from SciFi scenarios wherein alien beings have collective hive-mind consciousness, for the actual world as we know it, we assume that for any single mental state, such as a belief, there can be only one experiencer (not, for example, a sum of experiencers that have ‘co-experiencship’).

**(ii) agency of an event can be shared.** Agency of events such as making-a-statement events can be genuinely shared (e.g., via co-authorship of a written statement).

**(iii) contents of mental states can be shared.** Although there is some philosophical debate as to the extent to which the contents of the beliefs of any two experiencers can be identical, at the level of granularity of beliefs we assume here, we take it for granted that there is a genuine sense in which two experiencers can hold the same belief *qua* contents.

**(iv) sets of states such as  $\lambda s. belief(s)$  are not quantized relative to one experiencer.** States, such as belief-states hold true for experiencers at relatively long and vague time intervals. If an experiencer has a mental state (e.g. belief), with some propositional contents for some time interval, then they will be in that mental state at any subinterval and moment within that interval. For example, if Ann believes that it's raining during time interval  $i$ , then this belief persists at any subinterval and moment of  $i$ . Therefore the set of mental attitude states is not quantized, even if the experiencer and the content of these states is the same.

**(v) sets of events such as  $\lambda e. statement(e)$  are quantized relative to an agent, and a propositional contents.** Suppose that  $a$  states that  $p$  and that this eventuality,  $e$  has a run time  $t$ . There is no sub-eventuality in which  $a$  also states that  $p$ . (Of course, if there are sub-propositions of  $p$ , such as  $q$  and  $r$ , then there may be  $e', e'' \sqsubseteq e$  such that  $e'$  is a stating  $q$  event

and  $e''$  is a stating  $r$  event, mutatis mutandis for plural agents). In other words, predicates of accomplishments, restricted in this way are quantized.

**(vi) simplex complement clauses fix the context to one in which the proposition denoted counts as ‘one’.** If  $a$  states/believes that  $S$ , where  $S$  is a simplex sentence, then there is, in terms of logical possibility, no guarantee that what is expressed by  $S$  will count as one proposition. By assumption (§4.1), plural structures based on the domain of propositions are generated from atomic propositions (those that are equivalent to singleton sets of possible worlds), hence all but those sentences that express atomic propositions will express sums of elements of  $\mathbf{PL}_{\langle s,t \rangle}$  (which could count as one or as many propositions). However, the use of words to express a proposition intuitively *frames* that thought in some way, and simplex sentences such as ‘Alex is coming to the party’ seem to militate against individuating the propositions they express from counting as more than one: our choices to use certain words impose particular individuation schemas on that which we refer to. Formally, if  $S$  is a simplex sentence in a complement clause, we assume that the context of evaluation,  $c$  is such that:

$$\mu_{\#}(\llbracket S \rrbracket^c, \mathcal{Q}_c(\mathcal{P}(\mathcal{D}_{\langle s,t \rangle}) \setminus \emptyset)) = 1 \quad (18)$$

In words, the cardinality of the proposition(s) expressed by a simplex sentence  $S$  in context  $c$  is 1 relative to the domain of propositions under the individuation schema  $\mathcal{Q}_c$  (where  $\mathcal{Q}_c(\mathcal{P}(\mathcal{D}_{\langle s,t \rangle}) \setminus \emptyset)$  is a maximally quantized subset of the domain of (pluralities of) propositions).

**(vii) complex complement clauses are compatible with the proposition denoted counting as more than one.** The same restriction does not seem to apply to complex sentences. To take the sentence conjunction case, if  $a$  states/believes that  $S$  and  $S'$ , then the use of the conjunction seems to license a certain kind of conceptual freedom in whether we individuate what is expressed by ‘ $S$  and  $S'$ ’ as one proposition, or as two. Formally, then, if  $S$  is a complex sentence in a complement clause, then we assume that there is a licensed context of evaluation,  $c'$ , such that:

$$\mu_{\#}(\llbracket S \rrbracket^{c'}, \mathcal{Q}_{c'}(\mathcal{P}(\mathcal{D}_{\langle s,t \rangle}) \setminus \emptyset)) \geq 1 \quad (19)$$

**(viii) Propositions denoted by singular IONs do not cumulate with Agents/Experiencers.** Whether the *contents* conjunct is closed under sum as in (12) and (17) or is not as in (11) and (16) makes an important difference when a definite DP involving an ION refers to propositions. Just as singular definites like *the cat* denote single entities (atoms), when an ION is realised as a singular definite argument, the ION can only refer to a proposition that counts as ‘one ION’ (i.e. a single member of the set of propositions under  $\mathcal{Q}_c$ ). Thus, if the proposition that counts as ‘one ION’ is a sum of propositions in a given context, none of its proper parts can count as ‘one ION’ in that same context. Consequently, if the ION is in the definite singular, and the agent/experiencer is realised as a plural argument, individual agents/experiencers cannot be distributed over proper parts of the proposition. When the ION is plural, this is not the case, provided that the individual agents/experiencers are related, via believing/stating, to a proposition that is in the set of propositions that count as one or sums thereof in that context (i.e. cumulative readings are available).

#### 4.4 Deriving the available readings

**(A) SG IONs with PL genitives and simplex complement clauses (4-a, 5-a).** The fact that these constructions have simplex complement clauses ensures that the proposition

denoted by the complement clause counts as ‘one’ in the context, so on a definite interpretation, there is only a single proposition referred to (assumption vi). For STATE-IONs, and based on our common world knowledge that any single token mental state is experienced by only one experiencer (i.e., there is no ‘co-experience’ of a single token mental state by a sum of experiencers, assumption (i)), if the Experiencer argument is plural, we know that at least two belief states must be involved. However, the use of the singular *belief* (as opposed to the plural *beliefs*) in the subjective genitive construction in (4-a) militates against individuating propositions in terms of multiple belief states each tied to a different experiencer. This leaves only one interpretation: a definite DP denoting a single proposition, hence *my cousins’ belief that my family were snobs* denotes a single proposition that is the contents of some belief state that each of the cousins have. For EVENT-IONs, the plural Agent *can* be interpreted as multiple individuals with joint agency (assumption ii), and so there may be only a single event involved. Indeed, since the set of events that is denoted by the singular *statement* is quantized (assumption v), the definite DP *Franks and Vershbow’s statement*, if it refers to an EVENT, can only refer to a single event with joint agency. On its proposition denoting reading, *Franks and Vershbow’s statement that the US would produce the evidence* can only be drawn from a quantized set of propositions that each count as ‘one’ in the context (assumption vi), and so, if it refers to a proposition, it refers only to one proposition.

**(B) PL IONs with PL genitives and simplex complement clauses (4-b, 5-b).** Excluding downwards entailing contexts, these constructions cannot refer to pluralities of propositions, because the simplex complement specifies a proposition that counts as ‘one’ in the context (assumption vi). Hence, it would be pragmatically infelicitous to refer to multiple propositions (just as it is to say *There are cats are on the mat* when only one cat is). For STATE-IONs, *beliefs* should be able to denote pluralities of propositions (12), but the simplex complement clause rules this out (assumption vi). We, furthermore, have a plurality of experiencers, and we know that they cannot share a belief state (assumption i), although they can all share the contents of a belief (assumption iii). The pragmatic effect of the use of the plural is to try to identify some kind of plurality of things in the context. The only option left, therefore, is to try to individuate belief states. Problematically, however, this set is not quantized for STATE-IONs (assumption iv), and so no clear individuation criteria is provided by the meaning of *beliefs*. The solution, we propose, is to derive a quantized set of, e.g., belief states via anchoring each state to an experiencer (see Davidson (1969/1980) and Krifka (1989) for individuating eventualities with respect to participants, see Grimm (2014) for anchoring to experiencers). In other words, one infers that the speaker is referring to a plurality of mental states, each possessed by a different experiencer. Hence *my cousins’ beliefs that my family were snobs* has only one reading: reference to multiple belief states that are individuated with respect to individual cousins, all of whom share the content of a belief. EVENT-IONs also cannot refer to pluralities of propositions in these constructions for the same reasons just given for STATE-IONs. However, they can, straightforwardly, refer to pluralities of EVENTS as licensed by the fact that singular EVENT-IONs denote quantized sets of eventualities (assumption v) and plural EVENT-IONs denote pluralities of eventualities that are generated from quantized sets e.g., *\*statement* in (17). Hence *Franks and Vershbow’s statements that the US would produce the evidence* can only denote a plurality of stating-events, the propositional content of each is the same.

**(C) SG IONs with PL genitives and complex complement clauses (6-a, 7-a).** Even though these constructions have complex complements, because the ION is singular, the propositions that can be referred to count as ‘one’ and it is not possible for single Agents/Experiencers to believe/state only proper parts of the relevant proposition (assumption viii). The use of the singular form also implicates that what is being referred to is not a plurality of either proposi-

tions or eventualities. For STATE-IONs, a STATE referring reading is out, since the set of states is not generated from a quantized set (assumption iv). This just leaves one reading: reference to a single proposition (albeit one expressed by a complex complement clause), that is the contents of each of the mental states of the experiencers. Hence, *the complainants' belief that the ads were irresponsible and could encourage emulation by children* refers to the proposition expressed by the complement and frames it as counting as one. For EVENT-IONs, since the set of EVENTS is quantized (assumption v), there is a single EVENT-referring reading available, albeit one with joint agency (assumption ii). The single (albeit complex) proposition-referring reading is also available. Unlike the *belief*-case, this must be linked to a single stating-event. Hence, *the officials' statement that Iraq had weapon stockpiles and was close to having a nuclear weapon* refers either to the proposition expressed by the complement and frames it as counting as one, or refers to a single joint-agency event of stating this complex proposition.

**(D) PL IONs with PL genitives and complex complement clauses (6-b, 7-b).**

The use of the plural noun in these cases implicates that some plurality of things is being referred to. The difference between the interpretations of plural STATE-IONs and plural EVENT-IONs plays a role: the former licence the application of the \*-operator only on  $\lambda s.Q_c(\text{contents}(s))$ , the latter on both  $\lambda e.Q_c(\text{contents}(e))$  and e.g.,  $\lambda e.\text{statement}(e)$ . For STATE-IONs, there is a plurality of experiencers, and a plurality of propositions generated from a quantized set of propositions (by assumption (vii) the proposition expressed by the complex complement can count as more than one for the purposes of counting). As before, the set of STATES is not quantized (assumption iv). Given this, such DPs can refer to pluralities of propositions, but there is no impetus to try to derive a quantized set of states via anchoring states to experiencers. Hence, *the complainants' beliefs that the ads were irresponsible and could encourage emulation by children* denotes a plurality of propositions, e.g.,  $pl_{(s,t)}(\llbracket\text{the ads were irresponsible}\rrbracket) \sqcup pl_{(s,t)}(\llbracket\text{the ads could encourage emulation by children}\rrbracket)$ , and it suffices for each of these propositions to be believed by at least one of the complainants and for each of the complainants to believe at least one of those propositions (a cumulative reading). For EVENT-IONs, we have not only pluralities of propositions generated from a quantized set, but also pluralities of EVENTS generated from a quantized set (see 17). This allows for either of two different cumulative readings: individuation in terms of the propositions or individuation in terms of the events. Hence *the officials' statements that Iraq had weapon stockpiles and was close to having a nuclear weapon* can refer to either a plurality of propositions, each of which is the contents of at least one statement made by at least one official (such that each of the officials made a statement, individually or jointly, the contents of which was one of the propositions), or to a plurality of EVENTS, each of which was made by at least one of the officials and expressed a least one of the propositions (such that each of the propositions was stated by at least one of the officials).

Although, due to lack of space, we do not explicitly derive such cumulative truth conditions, we suspect an approach that closely tracks that of Schmitt and Haslinger (Haslinger and Schmitt, 2019; Schmitt, 2019) for sentences involving propositional attitude verbs such as *Alex, Billie and Charlie believe that p, q, and r* would be fruitful.

## 5 Concluding remarks

This paper brings together a number of different major topics in semantics: the count/mass distinction, telicity, plurality, and cumulativity and distributivity. Interestingly, the understudied class of abstract nouns seems to be a rich vein in which to explore the interactions of these topics. In order to try to make sense of the data, we have argued that a couple of theoretical

ingredients are crucial. First, that individuation, grounded in an extensional property such as quantization, must be sensitive to context such that what counts as one proposition or eventuality in some contexts may count as many in others. Second, that the aspectual class of verbs from which abstract nouns are derived impacts individuation. Put roughly, for STATE-IONS, if a proposition denoting reading is available, it will be the default reading. Individuating states (done indirectly via, e.g., Experiencer anchoring) will be a strategy of last resort. In contrast, for EVENT-IONS, if the combination of the number marking on the ION and the complexity of the complement allows it, we can as happily individuate the events as we can the propositions involved. Third, indeed the complexity of propositional complements seems to play a huge role in how we individuate propositions and eventualities denoted by IONS. This is surprising, from a semantic point of view, since the distinction cast here between simplex and complex complements is fundamentally one of form: a simplex complement can express the same proposition *qua* set of possible worlds as a complex one, and yet whether we view this proposition as one or many in the context of use turns on the words that we use, not merely on what they mean.

## References

- Davidson, D. (1969/1980). “The individuation of events”. *Essays in honor of Carl G. Hempel*. Springer, pp. 216–234.
- Filip, H. (2008). “Events and maximalization: The case of telicity and perfectivity”. *Theoretical and crosslinguistic approaches to the semantics of aspect*. Ed. by S. Rothstein. John Benjamins Publishing, pp. 217–256.
- Filip, H. and P. R. Sutton (2017). “Singular Count NPs in Measure Constructions”. *SALT* 27, pp. 340–357.
- Grimm, S. (2014). “Individuating the abstract”. *Proceedings of Sinn und Bedeutung* 18. Ed. by U. Etcheberria et al., pp. 182–200.
- Haslinger, N. and V. Schmitt (2019). “Asymmetrically distributive universals and plural projection”. Manuscript.
- Krifka, M. (1989). “Nominal Reference, Temporal Constitution and Quantification in Event Semantics”. *Semantics and Contextual Expression*. Ed. by R. Bartsch, J. van Benthem, and P. van Emde Boas. Foris, pp. 75–115.
- Landman, F. (2011). “Count Nouns – Mass Nouns – Neat Nouns – Mess Nouns”. *The Baltic International Yearbook of Cognition* 6, pp. 1–67.
- (2016). “Iceberg Semantics for Count Nouns and Mass Nouns: Classifiers, measures and portions”. *The Baltic International Yearbook of Cognition* 11, pp. 1–48.
- Nicholas, D. (2010). “Towards a semantics for mass expressions derived from gradable nominals”. *Recherches linguistiques de Vincennes* 39.10, pp. 163–198.
- Schmitt, V. (2013). “More pluralities”. PhD thesis. University of Vienna.
- (2017). “Cross-categorial plurality and plural composition”. manuscript.
- (2019). “Pluralities in intensional contexts: Cumulativity and individuation”. Presented at SinFonIJA 2019.
- Sutton, P. R. and H. Filip (2019). “Informational Object Nouns and the mass/count distinction”. Presentation at Sinn und Bedeutung 24.
- Tovena, L. M. (2001). “Between Mass and Count”. *Proceedings of WCCFL* 20. Ed. by K. Megerdooimian and L. Bar-el, pp. 565–578.
- Zamparelli, R. (2018). “Countability Shifts and Abstract Nouns”. *Mass and Count in Linguistics, Philosophy, and Cognitive Science*. Ed. by F. Moltmann and L. Tovena. Benjamins (To appear).

# Existential Semantics in Equatives in Japanese and German\*

Eri Tanaka<sup>1</sup>, Kenta Mizutani<sup>2</sup>, and Stephanie Solt<sup>3</sup>

<sup>1</sup> Osaka University, Osaka Japan  
eri-tana@let.osaka-u.ac.jp

<sup>2</sup> Osaka University, Osaka Japan  
l.g.fuad0809@gmail.com

<sup>3</sup> Leibniz-Centre General Linguistics (ZAS), Berlin, Germany  
solt@leibniz-zas.de

## Abstract

This paper investigates the semantics of two lesser studied equative markers, Japanese *hodo* and German *dermaßen*, focusing on accounting for their polarity sensitivity and their presuppositions. We present an analysis according to which these items have weak existential semantics, producing a trivial meaning in certain configurations, and where presuppositional effects derive from competition with the unmodified positive form.

## 1 Introduction

Cross-linguistic variation in the semantics of equative constructions has been the subject of considerable recent interest (see e.g. [3],[12],[13],[16]). We contribute to this body of research with an investigation of the Japanese equative marker *hodo* and the German *dermaßen*, which have the interesting property that they are polarity sensitive in some but not all of their uses, a pattern that has not to our knowledge been previously observed. In what follows, we present the relevant data, develop an analysis that derives the polarity sensitivity of these items from a semantics based on existential quantification over degrees, and discuss how the analysis can be refined to account for the presuppositions of *hodo* and *dermaßen*.

## 2 Data

### 2.1 Japanese *hodo*

The examples in (1)-(2) illustrate a use of *hodo* that corresponds to English ‘*as . . . as*’, where (1) features a phrasal standard and (2) a clausal standard. Here *hodo* appears to be a negative polarity item, being grammatical in negative sentences but not their positive counterparts. In positive contexts, *hodo* must be replaced with another equative marker, *kurai*.

- (1) *Taro-wa Jiro-hodo se-ga \*takai/takaku-nai.*  
Taro-TOP Jiro-hodo height-NOM tall/tall-NEG

‘Taro \*is/is not as tall as Jiro.’

---

\*This work has been supported by “On Development of Logical Language and Mathematical Concepts”, Osaka University International Joint Research Program (A), (Principal Investigator: Yoichi Miyamoto). Additional support was provided by the German Science Foundation (DFG) via grant SO1157-1/2 to the third author. We thank the audience at LENLS16 as well as the three anonymous reviewers for the Amsterdam Colloquium for comments that have helped improve this work. Our thanks also to Luka Crnić for very helpful discussion.

- (2) *Taro-wa Jiro-ga nonda-hodo biiru-o \*nonda/noma-nakat-ta.*  
 Taro-TOP Jiro-NOM drank-hodo beer-ACC drank/drink-NEG-past  
 ‘Taro \*drank/didn’t drink as much beer as Jiro did.’

*Hodo*, however, is not a negative polarity item in a standard sense. First, it is licensed not only by negation in the matrix clause but also by negation in the clausal complement. In this case, the matrix predicate has to be affirmative, as shown in (3), and the sentence yields a comparative interpretation.

- (3) *Taro-wa [Jiro-ga noma-nakat-ta-hodo] (takusan) biiru-o nonda/\*noma-nakat-ta.*  
 Taro-TOP [Jiro-NOM drink-NEG-Past-hodo] (much) beer-ACC drank/drink-NEG-Past  
 (Lit.) ‘Taro drank as much beer as Jiro didn’t drink.’  
 ‘Taro drank more beer than Jiro did.’

Secondly, *hodo* can be licensed by negation in a higher clause, and this is possible in some constructions that do not license other polarity items such as *nidoto* ‘again’. For example:

- (4) a. [*Taro-ga Jiro-hodo se-ga takai to-iu-koto*]-*wa nai.*  
 [Taro-NOM Jiro-hodo height-NOM tall COMP-say-fact]-TOP NEG  
 ‘It is not the case that Taro is as tall as Jiro.’  
 b. ??[*Taro-ga nidoto kuru to-iu-koto*]-*wa nai.*  
 [Taro-NOM again come COMP-say-fact]-TOP NEG  
 Intended: ‘It is not the case that Taro will come again.’

But most importantly, there is a distinct use of *hodo* which corresponds more closely to English ‘*so... that*’, as illustrated in (5). On this use, it is not polarity sensitive, being acceptable in positive as well as negative sentences.

- (5) *Taro-wa basukettobooru senshu-ni nar-eru-hodo se-ga takai/takaku-nai.*  
 Taro-TOP basketball player-to become-can-hodo height-NOM tall/tall-NEG  
 ‘Taro is/is not so tall that he could become a basketball player.’

Thus *hodo* is quite unlike ‘ordinary’ polarity items, but instead displays an interesting and variable pattern of polarity sensitivity.

In addition to the above-described patterns of acceptability, sentences with *hodo* exhibit presuppositional effects (cf. [4], [10] on similar patterns with equative *kurai* and comparative *izyoo(-ni)*). Specifically, *hodo* on its ‘*as*’ use introduces norm-related presuppositions on both the standard of comparison and the subject. In an example such as (1), the standard – here, Jiro – must count as a clear case of ‘tall’; this explains why a *hodo* comparison to the 209 cm tall Giant Baba is felicitous, whereas comparison to the 145 cm tall Ikeno Medaka is odd.

- (6) *Taro-wa Giant Baba/#Ikeno Medaka-hodo se-ga takaku-nai.*  
 Taro-TOP Giant Baba/Ikeno Medaka-hodo height-NOM tall-NEG  
 ‘Taro is not as tall as Giant Baba/#Ikeno Medaka.’



Likewise, the subject must also count as ‘tall’: (1) conveys that Taro is tall but not as tall as Jiro, and would be infelicitous if Taro’s being tall were not already part of the common ground.

In the case of ‘*so*’-*hodo*, there is similarly a presupposition on the standard of comparison; thus (5) would be odd if ‘basketball player’ were replaced with ‘jockey’. But there is no presupposition on the subject; (5) could be felicitously uttered in a context where nothing was known about Taro’s height.

These presuppositional effects go hand in hand with other patterns that distinguish *hodo* from equative constructions such as English ‘*as ... as*’. In particular, *hodo* cannot be used to express crisp comparisons (see [8]): (1) would be infelicitous if Taro were just slightly (e.g. a few millimeters) shorter than Jiro. It furthermore does not allow proportional modifiers such as ‘twice as’. Taken together with the norm-related (evaluative) presuppositions, these properties suggest that *hodo* sentences can be aligned to Rett’s [13] class of implicit equatives.

## 2.2 German *dermaßen*

German has an equative marker *dermaßen* ‘to such an extent’ that patterns very similarly to Japanese *hodo*. In the construction *dermaßen ... wie* ‘to such an extent as’ it is sensitive to polarity. It is unacceptable in positive sentences, per (7a). But with sufficient contextual support it is acceptable (for most speakers we have consulted) in the corresponding negative sentence, per (7b). Even more acceptable, and frequently found in corpus data, are examples with a negative quantifier in the matrix clause or the *wie* complement, as in the naturally occurring examples in (8).

- (7) *Hans ist groß ...*  
Hans is tall ...
- a. \**Er ist (sogar) dermaßen groß wie sein Vater.*  
He is (even) *dermaßen* tall as his father.
- b. *Er ist (aber) nicht dermaßen groß wie sein Vater.*  
He is (but) not *dermaßen* tall as his father.
- (8) a. *Nirgends auf der Welt ist die Artenvielfalt dermaßen gross wie hier.*  
Nowhere in the world is the biodiversity *dermaßen* large as here.
- b. *Die Panik ist dermaßen gross wie noch nie zuvor.*  
The panic is *dermaßen* large as never before.

Like Japanese *hodo*, *dermaßen* has a second use *dermaßen ... dass* ‘to such an extent that’. Just as in the Japanese case, on this use it is not polarity sensitive:

- (9) *Hans ist (nicht) dermaßen groß, dass er Basketballspieler sein könnte.*  
Hans is (not) *dermaßen* tall that he basketball-player be-INF could
- ‘Hans is / isn’t so tall that he could be a basketball player.’

Like *hodo*, *dermaßen* sentences have presuppositions on both the standard of comparison and the subject. In the *wie* example in (7b), both Hans and his father must be tall; in *dass* examples such as (9), it is necessary that the complement clause introduce a standard that exceeds that of the positive form *groß* ‘tall’.



The German data are significant because they demonstrate that the patterns characterizing Japanese *hodo* do not derive from some idiosyncratic property of that language. Japanese and German degree constructions are markedly different in their syntax: *Hodo* sentences, like other Japanese comparative constructions, are formed with only a standard marker, i.e. an element that combines with the constituent that introduces the standard of comparison. *Dermaßen* sentences, like other German comparative constructions, feature both a parameter marker (*dermaßen*) that precedes the gradable expression as well as a standard marker (*wie* or *dass*). Some authors have argued that there are corresponding semantic differences, namely that Japanese differs from languages such as German in lacking degree abstraction (see e.g. [2]). That we observe such similar patterns of distribution and interpretation in two such diverse systems is evidence that these do not derive from properties specific to Japanese (or German) but rather have some more general source.

### 3 Proposal

#### 3.1 Background

The starting point for our analysis is a recent proposal by Crnič & Fox [3], according to which cross-linguistic variation in equative constructions can be related to the obligatory versus optional presence of a maximality operator in their semantics. Standard degree-based semantic analyses treat equative markers as degree quantifiers that express a relation between two maximum degrees, as in the following (cf. [1] and references therein):

- (10) Taro is as tall as Jiro.  
 $\max\{d : \text{Taro is } d\text{-tall}\} \geq \max\{d : \text{Jiro is } d\text{-tall}\}$

However, on the basis of differences in the behavior of equative constructions in English and Slovenian, Crnič & Fox argue that maximality is not an inherent component of the semantics of the equative. Rather, they propose, equative semantics derive from the presence of separate existential and maximality operators, the latter of which is optional in some languages (in particular Slovenian), but is inserted when needed to avoid a trivial meaning.

The crucial data are the following: Both English *as . . . as* and Slovenian *tako . . . kot* can be used with a positive clausal standard, as in (11). The English example is ungrammatical with negation in the clausal standard; but surprisingly, its Slovenian counterpart is acceptable.

- (11) a. John drove as fast [as Mary did / \*didn't].  
 b. *Janez se je peljal tako hitro [kot se je Marija (ni)].*  
 Janez self aux drive dem fast than self aux Mary (neg.aux)  
 'John drove as fast as Mary did(n't).'

Crnič & Fox propose that in both languages, the positive sentences involve a maximality operator, since otherwise the meaning would be trivial (there is always some degree  $d$  such that both John and Mary drove  $d$  fast). With negation in the standard clause maximality fails (there is no maximal degree  $d$  such that Mary didn't drive  $d$  fast). In English, maximality is obligatory, resulting in ungrammaticality. In Slovenian, however, maximality may be optionally omitted, allowing the negative version of (11b) to surface.

The central insight that we pursue here is that Japanese *hodo* and German *dermaßen* instantiate a third possibility: whereas maximality is mandatory in English and optional in Slovenian,

our claim is that *hodo* and *dermaßen* never introduce maximality, but rather have simpler existential semantics. Polarity-based distributional restrictions can then be shown to result from triviality of meaning.

In the next subsection we develop this analysis and apply it to Japanese *hodo*, and then briefly outline how it can be adapted to account for the different structure of German *dermaßen* constructions. In the following subsection we discuss how the analysis can be refined to also capture presuppositional effects, and other characteristics of implicit comparatives.

### 3.2 Basic analysis: explaining polarity sensitivity

To start, we assume that gradable predicates such as Japanese *se-ga takai* and German *groß* ‘tall’ relate individuals to degrees (as in [6]), and are monotonic, meaning that if Taro is 180 cm tall, he is also 170 cm tall, 160 cm tall, etc.:

$$(12) \quad \llbracket se-ga\ takai \rrbracket = \lambda d \lambda x. \mu_{HEIGHT}(x) \geq d$$

Starting with the Japanese case, we propose the following lexical entry for *hodo*, on which it takes as arguments a set of degrees  $D$ , a gradable predicate  $P$ , and an individual  $x$ , and introduces a variable over degrees  $d^*$  which is constrained to be an element of  $D$ , and which is subsequently existentially bound, per (13):

$$(13) \quad \llbracket hodo \rrbracket = \lambda D_{\langle dt \rangle} \lambda P_{\langle d, et \rangle} \lambda x. P(d^*)(x), \text{ where } d^* \in D$$

We apply this first to ‘*as*’-*hodo*, i.e. the use of *hodo* on which it may be paraphrased by English *as . . . as*. Consider first the ungrammatical positive sentence:

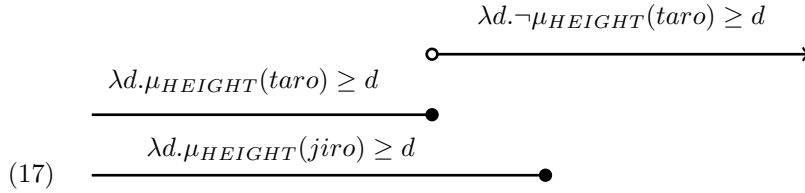
$$(14) \quad *Taro-wa \llbracket [Jiro-hodo] se-ga\ takai \rrbracket \quad \text{Intended: ‘Taro is as tall as Jiro.’}$$

Here the first argument of *hodo* is provided by the proper name *Jiro*. On the surface this is not of the right semantic type, being of type  $e$ , whereas *hodo* requires an argument of type  $\langle dt \rangle$ . Depending on one’s assumptions about the semantics of Japanese comparative constructions, the type mismatch might be resolved in one of two ways. As one option (cf. [5] on *yori* comparatives), an expression of type  $\langle dt \rangle$  might be contextually derived on the basis of the denotation of the complement of *hodo*, as in (15a). Alternately, we might take the standard in (14) to be covertly clausal (see again [1] and references therein), including an elided copy of the gradable predicate and null operator movement, as in (15b). Nothing in what follows depends crucially on the choice between these two options.

$$(15) \quad \begin{array}{l} \text{a. } f(\llbracket jiro \rrbracket) = \lambda d. \mu_{HEIGHT}(jiro) \geq d \\ \text{b. } \llbracket Op_i\ jiro\ t_i\ se-ga\ takai \rrbracket = \lambda d_i. \mu_{HEIGHT}(jiro) \geq d_i \end{array}$$

The following then presents the full derivation for (14). After existential closure over the variable  $d^*$ , the meaning we derive is that there is **some** degree of height that Jiro has that Taro also has. But with the monotonic semantics for *se-ga takai* ‘tall’ in (12), this meaning is entirely trivial: as illustrated in (17), there will always be some degree of height that the two individuals share. We take this to be the source of ungrammaticality.

- (16) a.  $\llbracket \text{jiro-hodo se-ga takai} \rrbracket = \lambda x. \mu_{\text{HEIGHT}}(x) \geq d^*$ ,  
 where  $\mu_{\text{HEIGHT}}(\text{jiro}) \geq d^*$
- b.  $\llbracket \text{taro-wa jiro-hodo se-ga takai} \rrbracket = \mu_{\text{HEIGHT}}(\text{taro}) \geq d^*$   
 After existential closure:  
 $\exists d^* : \mu_{\text{HEIGHT}}(\text{jiro}) \geq d^* [\mu_{\text{HEIGHT}}(\text{taro}) \geq d^*]$



In (18) and (19) we present the corresponding constituent structure and semantic interpretation for the negative version of (1).

- (18) Taro-wa  $\llbracket \text{Jiro-hodo} \rrbracket$  se-ga takaku-nai ‘Taro isn’t as tall as Jiro.’

- (19)  $\llbracket \text{taro-wa jiro-hodo se-ga takaku-nai} \rrbracket$   
 $= \exists d^* : \mu_{\text{HEIGHT}}(\text{jiro}) \geq d^* [\neg \mu_{\text{HEIGHT}}(\text{taro}) \geq d^*]$

Referring back to the illustration in (17), the effect of negation in the matrix clause is to invert the set of degrees it introduces. The sentence thus expresses a relation between an upper-bounded set of degrees (the set of Jiro’s heights) and a lower-bounded one (the set of heights that Taro doesn’t have). In this configuration, an ‘*as*’-*hodo* sentence is not trivial: (19) says that there is some degree of height that Jiro has that Taro **doesn’t** have, i.e. that Taro is shorter than Jiro.

Observe that in (19), existential closure takes scope over the negation operator introduced in the matrix clause. We assume that the opposite scope relationship is also in principle possible, but is blocked on account of triviality, being the negation of the trivially true (16b).

The analysis developed here also extends to clausal examples such as (2), with a similar choice regarding how to derive a first argument of the right semantic type for *hodo*. It can also capture cases involving negation in a higher clause, as in the earlier example (4a): *hodo* composes in situ and the composition proceeds as usual, with existential closure coming in at the end to bind the degree variable  $d^*$ .

We furthermore derive a prediction. Negation in the matrix clause had the effect of reversing the set of degrees it introduces, creating a configuration on which the resulting meaning is non-trivial. We then predict a parallel effect when negation is present in a clausal standard, such that it (rather than the matrix clause) introduces a lower-bounded set of degrees. This prediction is borne out, as illustrated by the previously discussed (3), which demonstrates that in the case of a negated clausal standard for *hodo*, it is the positive sentence that is grammatical, while the negated one is ill-formed.

We turn now to ‘*so*’-*hodo*, that is, the use of *hodo* on which it would be paraphrased with English *so ... that*. Here, we draw on Meier’s [11] analysis of *so ... that*, by taking the clausal complement of ‘*so*’-*hodo* to be covertly conditionalized, with the set of degrees derived as the standard of comparison being those degrees that are sufficient for the referenced state of affairs to obtain. In (5), whose structure is given in (20), the conditional proposition is as in (21a),

and the corresponding set of degrees is the set of heights that would be sufficient for one to be a basketball player, per (21b). Importantly, this set is lower bounded, as illustrated in (22); for example, if the minimum height to play basketball is 2 meters, the relevant set of degrees is  $\{d : d \geq 2m\}$ .

- (20) Taro-wa [[basukettobooru senshu-ni nar-eru-hodo] se-ga takai]  
‘Taro is so tall that he could become a basketball player.’

- (21) a. PRO is  $d$  tall in  $w \rightarrow$  PRO can <sub>$w,h$</sub>  become a basketball player in  $w$   
b.  $\lambda d. \text{sufficient-to-play-basketball}(d)$

$$(22) \quad \frac{\lambda d. \mu_{\text{HEIGHT}}(\text{taro}) \geq d}{\lambda d. \text{sufficient-to-play-basketball}(d)}$$

On this basis we derive the following as the interpretation for (20):

- (23)  $\exists d^* : \text{sufficient-to-play-basketball}(d^*)[\mu_{\text{HEIGHT}}(\text{taro}) \geq d^*]$

Crucially, (23) is not trivial but rather expresses the contingent proposition that Taro has some degree of height that would be sufficient for him to be a basketball player. In contrast to the case with ‘*as*’-*hodo* in a positive context, the sentence is therefore felicitous.

A ‘*so*’-*hodo* sentence can be felicitously negated, as in (24). Here in contrast to the case of negated ‘*as*’-*hodo* we take existential quantification to scope under negation, as in (25a). Just as before we assume that the opposite scope (25b) is also in principle possible, but here would result in a trivial meaning (trivially true, since assuming that Taro has finite height there will necessarily be some degree of height that he doesn’t have that would be sufficient to be a basketball player).

- (24) Taro-wa [[basukettobooru senshu-ni nar-eru-hodo] se-ga takaku-nai]  
‘Taro isn’t so tall that he could become a basketball player.’

- (25) a.  $\neg \exists d^* : \text{sufficient-to-play-basketball}(d^*)[\mu_{\text{HEIGHT}}(\text{taro}) \geq d^*]$  ✓  
b.  $\exists d^* : \text{sufficient-to-play-basketball}(d^*)[\neg \mu_{\text{HEIGHT}}(\text{taro}) \geq d^*]$  ✗

To summarize, the variable polarity sensitivity of *hodo* on its ‘*as*’ versus ‘*so*’ uses can be related to a difference between a standard that is an upper-bounded set of degrees and one that is a lower-bounded set.

To adapt this analysis to German *dermaßen*, it is necessary to factor in the distinct structure of comparison constructions in German versus Japanese. This might be done by analyzing *dermaßen* as a degree quantifier, similarly to standard treatments of degree morphemes such as comparative *-er*. The lexical entry is that in (26).

$$(26) \quad \llbracket \textit{derma\ss}en \rrbracket = \lambda D_{\langle dt \rangle} \lambda D'_{\langle dt \rangle} . D'(d^*), \text{ where } d^* \in D$$

On this interpretation *derma\ss*en undergoes quantifier raising for type reasons; degree abstraction in the *wie* or *dass* constituent provides its first argument, while abstraction over the type *d* trace in its base position provides the second argument. A simple (ungrammatical) positive *derma\ss*en . . . *wie* sentence thus has the LF in (27); the resulting interpretation is equivalent to that derived for the corresponding *hodo* sentence in (16b).

- (27) a. Hans ist *derma\ss*en gro\ss wie sein Vater.      Intended: ‘Hans is as tall as his father.’  
 b. [Derma\ss

en wie [2 [sein Vater ist t2 gro\ss]] [1 [Hans ist t1 gro\ss]]

With this modification, the account applied to the variable polarity sensitivity of Japanese *hodo* can be extended to German.

### 3.3 Refinement: accounting for norm-related presuppositions

The analysis developed in the preceding section provides an explanation for the variable polarity sensitivity of *hodo* and *derma\ss*en. However, we have not yet explained the presuppositional effects characterizing these items, nor the other similarities to implicit equatives. To briefly recap the relevant pattern, negated ‘*as*’-*hodo* introduces norm-related presuppositions on both the subject and the standard of comparison. By contrast, ‘*so*’-*hodo* has a presupposition on the standard but not on the subject. German *derma\ss*en behaves similarly.

Our proposal to account for these patterns is that the degrees over which existential quantification operates in *hodo* / *derma\ss*en sentences should be construed not simply as degrees but more specifically as possible thresholds  $\theta$  for the positive form of the gradable expression. *Hodo* and *derma\ss*en sentences are on this view a variety of positive constructions, and compete with the unmodified positive form, the result being presupposition-like effects. This approach is in the tradition of Simons [15], according to which presupposition-like interpretive patterns are analyzed as manner implicatures relative to simpler alternatives.

Focusing on the Japanese case for concreteness, the revised entry for *hodo* is (28), and the interpretation of a (grammatical) negative ‘*as*’-*hodo* sentence can be restated as in (29).

$$(28) \quad \llbracket \textit{hodo} \rrbracket = \lambda D_{\langle dt \rangle} \lambda P_{\langle d, et \rangle} \lambda x . P(\theta)(x), \text{ where } \theta \text{ is a threshold for } P \text{ and } \theta \in D$$

- (29) a. Taro-wa [[Jiro-hodo] se-ga takaku-nai]      ‘Taro isn’t as tall as Jiro.’  
 b.  $\exists \theta : \mu_{\text{HEIGHT}}(\textit{jiro}) \geq \theta [\neg \mu_{\text{HEIGHT}}(\textit{taro}) \geq \theta]$

Understanding *hodo* in this way first of all allows us to understand why it cannot be used to express crisp comparisons (see Section 2). Thresholds are inherently vague and context-dependent. Given that it is not possible to establish a sharp cut-off that divides individuals that are tall from those that are not, it is also not possible to establish such a precise threshold such that Jiro’s height is above it but Taro’s height (only a few millimeters shorter) is not.

As varieties of positive constructions, we propose that *hodo* sentences necessarily compete with the structurally simpler bare positive forms (cf. [7] on structurally determined alternatives). Following current practice, we analyze the latter as involving a null ‘positive’ morpheme *pos*, which introduces a contextually determined threshold  $\theta_c$ , as in (30). Thus we have (31) as the semantics of the simpler alternatives to *hodo* sentences.

$$(30) \quad \llbracket \textit{pos} \rrbracket = \lambda P_{\langle d, et \rangle} \lambda x . P(\theta_c)(x)$$

- (31) a. *Taro-wa se-ga takai / takaku-nai.*  
 Taro-TOP height-NOM tall / tall-NEG

‘Taro is / isn’t tall’

- b.  $\mu_{\text{HEIGHT}}(\text{taro}) \geq \theta_c / \neg \mu_{\text{HEIGHT}}(\text{taro}) \geq \theta_c$

Consider the ‘*as*’-*hodo* sentence in (29). If Jiro’s height were not distinct from the contextual standard for ‘tall’  $\theta_c$ , the truth conditions of the *hodo* sentence would be equivalent to those of the simpler bare positive form (i.e. the negated (31b)). If Jiro has a greater degree of height than  $\theta_c$ , then we can have a situation where (29) is true but (31b) is not. This is a situation where both Jiro and Taro have a greater degree of height than  $\theta_c$ , because in this case, the negated (31b) is false but (29) is true. Thus the standard-oriented inference regarding Jiro’s and Taro’s heights can be understood as accommodation of a situation in which the truth value of the sentence is distinct from that of the bare form. Put differently, the ‘*as*’-*hodo* sentence is blocked by the bare positive form except in the case that both Jiro and Taro have heights exceeding the contextual standard  $\theta_c$  (cf. [14] for a similar account of the implicatures of implicit comparatives formed with ‘compared to’).

A parallel explanation can be applied to the standard of comparison in ‘*so*’-*hodo* sentences: (20) is felicitous because ‘basketball player’ introduces a higher standard than simply ‘pos tall’; if this were not the case, the simpler positive form could have been used instead. But since the *hodo* sentence in this case produces a more informative assertion about the subject (Taro) than its simpler alternative, it is not blocked by the latter; there are therefore no presupposition-like effects with respect to the subject.

Here a question remains as to why a *hodo* sentence cannot be used to introduce a lower standard than the contextually salient one. That is, why is a very short individual infelicitous as a standard in an ‘*as*’-*hodo* sentence, and why is ‘become-a-jockey-hodo tall’ odd? While we do not have a full explanation for this, we note that a lower standard is possible when the *hodo* constituent is marked with the particle *mo* ‘even’:

- (32) *Taro-wa #Giant Baba/Ikeno Medaka-hodo-mo se-ga takaku-nai.*  
 Taro-TOP Giant Baba/Ikeno Medaka-hodo-even height-NOM tall-NEG

‘Taro isn’t even as tall as #Giant Baba / Ikeno Medaka’

Thus there appears to be an asymmetry between raising and lowering standards, with the latter being marked and needing to be overtly signaled. We also note similar patterns in other constructions that do not share the specific properties of *hodo* and *dermaßen* sentences (e.g. *John is so tall that he could be a ??jockey/basketball player*), suggesting that additional factors may be in play.

In concluding this section, we observe a connection to Klein’s [9] theory of comparatives, according to which *Taro is taller than Jiro* is analyzed essentially as expressing ‘there is some way of construing *tall* such that Taro is tall and Jiro is not tall’. The difference in the present case is that to say that ‘there is some way of construing *tall* such that both Taro and Jiro are tall’ is trivially true, resulting in ill-formedness.

## 4 Conclusions

We have shown that the distributional and interpretive effects characterizing *hodo* can be explained on the basis of a weak existential semantics, which yields a trivial interpretation in

certain configurations, coupled with pragmatic competition with the simpler positive form. Previous work by Crnić & Fox has shown that the obligatory versus optional presence of a maximality operator is a dimension along which the semantics of equative constructions may vary cross-linguistically. We have argued that Japanese *hodo* and German *dermaßen* instantiate a third possibility: these items never introduce maximality, the consequence being a more restricted and seemingly idiosyncratic distribution relative to better-studied equative markers. Our findings thus contribute to a fuller picture of variation in the semantics of degree constructions across languages.

## References

- [1] Beck, S.: Comparative constructions. In: Maienborn, C., von Stechow, C., Portner, P. (eds.) *Semantics: An international handbook of natural language meaning*, Vol. 2, pp. 1341–1389. Mouton de Gruyter, Berlin (2011)
- [2] Beck, S., Oda, T. and Sugisaki, K.: Parameteric variation in the semantics of comparison: Japanese vs. English. *Journal of East Asian Linguistics* **13**, 289–244 (2004)
- [3] Crnić, L., Fox, D.: Equatives and maximality. In: Altshuler, D., Rett, J. (eds.) *The Semantics of Plurals, Focus, Degrees, and Times*, pp. 163–184. Springer, Cham (2019)
- [4] Hayashishita, J.R.: *Izyoo(-ni)*-and *gurai*-comparatives: Comparisons of deviation in Japanese. *Gengo Kenkyu* **132**, 77–109 (2007)
- [5] Hayashishita, J.R.: *Yori*-comparatives: A reply to Beck et al. (2004). *Journal of East Asian Linguistics* **18**(2), 65–100 (2009) <https://doi.org/10.1007/s10831-009-9040-5>
- [6] Heim, I.: Degree operators and scope. In: Jackson, B., Matthews, T. (eds.) *Proceedings of the 10th Semantics and Linguistic Theory Conference (SALT10)*, pp. 60–84. CLC Publications, Ithaca, NY (2000). <http://dx.doi.org/10.3765/salt.v10i0.3102>
- [7] Katzir, R.: Structurally-defined alternatives. *Linguistics and Philosophy* **30**(6), 669–690 (2007). <https://doi.org/10.1007/s10988-010-9074-1>
- [8] Kennedy, C.: Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* **30**(1), 1–45 (2007). <https://doi.org/10.1007/s10988-006-9008-0>
- [9] Klein, E.: A semantics for positive and comparative adjectives. *Linguistics and Philosophy* **4**(1), 1–45 (1980). <https://doi.org/10.1007/BF00351812>
- [10] Kubota, Y.: The presuppositional nature of *izyoo(-ni)* and *gurai* comparatives: A note on Hayashishita (2007). *Gengo Kenkyu* **141**, 33–46 (2012)
- [11] Meier, C.: The meaning of *too*, *enough* and *so... that*. *Natural Language Semantics* **11**(1), 69–107 (2003). <https://doi.org/10.1023/A:1023002608785>
- [12] Penka, D.: Degree equatives - the same as comparatives? Workshop on Equative Constructions. University of Cologne (2016)
- [13] Rett, J.: Separate but equal: a typology of equative constructions. In: Hallman, P. (ed.) *Degree and quantification*. Brill, Leiden (in press)
- [14] Sawada, O.: Pragmatic aspects of implicit comparison: An economy-based approach. *Journal of Pragmatics* **41**(6), 1079–1103 (2009). <https://doi.org/10.1016/j.pragma.2008.12.004>
- [15] Simons, M.: On the conversational basis of some presuppositions. In: Capone, A., Lo Piparo, F., Carapezza, M. (eds.) *Perspectives on linguistic pragmatics*, pp. 329–348. Springer, Cham (2001/2013)
- [16] Umbach, C., Özge, U.: Scalar and non-scalar comparison across categories: The case of Turkish equatives. TbiLLC 2019 (2019)

# The role of “fake” past tense in acquiring counterfactuals\*

Maxime A. Tulling\*<sup>1</sup> and Ailís Cournane<sup>1</sup>

New York University Department of Linguistics, New York, NY, U.S.A.

\*corresponding author: [mat646@nyu.edu](mailto:mat646@nyu.edu)

## Abstract

Prior research on the acquisition of counterfactuals has not considered the mapping challenges associated with past tense morphology, which refers either to the actual past or a “fake” past expressing counterfactuality. In a corpus study of children’s spontaneous productions of counterfactual constructions, we found that *wish*-constructions are acquired before counterfactual conditionals, and that children make productive tense errors in counterfactuals producing present tense marking instead of past. These errors cease around the time children start producing *wish*-constructions that unequivocally display counterfactual reasoning, and could reflect a stage where children are still figuring out that counterfactual past tense does not signal a past event on the timeline, but rather a present non-actuality.

## 1 Introduction

In a distributional learning approach to the acquisition of semantic meaning (Landau et al., 2009), the syntactic context in which a word (e.g. an attitude verb like *think*) appears is taken to help learners bootstrap abstract semantic concepts and categories that are not directly observable in the world of reference (Hacquard and Lidz, 2018). However, morphosyntactic cues can be misleading. The past tense marking in counterfactual (CF) wishes (1) and conditionals (2) is considered to be “fake” (Iatridou, 2000), because this past tense form appears even when the temporal orientation of the embedded clause is present (1a/2a). Counterfactuals with true past temporal orientation require double past marking (1b/2b). This pattern is attested cross-linguistically (Bjorkman and Halpert, 2017; James, 1982).

- (1) a. I wish I **had** a car (right now/\*yesterday). PRESENT CF WISH  
b. I wish I **had** had a car back then. PAST CF WISH  
(Iatridou, 2000, 25-26)
- (2) a. If I **had** a car (right now/\*yesterday), I would drive. PRESENT CF CONDITIONAL  
b. If I **had** had a car back then, I would have driven. PAST CF CONDITIONAL  
(based on Ritter and Wiltschko, 2014, 62)

How do children figure out the counterfactual past tense is fake, and refers to a counterfactual world rather than an actual past world? This task is by no means trivial, from neither a linguistic nor cognitive perspective (Reilly, 1982). Prior research on the first language acquisition of counterfactuality shows that children start acquiring counterfactual constructions between age 3-5, after they have developed the ability to refer to hypothetical future events (e.g. Bowerman, 1986; Guajardo et al., 2009; Nyhout and Ganea, 2019; Reilly, 1982). This asymmetry between the acquisition of hypothetical future and counterfactuality has mainly

---

\*Thanks to Sam Mitchell and Stacy Gerchick for intercoder reliability and Sudha Arunachalam and Stephanie Harves for their comments and suggestions on an earlier version of this work. Also thanks to Annemarie van Dooren and Yu’an Yang for their help and support.



been attributed to the additional cognitive load demanded by counterfactual reasoning, which does not only depend on holding multiple possibilities in mind, but also requires considering a false possibility temporarily to be true (Beck et al., 2009; Byrne, 2007).

These studies do not consider grammatical complexity, but the mismatch between the morphological tense and the temporal orientation of the embedded clause poses a challenge to the grammatical mapping of counterfactual constructions. Rather than expressing regular past tense meaning, semantic accounts of counterfactuality assume that the past tense morpheme in counterfactual constructions plays a direct role in obtaining a counterfactual interpretation. Either by going back in time to gain access to an alternative world (Dudman, 1983; Ippolito, 2006; Ogihara, 2000; Romero, 2014) or by scoping over possible worlds rather than over time (Iatridou, 2000; Karawani and Zeijlstra, 2013; Ritter and Wiltschko, 2014). In both cases, acquiring counterfactual constructions requires the child to see through the “fakeness” of the past tense morphology, and learn to map this morpheme to a semantic operation supporting counterfactual meaning instead.

In this paper, we investigated the role of the “fake” past tense in the development of children’s natural productions of counterfactuals in English, searching through corpora of recorded and transcribed child’s speech. We investigated whether children go through a stage where they seem to think that the “fake” past tense is real. To our knowledge, no prior naturalistic study has looked specifically at counterfactuality in English, nor has the “fake” past tense been considered as an acquisition problem. Additionally, we examined whether counterfactual wishes are acquired before counterfactual conditionals. Although several studies included counterfactual conditionals in their overview of the acquisition of conditionals in general (Bowerman, 1986; Katis, 1997; Reilly, 1982), the development of counterfactual wishes has not yet been considered even though, unlike conditional structures, they are dedicated to express counterfactuality.

## 1.1 Hypotheses

Detecting that the past tense morpheme can be “fake” is not an obvious task. For example, past conditionals and counterfactual present/future counterfactual conditionals have identical antecedents (3), and would require the child to detect the *would* in the consequent in order to even start distinguishing it from other conditional structures.

- (3) a. If they went to the market, then they **would get** milk. CF CONDITIONAL  
 b. If they went to the market, then they **got** milk. PAST CONDITIONAL  
 (Karawani, 2014, 25)

While acquiring counterfactual wishes comes with its own challenges, i.e. the child will have to learn that desires expressed by *wish* are different from those expressed by *want*, we suspect that the *wish*-construction in English is easier to acquire than its conditional counterpart. Once the child has figured out *wish* differs from *want* in its counterfactual implication (4), the verb *wish* functions as a dedicated counterfactual marker that cannot occur with a present tense complement, even when the temporal orientation of the wish is present tense (5).

- (4) a. I live in Bolivia because I **want** to live in Bolivia.  
 b. \*I live in Bolivia because I **wish** *I lived*/to live in Bolivia. [cursive ours]  
 (Iatridou, 2000, 38/40)
- (5) a. \*I wish I have a car. PRESENT CF WISH  
 b. I wish I had a car. (Iatridou, 2000, 25)



was coded by the first author (non-native speaker). An inter-rater reliability analysis using the Kappa statistic was performed to determine consistency among raters in detecting errors and ‘clear’ counterfactual reasoning. A random subset of 100 child wishes were double-coded, by a native speaker of English. Since coding involves evaluation of the transcripts and assessments of grammatical and situational contexts, coders discussed all disagreements and came to a consensus for items where either coder missed contextual or grammatical cues in their original rating. A subset of disagreements remained where coders diverged on their answers and contextual cues could be interpreted in different ways. The inter-rater reliability was found to be  $\kappa = 0.69$  (substantial agreement) for judging clarity of counterfactual reasoning and  $\kappa = 0.85$  (nearly perfect agreement) for detecting children’s tense errors (Landis and Koch, 1977).

To get some more insight into the individual longitudinal development of children, we specifically looked at children that produced more than 10 wishes total. From the complete dataset of 68 different children, the data of 6 children, Abe - Kuczaj corpus (Kuczaj, 1977), Adam - Brown corpus (MacWhinney, 1991), Laura - Braunwald corpus (Braunwald, 1971), Mark & Ross - MacWhinney corpus (Brown, 1973) and Thomas - Thomas corpus (Lieven et al., 2009), remained with 175 wishes overall. For these 6 children we documented age of first wish, age of first presumed counterfactual wish (based on inference from the context), age of first definite counterfactual (based on lexicalized reality contrast), age of first tense error and age of last tense error. Additionally, for these six children we also extracted all counterfactual conditionals from the corpora by searching for utterances containing *if* in combination with *would*, *should* and *could*, and noted down age of first counterfactual conditional and any observed tense errors. Just as for the wishes, we coded for temporal orientation. We only recorded errors in present or past antecedents with *would*, since they unequivocally signal counterfactuality. *Could*-, *should*- and future *would*-conditionals have non-counterfactual meanings with non-past marking that confound coder ability to see present-for-past tense errors, e.g. “Maybe you shouldn’t be there, if you scare Ellen” or “If I saw a real wolf, I would kick the real wolf.”

We extracted 398 child wishes (0.026% of 1540,139 total utterances) in total. Since wish is counterfactual only if its complement is a full proposition (Iatridou, 2000, p.241), we excluded utterances where *wish* did not embed a proposition. 39 utterances were excluded this way for being non-counterfactual (e.g. *wish* used as a noun or wishing in utterances such as “wish you a merry Christmas”). We also excluded an additional 30 for being a repetition of either themselves or a parent and 15 sentences that were without a complement (either “I/you wish” on its own (in response to a prior) or cut off utterances). The 6 children with longitudinal data were responsible for 173 of the wishes. For those 6 children, we extracted 381 conditionals with *would*, *should* or *could*. We excluded 115 conditionals with *could* and 9 conditionals with *should* for not being non-counterfactual usages. An additional 17 counterfactual conditionals were excluded for being incomplete, and 42 for being a repetition, This left 198 counterfactual conditionals in total, 89 of which with present or past tense orientation.

Since present tense errors in counterfactual constructions could also be due to overall underdeveloped past tense marking, we estimated children’s overall past tense competence during the period where they made past tense errors in counterfactual constructions. From the six children, we extracted all utterances containing *yesterday*. From the 119 hits we excluded 24 utterances that lacked a verb and 6 utterances that were self repetitions. For the remaining 89 utterances we coded for present tense errors. Additionally, we investigated whether children productively used past tense morphology by searching for instances of overregularization (e.g. “I **telled** daddy something”) during the error period. We extracted all 4281 utterances containing past tense morphology within the counterfactual present tense error period of each individual child and noted down instances of regular past tense marking on irregular conjugating verbs.

### 3 Results

#### 3.1 Children’s Wishes

Most children produce their first wishes at the end of age 2 or beginning of age 3, and often the first wishes are followed by more occurrences in subsequent recordings. Repeated uses of a new construction within a short period of time is considered to be a signal of productivity (Snyder, 2007; Stromswold, 1990). Most wishes were spontaneous, and we found few instances (n=2) where children directly repeated a wish from their caregivers. Most early wishes are desires about things in direct proximity, e.g. wishing for a horse when seeing a horse (7).

- (7) a. EARLY WISHES (LIKE DESIRES)  
 \*CHI: I wish I had a horsie. (Becky, 2;07, Theakston et al., Manchester)  
 b. \*CHI: I wish that you stop talking. (Abe, 2;10, Kuczaj)

From their early uses, it is not obvious that children know that *wish* can only be used counterfactually, i.e. the desire is either established to be impossible or unlikely to be fulfilled. Evidence supporting are instances when parents correct their children’s wishes, or comment on the non-counterfactuality of them. Throughout the different corpora we found 9 conversations where this happened, e.g. (8).

- (8) \*CHI: I wish you were my mommy.  
 \*MOT: I am your mommy. (Laura, 3;02, Braunwald)

For this reason we separated wishes (n=314) into three separate categories: wishes that are undeniably counterfactual based on context and lexical content, i.e., past, negation and contrasting adverbs (n=40, 12.7%), wishes that are presumably counterfactual (based on context) (n=66, 21.0%) and wishes which may or may not express counterfactual reasoning (n=208, 66.2%). The first undeniably counterfactual wishes (based on lexical contrast) occur between age 3 and 4. Examples of these are past counterfactual wishes (9a), negated wishes (9b) and wishes that contain a contrast with reality (9c). An example of a wish where contextual information contains a contrast between reality and desire are displayed in (10).

- (9) a. LEXICAL EVIDENCE FOR COUNTERFACTUAL: PAST  
 [hearing a train in the distance] (Thomas, 3;01, Thomas)  
 \*CHI: I wish gone Burnage Station watch that train.  
 <later in recording Thomas comments “I’m missing all the trains” >  
 b. LEXICAL EVIDENCE FOR COUNTERFACTUAL: NEGATION  
 \*CHI: I wish humans were not humans. (Ross, 4;02, MacWhinney)  
 c. LEXICAL EVIDENCE FOR COUNTERFACTUAL: CONTRASTING ADVERB  
 \*CHI: oh I wish it was my birthday today really. (Thomas, 4;02, Thomas)

- (10) CONTEXTUAL EVIDENCE FOR COUNTERFACTUAL  
 \*FAT: You don’t see bumblebees in the dark at all.  
 \*CHI: I wish that the lights were on. (Mark, 5;10, MacWhinney)

#### 3.2 Tense Errors

Children were found to make several tense errors in the complement of *wish*. In total we found 40 errors (13.2% of the 303 total wishes with complements) that inflected present rather than

past tense. For 16 of these errors, it is not entirely clear whether they are really marking present tense as they are indistinguishable from a bare verb construction (11a). For the remaining 24 it was clear, i.e. due to inflection (11b).

## (11) PRESENT TENSE ERRORS

- a. \*CHI: I wish I **have** a banjo like dis [this]. (Adam, 5;02, Brown)  
 b. \*CHI: I wish it’s valentine. (Sarah, 3;06, Brown)

When we group the error counts by age group (per year) we find that the vast majority of errors are made before age 4, and after this the amount of errors drops off to until it reaches 0 after age 6. This decrease in error rate is displayed in Table 1. Following exclusions, there were 302 child wishes for which we modeled the present-instead-of-past tense error rate (excluding 12 other errors) with a generalized linear mixed-effect model (lme4 package Bates et al., 2014; Team, 2015) to investigate whether the factor Age (in months) is a significant predictor of Error Rate, including different child corpora as a random factor (glmer, ErrorRate  $\sim$  Age+(1|corpus), (1|child), family=binominal). This model confirmed that age is a significant predictor of tense error rates in counterfactual wishes,  $\beta = -0.11$ ,  $z = -6.21$ ,  $p = < 0.001$ .

Age (years)	sample size per age window	# errors	% of total
2 - 3	15	18	56.2
3 - 4	15	14	19.4
4 - 5	42	5	3.94
5 - 6	26	3	4.55
6+	2	0	0
Total	76	40	100

Table 1: Past tense errors (present instead of past) per age window

### 3.3 Individual Development of Counterfactual Constructions

To get an idea of the counterfactual development of individual children, we investigated the age of first counterfactual conditional, first suspected counterfactual wish (based on context), 1st certain counterfactual wish (lexical evidence), and tense errors for the 6 children we had enough longitudinal *wish* data for. The individual development of each child is displayed in Figure 1, where one can observe each instance of *wish* grouped by counterfactual type, the age of their first counterfactual conditional (12) and any tense errors made in the counterfactual construction.

## (12) FIRST COUNTERFACTUAL CONDITIONALS

- a. \*CHI: if a really hole was in here then I would cry for new pants. (Laura, 2;08)  
 b. \*CHI: we could fly if we had wings. (Mark, 3;07, MacWhinney)

The age of first *wish* uses vary from age 2;01 (25 months) to 3;08 (44 months), with subsequent usages following within the same file or month. The age of the first clearly counterfactual *wish* usages, however, vary widely by child, ranging from 2;10 (34 months) - 5;02 (62 months). We found that the acquisition of wishes precedes (4/6 children) or coincides (2/6) with first counterfactual conditionals. An interesting pattern seems to emerge. We observed that most tense error mistakes are made in the early stages of the emergence of the *wish*-construction,

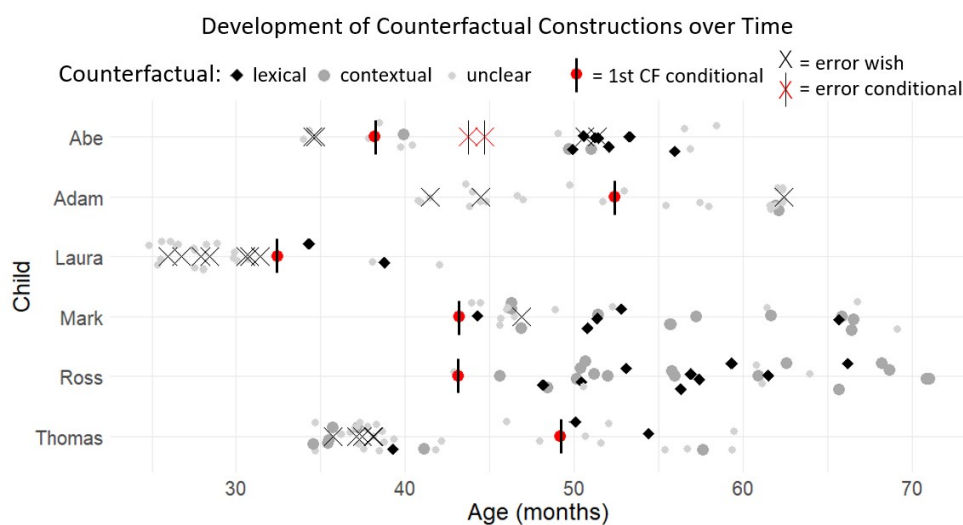


Figure 1: Wishes grouped by meaning (black diamond: clearly counterfactual based on lexical info, dark grey dot: suspected counterfactual based on contextual info, small grey dot: unclear whether counterfactual) and present-for-past-tense errors (crosses) over time (in months) for 6 children. Dot with black line indicates first clear counterfactual conditional.

regardless of at which age the child started using the construction. Moreover, the age of the last error seems to coincide with the age that children start using wishes that are clear lexical counterfactuals. This pattern is clear for 3/6 children (Abe, Laura and Thomas) and inconclusive for Ross (as we have no documented tense errors) and Adam (as the age of last error is also the last recording). Half of the children (Laura, Ross and Thomas) start using counterfactual conditionals around the time that they display clear counterfactual reasoning in their wishes, and after their last tense errors. Mark’s trajectory is almost the same, except that we found one tense-error soon after his first clear counterfactual wish and conditional. Abe and Adam, start using counterfactual conditional constructions during the period in which they make tense errors in their wishes. Abe is the only child for whom we found tense marking errors in their conditional constructions as well. These errors stop around the time he produces clearly counterfactual wishes.

### 3.4 Productive Tense Marking

We investigated children’s overall productive past tense usage during the period where they made past tense errors in counterfactual constructions, to make sure their errors in these counterfactual contexts are not due to a lack of past tense marking in general. For each child we recorded their (first) past tense marking in the context of the temporal adverb *yesterday* (YD), and (first) instances of past tense marking overregularization (OR) during the period they made mistakes in counterfactuals, displayed in Table 2. For all children, we found indications of productive past tense usage (both in overregularization and past tense usage with *yesterday*) outside counterfactual contexts during their error period. Only for Laura did we find present tense mistakes with *yesterday* before 28 months, indicating that some of her earliest mistakes could be due to overall lack of past tense marking.



Child	Age Range (months)	Error Range (months)	Age 1st YD Past	# Correct YD Past	Age 1st OR	# OR
Abe	28-58	35-51	31.4	17/17	34.1	109
Adam	27-62	42-62	41.0	3/3	42.3	7
Laura	15-84	26-31	28.0	2/5	22.6	10
Mark	7-66	47-47	44.9	2/2	48	4
Ross	6-92	NA	NA	NA	NA	NA
Thomas	24-59	36-38	32.8	1/1	33.6	2

Table 2: Overview Individual’s Present-for-past Errors over Time

## 4 Discussion

In this paper we examined the acquisition of counterfactual constructions, focusing mainly on the development of wishes. Since *wish* is a dedicated marker of counterfactuality in English, we hypothesized that counterfactual wishes would be easier to acquire than counterfactual conditional constructions. Indeed we found that children seem to produce the *wish*-construction before (or simultaneously with) counterfactual conditionals. Counterfactual wishes seem to be productive between age 2 - 3, while counterfactual conditionals emerge age 2.5 - 4.

The second question addressed in this study was whether children go through a phase where they make tense-marking mistakes in the complement of counterfactuals. Acquiring counterfactual utterances requires discovering that the past tense in its complement/antecedent is “fake” and actually marks counterfactuality. This mapping between counterfactuality and the past tense morpheme is thought to require complex semantic operations (Iatridou, 2000; Ogihara, 2000; Karawani and Zeijlstra, 2013). Since children have to see through the “fakeness” of the past tense in order to learn this mapping, we hypothesized that children would productively form present counterfactual wishes, that have a present tense (rather than past tense) marking on the embedded matrix verb, as this aligns with the temporal orientation of a present wish. Indeed, we found that children make a substantial amount of past tense errors (13% of total wishes), most of them between age 2-4 (75.6%), with a clear drop by the beginning of age 4.

One may wonder whether the tense errors found in the complement of *wish* are due to children not yet having acquired the past tense form in general. This seems unlikely, as children generally have productive past tense usage before age 3 (e.g. Brown, 1973; Kuczaj, 1977). For six children, we showed that they display clear signs of productive tense marking during the period in which they make tense marking errors in counterfactual constructions. They use past tense in utterances with *yesterday* and overregularize the past tense morpheme on irregular verbs, showing productive usage. Only for the youngest child, Laura, do we find some tense marking errors outside counterfactual constructions, suggesting that her earliest errors (before 28 months) might be partially due to a general problem with the past tense morpheme.

Zooming into the individual development of wishes, we found that the age at which children stop making past tense errors seems to be very close to the age when they start using concrete counterfactuals. While we do not have enough data to make any clear generalizations, this finding is compatible with a view where children’s mastery of the mapping between past tense and counterfactuality correlates with mastery of counterfactual reasoning. Perhaps it is only after they have fully mastered counterfactual reasoning that they realize how the past tense morpheme in counterfactuals is separate from the regular past tense morpheme. Most children start using counterfactual conditionals around or after the time they use clear counterfactual wishes. This suggests that figuring out the mapping between the past tense morpheme and

counterfactuality in wishes helps with acquiring counterfactual conditionals, but there are children (i.e. Abe and Adam) that do not follow this pattern. In accordance with this idea, we found almost no tense marking errors in the conditional constructions, suggesting children already know the mapping when using them. The only tense errors we found came from Abe, who actually participated in a longitudinal study investigating the development of hypothetical conditionals (Kuczaj and Daly, 1979), so this could have influenced his development.

All in all, these findings show that counterfactual constructions are not only challenging because they require complex reasoning, but also because they involve complex meaning-to-form mapping. Children productively make tense errors predicted by semantic accounts of counterfactuality that analyze the past tense in these constructions as “fake”, and acquire the more transparent counterfactual *wish*-construction before counterfactual conditionals.

## References

- Bates, D., Mächler, M., Bolker, B. and Walker, S. (2014), ‘Fitting linear mixed-effects models using lme4’, *arXiv preprint arXiv:1406.5823*.
- Beck, S. R., Riggs, K. J. and Gorniak, S. L. (2009), ‘Relating developments in children’s counterfactual thinking and executive functions’, *Thinking & Reasoning* **15**(4), 337–354.
- Bjorkman, B. M. and Halpert, C. (2017), In an imperfect world: Deriving the typology of counterfactual marking, in ‘Modality Across Syntactic Categories.’, Vol. 1 of *Arregui, A., Rivero, M., & Salanova, A. (Eds.)*, Oxford University Press.
- Bowerman, M. (1986), First steps in acquiring conditionals, in ‘On conditionals’, Cambridge University Press, pp. 285–308.
- Braunwald, S. R. (1971), ‘Mother-child communication: the function of maternal-language input’, *Word* **27**(1-3), 28–50.
- Brown, R. (1973), *A first language: The early stages.*, Harvard U. Press.
- Byrne, R. M. J. (2007), ‘Précis of *The Rational Imagination: How People Create Alternatives to Reality*’, *Behavioral and Brain Sciences* **30**(5-6), 439–453.
- Dowty, D. R. (1986), ‘The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics?’, *Linguistics and philosophy* **9**(1), 37–61.
- Dudman, V. H. (1983), ‘Tense and time in English verb clusters of the primary pattern’, *Australian Journal of Linguistics* **3**(1), 25–44.
- Guajardo, N. R., Parker, J. and Turley-Ames, K. (2009), ‘Associations among false belief understanding, counterfactual reasoning, and executive function’, *British Journal of Developmental Psychology* **27**(3), 681–702.
- Hacquard, V. and Lidz, J. (2018), ‘Children’s attitude problems: Bootstrapping verb meaning from syntax and pragmatics’, *Mind & Language*.
- Iatridou, S. (2000), ‘The grammatical ingredients of counterfactuality’, *Linguistic inquiry* **31**(2), 231–270.
- Ippolito, M. (2006), ‘Semantic Composition and Presupposition Projection in Subjunctive Conditionals’, *Linguistics and Philosophy* **29**(6), 631–672.



- James, D. (1982), ‘Past Tense and the Hypothetical a Cross-Linguistic Study’, *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”* **6**(3), 375–403.
- Karawani, H. (2014), *The real, the fake, and the fake fake in counterfactual conditionals, crosslinguistically*, number 357 in ‘LOT’, LOT, Utrecht. OCLC: ocn881612264.
- Karawani, H. and Zeijlstra, H. (2013), ‘The semantic contribution of the past tense morpheme *kaan* in Palestinian counterfactuals’, *Journal of Portuguese Linguistics* **12**(1), 105–119.
- Katis, D. (1997), ‘The emergence of conditionals in child language: Are they really so late?’, *Current Issues in Linguistic Theory* **4**(143), 355–386.
- Kuczaj, S. A. (1977), ‘The acquisition of regular and irregular past tense forms’, *Journal of Verbal Learning and Verbal Behavior* **16**(5), 589–600.
- Kuczaj, S. A. and Daly, M. J. (1979), ‘The development of hypothetical reference in the speech of young children\*’, *Journal of Child Language* **6**(3), 563–579.
- Landau, B., Gleitman, L. R. and Landau, B. (2009), *Language and experience: Evidence from the blind child*, Vol. 8, Harvard University Press.
- Landis, J. R. and Koch, G. G. (1977), ‘The measurement of observer agreement for categorical data’, *biometrics* pp. 159–174.
- Lieven, E., Salomo, D. and Tomasello, M. (2009), ‘Two-year-old children’s production of multiword utterances: A usage-based analysis’, *Cognitive Linguistics* **20**(3), 481–507.
- MacWhinney, B. (1991), ‘The childe project: Tools for analyzing talk. hillsdale, nj, us’.
- MacWhinney, B. (2000), *The CHILDES project: The database*, Vol. 2, Psychology Press.
- Nyhout, A. and Ganea, P. A. (2019), ‘Mature counterfactual reasoning in 4- and 5-year-olds’, *Cognition* **183**, 57–66.
- Ogihara, T. (2000), ‘Counterfactuals, Temporal Adverbs, and Association with Focus’, *Semantics and Linguistic Theory* **10**(0), 115–131.
- Reilly, Snitzer, J. (1982), *The Acquisition of Conditionals in English*, Unpublished PhD dissertation, University of California, Los Angeles, United States.
- Ritter, E. and Wiltschko, M. (2014), ‘The composition of INFL’, *Natural Language & Linguistic Theory* **32**(4), 1331–1386.
- Romero, M. (2014), ‘“Fake Tense” in Counterfactuals: A Temporal Remoteness Approach’, **2**, 47–63.
- Snyder, W. (2007), *Child Language: The Parametric Approach*, OUP Oxford.
- Stromswold, K. J. (1990), *Learnability and the acquisition of auxiliaries*, Thesis, Massachusetts Institute of Technology.
- Team, R. (2015), *RStudio: Integrated Development Environment for R*, RStudio, Inc., Boston, MA, <http://www.rstudio.com/>.
- Theakston, A. L., Lieven, E. V., Pine, J. M. and Rowland, C. F. (2001), ‘The role of performance limitations in the acquisition of verb-argument structure: An alternative account’, *Journal of child language* **28**(1), 127–152.

# Asking between the lines: Elicitation of evoked questions in text

Matthijs Westera<sup>1</sup> and Hannah Rohde<sup>2</sup>

<sup>1</sup> Universitat Pompeu Fabra  
matthijs.westera@gmail.com

<sup>2</sup> University of Edinburgh  
hannah.rohde@ed.ac.uk

## Abstract

We introduce a novel, scalable method aimed at annotating potential and actual Questions Under Discussion (QUDs) in naturalistic discourse. It consists of asking naive participants first what questions a certain portion of the discourse evokes for them and subsequently which of those end up being answered as the discourse proceeds. This paper outlines the method and design decisions that went into it and on characterizing high-level properties of the resulting data. We highlight ways in which the data gathered via our method could inform our understanding of QUD-driven phenomena and QUD models themselves. We also provide access to a visualization tool for viewing the evoked questions we gathered using this method (N=4765 from 111 crowdsourced annotators).

## 1 Introduction

A piece of discourse can evoke certain questions. For instance, the utterance “Latisha is worried.” will quite reliably evoke, in an addressee’s or other observer’s mind, the question of why Latisha is worried – and an addressee is likely to respond accordingly: “Why?”, “How come?”, or “Worried about what?”. The questions that a piece of discourse evokes provide a window into comprehenders’ expectations about both the underlying situations being described by a text as well as the way in which the discourse itself is unfolding. They reflect what humans generally care about (e.g., causes of mental states) as well as what they understand or expect to be relevant in the particular discourse at hand, say, what they anticipate may be the next Question Under Discussion as intended by the original speaker (QUD [6, 21, 19, 13, 10]).

This paper summarizes our work-in-progress on annotating discourses with evoked questions using a crowdsourcing approach. More concretely, we present people snippets from a number of texts, and ask them to enter a question that the text evokes. Subsequently, we let them read how the text continues and ask them if their question has been answered. Moreover, we ask our participants to highlight the word or short phrase in the discourse which primarily evoked the question, and likewise for the words which primarily provided the answer. We repeat this process for multiple snippets per text to obtain evoked questions and, where available, their answers, after basically every sentence, in a number of texts. Altogether, this results in a rich dataset of evoked questions, their triggers in the text, and their associated answers.

This paper explains and motivates our data gathering method, presents some basic statistics on the resulting dataset as well as some coarse-grained analysis aimed merely at demonstrating the promise of our approach. Our main motivation for this undertaking is not the evoked questions as such – although we think they represent an important, hitherto unexploited cognitive/linguistic resource – but the window they provide on QUDs. We think of our method for eliciting evoked questions as the first step in a scaleable pipeline for annotating QUDs: After all, questions which were both evoked and subsequently answered are plausible candidates to be the

original speaker’s intended QUD. Let us briefly clarify the need for a scalable QUD-annotation pipeline: QUDs form the backbone of many current semantic/pragmatic theories. Such QUD-based theories often make falsifiable predictions only once the assumption of a certain type of QUD (or a certain set of ‘alternatives’) is granted (just as the Gricean maxims [11] presupposed a notion of ‘relevance’). Now, this isn’t specifically problematic when merely assessing such theories intuitively using constructed linguistic examples, or when testing them in a controlled empirical experiment with constructed stimuli, since in both cases the relevant linguistic examples can be presented against the background of an explicit interrogative sentence, thereby fixing the intended QUD. But in naturally occurring discourse explicit interrogatives are rare, rendering this domain primarily off-limits for large-scale hypothesis testing of QUD-based theories. Clearly this situation is unsatisfactory. The remedy would be a powerful, quantitative model of QUDs that can tell us, for any stretch of natural discourse, which QUDs are plausible there and which ones aren’t. In order to develop (train, test) such a model, we will need a large amount of naturalistic data annotated with QUDs. For this reason it is important that we come up with a scalable QUD annotation pipeline, one which can be crowdsourced rather than relying on expert annotators (who are expensive, scarce, and potentially theoretically biased).

Although our main aim is to work towards a scalable QUD-annotation pipeline, we also think, as we mentioned, that the notion of an evoked question is interesting in its own right. Just as benchmark datasets for something like word similarity or word association are frequently relied upon (e.g., given the word “snow”, write the first alternative words that come to mind), we think that a benchmark of evoked questions (given a piece of discourse, write the first question that comes to mind) will provide an important window on linguistic competence and on the human mind more generally. Where word similarity benchmarks have become indispensable in lexical semantics (especially computational approaches), we think an evoked questions benchmark could become indispensable for pragmatics.

## 2 Related work

Questions Under Discussion (QUDs) are frequently relied upon in semantics and pragmatics in theories of discourse structure and the wide range of phenomena that depend on it, such as implicature (e.g., [9, 14, 2, 23]), discourse particles (e.g., [20]), prosody (e.g., [19, 5, 3, 24]) and pronoun interpretation (e.g., [7, 12]). However work on annotating QUDs is limited: to our awareness the only attempt is [8, 18]. Compared to that approach, our annotation task is much more free: our participants do not receive detailed instructions and can basically enter any question they like, whereas [8, 18] carefully formulated a number of (theory-derived) constraints on QUDs, such as the constraint that the QUDs of a discourse should form a tree. As a consequence, whereas they relied on instructed expert annotators, our procedure is sufficiently simple and natural for naive annotators, hence for a scalable crowdsourcing approach – though of course whether annotations obtained from the two different procedures share some features of interest remains to be seen. Another major difference is that their annotators were given the full discourse, whereas for our participants the texts are revealed one snippet at a time, at each point eliciting an evoked question (and asking if previous questions have been answered yet), prior to revealing how the text continues. One reason for this is that considering which questions (potential QUDs) a discourse *evokes* is a simpler, far more natural task than deciding which question an utterance *addresses* (actual QUD), and therefore that it lends itself better to crowdsourcing. Another reason is our interest in the forward-looking process itself, i.e., the anticipation/predictability of QUDs, drawing inspiration from [15].

While QUD annotation has only recently been undertaken, and on limited data, the an-

notation of discourse relations has a long history: e.g., the Penn Discourse TreeBank (PDTB; [16]). Contrary to QUD-theory – where the range of possible QUDs is in principle open-ended – the PDTB and other coherence annotation schemes recognize only a limited number of types of discourse relations. Moreover, for the PDTB the different types of relations are tied to particular natural language connectives such as “and then” and “because”, such that annotating discourse relations is a matter of (i) recognizing and categorizing explicit connectives and (ii) recognizing places where such connectives are arguably implicit, say, where they could naturally be inserted. This setup greatly facilitates annotation, though at the cost of relying on an a priori decided, fixed repertoire of relation types. It will be interesting to compare existing discourse relation annotations to the evoked questions we elicit – one would of course expect some kind of alignment, e.g., a ‘causal’ discourse relation may tend to occur after snippets that evoke questions such as “Why?” and “How come?”. To enable this kind of research we decided to elicit evoked questions for two small datasets that have already been annotated in PDTB style: TED-MDB (Multilingual Discourse Bank, six TED talk transcripts in different languages, though we’ll be using only the English portion; [26]) and DISCO-SPICE (dialogue and interview transcripts in Irish English; [17]). Although in the present paper we will not yet look at the alignment of discourse relations and evoked questions, we have recently conducted this research and report on it in a manuscript under review [25].

QUDs are variably regarded as a type of discourse move or as a representation of the discourse goal, though often a mix of both. The first perspective regards QUDs as a questioning speech act which can be either explicit (i.e., realized by an interrogative sentence in the discourse) or implicit. The second perspective regards them as a model of discourse goals, and goals are strictly speaking different creatures from the speech acts intended to serve them; both questioning and asserting speech acts must relate to the discourse goal, but are conceptually distinct from this goal, and also formally not necessarily equivalent to it. We think the second perspective is ultimately the more adequate/fruitful one: we need a notion of discourse goals in our theories, and they must be distinguished from the acts by means of which speakers try to achieve them. Nevertheless, the two perspectives are often quite harmlessly conflated: the main purpose of questioning speech acts, and of the interrogative sentences that express them, is to set the discourse goal for the next speaker; and any discourse goal can be fairly naturally paraphrased by means of an interrogative. We too rely on interrogative sentences – namely, the evoked questions we elicit from our participants – as a natural way of annotating the potential and actual QUDs of the discourse. But we must remain aware that this is a simplification, as interrogative sentences often explicate only a part of their QUD: just as a declarative utterance typically asserts only some of the propositions in the QUD, so too may an interrogative utterance highlight some of them while backgrounding others (e.g., [4, 22]).

### 3 Method

For concreteness we will first show what our elicitation task looked like, before going into details about how we cut up the texts to serve them in portions to our participants. The core of our method is to present participants a piece of text, ask them which question it evokes, then reveal how the text continues and ask whether their question has been answered. We set up our task in Ibex (Internet-based experiments, <https://github.com/addrummond/ibex>), although our incremental (chunk-wise) display of text and giving participants the ability to highlight text required a lot of customization. The two main screens of the task are shown in figure 1. As can be seen in the left figure, we simply asked participants to “enter a question the text evokes at this point”, and (motivated by pilot results) we included a clarification that it was to be

<p>Person A:</p> <p>Och the wee mite</p> <p>Person B:</p> <p>I know she's uh</p> <p>And also as well I'm getting really worried</p> <p>Everybody keeps going on <b>how wee she is</b></p> <hr/> <p>► Please enter a question the text evokes for you at this point. (The text so far must <i>not</i> yet contain an answer to the question!)</p> <p><b>Is she really small?</b></p> <p>► In the text, <b>highlight</b> the main word or short phrase that evokes this question.</p>	<p>Why is that hard? Well to see, let's imagine we take the Hubble Space Telescope and we turn it around and we move it out to the orbit of Mars. We'll see something like that, <b>a slightly blurry picture of the Earth</b>, because we're a fairly small telescope out at the orbit of Mars. Now let's move ten times further away. <b>Here we are at the orbit of Uranus</b>. It's gotten smaller, <b>it's got less detail, less resolve</b>. We can still see the little moon, but let's go ten times further away again.</p> <hr/> <p>Earlier, you also entered the following question:</p> <p><b>How can the picture be improved?</b></p> <p>► Was that question answered in the new piece of text?</p> <p>Not answered at all. 1 2 3 4 5 Completely answered.</p> <p>► Enter the (complete/partial) answer in your own words:</p> <p><b></b></p> <p>► In the new piece of text, <b>highlight</b> the main word or short phrase suggesting this answer.</p>
---	---

Figure 1: Our elicitation tool, asking for an evoked question (left) and whether a prior question has been answered (right); dialogue (left) and monologue (right) are rendered differently.

a question that hadn't yet been answered at that point. Other than this we did not provide instructions about what counts as an 'evoked question', trusting that it would be a sufficiently natural task. In particular, we did not instruct them to try to guess where the discourse was heading, although we do believe that our task biases them towards this, by repeatedly asking whether a question has been answered yet.

After a small pilot study which resulted in very few elicited questions, we decided to make entering a question mandatory, because we wanted our workers to make an effort, and there's always *something* that can be asked. (Although we may be 'overeliciting' questions in this way, questions that are too random/too desperate could in principle be filtered out by looking at inter-annotator agreement; cf. section 4.) After eliciting an evoked question, we would reveal the next chunk of text and ask them whether their question was answered at that point. This is shown in figure 1 on the right (though for a different text than the one on the left): we repeated the question they entered and asked for a rating on a 5-point Scale. For both questions and answers, participants were asked to highlight the main word or short phrase in the text that primarily evoked the question, or provided the answer, respectively. They could do this by dragging a selection in the newest chunk of text, where they could highlight a single span of up to 10 words. Highlights for different questions were given different colors, and highlights for answers were given the same color as the question they were answers too, which was intended as an additional visual guide to decrease cognitive load. We would repeat this procedure, incrementally taking each participant through a text, at each point asking for evoked questions as well as possible answers to questions evoked earlier.

We hosted our Ibex experiment on IbexFarm (<http://spellout.net/ibexfarm/>), and recruited participants from Amazon Mechanical Turk (<http://mturk.com>). Each participant could do the task once, where a task would consist of 6 excerpts, where each excerpt is comprised of 8 chunks of text (i.e., 8 probe points asking them for questions/answers). We estimated that the task would take about 50 minutes, and offered a compensation of \$8.50. Comments from our workers suggest that the task was a bit long and became boring after a while, sug-

gesting that we should have offered perhaps only one excerpt per task – this is something we’d do differently in the future when getting additional annotations for the same texts and/or annotating different texts. We aimed to have at least 5 participants for every probe point, but because we let the excerpts overlap many probe points have more than that.

As for the texts we used, as mentioned earlier we rely on two existing datasets, to which we contribute the elicited evoked questions as an additional layer of annotations. These datasets are TED-MDB [26] and DISCO-SPICE [17]. TED-MDB consists of transcripts of six scripted presentations from the TED Talks franchise, in multiple languages, from which we use the English portion (6975 words total). DISCO-SPICE consists of transcripts of Irish conversational English, of which at present we used only two transcripts (3807 words total; but comparable to TED-MDB in number of sentences). In addition we included a simple short story (56 words), which we ourselves composed with certain obvious evoked questions and subsequent answers in mind, as an introductory item for our participants and as a sanity check for our method.

Our aim was get full coverage of each text, namely, to elicit evoked questions and check for answers either after every sentence or at certain points within long sentences. To that end we first cut each text into sentences (or sentence-like fragments). For DISCO-SPICE we relied on the existing division into utterances; for TED-MDB we ran NLTK’s sentence tokenizer. Any remaining long sentences (> 150 words) were further automatically cut up (as necessary) at commas, colons or semicolons, both in order to elicit questions at those points, and to make all chunks more or less the same length. We also manually inserted cuts just prior to occurrences of “because”, anticipating that it could be interesting to investigate which questions are evoked there. For convenience we will refer to the resulting sentence-like pieces of text as sentences.

Although we wanted to elicit evoked questions at every sentence, unfolding the discourse one sentence at a time and asking participants to enter a question at each point will get tedious quickly and break the flow of the discourse – and often a single sentence will not contain much new information anyway to evoke a meaningful question or provide an answer. Therefore, we decided to unfold each excerpt in chunks of 2 sentences each, and probe each participant only after every such chunk. We rotated participants through the different ways of dividing a text into chunks to get full coverage this way. Altogether, this resulted in 460 probe points covering the six texts of TED-MDB and 397 probe points covering the two texts of DISCO-SPICE.

We decided not to present full texts to our participants, but only excerpts of around 10 of the aforementioned chunks, before letting them move on to an excerpt from a different text. This decision aligned with our aim to make our elicitation process part of a scalable pipeline for QUD-annotation, a pipeline which should be able in principle to target very long texts – texts too long for a single participant to cover in a timespan reasonable for crowdsourcing. It does come with a downside, namely, that we will be able to capture only relatively ‘local’ discourse structure. But we think that eliciting evoked questions through crowdsourcing is already biased towards local discourse structure to begin with, by virtue of the fact that coming up with local questions and finding their answers is an easier way of completing the task than asking high-level questions, which would require keeping the whole text in mind. We think that capturing more ‘global’ discourse structure would require a different type of task.

With the choice to give participants only excerpts instead of full texts, there is a risk that at the start of an excerpt they will lack the context necessary to understand what’s going on, such that the evoked questions will mostly be clarification questions (“who are the speakers?”, “what are they talking about?”) that do not reflect the actual discourse structure for those who were present. For this reason we let the excerpts of a given text overlap, such that what are the first chunks for some participants (who may lack context) will be chunks at the end for other participants. We can then in principle check whether there are systematic



between the questions evoked after excerpt-initial and excerpt-final chunks and, if necessary, discard excerpt-initial chunks without losing coverage of those pieces of discourse altogether.

Another decision that bears on whether we get to capture local or global structure is the following. Recall that, in addition to asking for a question evoked by the text, we ask our participants whether their previous questions have been answered yet. This raises the question of how long we should keep pestering workers with questions they entered that remain unanswered: if we keep at this for too long then the task will get tedious very quickly – “is it answered yet?”, “and now?”, “what about now?” – especially given the fact that most evoked questions will never be answered. We conducted a pilot where we prompted participants with the questions they entered for up to three chunks after the chunk that evoked them. We found that questions that didn’t get answered after one chunk were practically never answered within three chunks, either. Therefore, to cut task duration (hence cost) without risking substantial data loss we decided to decrease this limit to two: participants are asked at most twice (i.e., after two subsequent chunks of text) whether any given question they entered had been answered. This also meant that the highest number of unanswered questions a participant ever had to consider at any particular chunk was two. In any case, this is another respect in which our approach is biased towards local discourse structure; it is not suitable for getting at questions that are answered only after a longer stretch of discourse unless those questions are re-evoked at a subsequent probe point.

To summarize: we cut the source texts into overlapping excerpts of up to 8 chunks each, where each chunk is made up of two sentence-like parts (getting full coverage by rotating our participants through the different ways of cutting each text into chunks). During the task, unanswered evoked questions were kept around for up to two chunks. Our procedure is biased towards eliciting questions that reflect ‘local’ QUD structure. For the intended comparison to the discourse relation annotations, which we leave to [25], this restriction to local structure is quite alright: discourse relations are likewise predominantly local, since they are based around explicit or implicit connectives between clauses. But ultimately a complementary task would be required to get at the more global structure.

## 4 Bird’s-eye overview of the data

Although our focus in this paper is on the method and motivations for our design decisions, we give a very general overview of the data collected. For qualitative, manual inspection of the resulting data we have published the texts annotated with the evoked questions we elicited, with color-coding of inter-annotator agreement as discussed below, at <https://evoque-data.github.io>. The full dataset, including answers and highlights, and properly (re-)aligned with the existing discourse relation annotations, will be made available at a later date.

We recruited 111 participants from Amazon Mechanical Turk, who gave us a total of 4765 evoked questions for our 863 probe points. These and other numbers are summarized in Table 1. Overall, participants seemed to engage with the task and posed questions which showed interest and anticipated the text’s subsequent discourse moves. We compute an ANSWERED rating for evoked questions as the highest ANSWERED rating (on a scale from 1 to 5) assigned to that question in the two chunks following it, i.e., the quality of its best answer. Computed thus, the mean ANSWERED rating over all questions is 2.38. Less than half of the evoked questions (2251) were rated to be ‘not answered at all’ (rating 1), with the rest quite evenly divided between ratings 2, 3, 4 and 5 (‘completely answered’). So, participants do ask questions that anticipate speakers’ upcoming discourse moves, i.e., questions that end up being at least partially answered hence are plausibly QUDs, but also many questions that don’t get resolved at all.

Source texts:	8	Participants:	111	Probe points:	863
Word count:	10782	Excerpts/participant:	6	Questions:	4765
Excerpts:	142	Probe points/excerpt:	8	Answers:	1965
		Participants/probe point:	$\geq 5$		

Table 1: Summary statistics (excluding the constructed story all participants saw as practice).

Genre	ANSWERED	SIF-SIMILARITY	Same question type	# questions
DISCO-SPICE dialogue	2.11	.22	.19	2131
TED talk presentation	2.50	.27	.24	2412
Constructed story	2.89	.29	.27	222

Table 2: Inter-annotator agreement scores and ‘answered’ ratings for questions

As an example of a probe point with high ‘answered’ ratings, the following example from a spoken dialogue elicited questions from 11 participants:

- (1) He was uh he was a bit upset on uh uhm first day the Friday DISCO-SPICEp1a-094:line37  
**Elicited questions:** Why was he upset on his first day? Why was he upset? He was upset about what? Why was he upset? What happened to him? What happened to upset him? Is he better now? Why was he upset? Why was he upset? Why is he upset? Why was he upset?  
**Highlighted answer:**  
 [...] The oul side-effects of the medication. DISCO-SPICEp1a-094:line39

Besides most questions elicited for this chunk indeed being answered, there is also high agreement among our participants about what question was evoked. Moreover, all of them highlighted the same five-word topic (“he was a bit upset”) as evoking the question, and most of those whose questions were answered highlighted the same subsequent answer. We will attempt to quantify inter-annotator agreement shortly below.

First let us consider what types of questions were evoked. Here we define the ‘question type’ essentially by the first word of the question (wh-word or auxiliary verb), although we also took some multi-word expressions into account; for instance, we treat “how come” as the same type as “why”, not as “how”. Based on our coarse-grained classification of question types, *what*-questions were the most frequent, across genres, likely due to the flexibility of this wh-word. Auxiliary-initial polar questions were next, followed by *how/why*-questions. The DISCO-SPICE excerpts yielded more *who/where*-questions compared to TED talks and the constructed story, reflecting a higher proportion of clarifying and situating questions in these unscripted dialogues (e.g., “Who are they talking about?”, “Where are they?”). Breakdown of ANSWERED ratings (along with a measure of inter-annotator agreement, see below) by question type is shown in figure 2. It shows that the *who/where*-questions – many of which are clarifying/situating questions – are also the least answered, as expected: these meta questions were not at-issue for the original interlocutors of the texts – while *why/how*-questions were the most answered, suggesting more reliable anticipation of QUDs.

Breakdown of ANSWERED ratings by genre is shown in the first column of Table 2. The ANSWERED ratings show that, in line with expectation, anticipation of QUDs is easiest in our constructed story (as intended), hardest in unscripted dialogue, with TED talks in between. The other columns in Table 2 represent two coarse-grained measures of inter-annotator agreement,



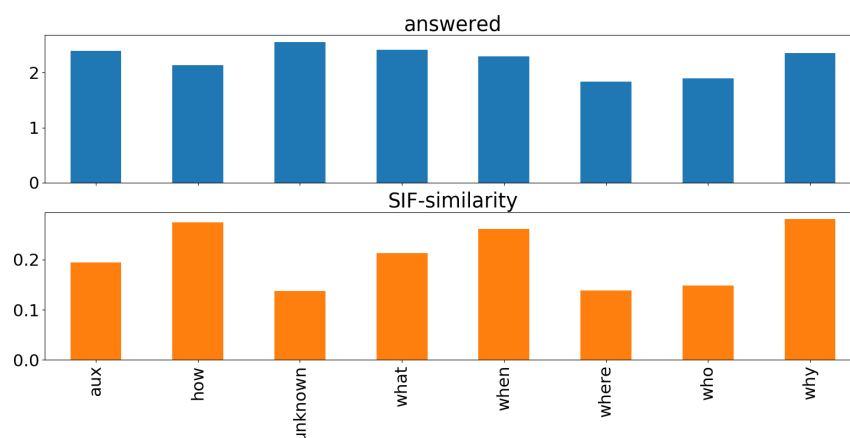


Figure 2: ANSWERED rating and SIF-SIMILARITY (inter-annotator agreement) by question type.

and ‘# questions’ the total number of questions in each subset of the data. The ‘same question type’ column represents simply the average proportion of pairs of questions (evoked by a given chunk) that are of the same question type as defined above, i.e., based simply on the first token(s) of the questions. The SIF-SIMILARITY column represents the average of a notion of semantic similarity computed between all pairs of questions evoked by a given chunk. It is computed as the cosine similarity of Smooth Inverse Frequency (SIF) sentence embeddings ([1]) of the elicited questions, which is a measure of (composed) distributional semantic similarity. Both measures of inter-annotator agreement show that TED talks and the constructed story yielded questions with higher inter-annotator agreement compared to the unscripted DISCO-SPICE dialogues, as expected, and in line with the pattern exhibited by the ANSWERED scores.

A quick note about the reliability of SIF-SIMILARITY is in order. In a smaller pilot study we ourselves annotated ‘question equivalence’ by hand (e.g., all questions for (1) except “Is he better now?” would be assessed as equivalent). This resulted in a Spearman correlation of 0.35 between our annotations and the SIF-SIMILARITY score, suggesting that the latter may indeed be sufficiently reliable for the coarse-grained statistics reported here, but also that a more accurate (more human) assessment of question equivalence, or inter-annotator agreement more generally, is clearly called for. In particular, the kinds of question pairs that SIF-SIMILARITY misclassifies as unrelated are ones that happen to use very different words, or cases where one question is very explicit and the other is relying more on context (pronouns, ellipsis). To remedy this, we have recently fed our evoked questions through another round of crowdsourcing, where we give workers a snippet with all the questions it evoked and ask them to indicate how related the different pairs of questions are (‘equivalence’ being one extreme). We describe this in a manuscript currently under review [25].

We have shown only some very coarse ways of looking at the data we collected. We have hardly looked, for instance, at which tokens our participants chose to highlight. Just to give an example of the detailed findings this rich data supports: we found that ANSWERED scores are the highest by far when the highlighted words include a wh-pronoun, reflecting the fact that explicit questions in a discourse are very likely to end up as the next QUD (and then be answered). Lastly, manual inspection of the annotated texts remains valuable, too, and our simple online visualization makes this easier: See <https://evoque-data.github.io>.

## 5 Outlook

We have introduced a novel, scalable method aimed at annotating potential and actual Questions Under Discussion (QUDs) in naturalistic discourse. It consists in asking naive participants first what questions a certain portion of the discourse evokes for them and subsequently which of those end up being answered as the discourse proceeds. Our focus in this paper has been on outlining the method and design decisions that went into it, and on exploring some high-level properties of the resulting data with the modest hope of showing the promise of this approach.

More analysis of our elicited data is ongoing, as well as alignment of these data with the discourse coherence relation annotations already available for DISCO-SPICE and TED-MDB; analysis in this respect of the latter is reported in a paper under review [25]. This was one of our main reasons for choosing these two datasets: adding an evoked questions tier to the existing discourse relation annotations allows us to address interesting new questions about the relation between QUDs and coherence relations, two key notions in pragmatics, e.g.: Do coherence-signaling devices overlap with participants' highlighted topics and answers? Do speakers omit discourse connectives at a higher rate for utterances that answer more predictable QUDs?

The data we elicited for the present paper represents the first batch of a larger scale collection, analysis, and ultimate release of a corpus annotated with potential and actual QUDs. Our hope is that this will be a valuable tool for developing and testing both semantic/pragmatic theories and computational models of QUDs. In addition, we think that the notion of an evoked question is cognitively/linguistically interesting in its own right, independently of their relation to QUDs, and we hope that our data can serve as an 'evoked questions benchmark', perhaps akin to 'word association norms' and other existing human judgment benchmarks.

## 6 Acknowledgments

This work was supported in part by a Leverhulme Trust Prize in Languages and Literatures to H. Rohde. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 715154). This paper reflects the authors' view only, and the EU is not responsible for any use that may be made of the information it contains.



## References

- [1] Sanjeev Arora, Yingyu Liang, and Tengyu Ma. A simple but tough-to-beat baseline for sentence embeddings. In *International Conference on Learning Representations*. 2017.
- [2] David Beaver, Mandy Simons, Craige Roberts, and Judith Tonhauser. Questions under discussion: Where information structure meets projective content. *Annual Review of Linguistics*, 3:265–284, 2017.
- [3] David I. Beaver and Brady Z. Clark. *Sense and Sensitivity: How Focus Determines Meaning*. Wiley-Blackwell, West Sussex, UK, 2008.
- [4] María Biezma and Kyle Rawlins. Responding to alternative and polar questions. *Linguistics and Philosophy*, 35(5):361–406, 2012.
- [5] Daniel Buring. On d-trees, beans, and b-accent. *Linguistics and Philosophy*, 26:511–545, 2003.
- [6] Lauri Carlson. *Dialogue Games: An Approach to Discourse Analysis*. Reidel, Dordrecht, 1983.

- [7] Chris Cummins and Hannah Rohde. Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in Psychology, Special issue on Context in communication: A cognitive view*, 6:1–11, 2015.
- [8] Kordula De Kuthy, Nils Reiter, and Arndt Riester. QUD-based annotation of discourse structure and information structure: Tool and evaluation. In Nicoletta Calzolari et al., editor, *Proceedings of the 11th Language Resources and Evaluation Conference (LREC)*, pages 1932–1938, 2018.
- [9] Judith Degen. Investigating the distribution of some (but not all) implicatures using corpora and web-based methods. *Semantics & Pragmatics*, 8, article 11:1–55, 2015.
- [10] Jonathan Ginzburg and Ivan Sag. *Interrogative Investigations*. CSLI Publications, Stanford, 2000.
- [11] H. P. Grice. Logic and conversation. In Peter Cole and Jerry L. Morgan, editors, *Syntax and semantics, vol. 3: Speech Acts*, pages 41–58. Academic Press, 1975.
- [12] Andrew Kehler and Hannah Rohde. Evaluating an expectation-driven QUD model of discourse interpretation. *Discourse Processes*, 54:219–238, 2017.
- [13] Staffan Larsson. Questions under discussion and dialogue moves. In *Proceedings of TWLT 13/Tuendial '98: Formal Semantics and Pragmatics of Dialogue*, 1998.
- [14] Laia Mayol and Elena Castroviejo. Projective meaning and implicature cancellation. In G. Kierstead, editor, *Proceedings of the ESSLLI 2011 Workshop on Projective Meaning*, 2011.
- [15] Edgar Onea. *Potential questions at the semantics-pragmatics interface*. Brill, 2016.
- [16] Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The Penn Discourse TreeBank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation*, pages 2961–2968, 2008.
- [17] Ines Rehbein, Merel Scholman, and Vera Demberg. Disco-spice (spoken conversations from the spice-ireland corpus annotated with discourse relations). In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 16)*, 2016.
- [18] Arndt Riester. Constructing QUD trees. In Klaus v. Heusinger, Edgar Onea, and Malte Zimmermann, editors, *Questions in Discourse*, volume 2. Brill, Leiden, 2019.
- [19] Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics*, 49: Papers in Semantics, 1996.
- [20] Tania Rojas-Esponda. A QUD account of german doch. In *Proceedings of the 18th Sinn und Bedeutung*, page 35976, 2014.
- [21] Jan van Kuppevelt. Discourse structure, topicality, and questioning. *Journal of Linguistics*, 31:109–147, 1995.
- [22] Matthijs Westera. *Exhaustivity and intonation: a unified theory*. PhD thesis, University of Amsterdam, 2017.
- [23] Matthijs Westera. An attention-based explanation for some exhaustivity operators. In *Proceedings of Sinn und Bedeutung*, volume 21, pages 1307–1324, 2018.
- [24] Matthijs Westera. Rise-fall-rise as a marker of secondary QUDs. In Daniel Gutzmann and Katharina Turgay, editors, *Secondary content: The semantics and pragmatics of side issues*. Brill, Leiden, 2019.
- [25] Matthijs Westera, Laia Mayol, and Hannah Rohde. TED-Q: a dataset of TED talks with the Questions they evoke. In , under review.
- [26] Deniz Zeyrek, Amália Mendes, Yulia Grishina, Murathan Kurfah, Samuel Gibbon, and Maciej Ogrodniczuk. TED multilingual discourse bank (TED-MDB): a parallel corpus annotated in the PDTB style. In *Proceedings of the 11th Language Resources and Evaluation (LREC)*, 2018.

# Negating Conditionals in Bilateral Semantics

Malte Willer

University of Chicago  
willer@uchicago.edu

## Abstract

A recurring narrative in the literature on conditionals is that the empirical facts about negated *ifs* provide compelling evidence for the principle of CONDITIONAL EXCLUDED MIDDLE and sit uncomfortably with a large family of analyses of conditionals as universal quantifiers over possible worlds. I show that both parts of the narrative are in need of a rewrite by articulating a bilateral update semantics for conditionals that distinguishes itself from previous frameworks by giving separate acceptance and rejection conditions for conditionals. The resulting framework shows that CEM is inessential for explaining the empirical facts about negated *ifs* but also how the principle can live happily in a strict analysis of conditionals.

## 1 The Plot

A recurring narrative in the literature on conditionals is that the empirical facts about negated *ifs* provide strong evidence for the principle of CONDITIONAL EXCLUDED MIDDLE (CEM) and sit uncomfortably with an analysis of conditionals as universal quantifiers over possible worlds.

$$(CEM) \quad \models (\phi > \psi) \vee (\phi > \neg\psi)$$

It is a familiar fact that CEM fails to be valid in a classical universalist setting, since whenever some of the relevant  $\phi$ -worlds make  $\psi$  true while others make  $\psi$  false, neither ' $\phi > \psi$ ' nor ' $\phi > \neg\psi$ ' will turn out to be true. The no less familiar arguments to the conclusion that this is a problem rather than a boon reliably appeal to the empirical fact that natural language *ifs* fail to enter into scope relations that seemingly are available if CEM is rejected—the interaction between conditionals and negation being the key case in point.

The basic observation is that conditionals such as (1a) and (1b) ring equivalent:

- (1) a. It's not the case that if John takes the exam, he will pass.
- b. If John takes the exam, he won't pass.

The challenge is to explain why (1a) should entail (1b): given CEM, this is just an instance of disjunctive syllogism. In contrast, the inference is a bit of a puzzler if the truth (or acceptance) of a conditional requires all relevant antecedent-verifying worlds to be consequent-verifying worlds (as in, among many others, Lewis 1973; Kratzer 1986; Veltman 1985), since the existence of a relevant world at which John takes the exam and fails is consistent with the existence of a relevant world at which John takes the exam and passes.

That negation causes trouble in a universalist setting is already noted by Lewis (1973), who ends up rejecting CEM but also bemoans that doing so precludes fully accounting for how conditionals play with negation in natural language. Starting with Stalnaker (1981), CEM enthusiasts have repeatedly pressed the latter point in the literature on conditionals, with a variety of glosses. One recent version of the story puts items that lexicalize negation into the spotlight (see Cariani and Goldstein Forthcoming and Santorio 2017). Consider the following pair:

- (2) a. I doubt that if John takes the exam, he will pass.  
 b. I believe that if John takes the exam, he won't pass.

(2a) and (2b) have the air of equivalence. That makes good sense if CEM is true, for then rejecting the claim that John will pass if he takes the exam immediately amounts to accepting the claim that he will not pass if he takes the exam. But it is not at all obvious how the equivalence can be explained if CEM is rejected. Other natural language data suggesting that negated conditionals entail their corresponding conditional negations involve the interaction between quantifiers and conditionals (von Fintel and Iatridou 2002; Higginbotham 2003) and the focus sensitive expression *only* (von Fintel 1997).

This paper demonstrates, first, that CEM is inessential for explaining the empirical facts about negated *ifs* but also, second, that CEM can live happily in a strict analysis of conditionals. I do so by developing a bilateral update semantics for conditionals that distinguishes itself from previous frameworks by giving separate acceptance and rejection conditions for conditionals.

## 2 Negating Conditionals

I couch my analysis in a bilateral semantic setting that distinguishes between two distinct foundational semantic concepts, for instance: truth-making and false-making, support and rejection, or affirmation and denial. Here I will choose an update-based route and recursively define a positive, acceptance inducing update function  $[\cdot]^+$  and in addition a negative, rejection inducing update function  $[\cdot]^-$  on a set of possible worlds  $s$ .<sup>1</sup> The target language  $\mathcal{L}$  is the smallest set that contains a set of propositional atoms  $\mathcal{A} = \{p, q, r, \dots\}$ , and is closed under negation ( $\neg$ ), conjunction ( $\wedge$ ), disjunction ( $\vee$ ), and the conditional connective ( $\triangleright$ ). I begin by stating the bilateral approach and its immediate payoff (Section 2.1) and conclude by addressing a foundational issue about the analysis of negation proposed here (Section 2.2).

### 2.1 Framework

Start with the obvious entries for atomic sentences and negation.

$$(A) \quad \begin{array}{l} s[p]^+ = \{w \in s : w(p) = 1\} \\ s[p]^- = \{w \in s : w(p) = 0\} \end{array} \quad (\neg) \quad \begin{array}{l} s[\neg\phi]^+ = s[\phi]^- \\ s[\neg\phi]^- = s[\phi]^+ \end{array}$$

A positive update with  $p$  eliminates from the input all possible worlds at which  $p$  is false, while a negative update eliminates all possible worlds at which  $p$  is true. A positive update with  $\neg\phi$  is a negative update with  $\phi$ ; a negative update with  $\neg\phi$  is a positive update with  $\phi$ .

Conjunction is commonly modeled as sequential updating in dynamic semantics, but here we do not need to worry about internal dynamics, and so a simple intersective approach à la Veltman 1996 will do. A negative update with a conjunction amounts to taking the union of (i) the result of updating with the negation of the first conjunct and (ii) the result of updating with the negation of the second conjunct. Disjunction receives its standard definition in terms of negation and conjunction.

$$(\wedge) \quad \begin{array}{l} s[\phi \wedge \psi]^+ = s[\phi]^+ \cap s[\psi]^+ \\ s[\phi \wedge \psi]^- = s[\phi]^- \cup s[\psi]^- \end{array} \quad (\vee) \quad \begin{array}{l} s[\phi \vee \psi]^+ = s[\phi]^+ \cup s[\psi]^+ \\ s[\phi \vee \psi]^- = s[\phi]^- \cap s[\psi]^- \end{array}$$

<sup>1</sup>Thus the upcoming proposal is in the spirit of Veltman 1996 but defines two update functions instead of just one. For some alternative approaches to the bilateral program in semantics, see, among others, Aloni ms.; Fine 2017; Francez 2014; Incurvati and Schlöder Forthcoming; Rumfitt 2000.

More interesting is the analysis of conditionals. Ramsey (1931) famously suggested that a conditional is to be accepted in case its consequent is (hypothetically) accepted under the assumption of the antecedent. Virtually every semantic analysis of conditionals owes inspiration to Ramsey’s suggestion, and so does the story I am about to tell here. The key idea of this paper, however, starts with the intuition that there is something important missing from Ramsey’s proposal. In saying what it takes to accept a conditional, Ramsey of course also specifies what it takes to *fail* to accept a conditional: do so if you fail to (hypothetically) accept the consequent under the assumption of the antecedent. But he does not—or at least not obviously—state what it takes to *reject* a conditional, for the plain reason that failing to accept a conditional is not trivially the same as rejecting it. It is exactly this omission that I suggest leaves room for an implementation of the Ramsey test for conditionals that lives happily with the empirical facts about negated conditionals. Specifically, I suggest that Ramsey’s dictum about accepting conditionals is compatible with the following proposal about what it takes to *reject* a conditional: do so if you (hypothetically) reject the consequent under the assumption of the antecedent.

To make the proposal outlined in the previous paragraph more precise, say that a state  $s$  *accepts*  $\phi$ ,  $s \models^+ \phi$ , just in case  $s[\phi]^+ = s$ ; and that  $s$  *rejects*  $\phi$ ,  $s \models^- \phi$ , just in case  $s[\phi]^- = s$ . The basic proposal would then be that a conditional of the form ‘ $\phi > \psi$ ’ is accepted by  $s$  just in case  $s[\phi]^+ \models^+ \psi$  and rejected just in case  $s[\phi]^+ \models^- \psi$ . One minor wrinkle: following standard protocol (see, e.g., von Fintel 2001; Gillies 2007; Willer 2017) we shall assume that conditionals presuppose that their antecedent be compatible with the relevant modal domain and since—for current purposes anyway—conditionals impose tests on the input context  $s$ , this is just to presuppose that the conditional antecedent is a possibility in the input state. Putting all of this together, we say:

$$\begin{aligned} (>) \quad & s[\phi > \psi]^+ = \{w \in s : s[\phi] \models^+ \psi\} \text{ iff } s[\phi]^+ \neq \emptyset \\ & s[\phi > \psi]^- = \{w \in s : s[\phi] \models^- \psi\} \text{ iff } s[\phi]^+ \neq \emptyset \end{aligned}$$

Positive and negative updates with conditionals both require that their antecedent be compatible with the input state. If defined, a positive update with a conditional tests whether its consequent is accepted once the input state is strengthened with the antecedent. If defined, a negative update with a conditional tests whether its consequent is rejected once the input state is strengthened with the antecedent. A passed test returns the original state; a failed test results in the absurd state ( $\emptyset$ ).

Bilateral setups such as the one we are exploring here allow for a variety of notions of logical consequence, and we will exploit this flexibility momentarily (and use subscripts to keep track of the options). For now, one way to go is to think of valid inferences as those whose premises induce acceptance of the conclusion. Again following standard protocol, we take presupposition failures as constraining the states of information that we have to consider in evaluating an argument for validity: only states for which updating with the premises and the conclusion is defined matter when we check for validity (see e.g. Beaver 2001).

A sequence  $\phi_1, \dots, \phi_n$  *induces acceptance* of  $\psi$ ,  $\phi_1, \dots, \phi_n \models_1 \psi$ , just in case for all  $s$  such that  $s[\phi_1]^+ \dots [\phi_n]^+ [\psi]^+$  is defined,  $s[\phi_1]^+ \dots [\phi_n]^+ \models^+ \psi$ .

Thinking of entailment as acceptance inducing is to recreate update-to-test consequence (with the possibility of presupposition failures) in a bilateral setting. A special case: say that a state  $s$  *admits*  $\phi$  just in case updating  $s$  with  $\phi$  (positively or negatively) is defined. Then  $\phi$  is a *validity* just in case every state that admits  $\phi$  also accepts  $\phi$ .

The resulting framework then predicts that negated conditionals entail conditional negations (and vice versa):

**Fact 1**  $\neg(\phi > \psi) \models_1 \phi > \neg\psi$

This is easy to see since every positive update with  $\neg(\phi > \psi)$  is a negative update with  $\phi > \psi$ , which is testing whether  $\psi$  is rejected—and thus whether  $\neg\psi$  is accepted—under the assumption of  $\phi$ . But this is just what  $\phi > \neg\psi$  is asking, which establishes the fact.

At the same time, CEM turns out to be invalid if validity amounts to guaranteed acceptance by every admitting state:

**Fact 2**  $\not\models_1 (\phi > \psi) \vee (\phi > \neg\psi)$

A very simple counterexample is a state that accepts  $p$  but is agnostic about  $q$ , say  $s = \{w_1, w_2\}$  such that  $w_1(p) = w_2(p) = 1$  and  $w_1(q) = 1$  while  $w_2(q) = 0$ . Then  $s[p]^+ = s$  but clearly  $s \not\models^+ q$  and  $s \not\models^- q$  and hence  $s[p > q]^+ = s[p > \neg q]^+ = \emptyset$ . So  $s[(p > q) \vee (p > \neg q)]^+ = \emptyset$  and since  $s$  is not empty we have  $s[(p > q) \vee (p > \neg q)]^+ \neq s$  and so  $s \not\models^+ (p > q) \vee (p > \neg q)$ , which establishes the fact.

This concludes the first key lesson of the current exercise: that there need not be a trade-off between signing up for a universalist semantics that rejects CONDITIONAL EXCLUDED MIDDLE and accounting for how conditionals play with negation in natural language. What allows us to see that—if we adopt a Ramseyan perspective anyway—is the realization that an account of what it takes to *accept* a conditional lives happily with the existence of a separate characterization of what it takes to *reject* a conditional. Once we couch our analysis of conditionals in a *bilateral* setting that is grounded in two primitive basic semantic concepts—truth and falsity, acceptance and rejection, or positive and negative updating—it becomes straightforward to see that the problem with negating conditionals is really about getting the meaning of *negation* straight rather than coming up with a theory of conditionals that validates CEM.

## 2.2 Frege-Geach Style Worries

Bilateral stories like the one told here inevitably raise the spectre of the Frege-Geach problem. Textbook metaethical expressivism starts with the natural intuition that a sentence like (3a) “Stealing is wrong” is used to express the attitude of disapproval of stealing.<sup>2</sup> Natural as this may sound, it raises the question of what attitude the negation of (3a), i.e. (3b), expresses. Consider:

- (3) a. Stealing is wrong.  
 b. Stealing is not wrong.  
 c. Not stealing is wrong.

The initial observation here is that (3b) cannot express disapproval of not stealing, since this is what (3c) expresses; nor can it simply express the absence of the attitude of disapproval of stealing, since someone who is thoroughly agnostic or undecided about the moral status of stealing fails to disapprove of stealing but does not endorse (3b). The obvious conclusion to draw here is that an expressivist analysis of the language of morals must appeal to a second basic attitude besides disapproval—say, the attitude of *tolerance*—and that (3b) expresses that very attitude toward stealing.

The challenge then is to explain why (3a) and (3b) seem to express *incompatible* attitudes. This is especially pressing for expressivists, who aim at explaining the inconsistency of sentences

<sup>2</sup>See Schroeder 2010 and references therein for a discussion of metaethical expressivism and its history.



in terms of an incompatibility between the states of mind expressed by these sentences: so (3a) and (3b) must be inconsistent *because* they express incompatible states of mind. But the issue is perfectly general, and the problem, as Schroeder (2008) explains, is that all that the story sketched so far delivers is that (3a) and (3b) express distinct attitudes toward the same kind of action that—as far as the analysis is concerned—might very well be logically unrelated. If all we have are two basic attitudes—tolerance and disapproval—the incompatibility between the states of mind expressed by (3a) and (3b) remains a matter of pure stipulation.

Some have insisted that there is nothing wrong with grounding one’s semantics in a primitive notion of incompatibility between basic attitudes (see e.g. Gibbard 2013). My goal here is not to referee this issue or to dive into a detailed discussion of the language of morals but simply to highlight that the current framework has no need for such maneuvers. It would, on first sight anyway, be dissatisfying if we had to *stipulate* that no single consistent state of information can both accept and reject a single sentence at the same time. Fortunately, however, it is easy to verify that the attitudes of acceptance and rejection are incompatible in the following sense:

**Fact 3** For all  $s$  and  $\phi$ : if  $s \models^+ \phi$  and  $s \models^- \phi$ , then  $s = \emptyset$ .<sup>3</sup>

No non-absurd state can both accept and reject a sentence. For instance, suppose that  $s$  is non-absurd and accepts the conditional “If Mary is in Chicago, then Jack is in Rome” ( $c > r$ ). Then  $s[c]^+$  is a non-absurd state and, moreover,  $s[c]^+ \models^+ r$ . Since no non-absurd state can both accept and reject that Jack is in Rome (that is, exclusively consist of  $r$ -worlds and of  $\neg r$ -worlds), it follows that  $s[c]^+ \not\models^- r$  and so that  $s \not\models^- c > r$ , that is,  $s$  fails to reject “If Mary is in Chicago, then Jack is in Rome.” For parallel reasons, if  $s$  is non-absurd and rejects “If Mary is in Chicago, then Jack is in Rome,” then the state must fail to accept the conditional.

The previous result depends, of course, on the details of the semantics of negation: we *could* have provided negative entries for some or all of our connectives (or atomic sentences) that make the attitudes of acceptance and rejection compatible with each other. Nonetheless, the fact remains that there is nothing *per se* dubious about a framework that appeals to two primitive semantic notions in semantic theorizing, and that the incompatibility between the attitudes of acceptance and rejection is not a mere matter of stipulation but something that follows from our semantics. Let us explore the framework a bit further.

### 3 Toward Conditional Excluded Middle

While the story told so far has taught us something important about conditionals—namely, that the challenge of validating CONDITIONAL EXCLUDED MIDDLE is separate from the task of getting the facts about negated *ifs* in natural language straight—there is good reason to think that more needs to be said about the interplay between *if*, *or*, and *not*. To see this, consider again the case of a perfectly fair coin flip:

- (4) a. If Maria flipped the coin, it landed tails.  
b. If Maria flipped the coin, it landed heads.

The simple intuition here is that it makes good sense to assign .5 probability to (4a) and .5 probability to (4b); since (4a) and (4b) are mutually exclusive, their disjunction has a probability of 1 and thus the distinct air of a validity after all (cf. Santorio 2017). So while the

<sup>3</sup>The key underlying observation here is that if defined, the intersection of  $s[\phi]^+$  and  $s[\phi]^-$  is guaranteed to be empty (the proof is straightforward via induction). But if  $s \models^+ \phi$  and  $s \models^- \phi$ , then  $s[\phi]^+ = s[\phi]^- = s[\phi]^+ \cap s[\phi]^- = s$ , and so it follows that  $s$  must be the absurd state.



empirical data about negated conditionals do nothing to demonstrate that CEM is valid, other considerations might give it some (if perhaps not irresistible) appeal. In this section I will show how to accommodate this point, and I will do so by exploiting the fact (already indicated earlier) that a bilateral semantics provides us with some interesting options for how to think about logical consequence.

Start again with the difference between acceptance failures and rejection. Earlier we saw why a conditional like “ $(p > q) \vee (p > \neg q)$ ” need not be accepted by an information carrier: if one is certain about  $p$  but agnostic about  $q$ , one would not accept “ $p > q$ ” and one would not accept “ $p > \neg q$ ,” and so the whole disjunction would not be accepted either. But again failing to accept  $\phi$  is not the same as rejecting  $\phi$ , and in particular we can observe that while a state  $s$  need not accept “ $(p > q) \vee (p > \neg q)$ ,” it cannot reject it either. For doing so requires that  $s[(p > q) \vee (p > \neg q)]^- = s$  and so  $s[p > q]^- \cap s[p > \neg q]^- = s$ , which in turn requires that  $s$  must reject both “ $p > q$ ” and “ $p > \neg q$ .” And that, in turn, requires that  $s[p]^+$  is consistent but rejects both  $q$  and its negation, which simply cannot happen. The case, indeed, generalizes:

**Fact 4** For all  $s, \phi$  and  $\psi$ :  $s[(\phi > \psi) \vee (\phi > \neg\psi)]^- = \emptyset$ .

This is an immediate consequence of the fact that no non-absurd state can accept both  $\psi$  and its negation (recall Fact 3).

The previous fact is suggestive, since it highlights an alternative way of thinking about logical consequence and thus of validities. Earlier we suggested that an argument is valid just in case updating with the premises rationally commits one to accepting the conclusion; but instead we may think of an argument as valid just in case one could not rationally update with the premises and also reject the conclusion. More precisely:

A sequence  $\phi_1, \dots, \phi_n$  *excludes rejection* of  $\psi$ ,  $\phi_1, \dots, \phi_n \models_2 \psi$ , just in case for all  $s$  such that  $s[\phi_1]^+ \dots [\phi_n]^+[\psi]^+$  is defined,  $s[\phi_1]^+ \dots [\phi_n]^+[\psi]^- \models \perp$ .

Here “ $\perp$ ” represents an arbitrary contradiction. This proposal then treats  $\phi$  as a validity just in case  $s[\phi]^- = \emptyset$  for all states that admit  $\phi$ .

In a classical setup, it does not matter whether we say that the premises of an argument induce acceptance of a conclusion or that they exclude its rejection, for whenever  $\phi_1, \dots, \phi_n$  entails  $\psi$ , then  $\phi_1, \dots, \phi_n, \neg\psi$  entails  $\perp$ , and vice versa. But it does matter in the current setting. Specifically, it follows immediately from Fact 4 that CONDITIONAL EXCLUDED MIDDLE is valid if entailment requires that updating with the premises exclude rejection of the conclusion.

**Fact 5**  $\models_2 (\phi > \psi) \vee (\phi > \neg\psi)$

If no consistent state can accept  $\psi$  and its negation, it cannot reject  $\psi$  and its negation either. So in particular, if consistent,  $s[\phi]^+$  cannot reject  $\psi$  and its negation, and so (if defined in the first place)  $s[\phi > \psi]^- = \emptyset$  or  $s[\phi > \neg\psi]^- = \emptyset$ .

Validity understood as exclusion of rejection (“validity<sub>2</sub>” for short) is thus an interesting alternative to the more familiar concept of validity as guaranteed acceptance (“validity<sub>1</sub>” for short). It also preserves the earlier made prediction about the interplay between negated conditionals and conditional negations, since whenever an argument is valid<sub>1</sub>, it is also valid<sub>2</sub>:

**Fact 6** If  $\phi_1, \dots, \phi_n \models_1 \psi$ , then  $\phi_1, \dots, \phi_n \models_2 \psi$ .

No state that accepts  $\psi$  can be updated with the negation of  $\psi$  without resulting in the absurd state. So whenever  $\phi_1, \dots, \phi_n$  guarantees acceptance of  $\psi$ , it also excludes rejection of  $\psi$ . Thus validity<sub>2</sub> preserves what we want about negated conditionals but also delivers CEM as valid.

The second important upshot of this discussion, then, is that **CONDITIONAL EXCLUDED MIDDLE** lives happily in a framework that treats conditionals as universal quantifiers over a contextually determined set of possible worlds. The key idea here, again, is that rejecting a conditional is not the same as just failing to accept it. This, recall, allowed us to get the facts about negated conditionals straight without endorsing CEM. But it also turns out that while not every instance of CEM needs to be accepted, all of them are resistant to rejection, and this inspires a notion of validity as excluded rejection that delivers CEM as a tautology after all. Let me highlight another notable consequence before I conclude.

CEM is not the only principle that has some appeal but fails to hold in a large variety of conditional analyses. Another is **SIMPLIFICATION OF DISJUNCTIVE ANTECEDENTS (SDA)**:

$$(SDA) \quad (\phi \vee \psi) > \chi \models (\phi > \chi) \wedge (\psi > \chi)$$

While the case for CEM tends to be indirect, SDA has the air of an intuitive validity:

- (5) If you pay cash or with a debit card, you will receive a five percent discount.  
 $\rightsquigarrow$  If you pay cash, you will receive a five percent discount.  
 $\rightsquigarrow$  If you pay with a debit card, you will receive a five percent discount.

Intuitively, (5) suggests that you get five percent if you pay cash, *and* if you pay with a debit card. And indeed SDA is valid<sub>2</sub> in the framework developed here:<sup>4</sup>

**Fact 7**  $(\phi \vee \psi) > \chi \models_2 (\phi > \chi) \wedge (\psi > \chi)$

The underlying fact is that if  $\chi$  is accepted in some state  $s$ , then  $\chi$  it is not rejected by any consistent strengthened state  $s' \subseteq s$ . So if  $s[\phi \vee \psi]^+ \models^+ \chi$ , then both  $s[\phi]^+ \not\models^- \chi$  and  $s[\psi]^+ \not\models^- \chi$ . It follows that neither ' $\phi > \chi$ ' nor ' $\psi > \chi$ ' can be rejected by  $s$ .

Cariani and Goldstein (Forthcoming) show that, taken together, CEM and SDA have the potential of bringing a number of unwelcome consequences in their wake, given minimal additional assumptions. The perhaps most striking one is that the following principle turns out to be a validity, which Cariani and Goldstein label the **INTERCONNECTIVITY OF ALL THINGS (IAT)**:

$$(IAT) \quad \models ((\phi > \chi) \wedge (\psi > \chi)) \vee ((\phi > \neg\chi) \wedge (\psi > \neg\chi))$$

What makes IAT so unattractive is that it clashes with the uncontroversial fact that one may endorse the conjunction of (6a) and (6b):

- (6) a. If Maria comes to the party, it will be fun.  
 b. If Bill comes to the party, it won't be fun.

So IAT is no good. But (as Cariani and Goldstein lay out in more detail) if CEM is a validity, then so in particular is ' $((\phi \vee \psi) > \chi) \vee ((\phi \vee \psi) > \neg\chi)$ '. And if such conditionals simplify—as they should if SDA is really valid—it seems to follow right away that IAT is a validity as well.

And yet it is easy to verify that IAT is not a validity in the bilateral system that we have been exploring so far. Consider the instance " $((p > r) \wedge (q > r)) \vee ((p > \neg r) \wedge (q > \neg r))$ " and let  $s = \{w_1, w_2\}$  with the following distribution of truth-values:

<sup>4</sup>SDA is also valid<sub>1</sub> but for current purposes we will be specifically interested in a notion of validity that delivers both CEM and SDA.

	$p$	$q$	$r$
$w_1$	T	F	F
$w_2$	F	T	T

It is easy to see that  $s \models^+ p > \neg r$  and so  $s \models^- p > r$  and, moreover,  $s \models^+ q > r$  and so  $s \models^- q > \neg r$ . Accordingly,  $s[(p > r) \wedge (q > r)]^- = s$  and  $s[(p > \neg r) \wedge (q > \neg r)]^- = s$  and so  $s \models^- ((p > r) \wedge (q > r)) \vee ((p > \neg r) \wedge (q > \neg r))$ .

This is a surprising result: if CEM is valid and conditionals of all stripes simplify, how could IAT not be a validity as well? The answer is that validity<sub>2</sub> fails to be transitive.

**Fact 8** For some  $\phi$ ,  $\psi$ , and  $\chi$ :  $\phi \models_2 \psi$  and  $\psi \models_2 \chi$  but  $\phi \not\models_2 \chi$ .

Specifically, we have, with “ $\top$ ” being any tautology: (i)  $\top \models_2 ((p \vee q) > r) \vee ((p \vee q) > \neg r)$  and (ii)  $((p \vee q) > r) \vee ((p \vee q) > \neg r) \models_2 ((p > r) \wedge (q > r)) \vee ((p > \neg r) \wedge (q > \neg r))$ , but also (iii)  $\top \not\models_2 ((p > r) \wedge (q > r)) \vee ((p > \neg r) \wedge (q > \neg r))$ . The reason: (i) says that no state may consistently reject a certain CEM-instance and (ii) says that once some CEM-instance is accepted, the corresponding IAT instance cannot be rejected. But just because  $s$  cannot consistently reject  $\phi$  does not mean that  $s$  must accept  $\phi$ . Accordingly, just because some CEM-instance cannot be rejected does not mean that it must be accepted, and so certain IAT-instances may be rejected by non-absurd states of information.

Thinking of validity as exclusion of rejection, then, shows that CONDITIONAL EXCLUDED MIDDLE can live happily in Ramseyan setting, and that it can co-exist with SIMPLIFICATION OF DISJUNCTIVE ANTECEDENTS without negative side effects. More should—and can—be said about the intuitive foundations of this notion of validity; for now I submit that it is worthy of serious attention.

## 4 Conclusion and Outlook

The framework proposed here—like others before—has taken the Ramsey test as a source inspiration for its story about conditionals, but it adds to the classical Ramseyan question of what it takes to accept a conditional the one of what it takes to reject a conditional. The resulting bilateral semantic analysis of conditionals allows us to disentangle the question of how conditionals play with negation in natural language from the question of whether CONDITIONAL EXCLUDED MIDDLE is valid. It also demonstrates that the principle is consistent with the spirit of an analysis of conditionals as universal quantifiers over a contextually provided set of possible worlds. Other conditional heresies such as SIMPLIFICATION OF DISJUNCTIVE ANTECEDENTS may be adopted without running into triviality results.

Some important questions must be left to another day. I have not said how conditionals can induce non-trivial changes on the common ground if they articulate tests (in brief, by modeling the common ground as a set of sets of possible worlds); and I have remained silent on how to analyze *might*-conditionals if CEM holds (roughly, as conditionals with modalized consequents). But the perhaps most pressing technical issue is that, as things stand, validity<sub>2</sub> not only delivers CEM and SDA but also the unwelcome principle of ANTECEDENT STRENGTHENING:

$$(AS) \quad \phi > \chi \models (\phi \wedge \psi) > \chi$$

Lewis (1973) rejects AS since *Sobel sequences* appear to be perfectly consistent:

- (7) If Alice had come to the party, it would have been fun. But if Alice and Bert had come to the party, it would not have been fun.

(7) is a sequence of counterfactuals but the point applies to conditionals of all stripes (see e.g. Willer 2017). And yet it is easy to see that no state that treats  $\phi$  and  $\psi$  as possibilities can accept  $\chi$  if strengthened with  $\phi$  but reject  $\chi$  if strengthened with  $\phi$  and  $\psi$ .

Nothing about the antecedent strengthening facts undermines the important result that we can endorse CEM and SDA without IAT, but let me at least sketch how AS may be avoided in the current setting. The basic (and familiar) idea is that conditional antecedents may bring hitherto ignored possibilities into view and so while *ifs* are universal quantifiers over a set of possible worlds, this domain evolves dynamically as discourse proceeds (see von Stechow 2001 and Gillies 2007 for seminal discussion). So in entertaining the possibility of Alice coming to the party, we might ignore the possibility of Bert coming as well, and accept that the party will be fun on these grounds. Once we consider a conditional whose antecedent presupposes the possibility of Alice and Bill coming to the party, the modal horizon shifts, and it may very well be that all the possibilities thus brought into view are such that the party is not fun.

What underlies the invalidity of AS is then that one may entertain the possibility of  $\phi$  without entertaining the possibility of  $\phi \wedge \psi$ . Importantly, this response to the trouble with antecedent strengthening in a strict setting can be elaborated in such a way that we *preserve* the validity of SDA, the key observation concerning what it takes to entertain a disjunctive possibility (see Willer 2018 for details). One may, for sure, entertain the possibility of Alice coming to the party without Alice and Bert coming to the party. But in entertaining the possibility of Alice or Bert coming to the party, one is entertaining the possibility of Alice coming to the party, and the one of Bert coming to the party. The idea, in brief, then is that a conditional such as (8a) already brings into view the possibilities that matter for the antecedents of (8b) and (8c).

- (8) a. If Alice or Bill come to the party, it will be fun.  
 b. If Alice comes to the party, it will be fun.  
 c. If Bill comes to the party, it will be fun.

The basic observation is that disjunctive possibilities imply each of their disjuncts as possibilities (Kamp 1973). So if conditionals presuppose the possibility of their antecedent, we expect them to simplify: the possibilities of the simplified conditionals are already in view and so there is no horizon shift. This marks the crucial difference with AS, where a horizon shift is very well possible. Clearly, this explanatory strategy does not effect the story about CEM, since  $\neg(\phi > \psi)$  and  $\neg\phi > \psi$  share the same possibility presuppositions.

I have not said much about the conceptual foundations of a bilateral approach to semantic theorizing. It is worth repeating, however, that the story about negation allows us to derive acceptance and rejection as incompatible attitudes, rather than having to stipulate that they are. As such, the theory faces no Frege-Geach style problems with negation. We thus have every reason to be confident that the bilateral semantics outlined here provides a conceptually and technically stable foundation for exploring the interplay between negation and conditionals in natural language.

**Acknowledgments** Warm thanks to three anonymous reviewers for their comments.

## References

- Maria Aloni. FC disjunction in state-based semantics, ms. Manuscript, University of Amsterdam.
- David I. Beaver. *Presupposition and Assertion in Dynamic Semantics*. CSLI Publications, Stanford, 2001.
- Fabrizio Cariani and Simon Goldstein. Conditional heresies. *Philosophy and Phenomenological Research*, Forthcoming. <http://dx.doi.org/10.1111/phpr.12565>.
- Kit Fine. Truthmaker semantics. In Bob Hale, Crispin Wright, and Alexander Miller, editors, *A Companion to the Philosophy of Language*, pages 556–577. Wiley-Blackwell, Malden, MA, 2nd edition, 2017.
- Kai von Fintel. Bare plurals, bare conditionals, and *Only*. *Journal of Semantics*, 14(1):1–56, 1997. <http://dx.doi.org/10.1093/jos/14.1.1>.
- Kai von Fintel. Counterfactuals in a dynamic context. In Michael Kenstowicz, editor, *Ken Hale: A Life in Language*, pages 123–152. MIT Press, Cambridge, MA, 2001.
- Kai von Fintel and Sabine Iatridou. If and when *If*-clauses can restrict quantifiers. Manuscript, MIT, 2002.
- Nissim Francez. Bilateral relevant logic. *Review of Symbolic Logic*, 7(2):250–272, 2014. <http://dx.doi.org/10.1017/S1755020314000082>.
- Allan Gibbard. *Meaning and Normativity*. Oxford University Press, New York, 2013. <http://dx.doi.org/10.1093/acprof:oso/9780199646074.001.0001>.
- Anthony S. Gillies. Counterfactual scorekeeping. *Linguistics and Philosophy*, 30(3):329–360, 2007. <http://dx.doi.org/10.1007/s10988-007-9018-6>.
- James Higginbotham. Conditionals and compositionality. *Philosophical Perspectives*, 17(1):181–194, 2003. <http://dx.doi.org/10.1111/j.1520-8583.2003.00008.x>.
- Luca Incurvati and Julian J. Schlöder. Inferential expressivism and the negation problem. In Russ Shafer-Landau, editor, *Oxford Studies in Metaethics*, volume 16. Oxford University Press, New York, Forthcoming.
- Hans Kamp. Free choice permission. *Proceedings of the Aristotelian Society*, 74:57–74, 1973. URL <http://www.jstor.org/stable/4544849>.
- Angelika Kratzer. Conditionals. *Chicago Linguistics Society*, 22:1–15, 1986.
- David K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, MA, 1973.
- Frank P. Ramsey. *The Foundations of Mathematics*. Routledge and Kegan Paul, London, 1931.
- Iam Rumfitt. “Yes” and “No”. *Mind*, 109(436):781–823, 2000. <http://dx.doi.org/10.1093/mind/109.436.781>.
- Paolo Santorio. Conditional excluded middle in informational semantics. In Alexandre Cremers, Thom van Gestel, and Floris Roelofsen, editors, *Proceedings of the 21st Amsterdam Colloquium*, pages 385–394, Amsterdam, 2017. ILLC Publications.
- Mark A. Schroeder. *Being For: Evaluating The Semantic Program of Expressivism*. Oxford University Press, New York, 2008. <http://dx.doi.org/10.1093/acprof:oso/9780199534654.001.0001>.
- Mark A. Schroeder. *Noncognitivism in Ethics*. Routledge, New York, 2010.
- Robert C. Stalnaker. A defense of conditional excluded middle. In William L. Harper, Robert C. Stalnaker, and Glenn Pearce, editors, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pages 87–104. Reidel, Dordrecht, 1981. [http://dx.doi.org/10.1007/978-94-009-9117-0\\_4](http://dx.doi.org/10.1007/978-94-009-9117-0_4).
- Frank Veltman. *Logics for Conditionals*. PhD thesis, University of Amsterdam, 1985.
- Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996. <http://dx.doi.org/10.1007/BF00248150>.
- Malte Willer. Lessons from sobel sequences. *Semantics and Pragmatics*, 10(4):1–57, 2017. <http://dx.doi.org/10.3765/sp.10.4>.
- Malte Willer. Simplifying with free choice. *Topoi*, 37(3):379–392, 2018. <http://dx.doi.org/10.1007/s11245-016-9437-5>.

# Two types of higher-order readings of WH-questions\*

Yimei Xiang

Rutgers University, New Brunswick, New Jersey, USA  
yimei.xiang@rutgers.edu

## Abstract

Evidence from questions with modals and collective predicates suggests that sometimes WH-questions must be interpreted with higher-order readings. In such readings, questions expect answers naming generalized quantifiers rather than entities. This paper investigates the distribution and the derivation of those readings. First, I argue that the generalized quantifiers that can serve as semantic answers to questions are subject to two constraints — Positiveness and Homogeneity. Next, I present two ways to account for a distributional constraint with higher-order readings, that is, questions with a singular-marked or numeral-modified WH-complement admit elided disjunctive answers but not conjunctive answers.

## 1 Introduction

WHPs are commonly treated as functions (e.g., existential ( $\exists$ )-quantifiers or domain restrictors) over first-order (FO-)predicates. In this view, the domain for quantification or abstraction is the set denoted by the extension of the WH-complement. The following uses categorial approaches to illustrate this idea (framework chosen just for ease of presentation): in the question *which boy left*, the WHP *which boy* combines with a FO-predicate  $\lambda x_e.\text{left}(x)$  and restricts the domain to the set of atomic boys. The yielded question denotation is a FO-predicate  $\lambda x_e : \text{boy}(x).\text{left}(x)$ . It can serve as a function for an entity-denoting answer as in (1a), and an argument of a generalized quantifier (GQ)-denoting answer as in (1b).<sup>1</sup> I henceforth call this functional question denotation a “Q-function” and its domain a “Q-domain”.

- (1) a.  $\llbracket \text{Q} \rrbracket(\llbracket \text{Andy} \rrbracket) = (\lambda x_e : \text{boy}(x).\text{left}(x))(a) = \text{prof}(a).\text{left}(a)$   
b.  $\llbracket \text{Andy or Bill} \rrbracket(\llbracket \text{Q} \rrbracket) = (a^\uparrow \cup b^\uparrow)(\lambda x_e : \text{boy}(x).\text{left}(x)) = \text{boy}(a) \wedge \text{boy}(b).\text{left}(a) \vee \text{left}(b)$

In some cases, however, a WH-question can only be completely addressed by a GQ. This fact shows that WH-questions have also higher-order (HO-)readings, in which the yielded Q-functions take GQs as arguments. More strikingly, there seems to be two types of HO-readings, which admit different HO-answers and have different distributions. In particular, questions in which the WHP is number-unmarked (*who*, *what*) or the WH-complement is bare plural (*which children*) admit both conjunctions and disjunctions, while questions in which the WH-complement is SG-marked (*which child*) or NUM-modified (*which two children*) admit only disjunctions. I call these two readings ‘conjunction-admitting’ and ‘disjunction-only’, respectively.

The rest of this paper is organized as follows. Section 2 provides evidence for the existence of HO-readings. Section 3 discusses the constraints as to which HO-meanings can be included

---

\*For helpful discussions, I thank Luis Alonso-Ovalle, Lucas Champollion, Gennaro Chierchia, Danny Fox, Manuel Kriz, Benjamin Spector, Bernhard Schwarz, the audiences at Georg-August-Universität Göttingen and Ecole Normale Supérieure, and the abstract reviewers of Amsterdam Colloquium 2019. All errors are mine.

<sup>1</sup>Here ‘ $a^\uparrow \cup b^\uparrow$ ’ denotes the disjunction of two Montagovian individuals. Disjunctions are standardly treated as set **unions**. Since entities are not sets, they must be Montague-lifted into GQs before being disjoined: for any entity  $\alpha$ , the corresponding Montagovian entity is  $\alpha^\uparrow$  such that  $\alpha^\uparrow = \lambda m_{\langle e,t \rangle}.m(\alpha)$ . This idea follows a more general schema in Partee & Rooth 1983. The conjunctive *and* is usually treated as ambiguous between as denoting an **intersection** operator ‘ $\cap$ ’ or a **summation** operator ‘ $\oplus$ ’ (Link 1983; Hoeksema 1988; a.o).

in a Q-domain. Sections 4 and 5 explain the distributional constraints of the two HO-readings and derive them compositionally. Section 6 concludes.

## 2 Evidence for higher-order readings

Treating questions as FO-functions yield two predictions with respect to GQ-denoting answers. First, the specified GQs must be interpreted with a wide scope relative to the scopal elements in the question nucleus. Second, for any propositional answer  $p$  specifying a GQ  $\pi$ ,  $p$  and  $\pi$  are not in the answer space (viz., Hamblin set) and the Q-domain of this question; instead, the derivation of  $p$  involves additional operations to the propositions in the answer space. This section presents counterexamples to both predictions, showing that FO-reading is insufficient.

**Evidence for disjunctions and  $\exists$ -quantifiers** Spector (2007, 2008) observes that elided disjunctions can completely address WH-questions with a necessity modal (“ $\square$ -questions” henceforth). As seen in (2b), interpreted with a narrow scope, the disjunction can be the complete specification of John’s reading obligations — there is no specific book that John has to read, he just has to choose between the French novels and the Russian novels. The same complete answer reading is available with  $\exists$ -quantifiers, as in (3).

- (2) “Which books does John have to read?” “The French novels or the Russian novels.”  
 a. ‘John has to read F or R, I do not know which exactly.’ (Partial:  $or \gg \square$ )  
 b. ‘John has to read F or R, and the choice is up to him.’ (Complete:  $\square \gg or$ )
- (3) “Which books does John have to read?”  
 “At least/ More than/ Exactly two books by Balzac.” ( ${}^{OK}\exists \gg \square, {}^{OK}\square \gg \exists$ )

Spector derives the complete answer reading of (2b) by interpreting the disjunctive answer as a GQ  $f^\uparrow \cup r^\uparrow$  (see definitions in footnote 1) and assuming semantic reconstruction in the nucleus (*a la* Cresti 1995). Adapting this analysis to the categorial approach, I assume the LF and Q-function as in (4a-b). In the LF, WH-movement leaves a HO-trace  $\pi$  (of type  $\langle et, t \rangle$ ) under the  $\square$ -modal *have to*. The yielded Q-function is defined for GQs that range over a set of books.<sup>2</sup> (This domain condition is to be revised.) Combining with this HO-function, a GQ-denoting answer is interpreted at whichever scopal position that  $\pi$  takes, as seen in (4c).

- (4) a. [Which-books  $\lambda\pi$  [have-to [ $\pi_{\langle et, t \rangle}$   $\lambda x$  [John read  $x_e$ ]]]]  
 b.  $[[Q]] = \lambda\pi_{\langle et, t \rangle} : \text{SMLO}(\pi) \subseteq \text{*book}.\square[\lambda w.\pi(\lambda x.\text{read}_w(j, x))]$  (To be revised)  
 c.  $[[Q]](f^\uparrow \cup r^\uparrow) = \{f, r\} \subseteq \text{*book}.\square[\lambda w.\text{read}_w(j, f) \vee \text{read}_w(j, r)]$

$\square$ -questions provide evidence for ruling in disjunctions and  $\exists$ -quantifiers because they may yield Q-functions that are non-reducible relative to disjunctions and  $\exists$ -quantifiers, as in (5a). This diagnostic, however, does not extend to conjunctions —  $\lambda\pi.J$  has to read  $\pi$  is reducible relative to conjunctions, as seen in (5b).

- (5) A function  $\theta$  is reducible relative to a GQ  $\pi$  iff  $\theta \circ \pi = \pi(\lambda x_e.\theta \circ x^\uparrow)$  where ‘ $\circ$ ’ stands for the operation for combining  $\theta$  with a GQ. For example:  
 a.  $[\lambda\pi.J \text{ has to read } \pi](f^\uparrow \cup r^\uparrow) \neq J \text{ has to read } f \vee J \text{ has to read } r$   
 b.  $[\lambda\pi.J \text{ has to read } \pi](f^\uparrow \cap r^\uparrow) = J \text{ has to read } f \wedge J \text{ has to read } r$

<sup>2</sup>Here ‘SMLO( $\pi$ )’ stands for the smallest **live-on** set of  $\pi$ . For any  $\pi$  (of type  $\langle \tau t, t \rangle$ ) and  $A$  (of type  $\langle \tau, t \rangle$ ), we have  $\pi$  lives on  $A$  iff  $\forall B[\pi(B) \leftrightarrow \pi(A \cap B)]$  (Barwise & Cooper 1981), and  $\pi$  **ranges over**  $A$  iff  $A$  is the smallest live-on set of  $\pi$  (Szabolcsi 1997). In most cases (except for exceptives such as *every/no student but John*), the smallest live-on set of a GQ of the form ‘Det-NP’ is the set denoted by the extension of the complement NP.



**Evidence for conjunctions and  $\forall$ -quantifiers** The following presents the arguments from Xiang (2016: ch 1) for ruling in conjunctions, which draw on the lack of uniqueness effects in questions with a stubbornly collective predicate.

First, for stubborn collectivity, observe in (6) that the quantized collective predicates *formed a/one team* admits only a collective reading, in contrast to its plural counterpart *formed teams* and other collective predicates like *lifted the piano* which admit also cumulative/ non-atomic distributive readings. Next, for absence of uniqueness, compare (7a-b) in the same context. The declarative-embedding sentence (7a) suffers a presupposition failure — the factive *know* embeds a collective declarative that is false in the given context. In contrast, the corresponding question-embedding sentence (7b) is felicitous, and it implies the inference (7c) that John knows precisely the component members of all the teams formed by the considered kids.

- (6) (*w*: The four kids formed two teams in total:  $a + b$  formed one, and  $c + d$  formed one.)
- a. # The kids formed a/one team.
  - b. The kids formed teams.
- (7)
- a. # John knows [that **the kids** formed a/one team].
  - b. John knows [**which kids** formed a/one team].
  - c.  $\rightarrow$  John knows that  $a + b$  formed a team and  $c + d$  formed a team.

The conjunctive inference in (7c) is quite surprising — where does the conjunctive closure come from? Clearly, no matter how we analyze distributivity/collectivity, this closure cannot come from the predicate *formed a/one team* or anywhere within the question nucleus, otherwise (6a) would admit a distributive/covered reading. I propose that this conjunctive closure comes from the WHP: the WHP quantifies over a set of HO-meanings including the conjunction  $(a \oplus b)^\uparrow \cap (c \oplus d)^\uparrow$  (i.e., the intersection of Montagovian plurals). The LF and Q-function are as follows:

- (8) Which kids formed a team?
- a. [which-kids  $\lambda\pi$  [<sub>IP</sub>  $\pi_{\langle et,t \rangle}$   $\lambda x$  [<sub>VP</sub>  $x_e$  formed a team ]]]
  - b.  $[[Q]] = \lambda\pi_{\langle et,t \rangle} : \text{SMLO}(\pi) \subseteq \text{*kid}.\lambda w[\pi(\lambda x.\text{f-a-t}_w(x))]$  (To be revised)
  - c.  $[[Q]]((a \oplus b)^\uparrow \cap (c \oplus d)^\uparrow) = \{a \oplus b, c \oplus d\} \subseteq \text{*kid}.\lambda w[\text{f-a-t}_w(a \oplus b) \wedge \text{f-a-t}_w(c \oplus d)]$

One might suggest to ascribe this conjunctive closure to operations outside the question root. The answerhood-operator of Heim (1994), for example, returns the conjunction of the true propositions in the answer space. This idea, however, cannot account for the contrast in (9): (9b) is infelicitous because the NUM-modified WHP yields a false uniqueness inference.

- (9) (Context is the same as in (6).)
- a. John knows [**which kids** formed a/one team].
  - b. # John knows [**which two kids** formed a/one team].  
 $\rightarrow$  Among the relevant kids, exactly two of them formed one single team.

Uniqueness effects in questions are standardly explained by “Dayal’s (1996) presupposition” — a question is defined only if it has a strongest true answer. In the given context, the strongest true answer to the embedded question in (7b) is the conjunctive proposition (8c), derived based on the Boolean conjunction  $(a \oplus b)^\uparrow \cap (c \oplus d)^\uparrow$ . To account for the uniqueness effect in (9), one just needs to assume that the WHP *which two kids* (in contrast to *which kids*) does not quantify over HO-meanings and thus does not support HO-readings.



**Evidence for GQ-coordinations** The diagnostics based on non-reducibility and stubborn collectivity rule in GQs such as dis/con-junctions and  $\exists/\forall$ -quantifiers. Combinations of these two diagnostics also rule in the coordinations of these GQs. For example, the question (10a) involves both a necessity modal *have to* and a stubbornly collective predicate *present a paper together*. The answer should be interpreted as a Boolean disjunction over Boolean conjunctions of plurals (viz.,  $((a_1 \oplus b_1)^\uparrow \cap (c_1 \oplus d_1)^\uparrow) \cup ((a_2 \oplus b_2)^\uparrow \cap (c_2 \oplus d_2)^\uparrow)$ ) and be read with the following scopal pattern:  $\square \gg or \gg and \gg a\ paper$ . To obtain this scopal pattern, the question nucleus should involve a HO-WH-trace in between the modal and the collective predicate, as in (10c).

- (10) a. “Which students have to present a paper together this week?”  
 b. “The two junior linguists ( $a_1 \oplus b_1$ ) and the two senior linguists ( $c_1 \oplus d_1$ ), OR, the two junior philosophers ( $a_2 \oplus b_2$ ) and the two senior philosophers ( $c_2 \oplus d_2$ ).”  
 c. [which-students  $\lambda\pi$  [have to [ $\pi_{\langle et,t \rangle}$   $\lambda x$  [ $x_e$  present a paper together this week]]]]

### 3 Constraints of higher-order Q-domain

Section 2 has provided evidence for ruling in a variety of GQs and GQ-coordinations into the Q-domain. This section, in contrast, discusses GQs that should be excluded from the Q-domain.

Consider the truth conditions of the Q-embedding sentence (11b) under the context (11a). Strikingly, (11b) implies that Sue knows John’s reading obligation (i), but not that she knows (ii); Sue can be ignorant about whether John should read any books by Betty. Given this contrast, Spector (2008) proposes that the GQs involved in a Q-domain must be increasing: ‘ $x$  knows Q’ implies that  $x$  knows the complete/strongest true answer of Q; therefore, that Sue can be ignorant about (ii) excludes the decreasing GQ *no book by Betty* and the non-monotonic GQ-coordination *at least two books by Anne and no book by Betty* from the Q-domain.

- (11) a. Context: John’s reading obligations include the following:  
 (i) he must read at least two books by Anne; (ii) he must read no book by Betty.  
 b. Sue knows which books John must read.

The following example extends Spector’s diagnostic to a broader range of GQs. Among the list of GQs, (a-b) are increasing, (c-d) are decreasing, and (e) is non-monotonic. Intuitively, the question-embedding sentence (13) implies that Sue knows about not only (a-b) but also (e).

- (12) John is playing a board game. This game requires that him to play ...  
 a. {at least three, more than two} red spades;  
 b. every black spade except the smallest one in his hand;  
 c. {at most three, less than four} black diamonds;  
 d. no red diamond except largest one in his hand;  
 e. exactly two hearts;
- (13) Sue knows which cards John must play.

Spector’s increasing-ness constraint incorrectly rules out the non-monotonic GQ *exactly two hearts* and fails to predict that (13) implies that Sue knows the requirement (e). In contrast, I propose that whether a HO-meaning should be ruled in into a Q-domain is determined by its **positiveness** — only positive GQs and their Boolean coordinations can be in a Q-domain. A GQ being positive means that it ensures existence. For example, *at least two books* and *exactly two books* entail *some books*, while *no book* and *less than two books* do not entail *some book(s)*. A

formal definition of positiveness is as follows, where **some** is the basic  $\exists$ -quantifier, and  $\text{SMLO}(\pi)$  is the smallest live-on set of  $\pi$  (see definitions in footnote 2):

(14) For any  $\pi$  of type  $\langle et, t \rangle$ ,  $\pi$  is **positive** iff  $\pi \subseteq \text{some}(\text{SMLO}(\pi))$ .

Table 1 compares monotonicity and positiveness for GQs that range over books. Observe that increasing GQs are positive, while decreasing ( $\downarrow_{\text{MON}}$ ) and non-monotonic (N.M.) GQs are not.

GQ $\pi$	SMLO( $\pi$ )	Increasing?	Positive?
at least two books	books	Yes	Yes
every book except $a$	book - $\{a\}$	Yes	Yes
at most two books	books	No ( $\downarrow_{\text{MON}}$ )	No
no book except $a$	book - $\{a\}$	No ( $\downarrow_{\text{MON}}$ )	No
exactly two books	books	No (N.M.)	Yes
two to ten books	books	No (N.M.)	Yes

Table 1: Increasing-ness/monotonicity versus positiveness

Table 1 considers only simplex GQs of the form ‘Det+NP’. Benjamin Spector (p.c.) points out that positiveness does not exclude the unwanted non-monotonic GQ-coordinations such as *every article and no book*. Representing this GQ-conjunction  $\pi$  as  $\{E \mid A \subseteq E \wedge B \cap E = \emptyset\}$ , we have  $\text{SMLO}(\pi) = A \cup B$ ,<sup>3</sup> and  $\pi \subseteq \text{some}(A \cup B)$ . Benjamin Spector further points out that a GQ  $\pi$  can be semantically decomposed into the conjunction  $\pi^+ \cap \pi^-$ , where  $\pi^+$  is the strongest increasing GQ entailed by  $\pi$ , and  $\pi^-$  is the strongest decreasing GQ entailed by  $\pi$ . In particular,  $\pi$  is increasing iff  $\pi^- = D_{\langle e, t \rangle}$ , and  $\pi$  is decreasing iff  $\pi^+ = D_{\langle e, t \rangle}$ .

(15) For any  $\pi_{\langle et, t \rangle}$ :  $\pi^+ =_{\text{df}} \{P \mid \exists P' \subseteq P[\pi(P')]\}$ , and  $\pi^- =_{\text{df}} \{P \mid \exists P' \supseteq P[\pi(P')]\}$ .

To rule out non-monotonic GQ-coordinations like *every article and no book*, we just need to exclude the non-monotonic  $\pi$ s where the retrieved  $\pi^+$  and  $\pi^-$  range over different sets. For example, the simplex GQ *exactly two books* is not excluded since it is formed by conjoining two GQs ranging over the same set **books** (i.e., *at least two books* and *no more than two books*). I call such GQs ‘homogenous’, defined as in (16). Table 2 compares homogeneity and positiveness. Observe that *some A and/or no B* is positive but not homogenous.

(16) For any  $\pi$  of type  $\langle et, t \rangle$ ,  $\pi$  is **homogenous** iff  $\pi$  is monotonic or  $\text{SMLO}(\pi^+) = \text{SMLO}(\pi^-)$ .

GQ $\pi$	SMLO( $\pi$ )	SMLO( $\pi^+$ )	SMLO( $\pi^-$ )	Positive?	Homogenous?
at least two B	$B$	$B$	$D_e$	Yes	Yes
exactly two B	$B$	$B$	$B$	Yes	Yes
every A and no B	$A \cup B$	$A$	$B$	Yes	No

Table 2: Positiveness versus homogeneity

In conclusion, the Q-domain yielded by the phrase ‘wh-A’ in a HO-reading, if any, is the set consisting of the positive homogenous GQs ranging over a subset of  $A$ . I write this set as  ${}^{\text{H}}A$ .

<sup>3</sup> The following shows why  $A \cup B$  is the smallest live-on set of  $\pi$ : for any  $a$ , replacing  $E$  with  $E \cap (A \cup B - \{a\})$  in the set description makes no change to the defined set of  $E$  iff  $a \notin A$  and  $a \notin B$ .

$$\begin{aligned}
\text{(i)} \quad & \{E \mid [A \subseteq (E \cap ((A \cup B) - \{a\}))] \wedge [B \cap (E \cap ((A \cup B) - \{a\})) = \emptyset]\} \\
& = \{E \mid [A \subseteq E \wedge A \subseteq (A \cup B - \{a\})] \wedge [(B \cap ((A \cup B) - \{a\})) \cap E = \emptyset]\} \\
& = \{E \mid [A \subseteq E \wedge A \subseteq (A \cup B) \wedge a \notin A] \wedge [(B - \{a\}) \cap E = \emptyset]\} \quad (= \pi \text{ iff } a \notin A \cup B)
\end{aligned}$$

$$(17) \quad {}^H A = \{\pi_{\langle e,t \rangle} \mid \text{SMLO}(\pi) \subseteq A \wedge \pi \subseteq \text{some}(\text{SMLO}(\pi)) \wedge \pi \text{ is homogenous}\}$$

It is yet unclear where the positiveness and homogeneity constraints come from. They could be in the lexicon of a type-shifting operator, presupposed by the HO-WH-trace, or constraints on semantic reconstruction. For now, I treat ‘H’ as a syntactically presented operator that asserts positiveness and homogeneity.

## 4 Distributing the ‘conjunction-admitting’ reading

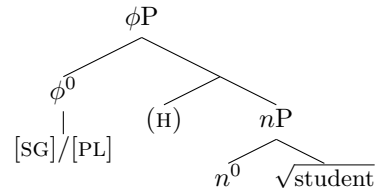
As argued in section 2, uniqueness effects in SG-marked and NUM-modified questions show that the Q-domain of these question do not contain Boolean conjunctions and further that these questions do not have HO-readings.<sup>4</sup> Interestingly, unlike NUM-modifiers, PP-modifiers such as *in a group of two* do not trigger uniqueness and do not block HO-readings. This contrast suggests that the availability of HO-reading is sensitive to the internal structure of the *wh*-complement.

- (18) (*w*: *a + b* presented a paper together, and *c + d* presented a paper together.)
- # I know which **two students** presented a paper together.
  - I know which **students in a group of two** presented a paper together.

To account for the above distributional constraints, I propose that the H-shifter is applied locally to the root *nP* within the *wh*-complement. In the following, I argue that the application of H is blocked in SG-marked nouns and NUM-modified nouns due to conflicts in meaning and types.

First, I assume the structure in (19) for a singular/plural bare noun. The denotation of the root *nP* is a semi-lattice structure (Harbour 2014). The number feature [SG]/[PL] is evaluated at  $\phi^0$ . Following Sauerland (2003), I interpret [SG] as an atomicity modifier while treating [PL] semantically vacuous. This analysis explains why SG-marked WHPs do not support HO-readings: applying H to *nP* returns a set of GQs, which are all non-atomic and are conflicting with the atomicity requirement of [SG]. In contrast, H-shifter can be freely used in simple PL-marked and number-neutral *whPs* because in those cases the  $\phi^0$  does not presuppose atomicity.

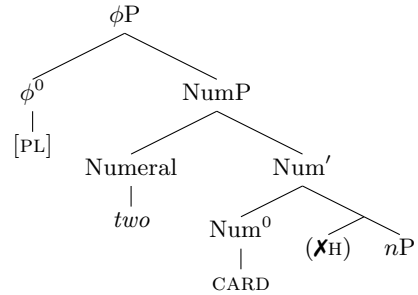
- (19) *student* (consider three students *abc*)
- $\llbracket [\text{SG}] \rrbracket = \lambda P_{\langle e,t \rangle} \lambda x_e. \text{ATOM}(x) \wedge P(x)$
  - $\llbracket [n\text{P}] \rrbracket = \{a, b, c, a \oplus b, a \oplus c, b \oplus c, a \oplus b \oplus c\}$
  - $\llbracket [\text{SG}](n\text{P}) \rrbracket = \{a, b, c\}$
  - $\llbracket [\text{SG}](\text{H}(n\text{P})) \rrbracket = \emptyset$



Next, for NUM-modified NPs, I place cardinal NUM-modifiers at [Spec, NumP] and assume that Num<sup>0</sup> is occupied by a cardinality predicate CARD (Scontras 2014). As defined in (20a), CARD takes a predicate *P* and a numeral *n* and returns the set of individuals in *P* each of which is constituted of exactly-*n* atoms. These assumptions automatically explain why the H-shifter cannot be used in a NUM-modified NP: the CARD-predicate at Num<sup>0</sup> requires to check the cardinality of the elements in the set it combines with and hence it cannot combine with a set of GQs. In contrast to NUM-modifiers, PP-modifiers are adjoined to the entire NP/ $\phi$ P. Hence, the H-shifter can be used in a PP-modified NP without causing a type-mismatch.

<sup>4</sup>Note that a SG-marked phrase can be number-neutral in semantics (e.g., Spanish *quién* ‘who.SG’). For details see Maldonado (2017), Elliot et. al. (2018), and Alonso-Ovalle & Rouillard (2019).

- (20) *two students*
- $\llbracket \text{CARD} \rrbracket = \lambda P \lambda n \lambda x. P(x) \wedge |x| = n$
  - Without the H-shifter  
 $\llbracket \text{Num}' \rrbracket = \lambda n \lambda x. * \text{stdt}(x) \wedge |x| = n$   
 $\llbracket \text{NumP} \rrbracket = \lambda x. * \text{stdt}(x) \wedge |x| = 2$   
 $= \{a \oplus b, b \oplus c, a \oplus c\}$
  - With the H-shifter:  
 Num' suffers type-mismatch



## 5 Deriving the ‘disjunction-only’ reading

Surprisingly, in responding to  $\square$ -questions with a SG-marked or NUM-modified WHP, narrow scope disjunctions are not as bad as conjunctions, as seen in (21).

- (21) “Which textbook should we read for this class?”  
 “*Heim&Kratzer* or/#and *Meaning&Grammar*, (the choice is up to you.)” ( $\square \gg \text{or}$ )

A similar contrast is observed with questions with a possibility modal (“ $\diamond$ -questions” henceforth). As in (22), multi-choice answers to  $\diamond$ -questions can be equivalently expressed with an elided conjunction or an elided disjunction. In particular, the disjunction is read as free choice.

- (22) “What can I use (as a textbook) for this class?”  
 “*Heim&Kratzer* or/and *Meaning&Grammar*, (the choice is up to you.)”

For SG-marked  $\diamond$ -questions, it was commonly thought that these questions are subject to uniqueness and hence cannot have multi-choice readings. Hirsch & Schwarz (2019) finds out that, however, multi-choice readings are available as long as the uniqueness inference evoked by the SG-marked WHP takes narrow scope, as seen in (23).

- (23) “Which letter could be missing in *fo\_m*?”  
 “The missing letter could be *a*, and the missing letter could be *r*.”

In Hirsch & Schwarz’s example (23), the multi-choice answer can be expressed as a conjunction over two full sentences. Interestingly, as seen in (24), if elided, multi-choice answers to SG-marked  $\diamond$ -questions must be disjunctions. NUM-modified  $\diamond$ -questions have the same behavior.<sup>5</sup>

- (24) a. “Which textbook can I use for this class?” “*H&K* or/ #and *M&G*.”  
 b. “Which letter could be missing in *fo\_m*?” “Letter *a* or/ #and letter *r*.”

Here arise two puzzles: first, why these questions admit and only admit disjunctions? second, why such a ‘disjunction-only’ HO-reading is available despite that the WHP is SG-marked or

<sup>5</sup>Gentile & Schwarz (2018) make a similar observation with *how many*-questions. *How many*-questions presuppose uniqueness: (i-a) cannot be felicitously responded by a multi-choice answer expressed by a conjunction of two cardinal numerals. Given that the predicate *solved this problem together* is stubbornly collective, Gentile & Schwarz conjecture that the Q-domain of this question does not include Boolean conjunctions over numerals. Further, Gentile & Schwarz observe that  $\diamond$ -modals can obviate violations of uniqueness in *how many*-questions: (i-b) admits multi-choice answers like “Two is ok and three is ok” and does not presuppose uniqueness. In analogy to (24), I add to (i-b) that the multi-choice answer cannot be expressed by an elided conjunction.

- (i) a. “How many students solved this problem together?” “#Two and three.”  
 b. “How many students are allowed to solve this problem together?” “Two or/#and three.”

NUM-modified? The following presents two analyses to these puzzles. One approach treats the ‘disjunction-only’ reading the very same reading as the regular HO-reading but gives a weaker semantics to singular and numeral-modified nouns. In this approach, the distributional difference between conjunctive and disjunctive answers comes from that atomicity and cardinality restrictors remove Boolean conjunctions but not disjunctions. The other approach assumes that the derivation of this reading involves reconstructing the WH-complement to the question nucleus and interpreting uniqueness locally. In this approach, conjunctive answers are unacceptable because applying conjunction directly over uniqueness yields a contradiction.

**Solution 1: A uniform approach** This approach treats the ‘disjunction-only’ reading the very same reading as the ‘conjunction-admitting’ reading.<sup>6</sup> To derive these two readings uniformly, all we need is to allow some of the Boolean disjunctions to be atomic and/or cardinal, just like entities. In the following definitions, the (a)-condition on minimal witness sets ensures the atomic/cardinal GQ to be a disjunction, an  $\exists$ -quantifier, or a Montagovian individual.<sup>7</sup>

- (25) A GQ  $\pi$  is atomic iff (a) the minimal witness sets of  $\pi$  are all singleton sets, and (b) every member in the smallest live-on set of  $\pi$  is atomic.
- (26) A GQ  $\pi$  has the cardinality  $n$  iff (a) the minimal witness sets of  $\pi$  are all singleton sets, and (b) every member in the smallest live-on set of  $\pi$  has the cardinality  $n$ .

With the above assumptions, I re-define the [SG] feature and the CARD predicate polymorphically as in (27) and (28). ‘MWS( $A, x$ )’ is read as ‘ $A$  is a minimal witness set of  $x$ ’.

$$(27) \quad \llbracket \text{[SG]} \rrbracket = \lambda P \lambda x. \begin{cases} P(x) \wedge \text{ATOM}(x) & \text{if } P \subseteq D_e \\ P(x) \wedge \forall A [\text{MWS}(A, x) \rightarrow |A| = 1] \wedge \forall y \in \text{SMLO}(x) [\text{ATOM}(y)] & \text{if } P \subseteq D_{\langle et, t \rangle} \end{cases}$$

$$(28) \quad \llbracket \text{[CARD]} \rrbracket = \lambda P \lambda n \lambda x. \begin{cases} P(x) \wedge |x| = n & \text{if } P \subseteq D_e \\ P(x) \wedge \forall A [\text{MWS}(A, x) \rightarrow |A| = 1] \wedge \forall y \in \text{SMLO}(x) [|y| = n] & \text{if } P \subseteq D_{\langle et, t \rangle} \end{cases}$$

With the new definitions, the H-shifter can be used in singular and NUM-modified nouns. With three students  $abc$ , *student* and *two students* are interpreted as in (a) for FO-readings and as in (b) for HO-readings. The NP structures are the same as in (19)/(20).

- (29) *student*:
- a. for FO:  $\llbracket [\phi_P \text{[SG]}(nP)] \rrbracket = \{a, b, c\}$
- b. for HO:  $\llbracket [\phi_P \text{[SG]}(\text{H}(nP))] \rrbracket = \{\sqcup A \mid A \subseteq \{x^\uparrow \mid x \in \{a, b, c\}\}$
- (30) *two students*: we have  $\llbracket [\phi_P \text{[PL]}(\text{NumP})] \rrbracket = \llbracket \text{[NumP]} \rrbracket$ , where ...
- a. for FO:  $\llbracket [\text{[NumP two CARD]}(nP)] \rrbracket = \{a \oplus b, b \oplus c, a \oplus c\}$
- b. for HO:  $\llbracket [\text{[NumP two CARD H]}(nP)] \rrbracket = \{\sqcup A \mid A \subseteq \{x^\uparrow \mid x \in \{a \oplus b, b \oplus c, a \oplus c\}\}$

In sum, in the uniform approach, the disjunction-only reading (for a SG-marked question) and the conjunction-admitting reading (for a number-neutral or PL-marked question) are derived uniformly as follows:

- (31) Which book(s) does John have to read?  
 $[\text{which-}^{\text{H}}\text{book}(s) \lambda \pi [\text{have-to } [\pi_{\langle et, t \rangle} \lambda x [\text{John read } x_e]]]]$

<sup>6</sup>I thank Manuel Križ for pointing out the direction of this analysis.

<sup>7</sup>Witness sets are defined in terms of the living-on property as follows: if a GQ  $\pi$  lives on a set  $B$ , then  $A$  is a witness set of  $\pi$  iff  $A \subseteq B$  and  $\pi(A)$ . For example, given a discourse domain including three students  $abc$ , the  $\forall$ -quantifier *every student* has a unique minimal witness set  $\{a, b, c\}$ , while the singular  $\exists$ -quantifier *some student* has three minimal witness sets  $\{a\}$ ,  $\{b\}$ , and  $\{c\}$ , each of which consists of one atomic student.

**Solution 2: A reconstruction approach** This approach makes use of syntactic reconstruction. As in (32a-b), the *wh*-complement *book* is reconstructed back to the nucleus. This reconstruction has two consequences. First, it leaves a semantically unmarked variable  $D$  as the restrictor of the WHP, which can be type-lifted by the H-shifter without causing a type-mismatch or a conflict w.r.t. atomicity. Thus, a HO-reading arises if the H-shifter is applied to the  $D$  variable and if *wh*-phrase binds a HO-trace, as in (32b). Second, uniqueness is evaluated at the scopal position where the reconstructed noun adjoins to. This prediction explains why a question in this reading rejects conjunctive answers: combining a conjunction with a predicate of uniqueness yields a contradiction, as seen in (33).

- (32) “Which book does John have to read?” “Book A or/#and Book B.”
- a. [which<sub>D</sub>  $\lambda x$  [have-to [ $x$  is the book John read]]] ( $\square \gg \iota$ )  
 $\llbracket Q_{FO} \rrbracket = \lambda x_e : x \in D. \square \lambda w [x = \iota y [\text{book}_w(y) \wedge \text{read}_w(y)]]$
- b. [which<sub>H<sub>D</sub></sub>  $\lambda \pi$  [have-to [ $\pi_{\langle et,t \rangle}$   $\lambda x. x_e$  is the book John read]]] ( $\square \gg \pi \gg \iota$ )  
 $\llbracket Q_{HO} \rrbracket = \lambda \pi_{\langle et,t \rangle} : \pi \in {}^H D. \square \lambda w [\pi(\lambda x_e. x = \iota y [\text{book}_w(y) \wedge \text{read}_w(j, y)])]$
- (33)  $\llbracket Q_{HO} \rrbracket (a^\uparrow \cap b^\uparrow) = \square \lambda w [a = \iota y [\text{book}_w(y) \wedge \text{read}_w(j, y)] \wedge [b = \iota y [\text{book}_w(y) \wedge \text{read}_w(j, y)]]$   
 (#It has to be the case that the book that John read is Book A and is Book B.)

This analysis also extends to  $\diamond$ -questions. In Xiang (2016: ch. 2), I argue that the mention-all (MA)-reading expecting a conjunctive answer and the reading expecting a disjunctive answer are derived with distinct LF structures, as shown in (34) [with simplification]. In particular, conjunctive MA arises if the HO-WH-trace  $\pi$  scopes over the  $\diamond$ -modal, making the conjunctive MA answer the strongest true answer. In contrast, disjunctive MA arises if the trace  $\pi$  (regardless of its scope) is associated with a free choice licensing operator DOU ( $\approx$  Mandarin FC-licensing particle *dou*, see details in Xiang (To appear)). Roughly, applying DOU to a disjunctive answer yields a free choice interpretation, making the disjunctive MA answer the strongest true answer.

- (34) “What can we use for this class?” “Book A or/and Book B.”
- a. Conjunctive MA ( $\pi \gg \diamond$ )  
 $\llbracket Q \rrbracket = \lambda \pi_{\langle et,t \rangle} [\pi \lambda x_e [\text{can} [\text{we use } x]]]$   $\llbracket Q \rrbracket = \lambda \pi_{\langle et,t \rangle} . \pi(\lambda x_e. \diamond f(x))$
- b. Disjunctive MA (DOU  $\gg \diamond \gg \pi$ )  
 $\llbracket Q \rrbracket = \lambda \pi_{\langle et,t \rangle} [\text{DOU} [\text{can } \pi \lambda x_e [\text{we use } x]]]$   $\llbracket Q \rrbracket = \lambda \pi_{\langle et,t \rangle} . \text{DOU} \diamond \pi(\lambda x_e. f(x))$

Next, consider the SG-marked  $\diamond$ -question in (35). Reconstructing the singular noun *book* and letting the trace  $\pi$  take scope above the  $\diamond$ -modal yield the following scopal pattern:  $\pi \gg \iota \gg \diamond$ . As shown in (35a), unless A and B are the same book, combining the derived Q-function with the Boolean conjunction  $a^\uparrow \cap b^\uparrow$  yields a contradiction. In contrast, as seen in (35b), if  $\pi$  takes a narrow scope relative to the  $\diamond$ -modal and is associated with DOU across the modal, the derived free choice inference is not contradictory and expresses the desired MA-answer.

- (35) “Which book can we use for this class?” “Book A or/#and Book B.”
- a. If  $\pi \gg \iota \gg \diamond$ :  
 $\llbracket Q \rrbracket = \lambda \pi_{\langle et,t \rangle} : \pi \in {}^H D. \lambda w [\pi(\lambda x_e. x = \iota y [\text{book}_w(y) \wedge \diamond_w f(y)])]$   
 $\llbracket Q \rrbracket (a^\uparrow \cap b^\uparrow) = \lambda w [a = \iota y [\text{book}_w(y) \wedge \diamond_w f(y)] \wedge [b = \iota y [\text{book}_w(y) \wedge \diamond_w f(y)]]$   
 (# $a$  is the unique book that we can use, and  $b$  is the unique book that we can use.)
- b. If DOU  $\gg \diamond \gg \pi \gg \iota$  (viz., FC  $\gg \diamond \gg \iota$ ):  
 $\llbracket Q \rrbracket = \lambda \pi_{\langle et,t \rangle} : \pi \in {}^H D. \text{DOU} \diamond [\lambda w. \pi(\lambda x_e. x = \iota y [\text{book}_w(y) \wedge f_w(y)])]$   
 $\llbracket Q \rrbracket (a^\uparrow \cup b^\uparrow) = \text{DOU} \diamond [\lambda w. (a^\uparrow \cup b^\uparrow)(\lambda x_e. x = \iota y [\text{book}_w(y) \wedge f_w(y)])]$   
 $= \diamond \lambda w [a = \iota y [\text{book}_w(y) \wedge f_w(y)] \cap \diamond \lambda w [b = \iota y [\text{book}_w(y) \wedge f_w(y)]]$   
 ( $a$  can be the unique book that we use, and  $b$  can be the unique book that we use.)

## 6 Conclusions

This paper investigates the higher-order (HO)-readings of *wh*-questions. First, drawing on evidence from questions with necessity modals or collective predicates, I showed that sometimes a *wh*-question can only be completely addressed by a GQ and must be interpreted with a HO-reading. Next, I argued that the GQs that can serve as complete answers of questions are subject to two constraints — positiveness and homogeneity. Incorporating these constraints into the meaning of a H-shifter, I proposed that HO-reading arises if the H-shifter converts the *wh*-restrictor into a set of HO-meanings and if the *wh*-phrase binds a HO-trace. Accordingly, HO-readings are unavailable if the application of the H-shifter is blocked, either by the atomicity constraint of [SG] in singular nouns, or by the cardinality constraint in NUM-modified nouns.

Further, I observed that SG-marked and NUM-modified questions admit only disjunctive answers. I provided two explanations to this distribution. In the uniform account, these questions admit disjunctions because disjunctions (but not conjunctions) may satisfy the atomicity/cardinality requirement. In the reconstruction account, the WH-complement is reconstructed, which gives rise to local uniqueness and yields contradictions for conjunctive answers.

## References

- [1] L. Alonso-Ovalle and V. Rouillard. Number inflection, spanish bare interrogatives, and higher-order quantification. In *Proceedings of NELS 49*, 2019.
- [2] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2):159–219, 1981.
- [3] D. Cresti. Extraction and reconstruction. *Natural Language Semantics*, 3(1):79–122, 1995.
- [4] P. D. Elliott, A. C. Nicolae, and U. Sauerland. Who and what do *who* and *what* range over cross-linguistically, 2018.
- [5] F. Gentile and B. Schwarz. A uniqueness puzzle: *how many*-questions and non-distributive predication. In *Proceedings of Sinn und Bedeutung 21*, 2018.
- [6] J. Groenendijk and M. Stokhof. Type-shifting rules and the semantics of interrogatives. In *Properties, types and meaning*, pages 21–68. Springer, 1989.
- [7] D. Harbour. Paucity, abundance, and the theory of number. *Language*, 90(1):185–229, 2014.
- [8] A. Hirsch and B. Schwarz. Singular which, mention-some, and variable scope uniqueness. Poster presented at Semantics and Linguistic Theory 29, 2019.
- [9] G. Link. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In C. S. Rainer Bäuerle and A. von Stechow, editors, *Meaning, use, and interpretation of language*, pages 302–323. De Gruyter, 1983.
- [10] M. Maldonado. Plural marking and d-linking in Spanish interrogatives, 2017.
- [11] U. Sauerland. A new semantics for number. In *Proceedings of SALT 13*, 2003.
- [12] G. Scontras. *The semantics of measurement*. PhD thesis, Harvard University, 2014.
- [13] B. Spector. Modalized questions and exhaustivity. In *Proceedings of SALT 17*, 2007.
- [14] B. Spector. An unnoticed reading for *wh*-questions: Elided answers and weak islands. *Linguistic Inquiry*, 39(4):677–686, 2008.
- [15] A. Szabolcsi. Background notions in lattice theory and generalized quantifiers. In A. Szabolcsi, editor, *Ways of Scope Taking*, pages 1–27. Springer Netherlands, Dordrecht, 1997.
- [16] Y. Xiang. *Interpreting questions with non-exhaustive answers*. PhD thesis, Harvard University Cambridge, Massachusetts, 2016.
- [17] Y. Xiang. Function alternations of the Mandarin particle *dou*: Distributor, free choice licenser, and ‘even’. *Journal of Semantics*, To appear.



# Bridging Distributivity and Free Choice: The Case of Mandarin *Dou*\*

Zhuoye Zhao

New York University, New York, USA  
zhuoye.zhao@nyu.edu

## Abstract

This paper presents a novel perspective approaching the semantics of mandarin particle *dou*. It leads to a unifying treatment of unconditionals and free choice constructions, which in turn motivates the meta-characterization of distributivity, free choice and unconditionals as all conveying **orthogonality**.

## 1 The Puzzle of *Dou*

The semantic contribution of the mandarin particle *dou* has attracted a substantial amount of attention in recent literatures (Liu, 2017; Yimei Xiang, 2019, a.o.), in particular because of its ability to co-occur with different types of expressions and activate various semantic effects. This paper focuses on two of them: the **distributive** use and the **free choice** use.

### 1.1 Distributivity

In its most common use in basic declarative sentences, *dou* is associated with a preceding noun phrase (enclosed in ‘[.]’) and distributes over its subparts with the remnant predicate, as shown in (1).

- (1) a. [Tamen] *dou* du -le san-ben shu.  
they **dou** read -ASP three-CL book.  
‘They *all* read three books.’
- b. *Context: On Sunday, Alice, Bob and Charlie rented a boat together and wandered around the canals in Amsterdam.*  
[Tamen] (#*dou*) zu -le yi-sou chuan.  
They (#**dou**) rent -ASP one-CL boat.  
‘They (#all) rented a boat.’

(1a) demonstrates the *distributivity effect* activated by *dou* with the absence of the cumulative/collective reading ‘*They read three books between them*’, while (1b) showcases a *plurality requirement* of *dou*: it requires at least two proper subparts of (the denotation of) its plural associate to satisfy the subsequent predicate. Since the plurality sum of Bill, Bob and Barbara is the only entity that can be truthfully applied to the predicate *rented a boat*, the plurality

---

\*Thanks to Maria Aloni and Alexandre Cremers for supervising the project, and to Floris Roelofsen, Robert van Rooij, Anna Szabolcsi, Lucas Champollion, Chris Barker, Magdalena Kaufmann, Rahul Balusu and the NYU semantics group for valuable comments and discussions.



requirement is not satisfied, and *dou* is not licensed<sup>1</sup>. Note that the plurality requirement effectively blocks the collective reading (and induces distributive reading) in (1a). This is not entirely surprising due to the close correlation observed between distributivity and plurality (Landman, 1989, 2012). Moreover, Lin (1998) noticed that the distributivity effect of *dou* is ‘generalized’ in that it does not necessarily distribute down to atomic subparts. (2a) shows that the distribution can be applied to a sub-atomic level, indicating that ‘*John ate every part of that apple*’; while (2b) shows that plural/non-atomic subparts are also acceptable, since the predicate *be friends* applies only to plural entities. Therefore, Lin (1998) proposed to treat *dou* as an overt counterpart of the generalized distributor (Schwarzschild, 1996). In §2.1, a detailed analysis of this use of *dou* building on Lin’s proposal will be laid out.

- (2) a. Yuehan ba [na-ge pingguo] dou chi-le.  
 John BA that-CL apple **dou** eat-ASP.  
 ‘John ate that apple.’ (Xiang, 2019)
- b. *Context: Alice, Bill and Charlie are mutual friends.*  
 [Tamen] dou shi pengyou.  
 They **dou** BE friends.  
 ‘They are *all* friends.’

## 1.2 Free Choice

As mentioned above, *dou* can be associated with pre-verbal *wh*-phrases (3b) or disjunctions etc. (3a) and gives rise to universal free choice ( $\forall$ -FC) readings. The basic data is given in (3):

- (3) a. Wulun Yuehan haishi Mali \*(dou) keyi canjia.  
 no-matter John or Mary **dou** may participate.  
 ‘Both John and Mary may participate.’
- b. (Wulun) [shenme] shuiguo Yuehan \*(dou) keyi chi.  
 (no-matter) what fruit John **dou** may eat.  
 ‘John may eat any fruit.’

Note that in these cases, the alternative-generating expressions associated with *dou* are optionally preceded by *wulun* (no-matter). These constructions share a very similar structure with mandarin *unconditionals*, with *wulun* as the unconditional head, and *dou* in the consequent:

- (4) a. (Wulun) paidui shi zai bier jia haishi baobo jia, Yuehan \*(dou) keyi qu.  
 No-matter party BE at Bill house or Bob house, John \*(**dou**) may go.  
 ‘Whether the party is at Bill’s house or Bob’s house, John may go.’
- b. (Wulun) paidui zai shui jia, Yuehan \*(dou) keyi qu.  
 No-matter party at who house, John \*(**dou**) may go.  
 ‘No matter whose house the party will be at, John may go.’

This observation motivates an *unconditional* analysis (Rawlins, 2013) of the  $\forall$ -FC reading, resonating with a recent trend (e.g. Szabolcsi, 2019). Moreover, as we will see in §2.3, such analysis calls for a schema that resembles the one used for capturing the distributivity effect (more in §2.2), which not only justifies the uniformity of the current proposal, but also provides breeding ground for the insight towards the connection between distributivity and free choice.

<sup>1</sup>The term *plurality requirement* is borrowed from Xiang (2019), but with a different sense. Xiang interpreted it literally as a requirement of plurality/non-atomicity over the associates of *dou* and argued that such requirement is an illusion (as shown in (2)). I agree with this judgment, but instead interpret the term as forcing the plurality on the set of subparts of the associates, and take it as a robust requirement here and henceforth.

## 2 From Distributivity to Free Choice

### 2.1 *Dou* as a generalized distributor

In §1.1 it is established that *dou* evokes a *generalized* distributivity effect. Lin (1998) captures such effect by equating *dou* with the generalized distributor (Schwarzschild, 1996). That is, *dou* distributes over a contextually determined plurality *cover* of the associated referent, and induce a universal reading w.r.t. the subsequent predicate. The definition of a plurality *cover* and a preliminary entry for the distributive *dou* is given in (5) and (6), resp.

- (5) a.  $\mathcal{C}$  is a plurality COVER of  $x$ , written as  $Cov(x, \mathcal{C})$ , iff  $\mathcal{C}$  covers  $x$  and no proper subset of  $\mathcal{C}$  covers  $x$ .  
 b.  $\mathcal{C}$  covers  $x$  iff (i)  $\mathcal{C}$  is a set of subparts of  $x$ , (ii) Every subpart of  $x$  belongs to some element of  $\mathcal{C}$ , and (iii)  $\emptyset \notin \mathcal{C}$ .

- (6) Semantics of *Dou* (Lin, 1998, to be revised):  

$$\llbracket dou \rrbracket = \lambda P_{\langle e, st \rangle} \lambda x_e \lambda w_s. \underbrace{\forall y \in \mathcal{C}. P(y)(w)}_{\text{distributivity effect}}, \text{ where } Cov(x, \mathcal{C})$$

Intuitively, then, a *cover* of a (plural) entity is a *minimal* set of its subparts whose plurality sum equals the entity itself. Since a cover is allowed to contain plural elements (and sub-atomic elements as well), the ‘generality’ of the distributivity effect is accounted for. Take (2b) as an example. Assume  $x = a \oplus b \oplus c$  denotes the plurality sum of the three individuals Alice, Bob and Charlie. Take  $\mathcal{C} = \{a \oplus b, a \oplus c, b \oplus c\}$ , then clearly  $\mathcal{C}$  covers  $x$ , and no proper subset of it does so, thus  $Cov(x, \mathcal{C})$ . Finally, applying a universal quantification over the elements in  $\mathcal{C}$  w.r.t. the predicate *be friends* yields the desired reading: Alice and Bob are friends, Bob and Charlie are friends, and Alice and Charlie are friends.

However, as noted by Xiang (2019) and others, the entry (6) leaves the *plurality requirement* unaccounted for. The missing part is a mechanism to rule out singleton covers. Consider the infelicitous sentence (1b). Again,  $x = a \oplus b \oplus c$  denotes the corresponding plurality sum. Now if  $\mathcal{C}$  is the singleton set  $\{a \oplus b \oplus c\}$ , the definition (5) still admits  $\mathcal{C}$  as a cover of  $x$ ; and applying this to the predicate *rented a boat*, we incorrectly predict the sentence to be true. To fix this, we insert the plurality requirement as a presupposition into the revised entry (7) for *dou*:

- (7) Semantics of *Dou* (revised)  

$$\llbracket dou \rrbracket = \lambda P_{\langle e, st \rangle} \lambda x_e \lambda w_s. \underbrace{\exists \mathcal{C}. Cov(x, \mathcal{C}) \wedge |\mathcal{C}| > 1}_{\text{plurality requirement}} \cdot \underbrace{\forall y \in \mathcal{C}. P(y)(w)}_{\text{distributivity effect}}$$

The new entry would rule out (1b) as a presupposition failure, since there is no non-singleton cover of  $x$  whose elements all satisfy the predicate ‘*rented a boat*’ in the given context. In the following, I will attempt to adapt this schema, i.e. the combination of a plurality requirement and a distributivity effect, in order to capture the semantics of *dou* when used in unconditionals and free choice constructions.

### 2.2 Lifting Conditionals

As discussed in §1.2, a detour through unconditionals would benefit our understanding of Mandarin free choice constructions and the role that *dou* plays in both environments. In Rawlins (2008, 2013), unconditional readings are derived via a universal closure  $[\forall]$  of a point-wisely composed Hamblin set of conditional sentences. For instance, given the sentence (4a), the

antecedent question “*whether the party is at Bill’s house or Bob’s house*” generates a Hamblin set of possible answers  $\{The\ party\ is\ at\ Bill’s\ house, The\ party\ is\ at\ Bob’s\ house\}$ , which are point-wisely composed with the consequent sentence, producing a set of conditionals  $\{The\ party\ is\ at\ Bill’s\ house \Rightarrow John\ may\ go, The\ party\ is\ at\ Bob’s\ house \Rightarrow John\ may\ go\}$ . ‘ $\Rightarrow$ ’ is the conditional operator to which one can feed her favorite conditional semantics<sup>2</sup>. Finally, a universal operator closes off the Hamblin set of conditionals, yielding the desired reading “*The party is at Bill’s house  $\Rightarrow$  John may go & The party is at Bob’s house  $\Rightarrow$  John may go*”. This account, although informally introduced here, showcases striking parallels with the schema in (7): the Hamblin set generated by the antecedent question corresponds to the contextual cover, the consequent corresponds to the ‘predicate’ (applied to each possible answer via conditional composition), and the distributivity effect induced by *dou* provides the universal closure. Moreover, the *plurality requirement* can be viewed as a requirement of the contextual *inquisitiveness* on the antecedent question. I claim that these parallels indeed characterize the semantics of *dou* in unconditionals, and I will flesh it out formally using the notion of *Lifted Conditionals* (Ciardelli, 2016) based on Inquisitive Semantics.

Inquisitive Semantics unifies the semantic denotations of declarative and interrogative sentences using the notion of *issues*. An issue is a non-empty, *downward-closed* set of *information states*, where an information state stands for a set of possible worlds. Therefore, the propositional semantics undergoes a set-theoretic *lift* from sets of possible worlds to sets of sets of possible worlds. An important notion that is used to distinguish between declarative and interrogative sentences – now that they have the same semantic type – is *alternatives*, defined as the maximal elements in an issue.

- (8) An information state  $\alpha$  is an **alternative** of an issue  $I$  iff (i)  $\alpha \in I$ , and (ii)  $\neg \exists \alpha'$  s.t.  $\alpha' \in I$  and  $\alpha \subsetneq \alpha'$ . The set of alternatives of an issue  $P$  is written as  $\text{alt}(I)$ .

As a result, we can distinguish between *inquisitive* issues with  $|\text{alt}(I)| > 1$  and *non-inquisitive* issues where  $|\text{alt}(I)| = 1$ . Finally, we will make use of the function *info* that retrieves the *informative content* of an issue as the *union* of its set of alternatives:

- (9)  $\text{info}(I) = \bigcup \{\alpha \mid \alpha \in \text{alt}(I)\}$

A lifted conditional operator ‘ $>$ ’ can be obtained by *lifting* the semantics of the conditional operator ‘ $\Rightarrow$ ’ into its inquisitive counterpart. Here  $\Rightarrow$  is defined simply as material implication (10), with the lifting result given in (11):

- (10) Given the set of possible worlds  $W$  and two information states  $s, t \subseteq W$ ,  
 $s \Rightarrow t := \{w \in W \mid w \in W \setminus s \text{ or } w \in Q\}$

- (11) Given two inquisitive issues  $I$  and  $I'$ ,  
 $I > I' := \{s \mid \forall \alpha \in \text{alt}(I) : \exists \beta \in \text{alt}(I'). s \subseteq \alpha \Rightarrow \beta\}$

Therefore, the issue  $I > I'$  contains the set of information states that support the classical material implications between *any* alternative  $\alpha$  of  $I$  and *some* alternative  $\beta$  of  $I'$ . For unconditionals, with inquisitive antecedents and non-inquisitive consequents, (11) can be further simplified<sup>3</sup>:

- (12) Given an inquisitive issue  $I$  and a non-inquisitive issue  $I'$ ,  
 $I > I' := \{s \mid \forall \alpha \in \text{alt}(I) : s \subseteq \alpha \Rightarrow \text{info}(I')\}$

<sup>2</sup>Rawlins followed the modal restriction tradition of Lewis and Keenan (1975); Kratzer (1981), a.o.

<sup>3</sup>Besides unconditionals, the lifted conditionals also captures standard *if*-conditionals (and the simplification of disjunctive antecedent), as well as conditional questions. For more details, see Ciardelli (2016).

This corresponds directly to the informal characterization of unconditional semantics given above, and the adaptation of (7) into the semantics of *dou* in unconditionals (written as  $dou_Q$  to avoid confusion) naturally follows:

$$(13) \quad \llbracket dou_Q \rrbracket = \lambda P_T \lambda Q_T \lambda s_{st}. \underbrace{|\text{alt}(Q)| > 1}_{\text{plurality}} \cdot \underbrace{\forall \alpha \in \text{alt}(Q_c) : s \subseteq [\alpha \Rightarrow \text{info}(P)]}_{\text{distributivity effect}}$$

As we can see, the schema of *plurality + distributivity* is preserved, the only difference being the types of arguments that *dou* takes: in (7) *dou* takes a predicate  $\langle\langle e, t \rangle\rangle$  and distributes it over the associated entity  $(e)$ , whereas in (13)  $dou_Q$  takes two issues  $(T = \langle s, \langle s, t \rangle \rangle)$  and connects them with the lifted conditional.

In the rest of the section, I will forge the  $\forall$ -FC constructions as in (3) into the same logical skeleton (12), so that (13) can be applied to derive the free choice effect.

### 2.3 Unconditional ▶ Free Choice

To apply the unconditional analysis to a  $\forall$ -FC structure, the first task is to reconstruct the ‘antecedent’ question and the ‘consequent’. To retrieve the antecedent question, we treat the *wh*-phrase as an *identity question* w.r.t. a type  $e$  variable  $u$  (written as ‘?u’), which will subsequently fill in a vacuous argument position in the following VP predicate and retrieve the consequent. As a result, a Mandarin  $\forall$ -FC sentence is reconstructed into an unconditional with a binding relation between the antecedent and the consequent. As an empirical support for such treatment, we observe that the Mandarin copula ‘*shi*’, which is used to impose a (real) identity question (14a), can be inserted between ‘*wln*’ and the *wh*-phrase as in (14b):

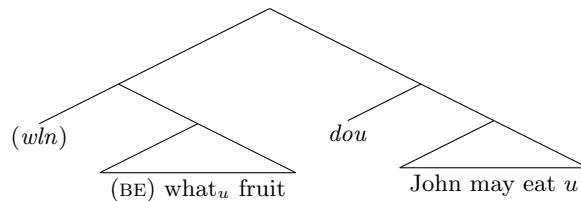
- (14) a. Ta shi shui?  
           He BE who  
           ‘Who is he?’  
       b. Wulu shi shui dou keneng yu-dao mafan.  
           No-matter BE who **dou** may run-into trouble  
           ‘Anyone can run into trouble.’

The derivation of  $\forall$ -FC effects is now in place. First, ?u is formally characterized as follows:

$$(15) \quad ?u := \forall x. ?(u = x), \text{ where } ?(u = x) := (u = x) \vee (u \neq x).$$

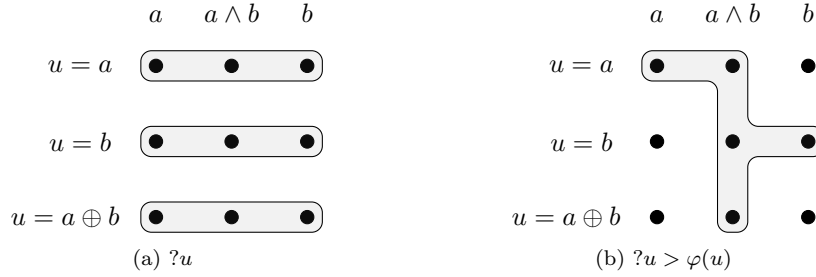
Let’s assume that the identity of  $u$  is a piece of information that helps pin down the actual world<sup>4</sup>. Now take the sentence (3b) (repeated below in (16)) as a working example, the unconditional analysis proceeds as follows:

- (16) (Wulun) [shenme] shuiguo Yuehan \*(dou) keyi chi.  
       (no-matter) what fruit John **dou** may eat.  
       ‘John may eat any fruit.’  
       a. Assuming the following toy structure:



<sup>4</sup>In a dynamic setting, this job can be transferred to the stack of discourse referents, but we won’t have space to discuss the detailed implementation

- b.  $\llbracket (\text{BE}) \text{ what}_u \text{ fruit} \rrbracket = ?u \wedge \llbracket \text{fruit} \rrbracket(u)$   
 Suppose the domain of individuals contains two atomic members  $a, b$  that are fruits, and their plurality sum  $a \oplus b$ , then the denotation of  $?u \wedge \llbracket \text{fruit} \rrbracket(u)$  can be visualized as in Fig. (1a).
- c.  $\varphi(u) := \llbracket \text{John may eat } u \rrbracket = \lambda s_{st}. \forall w \in s : \exists w' \in \text{MB}_d(w). \text{John eat } u \text{ at } w'$
- d. Apply the semantics of  $\text{dou}_Q$  (13), we get the following semantic representation of (3b):  
 $\llbracket (3b) \rrbracket = \llbracket \text{dou}_Q \rrbracket(?u)(\varphi(u)) = ?u > \varphi(u)$ , defined only if  $|\text{alt}(?u)| > 1$ .  
 Let  $a, b, a \wedge b$  represent the world information that ‘John may eat  $a$ ’, ‘John may eat  $b$ ’ and ‘John may eat  $a$  and  $b$ ’, the result is shown in Fig. (1b). The  $\forall$ -FC reading is successfully captured, since given any identity of  $u$ , the world in which John may eat  $u$  is always included.

Figure 1: Derivation of the  $\forall$ -FC reading

### 3 Orthogonality

For an unconditional such as (4a), applying a semantic analysis based on (12) gives rise to the same result as simply asserting the consequent ‘*John may go to the party*’. One may ask, then, what else is expressed when an unconditional is uttered? A key observation made by Rawlins (2013) is that unconditionals convey **orthogonality** between the antecedent issue and the consequent, which means, informally, that asserting the consequent doesn’t help (even partially) resolve the antecedent issue. Rawlins further noted that the notion of orthogonality ‘provides a powerful and useful unifying meta-characterization of many free choice effects’. The vision will be formally realized in this section. Meanwhile, I argue that the orthogonality effect may be carried to mandarin distributive structures by *dou* as well.

#### 3.1 Orthogonality in Free Choice

The notion of orthogonality can receive a formal characterization from inquisitive semantics, which basically says that for two orthogonal issues  $I_1, I_2$ , any alternative of  $I_1$  *cuts across* all the alternatives of  $I_2$ , and *vice versa*.

- (17) Two issues  $I_1, I_2$  are *orthogonal* if and only if:  
 For all  $p_1 \in \text{alt}(I_1)$ , there is no  $p_2 \in \text{alt}(I_2)$  s.t.  $p_1 \subseteq p_2$ , and for all  $p_2 \in \text{alt}(I_2)$ ,  $p_1 \cap p_2 \neq \emptyset$ ;  
 For all  $p_2 \in \text{alt}(I_2)$ , there is no  $p_1 \in \text{alt}(I_1)$  s.t.  $p_2 \subseteq p_1$ , and for all  $p_1 \in \text{alt}(I_2)$ ,  $p_2 \cap p_1 \neq \emptyset$ .

The orthogonality conveyed by (4a) is then illustrated in Fig. (2), where the antecedent question raises two alternatives “*the party is at Bill’s house*” and “*the party is at Bob’s house*” (enclosed in the gray areas in Fig. (2a)), and the issue denoted by the consequent ‘*John may go*’ (the rectangle in Fig. (2b)) cuts across both of them.

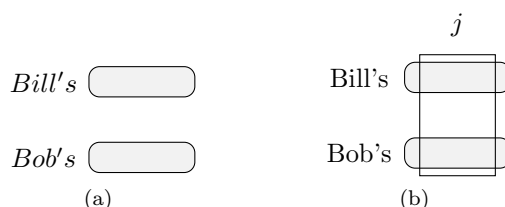


Figure 2: Orthogonality conveyed by (4a)

However, if we adhere to the idea that the orthogonality holds between the antecedent and the *consequent*, a problem occurs as we try to pinpoint the ones conveyed by  $\forall$ -FC constructions. Recall that in §2.3, the ‘antecedent’ of a  $\forall$ -FC construction is reconstructed as an identification question  $?u$ , while  $u$  is co-indexed in the ‘consequent’ and applied to the remnant predicate. Therefore, the semantic denotation of the ‘consequent’ actually *depends* on the resolution of the antecedent. This clearly goes against our intuition about orthogonality. Moreover, this problem extends beyond  $\forall$ -FC constructions to any unconditional with a binding relation between the antecedent and consequent, as we can see in (18): whenever a (complete) answer of the antecedent question ‘*who<sub>1</sub> comes?*’ is given, the truth condition of the consequent ‘*John will invite him/her<sub>1</sub>*’ is fixed.

(18) No matter who<sub>1</sub> comes, John will invite him/her<sub>1</sub> for dinner.

One might suggest that the orthogonality actually holds between the antecedent question and the *Question Under Discussion* (QUD) pertaining to the consequent, thus for (18) it is between the two questions ‘*who comes?*’ and ‘*who will John invite for dinner?*’. But it doesn’t solve the whole problem. In both Rawlins’ characterization and the definition (17), the orthogonality between two issues requires that the resolution of either one does not *partially* resolve the other. Now imagine a scenario where it is common knowledge that John only invites people for dinner if they come. (18) can still be felicitously uttered, but now a resolution to the QUD of the consequent, say, John will invite *Bill* for dinner, would provide at least a partial answer to the antecedent question, namely, *Bill* comes. The orthogonality is again sabotaged. Therefore, instead of resorting to QUD, I propose to relocate the orthogonality in between the antecedent and the *propositional content* of the unconditional. This modification pinpoints the orthogonality conveyed by  $\forall$ -FC constructions, while preserving the definition (17): recall that Fig. (1a) depicts the antecedent identity question raised in (3b), while Fig. (1b) illustrates its propositional content—and clearly the latter ‘cuts across’ the former. Moreover, as has already been noted in the beginning of the section that the semantic denotations of unconditionals *without* binding (thus unlike 18) is equivalent to its consequent, the new proposal still derives the orthogonality for cases like (4).

Finally, it is worth noting that at least one of the two issues standing in an orthogonal relation has to be *inquisitive*—otherwise, (17) would trivially characterize the logical independence between two non-inquisitive propositions. In unconditionals, such obligatory inquisitiveness is imposed on the antecedent, and as we can see from (13), corresponds directly to the plurality

requirement of  $dou_Q$ . This observation is crucial in revealing the connection between distributivity and free choice/unconditionals—as I will argue in §3.2, the plurality requirement of the distributive use of  $dou$  can also be construed as certain (sub-level) inquisitiveness, or *variation*.

### 3.2 Orthogonality in Distributivity?

We will base our discussion in a basic setting of Plural Predicate Logic (PPL, Champollion et al., 2017), which was inspired by Dynamic Plural Logic (van den Berg et al., 1996). PPL differs from Predicate Logic in that its formulae are evaluated relative to *sets of assignments*, which I refer to as *assignment matrices*, instead of single assignments. As a result, PPL provides formal mechanisms capable of encode fine-grained *plurality* information such as quantificational dependency (see, e.g. Brasoveanu, 2008; Henderson, 2014). Illustrations of an assignment matrix  $G$  consisting of three assignment functions  $g_1, g_2, g_3$  assigning values to a variable  $u$  are shown in Fig. (3).

$G$	...	$u$	...	$G$	...	$u$	...	$G$	...	$u$	...
$g_1$	...	$a \oplus b$	...	$g_1$	...	$a \oplus b$	...	$g_1$	...	$a$	...
$g_2$	...	$a \oplus b$	...	$g_2$	...	$b \oplus c$	...	$g_2$	...	$b$	...
$g_3$	...	$a \oplus b$	...	$g_3$	...	$a \oplus c$	...	$g_3$	...	$c$	...

(a) assignment-level plural, no variation, \**dou*

(b) assignment-level plural, with variation, ✓*dou*

(c) assignment-level singular, variation, ✓*dou*

Figure 3: Plurality requirement of the distributor use of  $dou$

When translating natural language utterances to PPL formulae, (plural) DP such as ‘*they*’ can pick up the variable  $u$  and denote the plurality sum of the values assigned by the functions in the assignment matrix, thus  $\llbracket they_u \rrbracket^G = a \oplus b \oplus c$ . However, the evaluation against a distributive predicate, such as ‘*come*’, still goes into the assignment level and checks if the individual assigned by *each* function satisfies the predicate. Therefore,  $\llbracket they_u come \rrbracket^G = \forall g \in G : Come(g(u))$ , i.e.  $a, b$  and  $c$  *each* comes. Finally, given the distributivity effect of  $dou$ , we assume that it forces the assignment-level evaluation to any predicate it modifies, even for a collective predicate such as ‘*be friends*’ in (2b)<sup>5</sup>.

Now back to plurality requirement. In §2.1, it is described as presupposing non-singleton covers, even though  $dou$  generally allows non-atomic elements in them. Transferring the description into a PPL setting, plurality boils down to a requirement of **variations** among different assignment functions within a matrix, as demonstrated by the contrast between Fig. (3a) and Fig. (3b). When applied to the sentence ‘*they\_u dou be friends*’, the matrix in Fig. (3a) suffers a failure even if  $a$  and  $b$  indeed are friends, whereas the one Fig. (3b) successfully yields the interpretation that ‘*a, b, c are mutual friends*’.

The requirement of *variation* pertaining to distributivity effect is not unfamiliar. It is brought out, for instance, by numeral reduplications in Telugu (Balusu, 2006), verbal pluractionals in Kaqchikel (Henderson, 2014) and dependent indefinites (Farkas, 1997). However, in

<sup>5</sup>However, distributive and collective predicates do not correspond one-to-one to assignment-level and matrix-level evaluations, since there are different types of collective predicates (Winter, 2002). For instance, *be friends* is a typical assignment-level collective predicates, as opposed to matrix-level collective predicates such as ‘*gather*’. In fact,  $dou$  can modify both types, but the limited space would not allow further discussions. For more details, I refer to Dotlačil and Roelofsen (2019) and Zhao (2019) ch.3,4.



the light of §3.1, we can further imagine an orthogonality-like characterization of distributivity effect—the orthogonality lies between the *identity* of each individual in the quantificational domain (note the similar illustration between identity questions in Fig. (1a) and Fig. (3)), and the fact that they *all* satisfy the remnant predicate. In fact, it is possible that the Cantonese distributivity marker ‘*saai*’ overtly realizes such effect, by imposing an additional *independent* interpretation over distributivity (Law, 2019).

## 4 Outlook

The endeavor towards a uniform analysis of the distributive and free choice uses of *dou* is not new (in particular, see Xiang, 2019). However, the paper provides a new perspective that reveals an underlying feature shared between them, namely to convey **orthogonality**, and argues the **plurality requirement** of *dou* guarantees its non-triviality.

I will end this paper with a few more puzzles in the family. First, an anonymous reviewer expressed worries on the **licensing** problems of  $\forall$ -FC containing *dou*<sub>Q</sub>. In §1.2, only two kinds of associates were introduced: the disjunction ‘*haishi*’ that is commonly used in alternative questions in (3a), and *wh*-phrases in (3b). However, *dou*<sub>Q</sub> is also able to combine with the polarity item ‘*renhe*’ (any) and declarative disjunctions with ‘*huozhe*’ (or), as shown in (19) and (20). However, compared to (3), they are much worse in episodic context. The analysis given in Xiang (2019) captures the infelicity in episodic context, but also predicts infelicity under necessity modals, as opposed to what empirical data shows (19b). A solution to this problem has to be left for future work, but to explore the interaction between orthogonality/plurality and modal context seems to be a promising start.

- (19) a. [Ren-he ren] dou keyi lai.                      b. [Ren-he ren] dou bixu canjia.  
       Any person **dou** may come.                      Any person **dou** must participate.  
       ‘Anyone may come.’                                      ‘Anyone must participate.’
- (20) [Yuehan huozhe Mali] dou keyi jiao jichu hanyu.  
       John or Mary **dou** may teach intro Chinese.  
       ‘Both John and Mary may teach intro Chinese.’ (Xiang, 2019)

Second, in this paper I left out yet another major use of *dou*, in which it associates with focused items and creates *even*-like readings (21).

- (21) (Lian) [YUEHAN]<sub>F</sub> dou lai-le.  
       (LIAN) JOHN **dou** come-LE.  
       ‘Even John came.’

An analysis using the same schema can be found in Zhao (2019). But a yet more interesting question to ask is whether the semantics of these constructions can also be characterized using orthogonality. I leave this also to future work.

Finally, though I adhered to Lin’s schema of treating *dou* as a distributor, it remains to be seen whether it is indeed one. An empirical argument against it is that *dou* co-occurs with real distributive quantifiers such as ‘*meige*’ (every). Zhao (2019) proposed an analysis that treats *dou* as imposing *only* the plurality requirement, and instead of a presupposition, I treated it as a *postsupposition* (cf. Brasoveanu, 2008; Henderson, 2014; Champollion, 2015). This alternative approach accommodates the co-occurrence of *dou* with other distributive quantifiers, while capturing most of the other data. However, more empirical investigation needs to be done to finally settle this issue.



## References

- Balusu, R. (2006). Distributive reduplication in telegu. In *PROCEEDINGS-NELS*, volume 36, page 39.
- Brasoveanu, A. (2008). Donkey pluralities: plural information states versus non-atomic individuals. *Linguistics and philosophy*, 31(2):129–209.
- Champollion, L. (2015). Every boy bought two sausages each: Distributivity and dependent numerals. In *Proceedings of the 32nd West Coast Conference on Formal Linguistics (WCCFL 32)*, pages 103–110. Cascadilla Proceedings Project Somerville, MA.
- Champollion, L., Bledin, J., and Li, H. (2017). Rigid and flexible quantification in plural predicate logic. In *Semantics and Linguistic Theory*, volume 27, pages 418–437.
- Ciardelli, I. (2016). Lifting conditionals to inquisitive semantics. In *Semantics and Linguistic Theory*, volume 26, pages 732–752.
- Dotlačil, J. and Roelofsen, F. (2019). Dynamic inquisitive semantics. In *Slides for the talk given at the 12th Paris-Amsterdam Logic Meetings of Young Researchers*.
- Farkas, D. F. (1997). Dependent indefinites. In *Empirical issues in formal syntax and semantics*. Citeseer.
- Henderson, R. (2014). Dependent indefinites and their post-suppositions. *Semantics and Pragmatics*, 7:6–1.
- Kratzer, A. (1981). The notional category of modality. words, worlds, and contexts: New approaches in word semantics, ed. by hj eikmeyer and h. reiser, 38–74.
- Landman, F. (1989). Groups, i. *linguistics and philosophy*, pages 559–605.
- Landman, F. (2012). *Events and plurality: The Jerusalem lectures*, volume 76. Springer Science & Business Media.
- Law, J. H.-k. (2019). Independence in distributivity. In *talk given at Semantics and Linguistics Theory (SALT) 29*.
- Lewis, D. and Keenan, E. (1975). Adverbs of quantification. *Formal semantics of natural language*, pages 178–188.
- Lin, J.-W. (1998). Distributivity in chinese and its implications. *Natural language semantics*, 6(2):201–243.
- Liu, M. (2017). Varieties of alternatives: Mandarin focus particles. *Linguistics and Philosophy*, 40(1):61–95.
- Rawlins, K. (2008). *(Un) conditionals: An investigation in the syntax and semantics of conditional structures*. University of California, Santa Cruz.
- Rawlins, K. (2013). (un) conditionals. *Natural language semantics*, 21(2):111–178.
- Schwarzschild, R. (1996). *Pluralities*, volume 61. Springer Science & Business Media.
- Szabolcsi, A. (2019). Unconditionals and free choice unified. In *Semantics and Linguistic Theory*, volume 30.
- van den Berg, M. H. et al. (1996). Some aspects of the internal structure of discourse. the dynamics of nominal anaphora.
- Winter, Y. (2002). *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*, volume 37. MIT press.
- Xiang, Y. (2019). Function alternations of the mandarin particle dou: Distributor, free choice licenser, and even.
- Zhao, Z. (2019). Varieties of distributivity: From mandarin dou to plurality, free choice and scalarity.

# On domain adjectives and the metaphors they modify

Joost Zwarts<sup>1</sup>

Utrecht University  
j.zwarts@uu.nl

## Abstract

The metaphorical meaning of a noun like *storm* is modeled as a function from domains to kinds of storms. The argument of this function can be set by a domain adjective (e.g., *a political storm*), but also adverbially or contextually. By making explicit how domain specification works, this proposal not only brings conceptual domains and metaphors into the heart of formal semantics, but it also opens avenues towards an account of non-metaphorical domain modifiers more generally (e.g., *botanical fruit*).

## 1 Introduction

The adjectives in (1) are examples of *domain adjectives*: they can specify the domain in which the noun they modify is interpreted (Ernst, 2004). An *academic crusade*, for instance, is a crusade in the domain of academia.

- (1) academic, budgetary, conceptual, culinary, economic, emotional, financial, political, psychological, spiritual

It has been observed that these adjectives can shift the meaning of their noun to a metaphorical meaning (Sullivan, 2013; Reijnerse, Burgers, Krennmayr, & Steen, 2018).

- (2) spiritual wealth, emotional scar, financial octopus, psychological jungle, political storm, academic crusade, economic midget, conceptual space, budgetary anorexia, culinary mecca

In the examples in (2), the noun gives the *source* of the metaphor (e.g., the many-tentacled octopus in *financial octopus*) that is mapped to a *target* from the adjectival domain (e.g., a large and diversified financial institution). This is a fully productive process (Hanks, 2004). We can have octopuses in as many domains as we have domain adjectives (3a) and the domain adjective *spiritual* can specify the target domain of an open-ended range of metaphorically used nouns (3b):

- (3) a. bureaucratic / cultural / electronic / financial / political / ... octopus  
b. spiritual food / growth / path / pollution / vacuum / warfare / ...

As section 2 will show, domain adjectives share properties with the more intensively studied *relational* adjectives, like in *solar heat*, *technical architect*, *musical instrument* (McNally & Boleda, 2004), but there are also important differences. A domain adjective can not simply be treated as an instance of a relational adjective, but a dedicated account is needed to explain how the meaning of the noun can shift from one domain to another depending on the adjective that modifies it, but also under the influence of the context or of a sentential adverb. This account has to address questions about how word meanings interface with

---

<sup>1</sup>I gratefully acknowledge (with the usual disclaimers) the comments given by the anonymous reviewers of the abstract for the Amsterdam Colloquium 2019.

compositional meaning (section 3). The proposal of this paper (worked out in section 4) is that the metaphorical meaning of a noun like *storm* is modeled as a function from domains to kinds of storms. The argument of this function can be set locally, by a domain adjective (e.g., a *political storm*), but also adverbially or contextually. Section 5 concludes the paper with a brief exploration of the wider perspectives. By making explicit how domain specification works, this proposal not only brings conceptual domains and metaphors into the compositional heart of formal semantics, but it also opens avenues towards an account of non-metaphorical domain modifiers more generally (e.g., *botanical fruit*).

## 2 Domain adjectives

In many respects, domain adjectives are similar to relational adjectives (McNally & Boleda, 2004). Like relational adjectives, domain adjectives do not allow for a predicative use (4). Without the noun, the adjectives sound odd for the same subject or they have a different meaning.

- (4) a. Chris is a technical architect. (relational adjective)  
 a'. ?Chris is technical.  
 b. Citicorp is a financial octopus. (domain adjective)  
 b'. ?Citicorp is financial.

Domain adjectives are also similar to relational adjectives in not allowing (degree) modification (5). Even though the adjectives *technical* or *bureaucratic* allow for such modification in other functions (5a''/b''), this modification is incompatible with their function as a relational or domain adjective (5a'/b').

- (5) a. Chris is a technical architect. (relational adjective)  
 a'. ?Chris is a very technical architect.  
 a''. This is a very technical job./The job is very technical.  
 b. The EU is a bureaucratic octopus. (domain adjective)  
 b'. ?The EU is a very bureaucratic octopus.  
 b''. This is a very bureaucratic process./The process is very bureaucratic.

Like a relational adjective, a domain adjective also needs to come closer to the noun than other adjectives (6).

- (6) a. a young technical architect/?a technical young architect (relational adjective)  
 b. a huge financial octopus/?a financial huge octopus (domain adjective)

Finally, like relational adjectives, domain adjectives are morphologically complex (7), with a derivational suffix that suggests a relation with a noun. This relation is usually not a simple matter of affixation (as usual in derivational morphology): *solar* relates to *sun*, *financial* to *finances*, *academic* to *academia*, and *culinary* to *kitchen*. Nevertheless, domain adjectives typically seem to be in a paradigmatic relation (in the morphological sense) with a noun that refers to the same domain. This is why an *academic crusade* can also be described as a *crusade in academia* or the *culinary mecca* the *mecca of the kitchen*.

- (7) a. *solar* heat, *technical* architect, *musical* instrument (relational adjective)  
 b. *financial* octopus, *academic* crusade, *culinary* mecca (domain adjective)

However, domain adjectives also differ from relational adjectives in a number of ways. First, domain adjectives do not establish a relation between the noun referent and another entity. As the label indicates, such a relation has often been pointed out for relational adjectives.

We can describe *solar heat* as ‘heat coming from the sun’ and a *financial advisor* as a ‘person who advises about finances’. This is different with domain adjectives. A financial octopus is not an octopus related to finances in the way that a financial advisor is an advisor related to (namely, advising about) finances. There is something about heat and advisors that provides the starting point for a relation with the sun and finances, respectively, but that is not the case with an octopus. The only prepositions that might be possible here are *of* and *in* (*the octopus of/in finances*), but these prepositions then play the more general role of relating the notion of octopus to a domain.

Domain adjectives are also not straightforwardly subsective, unlike relational adjectives. (8a) entails (8b) in every context, making *technical* a subsective adjective. It is harder to accept (9b) as an automatic entailment of (9a) in every context. The reason is that *octopus* has a metaphorical meaning in (9a) that it only has in (9b) in certain contexts.

- (8) a. Chris is a technical architect. (relational adjective)  
 b. Chris is an architect.
- (9) a. Citigroup is a financial octopus. (domain adjective)  
 b. Citigroup is an octopus.

Finally, domain adjectives are also different from relational adjectives in having an adverbial, sentential counterpart (10). It is not possible to paraphrase (8a) above, with the relational adjective, as (10a) below. The adverb *technically* in (10a) does not specify the *type* of architect Chris is. As pointed out in Ernst (2004), a domain adjective (9a) has a domain adverb (10b) as a counterpart. Such paraphrases are always possible for domain adjectives, but never for relational adjectives.

- (10) a. Technically, Chris is an architect. ( $\neq$ (8a)) (relational adverb)  
 b. Financially, Citigroup is an octopus. ( $=$ (9a)) (domain adverb)

### 3 The challenge of domain adjectives

We have seen that the phrase *financial octopus* refers to financial institutions that are similar to octopuses in certain respects. How is that interpretation derived from the interpretations of its two constituents, the domain adjective *financial* and the noun *octopus*? One might think that the easiest solution is to keep metaphor safely ‘in the lexicon’ as a part of lexical semantics. The combinatorial system takes a word like *octopus* from the lexicon with either a non-metaphorical meaning (say **octopus**) or a pre-packaged metaphorical meaning (say **m-octopus**) and it combines it with the meaning of the adjective, leading to either **financial(octopus)** or **financial(m-octopus)**. Only the second combination makes sense (because there is nothing financial about real octopuses) and this second combination therefore can have referents in the financial domain.

The problem with keeping metaphor entirely *inside* the lexicon is that the metaphorical meaning of the noun *octopus* depends on something *outside* the lexicon. This can either be an explicit local modifier (11a), a non-local modifier (11b), or the general context (11c). A financial octopus is something different than a bureaucratic or an electronic octopus and therefore it is hard to see how the choice of a *particular* metaphorical meaning for *octopus* can be made lexically, before it enters the compositional process.

- (11) a. It is a financial octopus.  
 b. Bureaucratically, it is an octopus.  
 c. How can we handle this octopus? (talking about a computer network)

One response might be that this problem does not arise when **m-octopus** corresponds to the widest possible sense of the word *octopus*, including anything that is similar to a octopus (like diversified financial institutions, complex government organizations, sprawling computer networks). What we see in (11) are simply ways to pick out specific subsets from this wide superset. The domain adjective in (11a) simply works intersectively, selecting those entities similar to octopuses that are at the same time financial entities.

$$(12) \quad \lambda x [ \mathbf{m-octopus}(x) \wedge \mathbf{financial}(x) ]$$

The problem with (12) is that it is incompatible with the non-predicative behaviour of domain adjectives that we mentioned in (4b'). Another problem with (12) is that this kind of local and explicit intersection does not work when the domain delimiter is not local (11b) or explicit (11c).

The kind-based analysis of relational adjectives (McNally & Boleda, 2004) goes only some way to address this problem. Extending their proposal, we could treat both the metaphorical noun as well as the domain adjective as predicates over *kinds* (variable  $k$ ), instead of ordinary individual entities (variable  $x$ ). The two kind denotations can be intersectively combined (13a), and related by a realization relation (**R**) to ordinary individuals (13b).

$$(13) \text{ a. } \text{financial octopus} \quad \lambda k [ \mathbf{m-octopus}(k) \wedge \mathbf{financial}(k) ]$$

$$\text{ b. } \text{be a financial octopus} \quad \lambda x \exists k [ \mathbf{R}(x,k) \wedge \mathbf{m-octopus}(k) \wedge \mathbf{financial}(k) ]$$

**m-octopus** denotes the set of kinds of things that are 'like' the octopus, in the metaphorical sense, including financial conglomerates, complex organizations, and electronic networks. The predicate **financial** picks out a proper subset of these kinds. This solves the problem of the non-predicative nature of domain adjectives: (4b') is ruled out because the subject *Citicorp* denotes an individual, while the predicate *financial* denotes a set of kinds (a sortal mismatch). However, we are still left with the question how a sentential adverb like *bureaucratically* can restrict the set of octopus kinds that are introduced through the noun in (11b) to the bureaucratic ones. More generally, the problem of the analysis in (13) is that it treats domain adjectives in the same way as relational adjectives. The notion of domain plays only an indirect and implicit role. As a result, the differences between domain adjectives and relational adjectives are left unaccounted for and the crucial question unanswered: what makes an adjective like *financial* a *domain* modifier in *financial octopus*?

## 4 The analysis of domain adjectives

The proposal of this paper is to make the role of metaphorical target domains explicit in the compositional semantics. If we want to understand how the meaning of a noun can be metaphorically shifted by domain adjectives, but also by other mechanisms (domain adverbs and context), then it makes sense to represent these domains somehow. They should not be confined to the lexicon, but made accessible for the compositional semantics. There are different ways to do this. This paper makes them maximally explicit as a sort of entities that variables in the logical interpretation language can refer to.

We assume a set  $D$  of *domains*, i.e.  $D = \{ \mathbf{finances}, \mathbf{bureaucracy}, \mathbf{electronics}, \dots \}$ . This set might have some structure, with one domain being a subdomain of another domain or the meet of two other domains, but that is not directly relevant for our purposes. They could be seen as 'pointers' to more complex and meaningful structures. From a cognitive perspective, domains could be seen as conceptual spaces in the sense of Gärdenfors (2000), that is, combinations of different quality dimensions that allow us to construct the concepts that belong to that domain. From the computational perspective of distributional semantics

they could be seen as areas of ‘topical similarity’ in a large, multidimensional space of word vectors (Clark, 2015). Or they could be seen as (constellations of) frames (Stickles, David, Dodge, & Hong, 2016).

We also assume a set  $K$  of *kinds*, with their own appropriate (taxonomic) structure but also with a particular relation to domains. This set will contain as elements such kinds as **octopus**, **storm**, and **scar**, each corresponding to their own (source) domain. The source domain that **storm** lives in would be **meteorology**, for instance. Kinds in a domain could be seen as regions in a conceptual space (Gärdenfors 2000), as word vectors (McNally 2017), or as nodes in a frame (Stickles, David, Dodge, & Hong, 2016).

What we need now is some way of implementing the idea of a conceptual metaphor as a mapping from a source domain to a target domain. Let us say that we can map a kind  $k$  to a domain  $d$ , resulting in a new kind  $k'$  that is the metaphorical counterpart of  $k$  in target domain  $d$ . For example, the result of mapping **octopus** to **finances**, represented as **octopus**  $\triangleright$  **finances** here, is the ‘counterpart’ of **octopus** in the **finances** domain. The kind **octopus**  $\triangleright$  **finances** is not an animal anymore, but a financial object that preserves properties of the octopus that are compatible with the financial domain. The operation  $\triangleright$  could be seen as a structure-preserving embedding of a concept in another conceptual space or as an algebraic transformation in a space of word vectors (Gutiérrez et al. 2016).

In principle, every kind has the potential now to be metaphorically instantiated in different domains (to the extent that its structure is preserved in those domains). It makes sense then to define, for a kind like **octopus**, a function  $\lambda d$  **octopus**  $\triangleright d$  that maps a target domain  $d$  to the metaphorical octopus **octopus**  $\triangleright d$  in that domain. If that function applies to **finances**, we get  $\lambda d$ .**octopus**  $\triangleright d$  (**finances**) = **octopus**  $\triangleright$  **finances**. This function gives us an individual concept, corresponding to the noun *octopus*, with the target domains have the role that *indices* play in intensional models: they allow us to relativize semantic values. In our case, metaphorical meanings can be relativized to target domains. Instead of the unanalyzed **m-octopus** that we had earlier, we now have an operator **m** that can be defined in terms of  $\triangleright$  and that lifts the non-metaphorical kind reference of a noun to its metaphorical, domain-dependent meaning: **m(octopus)** =  $\lambda k \lambda d k \triangleright d$  (**octopus**) =  $\lambda d$  **octopus**  $\triangleright d$ .

Building on this, a domain adjective can be analyzed as saturating the domain parameter of its noun. The adjective *financial*, for instance, takes the metaphorical meaning  $n$  of its nominal argument (a function from domains to kinds) and applies it to the domain **finances**, yielding the kind that corresponds to that domain. The lambda term of *financial* can be written as  $\lambda n.n$ (**finances**). This allows (4b) to be compositionally interpreted as in (14). The metaphorical lifting is specified in step (14b). A shortcut is taken in (14f): *is a* is directly linked to the lambda term  $\lambda k.\lambda x.\mathbf{R}(x,k)$  that maps kinds to their instantiations.

- |         |                             |   |
|---------|-----------------------------|---|
| (14) a. | octopus                     | <b>octopus</b>  |
| b.      | (metaphorical noun lifting) | <b>m</b> = $\lambda k \lambda d k \triangleright d$   |
| c.      | octopus                     | <b>m(octopus)</b><br>= $\lambda k \lambda d k \triangleright d$ ( <b>octopus</b> )<br>= $\lambda d$ <b>octopus</b> $\triangleright d$                                   |
| d.      | financial                   | $\lambda n n$ ( <b>finances</b> )   |
| e.      | financial octopus           | $\lambda n n$ ( <b>finances</b> ) ( $\lambda d$ <b>octopus</b> $\triangleright d$ )<br>= <b>octopus</b> $\triangleright$ <b>finances</b>                                |
| f.      | is a                        | $\lambda k \lambda x \mathbf{R}(x,k)$   |
| g.      | is a financial octopus      | $\lambda k \lambda x \mathbf{R}(x,k)$ ( <b>octopus</b> $\triangleright$ <b>finances</b> )<br>= $\lambda x \mathbf{R}(x, \text{octopus} \triangleright \text{finances})$ |
| h.      | Citicorp                    | <b>c</b>  |

- i. Citicorp is a financial octopus  $\lambda x \mathbf{R}(x, \mathbf{octopus} \triangleright \mathbf{finances})(\mathbf{c})$   
 $= \mathbf{R}(\mathbf{c}, \mathbf{octopus} \triangleright \mathbf{finances})$

The result of this derivation treats Citicorp as a realization of the financial counterpart of the octopus kind.

If there is no domain adjective, then the domain variable remains a free variable at the nominal level. One option is to specify this variable higher up in the structure by a domain adverb. (15) below shows how this is derived. The lambda term  $\lambda d. \mathbf{octopus} \triangleright d$  in (15a) that corresponds to the noun *octopus* (after metaphorical lifting) applies in (15b) to a variable  $d'$  that is abstracted over in step (15d) of the derivation, allowing the domain adverb *financially* (15e) to apply to the resulting function from domains to truth values (15f). (The variable  $s$  ranges over functions from domains to truth values, so the domain adverb *financially* is itself a function from such functions to truth values, or, equivalently, the set of all sets of domains that contain the **finances** domain.)

- |         |                                     |  |
|---------|-------------------------------------|--|
| (15) a. | octopus                             | $\lambda d \mathbf{octopus} \triangleright d$  |
| b.      | octopus                             | $\lambda d \mathbf{octopus} \triangleright d (d') = \mathbf{octopus} \triangleright d'$  |
| c.      | Citicorp is an octopus              | $\mathbf{R}(\mathbf{c}, \mathbf{octopus} \triangleright d')$   |
| d.      | (abstraction)                       | $\lambda d' \mathbf{R}(\mathbf{c}, \mathbf{octopus} \triangleright d')$  |
| e.      | financially                         | $\lambda s s(\mathbf{finances})$   |
| f.      | Financially, Citicorp is an octopus | $\lambda s s(\mathbf{finances})(\lambda d' \mathbf{R}(\mathbf{c}, \mathbf{octopus} \triangleright d'))$<br>$= \mathbf{R}(\mathbf{c}, \mathbf{octopus} \triangleright \mathbf{finances})$ |
| g.      | Citicorp is an octopus              | $\mathbf{R}(\mathbf{c}, \mathbf{octopus} \triangleright d')$   |

If there is no domain adverb, then the variable  $d'$  remains free (15g), with a domain value coming from the context.

This analysis explains the properties of domain adjectives that we discussed in section 3. The ‘nominal base’ in domain adjectives (*financ-ial*) corresponds to the domain identifier in the corresponding lambda term ( $\lambda n. n(\mathbf{finances})$ ). Their type (mapping domain-to-kind functions to kinds) is such that they can not apply an entity: hence the impossibility to use them predicatively. This type is also incompatible with degree modification and it also explains why domain adjectives are closer to the noun than other adjectives: if a noun is of the type domain-to-kind, then only domain adjectives (mapping domain-to-kind functions to kinds) can directly apply to them. Furthermore, we saw in (14) and (15) how the domain parameter in the noun allows for different ways of delimiting the domain. We understand now why domain adjectives are not straightforwardly substantive: it depends on the context that specifies domain  $d'$  in (15g) whether (15f) entails (15g).

## 5 The wider landscape of domain adjectives

With appropriate refinements, the proposal can scale up to a much wider set of examples with domain modifiers. There are cases where metaphorical adjectives (16a) and verbs (16b) are modified that work in the same way. There are also non-metaphorical cases of various types (Rawlins, 2004), like (16c) and (16d), and idiomatic examples like (16e) (Ernst, 1981; Gehrke & McNally, 2019):

- (16) a. Emotionally, he is damaged. / He is emotionally damaged.  
 b. Spiritually, she wandered. / She wandered spiritually.  
 c. Politically, California is irrelevant. / California is politically irrelevant.  
 d. Botanically, tomatoes are fruits. / Tomatoes are botanical fruits.

- e. Socially, he kicked the bucket. / He kicked the social bucket.

For *damaged* in (9a) and *wander* in (9c) we would have the functions  $\lambda d$  **damaged** $\triangleright d$  and  $\lambda d$  **wandering** $\triangleright d$ , respectively, with the same adverbs working either at the nominal or at the sentential level. What this would require is a way to talk about the kinds (or concepts) corresponding to adjectives like *damage* and verbs like *wander*. In (9c) and (9d) there are no metaphorical shifts, so we would have to generalize our idea of how predicates can depend on a domain:  $\lambda d$  **irrelevant** $\triangleright d$  and  $\lambda d$  **fruit** $\triangleright d$  are functions that relativize the concepts of irrelevance and fruit to different domains, making it possible to specify those domains at different points in the sentence. What makes the situation in (9e) special is that the domain adjective *social* modifies a noun (*bucket*) that is part of an idiom (*kick the bucket*) for ‘die’. A loosening of the relation between form and meaning, maybe along the lines of Gehkre & McNally (2019), is necessary to allow the adjective *social* to target the domain variable of the whole idiom’s meaning (that is,  $\lambda d$  **die** $\triangleright d$ ), even though it is the sister of the (meaningless) idiom part *bucket*.

Clearly, each of these types requires much more study before we can arrive at an integrated account of domain modification of different categories, with and without metaphorical shifts, with and without idiomatic structures. The analysis of domain modifiers given in this paper provides a promising starting point for such an account, because of the way it makes concepts and domains an explicit part of the semantics.

## 6 Conclusion

Domain adjectives with metaphorical nouns have never been properly analyzed in formal semantics. The need for such an analysis required us to take the role of domains seriously in compositional semantics. In this way, the intricate behaviour of domain adjectives received a natural account. It also opened up perspectives for a wider set of constructions with domain adjectives and domain adverbs. Moreover, giving domains their proper place in the compositional semantics allows that kind of semantics to connect to richer models of domains and metaphorical meanings, whether coming from a conceptual perspective (Gärdenfors, 2000), based on frames and mappings (like Stickers et al. 2016) or from a computational perspective, based on distributional vector spaces and transformations (Gutiérrez, Shutova, Marghetis, & Bergen, 2016). In this way, it helps us to better understand the amazing adaptability of word meanings across an open-ended space of conceptual domains.

## References

- Clark, S. (2015). Vector Space Models of Lexical Meaning. In S. Lappin, & C. Fox (Eds.), *The Handbook of Contemporary Semantic Theory* (pp. 493-522). Hoboken, NJ: Wiley.
- Ernst, T. (2004). Domain adverbs and the syntax of adjuncts. In J. Austin, S. Engelberg, & G. Rauh (Eds.), *Adverbials: the Interplay between Meaning, Context, and Syntactic Structure* (pp. 103-129). Amsterdam: John Benjamins.
- Ernst, T. (1981). Grist for the linguistic mill: Idioms and ‘extra’ adjectives. *Journal of Linguistic Research*, 1 (3), 51-68.
- Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press.
- Gehrke, B., & McNally, L. (2019). Idioms and the syntax/semantics interface of descriptive content vs. Reference. *Linguistics*, 57 (4), 769-814.



Gutiérrez, E., Shutova, E., Marghetis, T., & Bergen, B. (2016). Literal and metaphorical senses in compositional distributional semantic models. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (pp. 183-193).

Hanks, P. (2004). The syntagmatics of metaphor and idiom. *International Journal of Lexicography*, 17 (3), 245-274.

McNally, L., & Boleda, G. (2004). Relational adjectives as properties of kinds. In O. Bonami, & P. Cabredo Hoffherr (Eds.), *Empirical Issues in Syntax and Semantics 5* (pp. 179-196). Paris: CNRS.

Rawlins, K. (2004). Examining domain adverbs semantically. Handout, UC Santa Cruz.

Reijnierse, G., Burgers, C., Krennmayr, T., & Steen, G. (2018). On metaphorical views, dynamite, and doodlings. *Review of Cognitive Linguistics*, 16 (2), 431-454.

Stickles, E., David, O., Dodge, E., & Hong, J. (2016). Formalizing contemporary conceptual metaphor theory. *Constructions and Frames*, 8 (2), 166-213.

Sullivan, K. (2013). *Frames and Constructions in Metaphoric Language*. Amsterdam: John Benjamins.

# Monotonicity restored: *more* never means *purser*\*

Alan Bale<sup>1</sup>, Bernhard Schwarz<sup>2</sup>, and David Shanks<sup>2</sup>

<sup>1</sup> Concordia University, Montreal, Quebec, Canada  
alan.bale@concordia.ca

<sup>2</sup> McGill University, Montreal, Quebec, Canada  
bernhard.schwarz@mcgill.ca  
david.shanks4@mail.mcgill.ca

## Abstract

Bale and Barner (2009) suggested that comparatives with *more* plus mass noun, such as *There is more gold in the ring than there is in the bracelet*, can be interpreted as comparing degrees of purity of material. As Wellwood (2015) observes, this suggestion is incompatible with Schwarzschild’s (2006) *Monotonicity Constraint* which puts limits on the types of measure functions that can enter into the evaluation of such comparisons. Scrutinizing the truth conditions of a range of relevant cases, we resolve this tension by showing that comparisons of purity with *more* are only apparent, and that the correct analysis of such cases is one that respects the Monotonicity Constraint.

## 1 Introduction

To block unattested readings of comparatives and other degree constructions, Schwarzschild (2006) posits a Monotonicity Constraint (MC): measure functions that enter into the evaluation of certain constructions must be monotonic with respect to the part-whole ordering inherent in noun phrases. Despite the descriptive utility of MC, Bale and Barner (2009) discuss readings of comparative constructions which, in their (and Wellwood’s 2015) characterization, involve non-monotonic measurements of *purity*. We argue, however, that reference to such MC-violating functions is only apparent. We demonstrate that the cases that purportedly require purity measurements instead involve proportional measure functions that critically conform to MC.

Section 2 reviews MC and some of the empirical evidence that motivates it. Section 3 introduces apparent comparisons of *purity* first observed in Bale and Barner 2009, spelling out the MC-incompatible analysis that Bale and Barner hinted at. Section 4 presents an alternative account that is compatible with MC. Sections 5 and 6 argue that, independently of MC, this alternative is more adequate than Bale and Barner’s proposal. Section 7 concludes.

## 2 Measure functions and monotonicity

As discussed in Schwarzschild 2006, different types of modification seem to license different types of measurements. For example, although it is perfectly acceptable to talk about *three litres of water*—where the modification of *water* by *three litres* happens in a pseudo-partitive construction—it is very odd to talk about *three degrees of water*. In contrast, although one can talk about *three-degree water*—where *three-degree* modifies *water* by some kind of compounding

---

\*For comments and discussion, we thank the members of the Semantics Research Group at McGill University, especially Aron Hirsch, and three anonymous reviewers for *Amsterdam Colloquium 22*. The authors acknowledge support provided by the Social Sciences and Humanities Research Council of Canada (SSHRC; Bale, Schwarz: Insight Grant #435-2016-1376; Schwarz, Shanks: Insight Grants #435-2013-0592, #435-2016-1448) and by McGill University’s Faculty of Arts (Shanks: Arts Research Internship Award).

operation—it is very awkward to talk about *three-litre water*. [Schwarzschild](#) hypothesized that this contrast correlates with how measurement systems relate to certain types of partial orders.

In our examples with *water* above, there are two types of measurement systems: volume and temperature. The elements that *water* is true of are partially ordered by the material part-whole relation. The two measurement systems differ in terms of how they interact with this part-whole relation: if a portion of water  $x$  is a proper part of another portion of water  $y$ , then it necessarily follows that  $x$  is less in volume than  $y$ , however it does not necessarily follow that  $x$  is less in temperature than  $y$ . A function that measures volume is a *monotonic measure function*, in the sense defined in (1), while a function that measures temperature is not.

- (1) **Monotonic Measure Functions:** A measure function  $\mu$  is monotonic with respect to a strict partial order  $\sqsubset$  if and only if for all  $x$  and  $y$ ,  $x \sqsubset y \rightarrow \mu(x) <_{\mu} \mu(y)$ , where  $<_{\mu}$  is the strict ordering of degrees in the range of  $\mu$ .

According to [Schwarzschild](#), determining what counts as being “three litres” or “three degrees” requires measuring members of the nominal denotation. Different measure terms require different measure functions: *three litre(s)* requires a function that maps individuals to a measure of volume, whereas *three degrees* requires a function that maps individuals to a measure of temperature. Certain types of grammatical constructions, such as the pseudo-partitive above, only permit monotonic measure functions, whereas others exclude such functions.

Monotonicity not only plays a role in pseudo-partitives, but also in constructions where *much*, or its comparative counterpart *more*, combines directly with a noun. As discussed in [Cresswell 1976](#) and [Bale and Barner 2009](#), when *much* and *more* combine with mass nouns, the dimension of comparison can vary. For example, *having too much string* or *having more string than needed* usually means having too much in terms of length (but not weight or volume), whereas *having too much pudding* or *having more pudding than needed* usually means having too much in terms of weight or volume (but not length). In contrast, *witnessing too much celebration* or *witnessing more celebration than expected* usually means witnessing too much in terms of the time-span, number of participants, or intensity of emotion (but definitely not too much in terms of weight, length, or volume). Thus, *much* permits variable types of measurement. Nevertheless, not all measurement types are available. Consider the sentences in (2).

- (2) a. This tub has too much water in it.  
b. This tub has more water in it than that tub.

The sentence in (2a) cannot be true in a context where the weight/volume of the water is not *too much*, but where the water is too hot. Similarly, (2b) cannot be true in a context where both tubs have an equal amount of water in terms of weight/volume, but where the water in one tub is hotter than the water in the other. In other words, just like the pseudo-partitive constructions, the sentences in (2) are required to be evaluated using measurement systems that are monotonic with respect to the material part-whole relation of *water*.

To explain why sentences like (2a) cannot be construed as involving measurements of temperature, [Schwarzschild \(2006\)](#) adapted the constraint that he hypothesized for pseudo-partitive constructions so that it could also apply to the combination of *much* with mass nouns.

- (3) **Monotonicity Constraint (MC)** ([Schwarzschild 2006](#)): Only monotonic functions are available as values for the measure function in the meaning of *much*.

MC also applies to the comparative *more* as in (2b) if it is decomposed into *much+er* (cf. [Bresnan 1973](#)). We will in fact focus mainly on comparatives, as their truth conditions are less

vague and thus easier to assess. For concreteness, we will adopt a specific analysis of *much* and *-er*, although the generalizations discussed above hold for any reasonable analysis involving degrees. As (4a) shows, we assume that *much* is a determiner that applies to a degree and two sets of individuals (cf. Hackl 2000). Critically, the output that *much* returns depends on the value of an underspecified measure function (symbolized with the variable  $\mu$ ), which measures the supremum (i.e., the sum, symbolized with  $\sqcup$ ) of the elements in the input sets' intersection.<sup>1</sup>

$$(4) \quad \begin{array}{l} \text{a. } \llbracket \textit{much} \rrbracket = \lambda d_d. \lambda X_{\text{et.}}. \lambda Y_{\text{et.}}. \mu(\sqcup(X \cap Y)) \geq d \\ \text{b. } \llbracket \textit{-er} \rrbracket = \lambda f_{\text{dt.}}. \lambda g_{\text{dt.}}. \text{MAX}(g) > \text{MAX}(f) \end{array}$$

As (4b) shows, we take comparative *-er* to relate two sets of degrees. This meaning assumes that there are two instances of *much*, one in the main clause and another (which is usually silent because of ellipsis) in the *than*-clause. The degree argument in the *than*-clause is abstracted to form a degree predicate (the first argument to *-er*). The comparative morpheme combines with the *than*-clause creating a generalized degree quantifier. This entire phrase (*-er* plus *than*-clause) covertly moves to a propositional level binding a degree argument within the parameterized determiner (cf. Heim 2000). Thus, a sentence like (2b) would have an LF representation like the one in (5), resulting in truth conditions like those in (6).

$$(5) \quad \text{-er } \lambda d [\text{than } [ [d \textit{much}] \textit{water}] \lambda x [\text{that tub has } x \textit{ in it}]] \\ \lambda d [ [ [d \textit{much}] \textit{water}] \lambda x [\text{this tub has } x \textit{ in it}]]$$

$$(6) \quad \text{MAX}(\{d : \mu(\sqcup\{x : \text{WATER}(x) \ \& \ \text{IN-THIS-TUB}(x)\}) \geq d\}) > \\ \text{MAX}(\{d : \mu(\sqcup\{x : \text{WATER}(x) \ \& \ \text{IN-THAT-TUB}(x)\}) \geq d\})$$

We will adopt this general analysis of the comparative throughout, although it is important to remember that the compositional details are moot. What is critical for our purposes is how the measure function variable is valued. The correct truth conditions are obtained for (5) by setting the value of  $\mu$  in both the main-clause and *than*-clause to either of the monotonic values in (7), but critically not the non-monotonic value in (8).

$$(7) \quad \begin{array}{ll} \text{a. } \mu = \lambda x. \text{VOLUME}(x) & (\textit{than-clause} \ \& \ \textit{main clause}) \\ \text{b. } \mu = \lambda x. \text{WEIGHT}(x) & (\textit{than-clause} \ \& \ \textit{main clause}) \end{array}$$

$$(8) \quad \mu = \lambda x. \text{TEMPERATURE}(x) \quad (\textit{than-clause} \ \& \ \textit{main clause})$$

In summary, MC in (3) permits *much* to have a semantic value that is not fully fixed lexically (which is empirically required to account for its behaviour with respect to various mass nouns) while imposing limits on permissible values that prevent empirically unattested readings.

### 3 Relative readings: non-monotonic proportionality?

Given the data reviewed in the previous section, it is clear that something akin to MC is needed to prevent unattested meanings. However, Wellwood (2015) portrays certain comparatives with mass nouns like *gold* and *silver* as a potential challenge to MC. In this section, after first discussing data with such mass nouns that MC applies to correctly, we will present cases of readings, we will call them *relative*, that could be analyzed as inconsistent with this constraint.

<sup>1</sup>A familiar alternative to analyzing *much* as a projecting a determiner is to assume that it projects an intersective adjectival modifier (e.g., Schwarzschild 2006, Wellwood 2015). While lack of space prevents us from elaborating, we note that the existence of this alternative does not affect our overall conclusions, even though some of the details of our argumentation would change if couched in this alternative setting.

Intuitions indicate that there are a variety of different ways one could compare pieces of precious metals like gold and silver. One could compare them by weight, volume, beauty or shininess (a.k.a., gloss). If one worked with shaping these metals (where gauging temperature is relevant), it would even be natural to compare them in terms of heat. However, as discussed in [Schwarzschild 2006](#) and [Wellwood 2015](#), in many cases only certain types of comparisons are permitted. For example, (9a) can only be true in a situation where the gold is “too much” in terms of weight or volume; it cannot be true by virtue of the gold being too hot, shiny, or beautiful. Similarly, (9b) can only be true if Al’s gold has a greater weight or volume than Bill’s, independent of whether Al’s gold is hotter, shinier, or more beautiful.

- (9) a. He put too much gold in the ring. ([Schwarzschild 2006](#))  
 b. Al has more gold than Bill does. ([Wellwood 2015](#))

As [Schwarzschild](#) and [Wellwood](#) point out, MC applies again correctly, regulating the availability of these various dimensions of comparison. Like *water*, *gold* is implicitly associated with the material part-whole relation; and as before, *weight* and *volume* are monotonic with respect to this relation: for any piece of gold  $x$  that is a proper part of a piece  $y$ , it necessarily follows that  $x$  has less weight and less volume than  $y$ . In contrast, *beauty*, *shininess*, and *temperature* are not monotonic: for any piece of gold  $x$  that is a proper part of a piece of gold  $y$ , it does not necessarily follow that  $x$  is colder than  $y$  or that  $x$  has less beauty/shininess than  $y$ .

Although the range of readings in (9) is captured by underspecification restrained by MC, there are other constructions which may appear problematic. To illustrate, consider [Bale and Barner’s \(2009\)](#) example (10) below. [Bale and Barner](#) report that among the readings that (10) permits, there is one that they portray as a comparison of the silver and gold in terms of *purity*. Indeed, (10) can be true in a scenario where Esme’s ring is quite small but made of high quality, pure gold and Seymour’s necklace is quite large but made of a low-grade, impure silver. That is, the sentence can be true even when the total amount of gold in Esme’s ring is less in absolute weight and volume than the total amount of silver in Seymour’s necklace. This reading becomes even more prominent when *relatively speaking* is added to the front of the sentence, and for this reason we will refer to this type of reading as a *relative* reading.

- (10) Esme has more gold in her ring than Seymour has silver in his necklace.

As noted, [Bale and Barner](#) hint at an analysis, let us now label it *purity-based*, on which (10) in its relative reading is a comparison of purity measures. It will be useful for us to spell out this analysis in some detail. We begin by observing that mass nouns like *gold* and *silver* have both *loose* and *strict* usages. Under a strict usage, the nouns are only true of aggregates consisting of pure gold (the element Au) and pure silver (the element Ag) respectively. However, under a loose usage—the more common one by far—the nouns are not only true of pure gold and pure silver but also alloys consisting of these elements. It is this loose meaning that plausibly allows the phrase *12 karat gold* to be non-contradictory (only 24 karat gold is pure gold) or that permits the question *How pure is this silver?* to be non-trivial.

We can exploit this ambiguity to define a measure of purity: the purity of an aggregate of gold/silver, understood in its loose sense, is the proportion of pure gold/silver within that aggregate. Thus, in order to be consistent with the purity analysis, the occurrences of *gold* and *silver* in (10) must be understood in the loose sense. At the same time, the measure functions operative in the two occurrences of *much* must reference the strict sense of these nouns. (11) spells out such purity measures as functions mapping an aggregate of matter to the weight proportion of Au/Ag contained in it.

- (11) a.  $\text{PURITY}_{\text{Au}} := \lambda x. \text{WEIGHT}(\sqcup\{y: y \sqsubseteq x \wedge \text{Au}(y)\}) / \text{WEIGHT}(x)$   
 b.  $\text{PURITY}_{\text{Ag}} := \lambda x. \text{WEIGHT}(\sqcup\{y: y \sqsubseteq x \wedge \text{Ag}(y)\}) / \text{WEIGHT}(x)$

These measure functions can be used to derive plausible truth conditions for (10). The expected LF of (10) is shown in (12). Given the meaning of *much* and *-er* specified in the previous section, this LF ends up producing the general truth conditions in (13). There, we label the underspecified measure functions  $\mu_1$  and  $\mu_2$  respectively, using subscripts to indicate the connection between the measure functions and the two instances of *much*. The purity reading could then arise from the setting in (14). We note that this analysis predicts, correctly it seems, that the relative reading of (10) can receive a paraphrase like (15).

- (12) -er  $\lambda d$ [than [ [  $d$  much] silver]  $\lambda x$ [Seymour has  $x$  in is necklace]]  
 $\lambda d$ [ [  $d$  much] gold]  $\lambda x$ [Esme has  $x$  in her ring]]
- (13)  $\text{MAX}(\{d : \mu_1(\sqcup\{x : \text{GOLD}(x) \ \& \ \text{IN-ESME'S-RING}(x)\}) \geq d\}) >$   
 $\text{MAX}(\{d : \mu_2(\sqcup\{x : \text{SILVER}(x) \ \& \ \text{IN-SEYMOUR'S-NECKLACE}(x)\}) \geq d\})$
- (14) a.  $\mu_1 = \lambda x. \text{PURITY}_{\text{Ag}}(x)$  (than-clause)  
 b.  $\mu_2 = \lambda x. \text{PURITY}_{\text{Au}}(x)$  (main clause)
- (15) The gold in Esme's ring is purer than the silver in Seymour's necklace.

This purity-based analysis looks like an initially plausible account of the relative reading of (10). One of its benefits is that it naturally extends to other instances of relative readings. To illustrate with just one additional case, consider the sentence in (16).

- (16) This bottle has more alcohol in it than that bottle does.

This sentence too permits a relative reading, that is, one that does not involve a comparison of absolute weights or volumes. It can be judged true if the percentage of ethanol is higher in this bottle than in that bottle, but the absolute volume/weight of ethanol is not higher, for example in the following specific scenario: this bottle contains 30% ethanol within 750 millilitres of whisky but that bottle contains only 25% ethanol within 2 litres of whisky.

A purity-based analysis of (16) is fully parallel to the purity-based analysis of (10). Like *gold* and *silver*, the mass noun *alcohol* has both a strict and loose usage. In its strict usage, it is only true of ethanol. In its loose usage, it is true of liquids that have ethanol as a principal component. It is in this loose sense that statements like *this whisky is alcohol* can be true or questions like *How pure is this alcohol?* are non-trivial. The purity of an aggregate of alcohol in the loose sense can be characterized as the proportion of alcohol in the strict sense that this aggregate contains. Hence (16) can be analyzed by interpreting *alcohol* in the loose sense, and by positing a measure function that references alcohol in the strict sense. Specifically, the relevant purity function can be defined as in (17), as a function that maps any aggregate to the volume proportion of ethanol contained in it. The relative reading of (16) can now be captured assuming the LF in (18), deriving the general truth conditions in (19), and by setting the value of  $\mu$ , for both occurrences of *much*, to the purity function defined in (17).

- (17)  $\text{PURITY}_{\text{ethanol}} := \lambda x. \text{VOL}(\sqcup\{y: y \sqsubseteq x \wedge \text{ETHANOL}(y)\}) / \text{VOL}(x)$
- (18) -er  $\lambda d$ [than [ [  $d$  much] alcohol]  $\lambda x$ [that bottle has  $x$  in it]]  
 $\lambda d$ [ [  $d$  much] alcohol]  $\lambda x$ [that bottle has  $x$  in it]]
- (19)  $\text{MAX}(\{d : \mu(\sqcup\{x : \text{ALCOHOL}(x) \ \& \ \text{IN-THIS-BOTTLE}(x)\}) \geq d\}) >$   
 $\text{MAX}(\{d : \mu(\sqcup\{x : \text{ALCOHOL}(x) \ \& \ \text{IN-THAT-BOTTLE}(x)\}) \geq d\})$

So, positing that *much* can reference purity measure functions like those in (11) and (17) enables a seemingly natural analysis of relative readings with mass nouns like *gold*, *silver* and *alcohol*. However, as Wellwood (2015) notes, this analysis of relative readings is inconsistent with Schwarzschild’s MC. In virtue of capturing the intuitive notion of purity, the functions in (11) and (17) are non-monotonic. After all, if  $y$  is a piece of gold or silver alloy that has a piece  $x$  as a proper part, it does not necessarily follow that  $x$  is less pure than  $y$  in the sense of containing a lower proportion of Au or Ag. Similarly, a portion of whisky  $x$  that is a proper part of a portion of whisky  $y$  does not necessarily have a smaller ethanol proportion than  $y$ .

Does this mean that the existence of relative readings calls for a weakening of MC? We think not. The remainder of this paper is dedicated to detailing our justification for this answer.

## 4 Relative readings: monotonic proportionality?

This section articulates an account of relative readings which, minimally, appears like a credible alternative to the purity-based analysis. Given the availability of such an alternative, the mere existence of readings that qualify as relative (in virtue of not comparing absolute weights or volumes) is insufficient to establish that the interpretation of *much* can make reference to the non-monotonic measure functions like those in (11) and (17).

Returning to the relative reading of (10), we retain the LF in (12) and the resulting general truth conditions (13), but we now propose the value assignments in (20) as an alternative to (14). (In a somewhat different context, similar functions are posited in Bale and Schwarz 2019.)

$$(20) \quad \begin{array}{ll} \text{a. } \mu_1 = \lambda x. \text{ WEIGHT}(x)/\text{WEIGHT}(\text{Seymour's necklace}) & (\text{than-clause}) \\ \text{b. } \mu_2 = \lambda x. \text{ WEIGHT}(x)/\text{WEIGHT}(\text{Esme's ring}) & (\text{main clause}) \end{array}$$

For any input  $x$ , the functions in (20) output the proportion of  $x$ ’s weight to the weight of Seymour’s necklace and Esme’s ring, respectively. Given our assumptions, in the interpretation of (10) these functions take as inputs the supremum (i.e., the sum) of all the silver aggregates in Seymour’s necklace and of all the gold aggregates in Esme’s ring, respectively. (20) is compatible with interpreting *gold* and *silver* in either a strict or loose sense. If the strict reading is chosen, the truth conditions resulting from (20) can be stated as in (21). Similarly for the loose reading.

$$(21) \quad \frac{\text{WEIGHT}(\text{the Au in E's ring})/\text{WEIGHT}(\text{E's ring})}{\text{WEIGHT}(\text{the Ag in S's necklace})/\text{WEIGHT}(\text{S's necklace})} >$$

Such truth conditions indeed qualify as relative, as it is not absolute amounts that are being compared. In fact, under certain conditions they mimic the truth conditions based on (14), which we can state as in (22). Specifically, (21) and (22) become equivalent in a context where the ring and necklace referred to in (10) are taken to be made entirely of gold and silver (alloys), respectively — because then the weight of the gold (alloy) in Esme’s ring equals the weight of Esme’s ring, and likewise for Seymour’s necklace and the silver (alloy) it consists of.

$$(22) \quad \frac{\text{WEIGHT}(\text{the Au in E's ring})/\text{WEIGHT}(\text{the gold (alloy) in E's ring})}{\text{WEIGHT}(\text{the Ag in S's necklace})/\text{WEIGHT}(\text{the silver (alloy) in S's necklace})} >$$

The next two sections will provide a more thorough assessment of the relative merits of measure function settings like (14) and (20). Scrutinizing truth conditions, we will present evidence confirming that *much* can reference measure functions like those in (20), as well as evidence showing that *much* cannot reference measure functions like those in (11)/(14).



Before attending to this evidence, we note that the conclusions it supports bear on the assessment of Schwarzschild's (2006) MC. The reason is that, unlike the measure functions in (11)/(14), those in (20) are monotonic. To see this, we observe that in the fractions in (20), (i) WEIGHT, the measure function that appears in the numerator, is itself monotonic, and (ii) the denominator is a constant. The claim now follows from the general fact that for any measure function  $m$  and positive real number  $c$ , if  $m$  is monotonic, then so is  $\lambda x.m(x)/c$ . (So the crucial difference between the functions in (11)/(14) and (20) is that only the latter have this form, while the former instead have the form  $\lambda x.m(x)/n(x)$ , where the denominator can vary with the input.) The conclusion that the evidence presented below will support, then, is that the relative readings with *much* are compatible with MC, and that they in fact ultimately strengthen the case for MC, as there are truth-conditional reasons independent from MC for not letting *much* make reference to measure functions that are purity-based and hence non-monotonic.

## 5 Evidence for monotonic proportionality

We have said that measure functions like those in (20), which we now label *monotonic proportional*, under certain conditions replicate the effect of purity-based measure functions like those in (11). However, the two types of measure functions of course do not deliver identical truth conditions. We will now identify cases where the truth of a relative reading can be attributed to monotonic proportional measure functions but not to purity-based measure functions, thereby furnishing evidence for the availability of the former.

Returning again to (10), recall that the possible relative truth conditions in (21) and (22) become equivalent under the assumption that Esme's ring and Seymour's necklace are made up entirely of gold and silver (alloys), respectively. But it is of course not necessary to make this sort of assumption. A natural way of evaluating and comparing the adequacy of the meanings derived by different types of measure functions is to examine scenarios where the assumption does not hold. To pursue this avenue, we switch to the example in (23).

(23) There is more gold in the ring than there is in the bracelet.

This sentence permits the same sort of relative reading as (10) or (16). It is like (16), and unlike (10), in that the same mass noun featured in the main clause (viz. *gold*) is also the mass noun interpreted in the *than*-clause, which simplifies the exposition and judgments. Under our current assumptions, (23) has the LF (24) and the general truth conditions in (25).

(24) -er  $\lambda d [ [d \text{ much}] \text{ gold}] \lambda x [\text{there is } x \text{ in the bracelet}]$   
 $\lambda d [ [d \text{ much}] \text{ gold}] \lambda x [\text{there is } x \text{ in the ring}]$

(25)  $\text{MAX}(\{d : \mu_1(\sqcup\{x : \text{GOLD}(x) \ \& \ \text{IN-THE-RING}(x)\}) \geq d\}) >$   
 $\text{MAX}(\{d : \mu_2(\sqcup\{x : \text{GOLD}(x) \ \& \ \text{IN-THE-NECKLACE}(x)\}) \geq d\})$

On the purity-based analysis, the relative reading of (23) is credited to the setting of the measure function variables  $\mu_1$  and  $\mu_2$  in (26). The effect of this setting is to be compared to the effect of (27), the relevant setting under the monotonic proportionality analysis.

(26)  $\mu_1 = \mu_2 = \lambda x. \text{WEIGHT}(\sqcup\{y: y \sqsubseteq x \ \& \ \text{Au}(y)\})/\text{WEIGHT}(x)$  (*than*-clause & main cl.)

(27) a.  $\mu_1 = \lambda x. \text{WEIGHT}(x)/\text{WEIGHT}(\text{the bracelet})$  (*than*-clause)  
 b.  $\mu_2 = \lambda x. \text{WEIGHT}(x)/\text{WEIGHT}(\text{the ring})$  (main clause)

Consider now the following scenario. Both the ring and the bracelet mentioned in (23) are



made of copper plated with white gold. The white gold, a gold alloy, consists of 90 weight percent Au and 10 weight percent nickel. The ring consists of 10g of white gold and 10g of copper, while the bracelet consists of 10g of white gold and 30g of copper.

Note that this scenario is predicted to render (23) false in a reading that compares absolute amounts. Moreover, under the setting in (26), the truth conditions in (25) come out false too. After all, the weight proportion of Au in the white gold that the ring contains equals, hence does not exceed, the weight proportion of Au in the white gold that the bracelet contains. In contrast, (25) comes out true under the setting in (27), regardless of whether *gold* is interpreted loosely or strictly. Under a loose interpretation, (24) is true in virtue of the weight proportion of white gold in the ring (50 percent) exceeding the weight proportion of white gold in the bracelet (25 percent); under a strict interpretation (24) is true in virtue of the weight proportion of Au in the ring (45 percent) exceeding the weight proportion of Au in the bracelet (22.5 percent).

Aligned with similar judgments reported in Bale and Schwarz (2019), we take it that (23) can indeed be judged true in the scenario we have given. Intuitions indicate, moreover, that (23) is judged true in the scenario given on the grounds that the white gold/Au in the ring comprises a greater weight proportion of the ring than the white gold/Au in the bracelet does of the bracelet. This judgment can be captured by the monotonic proportional measure function setting in (27) but not by the non-monotonic, purity-based, setting in (26). It therefore provides non-confounded evidence for the availability of monotonic proportional measure functions like those in (20) and (27) in the interpretation of *much*.

To confirm and sharpen the relevant judgments, it will be useful to also consider examples where the purity-based analysis is not expected to apply in the first place. Recall that the purity-based analysis exploits the strict-loose ambiguity of the mass nouns that *much* combines with. This feature of the analysis invites us to consider comparatives where the relevant mass nouns do not participate in the strict-loose ambiguity. If such examples permit relative readings, they confirm that relative readings can have a source other than purity-based proportionality.

Let us compare comparatives with the mass nouns *alcohol* and *ethanol*. As noted in Section 3, the former is ambiguous between a loose meaning, which holds of portions of liquids that have ethanol as their principal component, and a strict meaning, which holds of portions of ethanol. We now add that, in contrast, the noun *ethanol* itself does not seem to have a comparable strict-loose ambiguity. With this in mind, we first present sentence (28a) below as a baseline. It allows for a relative reading of the same sort as the one attested for Bale and Barner's example (10). It can be read, in particular, as conveying that the volume proportion of ethanol in Sample 1 exceeds the volume proportion of ethanol in Sample 2, a reading on which the sentence's truth is independent of the absolute amounts of ethanol in the two samples. This interpretation could again be captured with reference to either a non-monotonic, purity-based, measure function (cf. (26)) or to a pair of monotonic proportional measure functions (cf. (27)).

- (28) a. There is more alcohol in Sample 1 than in Sample 2.  
 b. There is more ethanol in Sample 1 than in Sample 2.

But now compare (28a) to (28b), where *ethanol* substitutes for *alcohol*. *Ethanol* lacking the relevant strict-loose ambiguity, the purity-based analysis does not predict (28b) to have a relative reading. In particular, since this analysis requires that in the relative reading of (28a) *alcohol* have its loose interpretation (which *ethanol* does not share), the analysis does not predict that (28b) can be judged equivalent to (28a) in a relative reading. In contrast, the monotonic proportional analysis, which can derive a relative reading of (28a) while interpreting *alcohol* strictly (or loosely), predicts that (28b) can be read as equivalent to (28a) in its relative reading.

Again aligned with judgments reported in Bale and Schwarz 2019, speaker intuitions clearly

bear out the latter prediction. That is, (28a) and (28b) are indeed intuited to have equivalent relative readings. Again, this confirms that it is at least possible for relative readings in comparatives with *more* to arise from monotonic proportional measure functions.

The next section will strengthen this conclusion by showing that the purity-based analysis, apart from undergenerating relative readings, also overgenerates.

## 6 Evidence against non-monotonic proportionality

We return to the example (23), repeated in (29). Still assuming the general truth conditions in (25), we will continue to examine the effects of the measure function settings in (26) and (27).

(29) There is more gold in the ring than there is in the bracelet.

Consider now the following scenario. Both the ring and the bracelet are made of silver plated with white gold, an alloy composed of gold and nickel. The ring's white gold plating weighs 5g, and the ring's total weight is 20g; the bracelet's white gold plating weighs 20g, and the bracelet's total weight is 40g. The ring's white gold plating is composed of 90 percent Au and 10 percent nickel; the bracelet's white gold plating is 80 percent Au and 20 percent nickel.

This scenario again renders (29) false in a reading that compares absolute amounts. The scenario likewise fails to render true the truth conditions in (25) under the monotonic proportional measure function setting in (27), given that the ring's white gold plating (or the Au it contains) accounts for a much smaller proportion of the ring's weight than the bracelet's white gold plating (or the Au it contains) does of the bracelet's weight. However, the scenario satisfies the truth conditions that would arise from the purity-based measure function setting in (26), on the grounds that the ring's white gold plating contains a higher weight proportion of Au than the bracelet's white gold plating does.

As for speaker judgments about (29), it seems hard or impossible to interpret the sentence in a sense that is true in the scenario given. In this respect, (29) contrasts sharply with (30), whose truth in the scenario given is uncontroversial.

(30) The gold in the ring is purer than the gold in the bracelet.

This judgment about (29), and the intuited contrast with (30), are incompatible with the assumption that *much* can be evaluated using a non-monotonic, purity-based, measure function. In conjunction with the findings in the last section, they lead us to conclude that the purity-based analysis of relative readings hinted at in Bale and Barner 2009 is incorrect, and that the relative readings that may seem to call for such an analysis are instead to be analyzed as involving monotonic proportional measure functions.

## 7 Conclusions

We have argued that relative readings of sentences like Bale and Barner's (2009) example (10), repeated once more in (31), invoke reference to monotonic proportional measure functions, rather than non-monotonic purity-based measure function, and therefore pose no threat to Schwarzschild's (2006) Monotonicity Constraint (MC).

(31) Esme has more gold in her ring than Seymour has silver in his necklace.

Our conclusion that the interpretation of comparatives with *more* is compatible with MC is aligned with the assessment offered in Wellwood (2015). However, Wellwood bases this assessment on the assumption that relative readings do not exist in the first place. Without discussing examples like (31), Wellwood justifies this assumption with the observation that a comparative like (9b) above, repeated here as (32), does not permit a relative reading.

(32) Al has more gold than Bill does. (Wellwood 2015)

While the existence of relative readings for cases like (31) seems undeniable, we agree with Wellwood that no such reading is available for (32). It seems that (32) indeed can only be understood as a comparison of absolute amounts. The question is then why (31) and (32) should differ in this way. More generally, how are relative readings distributed, that is, how can we characterize the general conditions under which they can and cannot appear?

While we cannot properly address this question within the confines of this paper, we will end by reframing the issue in a way that may be helpful in going forward. The monotonic proportional measure functions that we have proposed figure into the relative reading of (31) feature fractions whose denominators, the weights of Esme's ring and Seymour's necklace, are measurements of entities picked out by referring expressions in the sentence, viz. by *her ring* and *his necklace*. It seems plausible that monotonic proportional measure functions can become available for interpretation of *much* only when they can be anchored in this way to suitable referring expressions in the sentence, and hence that (32) lacks a relative reading because no suitable anchors are available there. We hope to explore this intuition in future work.

## References

- Bale, Alan, and David Barber. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26:217–252.
- Bale, Alan, and Bernhard Schwarz. 2019. Proportional readings of *many* and *few*: the case for an underspecified measure function. *Linguistics and Philosophy* URL <https://doi.org/10.1007/s10988-019-09284-5>.
- Bresnan, Joan W. 1973. Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4:275–343.
- Cresswell, Max J. 1976. The semantics of degree. In *Montague Grammar*, ed. Barbara H Partee, 261–292. New York: Academic Press.
- Hackl, Martin. 2000. Comparative quantifiers. Doctoral Dissertation, Massachusetts Institute of Technology.
- Heim, Irene. 2000. Degree operators and scope. In *Proceedings of SALT X*, ed. Brendan Jackson and Tanya Matthews, 40–64. Cornell University, Ithaca, NY: CLC Publications.
- Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9:67–110.
- Wellwood, Alexis. 2015. On the semantics of comparison across categories. *Linguistics and Philosophy* 38:67–101.

# A Subregular Bound on the Complexity of Lexical Quantifiers

Thomas Graf

Stony Brook University, Stony Brook, NY, USA  
mail@thomasgraf.net

## Abstract

Semantic automata were developed to compare the complexity of generalized quantifiers in terms of the string languages that describe their truth conditions. An important point that has gone unnoticed so far is that these string languages are remarkably simple for most quantifiers, in particular those that can be realized by a single lexical item. Whereas complex quantifiers such as *an even number of* correspond to specific regular languages, the lexical quantifiers *every*, *no*, *some* as well as numerals do not reach this level of complexity. Instead, they all stay close to the bottom of the so-called subregular hierarchy. What more, the class of tier-based strictly local languages provides a remarkably tight characterization of the class of lexical quantifiers. A significant number of recent publications have also argued for the central role of tier-based strict locality in phonology, morphology and syntax. This suggests that subregularity in general and tier-based strict locality in particular may be a unifying property of natural language across all its submodules.

## 1 Introduction

Generalized quantifiers (GQs) have been a fruitful topic for mathematical investigation since Barwise and Cooper (1981) and Keenan and Faltz (1985). Among the many techniques to study GQs (see Peters and Westerståhl, 2006), semantic automata (van Benthem, 1986; Clark, 2001; Steinert-Threlkeld and Icard, 2013) are noteworthy because their grounding in formal language theory provides a bridge to phonology and syntax, which have also been studied from this perspective. However, computational phonology and syntax have recently undergone a shift towards particularly weak classes of formal languages that belong to the *subregular hierarchy* (see Graf 2018, Heinz 2018, and references therein). In this paper, I argue that the subregular hierarchy also provides a fruitful perspective on GQs.

The semantic automata approach characterizes GQs via string languages that encode their truth conditions. It is known that many GQs have string languages that belong to the class of regular languages. This result can be strengthened significantly, though, as many of the most common quantifiers belong to the subregular class of *tier-based strictly local* string languages (TSL; Heinz et al., 2011), which plays a central role in phonology and syntax. This includes *every*, *some*, *no*, numerals, *not all*, and *all but three*, but not *most* or *half*. In addition, TSL provides a new perspective on a long-standing puzzle in morphosemantics: Across all languages, only a small number of GQs are ever realized as a single lexical item. These *lexical quantifiers* include *every*, *some*, *no*, and numerals, to the exclusion of *not all*, *all but three*, or *an even number of*. Why only some GQs can be lexicalized in this manner has remained mysterious. TSL establishes a marked difference between these two groups, a difference that is connected to the semantically fertile notion of monotonicity. Lowering the bound of GQs to TSL thus has two major payoffs. It establishes a computational parallel between semantics on the one hand and syntax and phonology on the other, and it explains why only some GQs can be realized as a single lexical item.

The paper is organized as follows. I start with a brief introduction to generalized quantifiers and how they can be represented as string languages (§2). This section can be skipped by readers who are already familiar with semantic automata (but they might still want to take a look at the table in (1)). After that, I discuss TSL and show that this class subsumes many simple GQs while excluding more complicated ones like *an even number of* (§3). I then use monotonicity to define an even more restricted subclass of TSL, and I show that this subclass provides a tight approximation of the class of GQs that can be spelled-out as a single lexical item.

## 2 Generalized Quantifiers as String Languages

A GQ is a function that maps one or more relations to truth values. In linguistics, the focus of research has been mostly on quantifiers of type  $\langle 1 \rangle$  and  $\langle 1, 1 \rangle$ . A type  $\langle 1 \rangle$  quantifier maps a unary relation, i.e. a set, to a truth value. Examples are *everybody* or *nobody*, but adverbs such as *always* may also be viewed as type  $\langle 1 \rangle$  quantifiers. Type  $\langle 1, 1 \rangle$  quantifiers, on the other hand, take two sets as input and return a truth value. The prototypical quantifiers of English belong to this category: *every*, *some*, and *no*. But among its members are also *most*, *an even number of*, and numerals like *five*, *at most five*, and *at least five*. In the sentence *at most five dragons are chain smokers*, the  $\langle 1, 1 \rangle$  quantifier *at most five* takes the set of dragons and the set of chain smokers as its first and second argument, respectively, and returns true iff the intersection of the two sets contains at most five elements. Due to the prominence of  $\langle 1, 1 \rangle$  quantifiers across natural languages, they are the central object of study in GQ theory (another reason is that type  $\langle 1 \rangle$  quantifiers may be regarded as a special case of type  $\langle 1, 1 \rangle$  quantifiers where the first argument is fixed). This paper will also limit itself to  $\langle 1, 1 \rangle$  quantifiers, focusing exclusively on those that are determiners.

Given a type  $\langle 1, 1 \rangle$  quantifier  $Q$  and sets  $A$  and  $B$ ,  $Q(A, B)$  will be either true ( $= 1$ ) or false ( $= 0$ ), depending on the definition of  $Q$ . For example,  $every(A, B)$  is true iff  $A \subseteq B$ . So *every cat is a mammal* is true because the set of cats is a subset of the set of mammals, wherefore  $every(cat, mammal)$  holds. On the other hand, *no cat is a mammal* is false because  $no(A, B)$  holds iff  $A \cap B = \emptyset$ . The truth value of a  $\langle 1, 1 \rangle$  quantifier  $Q$  is thus contingent on the set relations that hold between  $A$  and  $B$ . For natural language quantifiers, additional properties hold that ensure that the meaning of a quantifier can be represented solely in terms of which members of  $A$  are members of  $B$ . This is the key to measuring GQ complexity via string languages. The basic idea is to convert the relation between  $A$  and  $B$  into a string of 0s and 1s so that  $Q$  can be viewed as a formal language of such strings.

Since the idea of representing a GQ as a string language is fairly abstract, a certain amount of rigor is required to avoid misinterpretation. Given some arbitrary enumeration  $e$  of all the elements of  $A$ ,  $f_B^A$  is a total function that converts  $e$  into a string over the alphabet  $\{0, 1\}$ . Every element of  $A$  that also belongs to  $B$  is replaced by 1, all others by 0. So with  $A := \{a, b, c\}$  and  $B := \{a, d\}$ , the enumeration  $e := bac$  of  $A$  would be rewritten as 010 by  $f_B^A$ . For a more concrete example, consider  $f_B^A$  with  $B$  the set of all males and  $A$  the set of all US presidents up to and including 2019. This function produces a string of length 45 where every symbol is 1 because every member of the set of US presidents is also in the set of male individuals. If  $A$  is the set of all humans, on the other hand,  $f$  produces a string with approximately 7 billion symbols, more than half of which are 0 because only 49% of the current world population are male. The order of 0s and 1s in this string depends on how one enumerates  $A$ , which is ultimately immaterial as any enumeration is a valid choice.

For any two (countably infinite) sets  $A$  and  $B$ , then,  $f_B^A$  encodes the relation between the

two as a (usually non-unique and possibly infinite) string over 0 and 1. I will call this a *binary string over A and B*. A GQ  $Q$ , in turn, is equivalent to a language  $L(Q)$  of binary strings. Throughout this paper, I refer to such sets of binary strings as *quantifier languages*.<sup>1</sup>

**Definition 1** (Binary Strings). *Let  $A$  and  $B$  be two (countably infinite) sets. The total function  $f_B^A$  maps each enumeration  $e$  of  $A$  to a (possibly infinite) string of 0s and 1s:*

$$f_B^A(e) := \begin{cases} f_B^A(a) \cdot f_B^A(e') & \text{if } e = a \cdot e', \text{ and } a \in A, \text{ and } e' \text{ is not the empty string } \varepsilon \\ 1 & \text{if } e \in A \cap B \\ 0 & \text{otherwise} \end{cases}$$

Here  $\cdot$  denotes string concatenation. We call  $s$  a binary string of  $A$  under  $B$  iff there is some enumeration  $e$  of  $A$  such that  $s = f_B^A(e)$ .

**Definition 2** (Quantifier Language). *Let  $Q$  be a type  $\langle 1, 1 \rangle$  quantifier. Then its quantifier language  $L(Q)$  is the unique set such that for all sets  $A$  and  $B$  and (possibly infinite) binary string  $s$  of  $A$  under  $B$ , it holds that  $s \in L(Q)$  iff  $Q(A, B)$  is true.*

Consider once more the case of *every*, and recall that  $every(A, B)$  is true iff  $A \subseteq B$ . If  $A \subseteq B$ , then every binary string of  $A$  under  $B$  contains only 1s, and never any 0s. On the other hand,  $A \subsetneq B$  implies that there is at least some  $a \in A$  that is not a member of  $B$ , and this  $a$  will be rewritten as 0 by  $f_B^A$ . This shows that the binary strings that encode a subset relation between  $A$  and  $B$  are exactly those that contain only 1s. The quantifier language  $L(every)$  thus consists of all strings, and only those, that do not contain any 0s, and  $every(A, B)$  is true iff  $f_B^A$  produces a binary string of this form. The same line of reasoning shows that  $L(no)$  consists of all and only those strings that contain only 0, while  $L(some)$  consists of exactly those strings with at least one 1.

The quantifier languages of *every*, *no*, *some*, *not all*, and numerals can be succinctly described in terms of their numerosity requirements on 1 and 0. The left table in (1) shows how this is done, with  $|s|_0$  and  $|s|_1$  denoting the number of 0s and 1s in a binary string  $s$ , respectively. The right table shows that other GQs have string languages for which the conditions are trickier to state (keep in mind that proportional GQs are only well-defined if  $A$  is finite).

	Quantifier	Language		Quantifier	Language
(1)	every	$ s _0 = 0$		between $m$ and $n$	$m \leq  s _1 \leq n$
	no	$ s _1 = 0$		an even number	$ s _1 \bmod 2 = 0$
	some	$ s _1 \geq 1$		half	$ s _1 =  s _0$
	not all	$ s _0 \geq 1$		most	$ s _1 >  s _0$
	at least $n$	$ s _1 \geq n$		at least one third	$3 s _1 \geq  s _0 +  s _1$
	at most $n$	$ s _1 \leq n$			
	exactly $n$	$ s _1 = n$			

The fact that some GQs have simpler descriptions of their string languages than others already suggests that certain complexity differences exist between them. Formal language theory provides the tools to state this explicitly in computational terms.

<sup>1</sup>Semantic automata theory usually assumes that all binary strings are finite because automata only generate finite strings. But this is not a necessity. The theory of  $\omega$ -automata (Perrin and Pin, 2004), for instance, allows for languages with infinite strings. More importantly, the grammar-based specification adopted in this paper is completely agnostic about whether strings are finite or infinite, as long as they are countably infinite.



### 3 Subregular Quantifier Languages: TSL

We just saw how the semantic automata approach recasts GQs as sets of binary strings. This has made it possible to establish important complexity differences between quantifiers. But as we will see next, the semantic automata approach has largely relied on very coarse classes that can hide important complexity differences (§3.1). The subregular notion of tier-based strict locality refines the existing picture in an insightful manner. I first describe the class in intuitive terms (§3.2) and show that it covers a wide range of simple GQs (§3.3) while excluding more complex ones (§3.4). The formal definition is provided in §3.5.

#### 3.1 From Regular to Subregular

Building directly on the Chomsky hierarchy (Chomsky, 1956), work on semantic automata has largely been concerned with the question which quantifier languages belong to the weakest level of that hierarchy, the *regular string languages*. It turns out that almost all natural language quantifiers are regular, including: *every*, *no*, *some*, *not all*, *at least 5*, *between 5 and 17*, and *an even number*. But proportional quantifiers like *most*, *half*, or *at least one third* are not regular — they have *context-free* quantifier languages.

While this is an important finding, it does not fully line up with the intuitive split between simple and complex quantifier languages in (1). In particular, the bifurcation into regular and context-free GQs suggests that *an even number* is just as simple as *some* because both are regular, which seems counter-intuitive.<sup>2</sup> It also misses an important typological fact about GQs: across languages, *every*, *no*, *some*, and numerals are the only GQs with regular quantifier languages that can be spelled out as a single lexical item, whereas *not all* and *an even number* always require multi-word expressions (Paperno, 2011). This is unexpected if all these quantifiers are regular and hence equally complex. The Chomsky hierarchy therefore provides an incomplete picture at best.

However, it has long been known that the class of regular languages can be decomposed into an elaborate hierarchy of subclasses (Schützenberger, 1965; McNaughton and Papert, 1971; Pin, 1997). This *subregular hierarchy* has recently found numerous applications in phonology, morphology, and syntax (Graf, 2018; Heinz, 2018). The class of *tier-based strictly local* languages (TSL; Heinz et al., 2011) has turned out to be particularly useful in all three domains. Perhaps it isn't too surprising, then, that TSL can also address the issues raised above.

#### 3.2 Tier-Based Strictly Local String Languages

TSL is an intuitively appealing class that is directly inspired by the notion of autosegmental tiers in phonology (Goldsmith, 1976). First one identifies a set of symbols that serves as the tier alphabet  $T$ . Given a string  $s$ , every symbol in  $s$  that belongs to  $T$  projects onto a dedicated tier. This tier is then checked for well-formedness against a finite list  $S$  of forbidden substrings (or *n-grams*). A string is well-formed iff its tier contains no forbidden substrings.

The formal definition of TSL is relegated to §3.5 at the end of this section. Instead, let us work through a concrete example by showing that  $L(\textit{some})$  is TSL. Remember that  $L(\textit{some})$  consists of all binary strings that contain at least one 1. Hence we have  $0010 \in L(\textit{some})$  but  $0000 \notin L(\textit{some})$ . We can capture this with a TSL grammar that I) projects a tier containing all 1s but no 0s, and II) requires this “1-tier” to be non-empty. This is illustrated below, with

<sup>2</sup>It is common practice to further distinguish between simple and complex quantifier languages based on the size of the semantic automata that generate these languages. But this makes even more counter-intuitive predictions because the automaton for *an even number* is smaller than the automaton for *at least three*.

$\bowtie$  and  $\bowtie$  as dedicated symbols for the left and right string/tier edge, respectively. Using these symbols, we can rule out empty tiers by banning the substring  $\bowtie\bowtie$ .

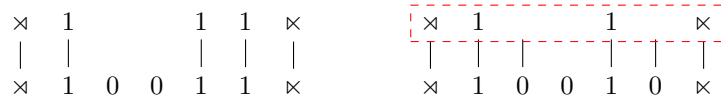


The left string is well-formed because  $\bowtie\bowtie$  is not a substring of the tier, thanks to the presence of the symbol 1. The right string, on the other hand, is illicit due to its tier lacking such a symbol and thus containing the forbidden bigram  $\bowtie\bowtie$ .

The TSL grammar for *some* can be extended to any numeral. For instance, replacing the bigram  $\bowtie\bowtie$  by the 4-gram 1111 will block all binary strings with four or more 1s, as illustrated below. In other words, this new grammar defines the quantifier language for *at most 3*.



For *at least 3*, one has to rule out tiers with less than three 1s via the following list of forbidden  $n$ -grams:  $\bowtie\bowtie$ ,  $\bowtie 1\bowtie$ ,  $\bowtie 11\bowtie$ . Any tier that contains at least one of those three substrings is ill-formed, and so is the binary string that the tier was projected from.



Adding the 4-gram 1111 to the list above yields a TSL grammar that generates  $L(\textit{exactly } 3)$ . If instead of 1111 we had added the 5-gram 11111, the resulting language would have been  $L(\textit{between } 3 \textit{ and } 4)$ .



We see, then, that TSL is powerful enough to handle *some* as well as numerals and related quantifiers.

### 3.3 Varying the Tier Alphabet

All the quantifiers in the previous section are modeled by TSL grammars with tier alphabet  $\{1\}$ . But  $\{0\}$ ,  $\{0, 1\}$ , and the empty set are also valid choices for the tier alphabet. By altering the tier alphabet, even more GQs can be described with TSL.

Changing the tier alphabet to  $\{0\}$  allows for the expression of some other generalized quantifiers.

	Forbidden Substrings	Quantifier
	$\bowtie\bowtie$	not all
(2)	$\bowtie\bowtie, 00$	all but one
	$\bowtie\bowtie, \bowtie 0 \bowtie, 000$	all but two
	⋮	



TSL grammars with an empty tier alphabet, on the other hand, are useless. An empty tier alphabet will produce an empty tier for every string, which means that the grammar cannot make any distinctions between strings.

This leaves us with a tier alphabet of  $\{0, 1\}$ . In this case, the tier is always identical to the binary string it is projected from. While this may seem useless, too, it is actually necessary to capture  $L(\text{every})$  and  $L(\text{no})$  with TSL — every symbol projects, and the tier must not contain the substring 0 for  $L(\text{every})$  or the substring 1 for  $L(\text{no})$ . Projecting only 0 or only 1 would not do in this case as every symbol in the string has to be inspected by the grammar. In sum, many GQs have TSL quantifier languages, although the tier alphabet may differ between them: *every*, *some*, *no*, *not all*, *all but 3*, and numerals.

### 3.4 Examples of Quantifiers that are not TSL

There are also many generalized quantifiers that are not TSL. Since every TSL language is also a regular language, proportional quantifiers are not TSL by virtue of not being regular. Moreover, some quantifiers are regular but not TSL. The most prominent example is *an even number*, for reasons that are easy to verify.

Every proposed TSL grammar for  $L(\text{an even number})$  must have  $\{1\}$  as its tier alphabet. Not projecting 1 makes it impossible to ensure that the binary string contains an even number of 1s, and projecting 0 only adds irrelevant material to the tier. Assume, then, that only 1 is projected. In that case, well-formed binary strings will have tiers of even length, whereas ill-formed binary strings have tiers of odd length. But a finite list of forbidden substrings cannot distinguish 1-tiers of even length from 1-tiers of odd length. The reasoning is as follows: In order to rule out a 1-tier of odd length, at least one of its substrings must be illicit. Since the list of forbidden substrings is finite, there is a longest forbidden substring  $s$  with  $k$  1s in a row. Now take some 1-tier of even length  $l > k + 2$  (we add 2 to account for  $\bowtie$  and  $\bowtie$ ). This tier should be well-formed, but as it contains  $s$  as a substring, it is mistakenly ruled out by the TSL grammar. Hence a TSL grammar for  $L(\text{an even number})$  either overgenerates by permitting some odd strings, or it forbids some even strings and thus undergenerates.

That TSL grammars cannot handle  $L(\text{an even number})$  is a welcome result. It shows that the step down from regular to TSL paints a more detailed picture of GQ complexity.

### 3.5 Formal Summary

With the intuition behind TSL firmly established, we can finally turn to the formal definition.

**Definition 3** (Tier-Based Strictly Local). *A tier-based strictly local grammar  $G$  over alphabet  $\Sigma$  is a pair  $\langle S, T \rangle$ , where  $T \subseteq \Sigma$  is a tier alphabet and  $S$  is a finite set of (finite) strings over  $T$ . We say that  $G$  is tier-based strictly  $n$ -local (TSL- $n$ ) iff the length of each string in  $S$  is at most  $n$ .*

*Given a string  $s$  over  $\Sigma$ , its  $T$ -tier is  $\bowtie T(s) \bowtie$ , where  $\bowtie, \bowtie \notin \Sigma$  are distinguished and  $T(s)$  is given by:*

$$T(s) := \begin{cases} T(u) \cdot T(s') & \text{if } u \in \Sigma \text{ and } s = u \cdot s' \\ s & \text{if } s \in T \\ \varepsilon & \text{otherwise} \end{cases}$$

*Here  $\varepsilon$  denotes the empty string — for all strings  $u$  over  $\Sigma$ ,  $u \cdot \varepsilon = \varepsilon \cdot u = u$ . The  $T$ -tier is well-formed iff none of its substrings are members of  $S$ . The language  $L(G)$  of strings generated by  $G$  contains all strings whose  $T$ -tier is well-formed, and only those. A string set is tier-based strictly local iff there is a tier-based strictly local grammar that generates it.*

Quantifier	String Condition	Complexity	Tier	Forbidden
<i>every</i>	$ s _0 = 0$	TSL	$\{0, 1\}$	$\{0\}$
<i>no</i>	$ s _1 = 0$	TSL	$\{0, 1\}$	$\{1\}$
<i>some</i>	$ s _1 \geq 1$	TSL	$\{1\}$	$\{\times \times\}$
<i>not all</i>	$ s _0 \geq 1$	TSL	$\{0\}$	$\{\times \times\}$
<i>(at least) n</i>	$ s _1 \geq n$	TSL	$\{1\}$	$\{\times 1^k \times\}_{k < n}$
<i>(at most) n</i>	$ s _1 \leq n$	TSL	$\{1\}$	$\{1^{n+1}\}$
<i>(exactly) n</i>	$ s _1 = n$	TSL	$\{1\}$	$\{1^{n+1}, \times 1^k \times\}_{k < n}$
<i>between m and n</i>	$m \leq  s _1 \leq n$	TSL	$\{1\}$	$\{1^{n+1}, \times 1^k \times\}_{k < m}$
<i>all but n</i>	$ s _0 = n$	TSL	$\{0\}$	$\{0^{n+1}, \times 0^k \times\}_{k < n}$
<i>an even number</i>	$ s _1 = 2n, n \geq 0$	regular		
<i>half</i>	$ s _1 =  s _0$	context-free		
<i>most</i>	$ s _1 >  s _0$	context-free		
<i>at least one third</i>	$3 s _1 \geq  s _0 +  s _1$	context-free		

Table 1: String complexity for some generalized quantifiers in English

We now have a more fine-grained view of quantifier complexity, which is summarized in Tab. 2. The table makes it very clear that TSL encompasses most of the quantifiers that linguists are inclined to consider natural. TSL gives us the desired complexity difference between those quantifiers and *an even number*. However, TSL itself still does not explain the typological fact that among all GQ with regular quantifier languages, only some can ever be spelled out as a single lexical item. In the next section, I show that TSL does in fact provide an answer to this puzzle once we look more carefully at the tier projection mechanism.

## 4 Lexical Quantifiers are Monotonic TSL

Our discussion so far has yielded a more refined classification of type  $\langle 1, 1 \rangle$  quantifiers according to their automata-theoretic complexity. While this is a novel result, it is not particularly useful in and of itself. That so many GQs are TSL is intriguing considering that this class is also commonly encountered in phonology, morphology, and syntax. But that still leaves us without any concrete empirical ramifications. In particular, there still seems to be no complexity difference between lexical quantifiers like *every* and multi-word quantifiers like *not all*. Why, then, is it that there is no known language that realizes *not all* as a single lexical item? The answer may lie in a property that is all too familiar to semanticists: monotonicity.

### 4.1 Monotonicity of Tier Alphabets

Consider once more the TSL quantifier languages in Tab. 1, and pay close attention to their tier alphabet. All lexical quantifiers use either  $\{1\}$  or  $\{0, 1\}$ , but never  $\{0\}$ . This sets them apart from quantifiers like *not all* and *all but 3*, which only project 0. The difference between the tier alphabets  $\{1\}$  and  $\{0, 1\}$  on the one hand and  $\{0\}$  on the other is that the former's characteristic functions are *monotonically increasing*.

**Definition 4.** Let  $A$  and  $B$  be two sets that are partially ordered by  $\leq_A$  and  $\leq_B$ , respectively. Then a function  $f$  from  $A$  to  $B$  is *monotonically increasing* iff  $x \leq_A y$  implies  $f(x) \leq_B f(y)$  for all  $x, y \in A$ .

Each tier alphabet  $T$  can be regarded as a characteristic function from the string alphabet  $\Sigma$  into the Boolean algebra  $\mathbf{2}$  of truth values, with  $0 < 1$ . But in the case of quantifier languages, the only alphabet symbols are 0 and 1, which are truth values. Consequently, the characteristic function of  $T$  is a mapping from  $\mathbf{2}$  to  $\mathbf{2}$ . The tier alphabets  $\{1\}$  and  $\{0, 1\}$  both correspond to monotonically increasing functions. With  $\{1\}$ ,  $0 \leq 1$  and  $f(0) = 0 \leq 1 = f(1)$ . With  $\{0, 1\}$ ,  $f(0) = 1 \leq 1 = f(1)$ . But with  $\{0\}$ , we have  $0 \leq 1$  yet  $f(0) = 1 \not\leq 0 = f(1)$ . Once we distinguish between arbitrary TSL languages and those with a monotonically increasing quantifier language, we get a much tighter approximation of the class of GQs that can be expressed by a single lexical item.

**Definition 5.** *A quantifier language  $L$  is monotonic TSL iff there is TSL grammar  $G := \langle S, T \rangle$  such that  $L = L(G)$  and the characteristic function of  $T$  is a monotonically increasing function from  $\mathbf{2}$  to  $\mathbf{2}$ .*

monotonic TSL	<	TSL	<	regular	<	context-free
<i>every</i>		<i>not all</i>		<i>an even number</i>		<i>most</i>
<i>no</i>		<i>all but n</i>				<i>half</i>
<i>some</i>						<i>at least one third</i>
<i>(at most) n</i>						
<i>(at least) n</i>						
<i>(exactly) n</i>						
<i>between m and n</i>						

Table 2: Complexity of natural language quantifier languages

**Proposition 1.** *If a GQ has a regular quantifier language, then this quantifier can be spelled out as a single lexical item only if its quantifier language is monotonic TSL.*

The restriction to monotonic TSL is not only natural from the perspective of semantics, where monotonicity is known to play a significant role, it is also a formal counterpart to the intuition that set membership (encoded by 1) is a simpler concept than non-membership (encoded by 0).

## 4.2 Open Ends

Proposition 1 is a new semantic universal on the phonetic exponents of GQs, and it marks a significant step towards a deeper understanding of the morphosemantics of quantifiers. That said, there are several loose ends.

First of all, the proposition only identifies a necessary condition, but not a sufficient one. There are GQs with monotonic TSL quantifier languages that nonetheless are never realized as a single lexical item. Table 2 already provides *between m and n* as one example of this. But there are also much more abstract quantifiers in this class. For instance, one can generalize *at least n* to a grammar with tier alphabet  $\{0, 1\}$  that requires every string to be at least  $n$  symbols long. This corresponds to a quantifier  $Q$  such that  $Q(A, B)$  is true iff  $A$  has at least  $n$  members. Monotonic TSL is still too large a class, and additional restrictions will be needed to identify the subclass of monotonic TSL quantifier languages that only encode quantifiers that are i) possible natural language quantifiers, and ii) can be expressed by a single lexical item.

In addition, the restriction of Prop. 1 to GQs with regular string languages only serves the purpose of excluding *most* and *half*. That these two quantifiers can be realized as a single lexical

item is even more puzzling now that our subregular analysis has widened the complexity gap between them and the other GQs. There are multiple answers to this that need to be explored in detail. One solution is to follow Hackl (2009) in assuming that *most* (and possibly *half*) is actually not a single lexical item. But this would require a very specific definition of what it means to be a single lexical item, which poses the risk of circular reasoning. Alternatively, one could reassess the complexity of *most* and *half*: these two quantifiers are context-free only if one is not allowed to impose any order on the binary strings.

If quantifier languages need not be closed under permutation, then *most* and *half* are both monotonic TSL. Suppose that we have a monotonic TSL grammar with tier alphabet  $\{0, 1\}$  and the following constraints on the tier: I) do not start with 0, II) do not end with 0, III) do not contain the substring 00. This only permits binary strings where every 0 is preceded by at least one 1 and followed by at least one 1. In other words, it holds for each one of these strings that more than half of its symbols are 1s. This is exactly the definition of *most*. One could speculate that this reliance on string orders might be a formal analog to match-up heuristics that seem to be used in verification tasks involving *most*. It remains to be seen whether this view of *most* is viable.<sup>3</sup>

## 5 Conclusion

I have argued that the semantic automata approach stands to gain a lot from incorporating recent work on subregular complexity. Not only does it allow for a more fine-grained classification of quantifiers, it also establishes computational parallels between phonology, morphology, syntax, and semantics, and it brings us closer to an explanation why some quantifiers can be realized as a single lexical item while others are always multi-word constructions.

I hope that this paper will serve as the starting point for a larger enterprise of subregular semantics. The obvious next step would be adverbial quantifiers, but any domain is viable as long as it can be modeled in terms of string languages (cf. the string-based view of tense in Fernando 2015).

## Acknowledgments

The work reported in this paper was supported by the National Science Foundation under Grant No. BCS-1845344. I would like to thank the anonymous reviewers as well as the audiences at Stony Brook's Mathematical Linguistics Reading Group and Rutgers's Linguistics colloquium. They all provided valuable feedback that led to significant improvements to this paper.

## References

- Jon Barwise and Robin Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.
- Noam Chomsky. Three models for the description of language. *IRE Transactions on Information Theory*, 2:113–124, 1956.

---

<sup>3</sup>No comparable strategy is available for *an even number*, so this quantifier would still not be monotonic TSL. However, *not all* and *all but n* can be made monotonically TSL by requiring all 0s to occur before all 1s.

- Robin Clark. Generalized quantifiers and semantic automata: A tutorial. Ms., University of Pennsylvania, 2001.
- Tim Fernando. The semantics of tense and aspect: a finite-state perspective. In Shalom Lappin and Chris Fox, editors, *The Handbook of Contemporary Semantic Theory, Second edition*, pages 203–236. Wiley, 2015.
- John Goldsmith. *Autosegmental Phonology*. PhD thesis, MIT, 1976.
- Thomas Graf. Why movement comes for free once you have adjunction. In Daniel Edmiston, Marina Ermolaeva, Emre Hakküder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit, and Miachel Tabatowski, editors, *Proceedings of CLS 53*, pages 117–136, 2018.
- Martin Hackl. On the grammar and processing of proportional quantifiers: Most versus more than half. *Natural Language Semantics*, 17(1):63–98, 2009. doi: 10.1007/s11050-008-9039-x.
- Jeffrey Heinz. The computational nature of phonological generalizations. In Larry Hyman and Frank Plank, editors, *Phonological Typology, Phonetics and Phonology*, chapter 5, pages 126–195. Mouton De Gruyter, 2018.
- Jeffrey Heinz, Chetan Rawal, and Herbert G. Tanner. Tier-based strictly local constraints in phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 58–64, 2011. URL <http://www.aclweb.org/anthology/P11-2011>.
- Edward L. Keenan and Leonard M. Faltz. *Boolean Semantics for Natural Language*. Reidel, Dordrecht, 1985.
- Robert McNaughton and Seymour Papert. *Counter-Free Automata*. MIT Press, Cambridge, MA, 1971.
- Denis Paperno. Learnable classes of natural language quantifiers: Two perspectives. Ms., UCLA, 2011. URL [http://paperno.bol.ucla.edu/q\\_learning.pdf](http://paperno.bol.ucla.edu/q_learning.pdf).
- Dominique Perrin and Jean-Éric Pin. *Infinite Words. Automata, Semigroups, Logic and Games*. Elsevier, Amsterdam, 2004.
- Stanley Peters and Dag Westerståhl. *Quantifiers in Language and Logic*. Oxford University Press, 2006.
- Jean-Eric Pin. Syntactic semigroups. In *Handbook of Language Theory*, pages 679–764. Springer, Berlin, 1997.
- M.P. Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.
- Shane Steinert-Threlkeld and Thomas F. Icard. Iterating semantic automata. *Linguistics and Philosophy*, 36(2):151–173, 2013. ISSN 1573-0549. doi: 10.1007/s10988-013-9132-6.
- Johan van Benthem. Semantic automata. In *Essays in Logical Semantics*, pages 151–176. Springer, Dordrecht, 1986. ISBN 978-94-009-4540-1. doi: 10.1007/978-94-009-4540-1\_8.

# Semantic Universals of Intonation and Particles \*

Yurie Hara<sup>1</sup> and Mengxi Yuan<sup>2</sup>

<sup>1</sup> Hokkaido University, Sapporo, Hokkaido, Japan  
hara@imc.hokkaido.ac.jp

<sup>2</sup> Jinan University, Guangzhou, Guangdong, China  
mengxiyuan49@hotmail.com

## Abstract

This paper proposes a new composition rule for discourse particles and prosodic morphemes that paratactically-associate with the main text. Furthermore, the data and analyses support the framework of inquisitive semantics since the morphemes in question can often embed both declarative and interrogative clauses.

## 1 Introduction

Discourse particles and prosodic morphemes often give rise to secondary meanings in addition to the meanings computed from the main text that they attach to. For instance, in Osaka Japanese, when a *wh*-question is uttered with a sentence-final particle *nen* with final fall ‘↓’ as in (1), the sentence seems to express two meanings. One is a plain question ‘What are you going to eat?’ and the other is the speaker’s irritation:

- (1) nani taberu nen↓  
what eat NEN  
‘What are you going to eat?!’ (You have to decide now!)

In the literature on the interpretation of prosodic morphemes (Bartels, 1999; Gunlogson, 2003) and discourse particles, it has been tacitly assumed that the morpheme/particle is somehow attached to the entire sentence and projects an expressive meaning independent of the meaning of the host sentence. This paper offers a more concrete compositional analysis of prosody and particles by introducing a new composition rule that instructs how to interpret paratactically-associated expressive morphemes.

Another hallmark of prosodic morphemes and particles is that they can often attach to both declarative and interrogative clauses. As an illustration, the same Osaka Japanese *nen*↓ can be attached to a declarative as in (2).

- (2) konban furansu ryoori taberu nen↓.  
tonight France cuisine eat NEN  
‘I’ll eat French cuisine tonight.’

The linguistic data and analyses offered in the current paper provide new evidence for the framework of inquisitive semantics (Ciardelli et al., 2019), which can deal with declaratives and interrogatives uniformly as a set of propositions.

---

\*This project is supported by JSPS Kiban (C) “Semantic-Pragmatic Interfaces at Left Periphery: a neuroscientific approach” (18K00589) awarded to the first author.

## 2 Proposal and Theoretical Background

This section presents two theoretical frameworks that are crucial to the semantic analysis of particles and intonation in question. First, we present a new type system for expressives, namely  $\mathcal{L}_{CI}^{+S,PA}$ , which enables us to compute paratactically associated expressions. Second, the framework of inquisitive semantics is briefly introduced to see how declaratives and interrogatives are given the same semantic type as a set of propositions.

### 2.1 Syntax and Composition of Paratactic Association, $\mathcal{L}_{CI}^{+S,PA}$

We propose that discourse particles and intonational morphemes are paratactically associated (Lyons, 1977; Bartels, 1999) to the main sentence. Syntactically, a prosodic morpheme or particle  $\beta$  is paratactically associated (indicated by ‘ $\otimes$ ’) to the head  $\alpha$  of the root clause, as depicted in (3).

- (3) Syntactic representation of paratactic association



Meanings that arise from intonation and particles are often analyzed as expressives or conventional implicatures (Potts, 2005b,a; Hara, 2006; Potts, 2012; McCready, 2008, among others). To assign a composition rule that corresponds to the structure in (3), this paper augments McCready’s (2010)  $\mathcal{L}_{CI}^{+S}$  type system for conventional implicatures, since the behaviors of linguistic items discussed in the current paper are different from that of expressive expressions discussed in Potts (2005b) in several respects. For example, the Japanese sentence-final auxiliary *daroo* only projects the expressive content but no at-issue content. The composition rule for expressives/conventional implicatures proposed by Potts’ (2005b), CI APPLICATION, consists of two functional applications, one which returns an expressive meaning  $\alpha(\beta) : \tau^c$  and the other which is an identity function that returns at-issue content  $\beta : \sigma^a$ :

- (4) CI APPLICATION
- $$\begin{array}{c} \beta : \sigma^a \bullet \alpha(\beta) : \tau^c \\ \swarrow \quad \searrow \\ \alpha : \langle \sigma^a, \tau^c \rangle \quad \beta : \sigma^a \end{array}$$

If we employed CI APPLICATION to *daroo* and a sentence it attaches to, it would return an illicit interpretation where the expressive content expresses a weaker meaning of the at-issue content, i.e., ‘*p* and probably *p*’.<sup>1</sup>

Thus, we adopt and modify McCready’s (2010)  $\mathcal{L}_{CI}^{+S}$  to give semantics to the structure proposed in (3).  $\mathcal{L}_{CI}^{+S}$  is an extension of Potts’ (2005b)  $\mathcal{L}_{CI}$  obtained by adding *shunting types* to the system. Expressions with shunting types shunt the meaning tier from at-issue to expressive, thereby generate expressive contents only without yielding at-issue ones. More concretely, when the function is of shunting type then the following rule is used instead of CI APPLICATION.

<sup>1</sup>See Hara (2006) for more discussions.

- (5) SHUNTING-TYPE FUNCTIONAL APPLICATION (McCready's (2010) R7)

$$\begin{array}{c} \alpha(\beta) : \tau^s \\ \swarrow \quad \searrow \\ \alpha : \langle \sigma^a, \tau^s \rangle \quad \beta : \sigma^a \end{array}$$

Now, we propose a new system  $\mathcal{L}_{CI}^{+S,PA}$ , which is obtained by adding the following type specification (6) and composition rule (7), PARATACTIC ASSOCIATION, to  $\mathcal{L}_{CI}^{+S}$ .<sup>2</sup>

- (6) A shunting product type
- 
- If
- $\sigma$
- and
- $\tau$
- are shunting types for
- $\mathcal{L}_{CI}^{+S,PA}$
- , then
- $\sigma \times \tau$
- is a shunting product type for
- $\mathcal{L}_{CI}^{+S,PA}$
- .

- (7) PARATACTIC ASSOCIATION

$$\begin{array}{c} \lambda\chi.\alpha(\chi)\blacklozenge\beta(\chi) : \langle \sigma, \tau \times \nu \rangle \\ \swarrow \quad \searrow \\ \lambda\chi.\alpha(\chi) : \langle \sigma, \tau \rangle \quad \lambda\chi.\beta(\chi) : \langle \sigma, \nu \rangle \end{array}$$

The PARATACTIC ASSOCIATION (7) merges two functions into one by abstracting over the argument type of the two functions ( $\blacklozenge$  is a metalogical operator that combines expressions of different types). The resulting function,  $\lambda\chi.\alpha(\chi)\blacklozenge\beta(\chi)$ , is combined with an at-issue expression  $\chi$  of type  $\sigma^a$  by McCready's SHUNTING-TYPE FUNCTIONAL APPLICATION (5) and outputs a pair of shunting-type expressions  $\alpha(\chi)\blacklozenge\beta(\chi)$  of type  $\tau^s \times \nu^s$ .

In summary, discourse particles and intonational morphemes that are paratactically associated to the main sentence are semantically composed by the PARATACTIC ASSOCIATION (7). The expression that results from the composition is a pair of shunting-type expressions.

## 2.2 Uniform treatment of declaratives and interrogatives

As we will see in section 3, many particles and prosodic morphemes can attach to both declarative and interrogative sentences. Inquisitive semantics is a suitable framework to analyze these items because if declaratives and interrogatives have the same semantic type, these items that embed them do not need to be ambiguously defined.

In inquisitive semantics, both declarative and interrogative sentences are treated as issues, which are downward closed sets of propositions, which in turn are sets of possible worlds:

- (8) a. A proposition
- $p$
- is a set of possible worlds, i.e.,
- $p \subseteq \mathcal{W}$
- .
- 
- b. An
- issue*
- $I \subseteq \wp(\mathcal{W})$
- is a non-empty, downward closed set of propositions.

In other words, whether it is a declarative or an interrogative, a sentence is a set of sets of possible worlds of type  $\langle\langle s, t \rangle, t\rangle$  (abbreviated as  $T$  in the following to avoid clutter).

To semantically distinguish declaratives and interrogatives, the notion of possibilities is introduced. The possibilities for a sentence  $\varphi$  are the maximal propositions in  $\llbracket \varphi \rrbracket$ :

- (9) POSSIBILITY(
- $\varphi$
- ) :=
- $\{p \mid p \in \llbracket \varphi \rrbracket \text{ and there is no } q \in \llbracket \varphi \rrbracket \text{ such that } p \subset q\}$
- .

In case of a declarative clause, the set only contains a single maximal element, i.e., it is a singleton set,  $|\text{POSSIBILITY}(\varphi)| = 1$ , while in case of an interrogative,  $|\text{POSSIBILITY}(\varphi)| \geq 2$ .

<sup>2</sup>See McCready (2010, 51-53) for the full type system of  $\mathcal{L}_{CI}^{+S}$ .



To illustrate, let us see how a disjunction, a polar interrogative and a *wh*-interrogative are semantically composed.<sup>3</sup> First of all, a simple declarative sentence such as *Marie drinks* is a downward closed set of propositions, written  $\{ |Marie\ drinks| \}^\downarrow$ .

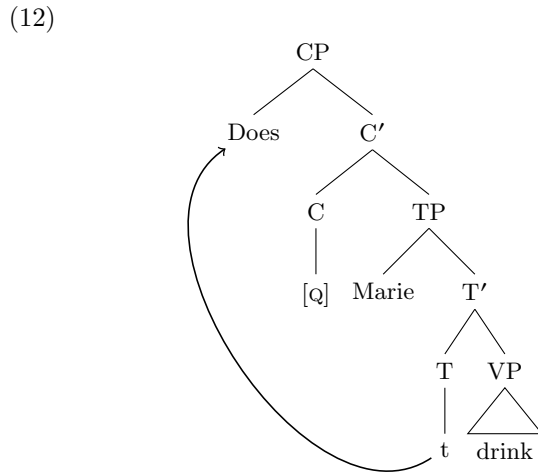
$$(10) \quad \llbracket \text{Marie drinks} \rrbracket = \{ p \mid \text{Marie drinks in every } w \in p \} = \{ |Marie\ drinks| \}^\downarrow = \llbracket \alpha \rrbracket$$

In the following illustrations, we use  $\llbracket \alpha \rrbracket$  for a denotation of *Marie drinks*.

Second, to compose a disjunction sentence like *Marie drinks or Bill eats*, we take union of two downward closed sets of propositions. Thus, the disjunction sentence is also a downward closed set of propositions. The set has two maximal propositions, ‘Marie drinks’ and ‘Bill eats’.

$$(11) \quad \llbracket \text{Marie drinks or Bill eats} \rrbracket = \llbracket \alpha \rrbracket \cup \llbracket \beta \rrbracket$$

Third, a polar interrogative is obtained by combining a declarative sentence with the question feature [Q]. In English, [Q] is realized by the auxiliary at Spec CP moved by the Subject-Aux inversion as in (12).



We assume that [Q] is an interrogative operator  $\langle ? \rangle$  proposed by [Roelofsen & Farkas \(2015\)](#). In order to define the semantics of  $\langle ? \rangle$ , the semantics of sentential negation needs to be defined as follows:

$$(13) \quad \llbracket \neg \varphi \rrbracket := \{ p \mid p \cap q = \emptyset \text{ for all } q \in \llbracket \varphi \rrbracket \}$$

Following [Roelofsen & Farkas \(2015\)](#),  $\langle ? \rangle$  is defined conditional on the status of its sister sentence. If its sister sentence  $\varphi$  is a declarative, that is, it is a singleton set of propositions,  $\langle ? \rangle$  takes a union of  $\llbracket \varphi \rrbracket$  and  $\llbracket \neg \varphi \rrbracket$ . If  $\varphi$  is already an interrogative sentence, i.e., contains multiple maximal propositions, it returns the same interrogative sentence.

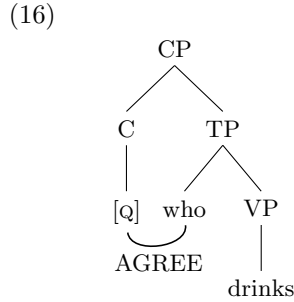
$$(14) \quad \begin{array}{l} \text{a. } \llbracket \langle ? \rangle \rrbracket \in D_{\langle T, T \rangle} \\ \text{b. } \llbracket \langle ? \rangle \varphi \rrbracket := \begin{cases} \llbracket \varphi \rrbracket \cup \llbracket \neg \varphi \rrbracket, & \text{if } |\text{POSSIBILITY}(\varphi)| = 1 \\ \llbracket \varphi \rrbracket, & \text{if } |\text{POSSIBILITY}(\varphi)| \geq 2 \end{cases} \end{array}$$

<sup>3</sup>See [Ciardelli et al. \(2017\)](#) for the fully compositional system for inquisitive semantics.

Thus, the polar interrogative, *Does Mary drink?* is also a union of two downward closed sets of propositions:

$$(15) \quad \llbracket \text{Does Mary drink} \rrbracket = \llbracket \alpha \rrbracket \cup \llbracket \neg \alpha \rrbracket$$

Finally, we assume that a *wh*-interrogative has the following structure in (16). The *wh*-pronoun agrees with [Q] at C.



The *wh*-clause denotes a downward closed set of propositions as in (17-a). This set then combines with ⟨?⟩ but it is not a singleton set so it returns the same set as in (17-b).

- (17) a.  $\llbracket \text{who drinks} \rrbracket = \{p \mid \exists x \in D_e . x \text{ is human} \ \& \ x \text{ drinks in every } w \in p\}$   
 b.  $\llbracket \langle ? \rangle \text{who drinks} \rrbracket = \{p \mid \exists x \in D_e . x \text{ is human} \ \& \ x \text{ drinks in every } w \in p\}$

In short, in inquisitive semantics, both declarative and interrogative clauses are issues, i.e., downward closed sets of propositions of type  $\langle \langle s, t \rangle, t \rangle = T$ .

### 2.3 Interim Summary

We have presented two frameworks necessary to analyze the semantics of discourse particles and prosodic morphemes. We first have proposed a new type system  $\mathcal{L}_{CI}^{+S,PA}$  which enables the semantic composition to output a pair of shunting-type expressives. Second, we have sketched how declaratives and interrogatives are uniformly treated as downward closed sets of propositions in inquisitive semantics. In other words, both have the same semantic type,  $\langle \langle s, t \rangle, t \rangle = T$ .

## 3 Deriving the interpretations

This section shows how the two systems introduced in the previous section can derive the meanings that arise from particles and prosodic morphemes.

### 3.1 Osaka Japanese *nen*↓

Osaka Japanese has a sentence-final particle *nen*↓ which has to be uttered with falling tone L%<sup>4</sup> (There is a phonological variant *en* after the past-tense morpheme *d/t* as in (21)). Hara & Kinuhata (2012) claim that *(n)en*↓ is an assertion marker since the implicit subject of (2), repeated here as (18) has to be the speaker and rendering (18) into a *yes-no* question by attaching a rising intonation (↑/LH%) results in ungrammaticality as in (19).

<sup>4</sup>See Ikeda (2001) for ToBI labelling of prosody in Osaka Japanese.

- (18) konban furansu ryoori taberu nen↓.  
 tonight France cuisine eat NEN  
 ‘I’ll eat French cuisine tonight.’
- (19) \*konban furansu ryoori taberu nen↑.  
 tonight France cuisine eat NEN  
 Intended: ‘Will you eat French cuisine tonight?’

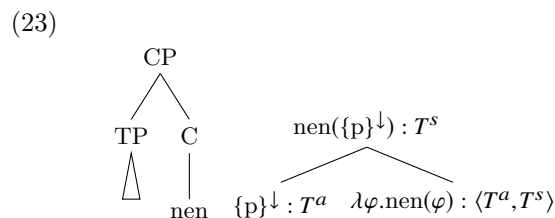
Interestingly, however, *nen*↓ can be attached to *wh*-interrogatives (though they still need to be uttered with falling intonation) and the constructions have emotive/discourse effects. In uttering (1), repeated here as (20), the speaker sounds irritated after waiting for the addressee to decide for a long time (n.b., it is still an information-seeking question). (21) can only be interpreted as a rhetorical question.

- (20) nani taberu nen↓  
 what eat NEN  
 ‘What are you going to eat?!’ (I’ve waited enough!)
- (21) dare-ga anta sodate-t-en↓  
 who-NOM you raise-PAST-NEN  
 ‘Who raised you up?!’ (Obviously, I did.)

To account for the data, we make two proposals: 1. *Nen*↓ is a complex lexical entry which is composed of phonemic segments /nen/ and prosodic segment (L%/↓). (In other words, *nen*↑ does not exist in the Osaka Japanese lexicon, hence (19) is ungrammatical.) 2. *Nen*↓ is an expressive morpheme which takes an at-issue set of propositions ( $T^a$ ) and returns an expressive set of propositions ( $T^s$ ), which denotes that one of the propositions in the set is true:

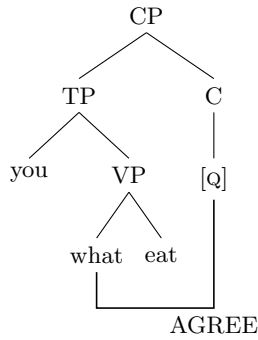
- (22) a.  $\llbracket \text{nen } \downarrow \rrbracket \in D_{\langle T^a, T^s \rangle}$   
 b.  $\llbracket \varphi \text{ nen } \downarrow \rrbracket := \{p\}$  for some  $q \in \llbracket \varphi \rrbracket : w \in q$  for every  $w \in p$

Thus, when  $(n)en$ ↓ attaches to a declarative as in (2), its argument is a downward closed set which contains a single maximal proposition  $p$ ,  $\{p\}^\downarrow$ . Thus, it simply asserts that the embedded proposition is true as depicted in (23).



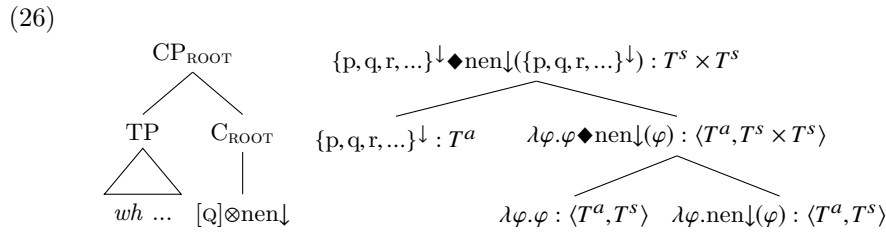
Turning to *wh*-interrogatives with  $(n)en$ ↓, as discussed above, a *wh*-pronoun agrees with a question feature [Q] at C:

- (24)



Furthermore, when [Q] occupies the root C, it renders an at-issue interrogative to an expressive one (25). The syntactic and composition trees of (20) are given in (26). *(N)en*↓ paratactically associates with this [Q], therefore the two expressive morphemes are combined by PARATACTIC ASSOCIATION (7), which yields a function that takes an at-issue meaning and returns a pair of expressive meanings,  $\{p, q, r, \dots\}^\downarrow \blacklozenge \text{nen} \downarrow (\{p, q, r, \dots\}^\downarrow)$ . Thus, it projects a question meaning and at the same time asserts that at least one of the propositions denoted by the interrogative clause is true. In (20), therefore, the speaker is urging the addressee to answer the question by asserting that one of the answers is true. In (21), the speaker knows which answer is true.

- (25) a.  $\llbracket [Q_{\text{ROOT}}] \rrbracket \in D_{\langle T^a, T^s \rangle}$
- b.  $\llbracket [Q_{\text{ROOT}}] \rrbracket = \lambda\varphi.\varphi$



### 3.2 Japanese rising *daroo*

Hara (2018) observed that a Japanese sentence-final auxiliary modal *daroo* has an intricate interaction pattern with clause types and prosody. In particular, a declarative that ends with a modal auxiliary *daroo* and a rising contour LH%/↑ yields an interpretation similar to a tag question as in (27).<sup>5</sup>

- (27) Marie-wa nomu daroo↑  
       Marie-TOP drink DAROO  
       ‘Marie drinks, right?’

Hara (2018, 2019) analyzes *daroo* as an expressive entertain modality  $E_{\text{Sp}}$  in inquisitive epistemic logic (Ciardelli & Roelofsen, 2015). When it is attached to a declarative, it indicates the speaker’s bias (28).

- (28) a.  $\llbracket \text{daroo} \rrbracket \in D_{\langle T^a, T^s \rangle}$

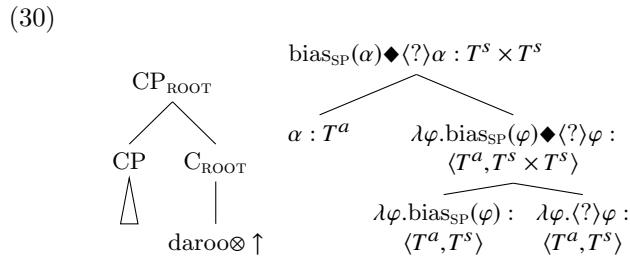
<sup>5</sup>See Venditti (2005) for the ToBI labelling for standard Japanese.

- b.  $\llbracket \varphi \text{ daroo} \rrbracket = E_{\text{SP}}(\varphi)$   
 If  $|\text{POSSIBILITY}(\varphi)| = 1$ ,  $\llbracket \varphi \text{ daroo} \rrbracket = \text{bias}_{\text{SP}}(\varphi)$

Furthermore,  $\uparrow$  is analyzed as an expressive polar question marker which denotes the interrogative operator defined above in (14).

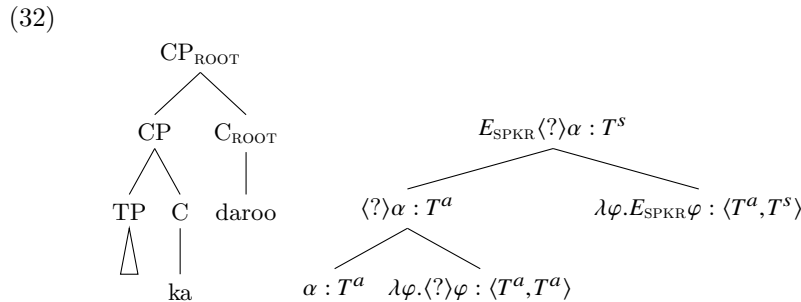
- (29) a.  $\llbracket \uparrow \rrbracket \in D_{\langle T^a, T^s \rangle}$
- b.  $\llbracket \uparrow \rrbracket = \lambda\varphi.\langle ? \rangle\varphi$

The syntactic and composition trees are given in (30). The two shunting-type morphemes are combined by PARATACTIC ASSOCIATION (7), which yields a function that takes an at-issue meaning and returns a pair of expressive meanings. As a result, (27) has two independent meanings, the speaker’s bias toward the single maximal proposition in  $\llbracket \alpha \rrbracket$  and her question  $\llbracket \langle ? \rangle \alpha \rrbracket = \llbracket \alpha \rrbracket \cup \llbracket \neg \alpha \rrbracket$ .



Furthermore, *daroo* can embed morphologically marked interrogatives, which supports the uniform approach for declaratives and interrogatives (see Hara, 2018, 2019).<sup>6</sup>

- (31) Marie-wa nomu daroo ka  
 Marie-TOP drink DAROO Q  
 ‘I wonder if Marie drinks.’



### 3.3 Final Fall in English and Mandarin

Zimmermann (2000) treats English Final Fall (H\*L-L%/↓) in disjunction declaratives like (33) as a closure operator which applies to a list in that it indicates that all and only items in the list have the “property in question”.

- (33) A: Which tube stations are one stop from Oxford Circus?

<sup>6</sup>Uegaki & Roelofsen (2018) give a similar analysis *daroo* using inquisitive epistemic logic, which makes different predictions. See Hara (2019) for the comparison.

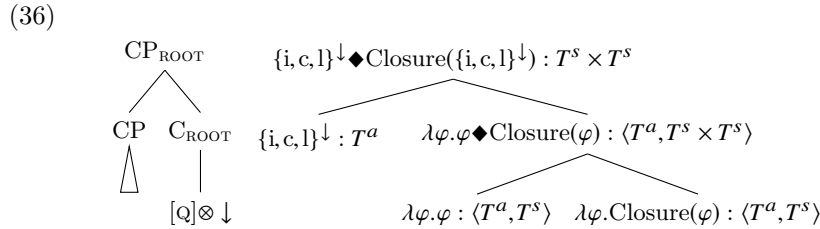
B: Piccadilly Circus, Bond Street, Tottenham Court Road, Green Park, Warren Street, Regent’s Park ↓

Biezma & Rawlins (2012) claim that the falling contour that accompanies alternative questions like (34) is the same closure operator as the one in (33), since they “offer unbiased choices” between the alternatives. We can derive the same interpretation by treating ↓ as an expressive closure operator defined in (35) based on Biezma & Rawlins (2012).

(34) Do you want iced tea, coffee, or lemonade↓

(35)  $\llbracket \downarrow \rrbracket = \llbracket \text{Closure} \rrbracket \in D_{\langle T^a, T^s \rangle}$ ,  
 $\llbracket \varphi \downarrow \rrbracket = \llbracket \text{Closure}(\varphi) \rrbracket := \{p \mid (\text{SalAlts} = \varphi) \text{ or } (\text{SalAlts} = \emptyset) \text{ in every } w \in p\}$ ,  
 where SalAlts is the set of propositional alternatives that are salient in the context of interpretation.

The [Q] (defined in (25)) and ↓/Closure are paratactically-associated as shown in (36). (34) raises a question  $\{i, c, l\}^\downarrow$  and expresses that all the alternatives are salient.



Mandarin A-not-A questions like (37) that end with Final Fall (L%/↓) seem to express a similar meaning, since they can be used only when the context is unbiased, i.e., both alternatives (p and ¬p) are equally salient (see also Yuan & Hara, 2013).

(37) Ni he-bu-he jiu↓  
 you drink-not-drink wine  
 ‘Do you drink wine or not?’

## 4 Conclusion

We have proposed a new type system  $\mathcal{L}_{CI}^{+S,PA}$  that includes the PARATACTIC ASSOCIATION rule.  $\mathcal{L}_{CI}^{+S,PA}$  can provide compositional analyses of expressive meanings that arise from prosodic morphemes and particles. Moreover, a wide range of cross-linguistic data show that prosodic morphemes and particles can embed both declaratives and interrogatives, which calls for a semantic platform that can uniformly deal with different clause types.

## References

Bartels, Christine. 1999. *The intonations of English statements and questions*. Garland Publishing.  
 Biezma, María & Kyle Rawlins. 2012. Responding to alternative and polar questions. *Linguistics and Philosophy* 35. 361–406.

- Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2019. *Inquisitive Semantics* Oxford Surveys in Semantics and Pragmatics. OUP Oxford. <https://books.google.co.jp/books?id=a8d-DwAAQBAJ>.
- Ciardelli, Ivano, Floris Roelofsen & Nadine Theiler. 2017. Composing alternatives. *Linguistics and Philosophy* 40. 1–36.
- Ciardelli, Ivano A. & Floris Roelofsen. 2015. Inquisitive dynamic epistemic logic. *Synthese* 192(6). 1643–1687.
- Gunlogson, Christine. 2003. *True to Form: Rising and Falling Declaratives as Questions in English*. New York: Routledge.
- Hara, Yurie. 2006. *Japanese Discourse Items at Interfaces*. Newark, DE: University of Delaware dissertation.
- Hara, Yurie. 2018. *Daroo* as an entertain modal: an inquisitive approach. In Shinichiro Fukuda, Mary Shin Kim, Mee-Jeong Park & Haruko Minegishi Cook (eds.), *Japanese/Korean Linguistics*, vol. 25, CSLI Publications.
- Hara, Yurie. 2019. *Daroo* ka↑: The interplay of deictic modality, sentence type, prosody and tier of meaning. Semantics Archive. <https://semanticsarchive.net/Archive/jA5ZWZhN/>.
- Hara, Yurie & Tomohide Kinuhata. 2012. Osaka Japanese *nen*: One-sided public belief and paratactic association. *Sprache und Datenverarbeitung: International Journal for Language Data Processing* 49–70.
- Ikeda, Youhei. 2001. *A proposal for ToBI labelling of Osaka dialect*. Nara Institute of Science and Technology MA thesis.
- Lyons, John. 1977. *Semantics*. Cambridge University Press.
- McCready, E. 2008. What man does. *Linguistics and Philosophy* 31(6). 671–724.
- McCready, E. 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3(8). 1–57. doi:10.3765/sp.3.8.
- Potts, Christopher. 2005a. Lexicalized intonational meaning. In Shigeto Kawahara (ed.), *UMOP 30: Papers on Prosody*, 129–146. Amherst, MA: GLSA.
- Potts, Christopher. 2005b. *The Logic of Conventional Implicatures* Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press. [Revised 2003 UC Santa Cruz PhD thesis].
- Potts, Christopher. 2012. Conventional implicature and expressive content. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 3, 2516–2536. Berlin: Mouton de Gruyter. This article was written in 2008.
- Roelofsen, Floris & Donka F. Farkas. 2015. Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91(2). 359–414.
- Uegaki, Wataru & Floris Roelofsen. 2018. Do modals take propositions or sets of propositions? evidence from Japanese *darou*. In *Proceedings of SALT 28*, .
- Venditti, J., Jennifer. 2005. The J-ToBI model of Japanese intonation. In Sun-Ah Jun (ed.), *Prosodic Typology*, 172–200. Oxford University Press.
- Yuan, Mengxi & Yurie Hara. 2013. Questioning and asserting at the same time: the L% tone in A-not-A questions. In Maria Aloni, Michael Franke & Floris Roelofsen (eds.), *Proceedings of the 19th Amsterdam Colloquium*, 265–272.
- Zimmermann, Thomas Ede. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8. 255–290.

# Explaining the Exceptive-Additive Ambiguity in Mandarin\*

Mingming Liu

Tsinghua University, Beijing, China  
markliu@scarletmail.rutgers.edu

## Abstract

As in many languages, the exceptive marker *chule* in Mandarin is ambiguous between ‘except’ and ‘in addition to’. The paper shows that the exceptive inference is an implicature while the additive one a presupposition. A unified analysis is sketched: *chule* encodes just subtraction; it removes something from a Roberts-style QUD [30]. Exceptive *chule* corresponds to removing an *individual* from the domain of the *wh* of the QUD, while additive *chule* removes a *proposition*. In the first case, reasoning about alternatives leads to the implicature [10, 16]. In the latter case, subtraction from a QUD is employed since the subtracted proposition is already known to be true, and not under discussion.

## 1 Introduction

In many languages one and the same item can be ambiguous between ‘except’ and ‘in addition to’, as has been reported in [25] for German *außer*, and recently in [38] for similar facts in Russian, Turkish, Hindi, Persian, Bulgarian, etc. Mandarin *chule* also exhibits this exceptive-additive ambiguity: when the matrix clause contains a universal quantifier as in (1)<sup>1</sup>, ‘*chule Lisi*’ adds an exceptive inference that Lisi didn’t pass, while it conveys additivity — Lisi passed — if the matrix clause contains an additive particle such as *ye* ‘also’ as in (2). Importantly, the additive particle is obligatory; without *ye*, (2) is infelicitous.

- |     |  |  |
|-----|--|--|
| (1) | <i>Chule Lisi, suoyou.ren dou guo.le</i><br>CHULE Lisi all.person all pass.ASP<br>‘Except for Lisi, everyone passed (the exam).’ | Exceptive inference<br>↪ Lisi didn’t pass. |
| (2) | <i>Chule Lisi, Zhangsan #(ye) guo.le</i><br>CHULE Lisi Zhangsan also pass.ASP<br>‘In addition to Lisi, Zhangsan also passed.’    | Additive inference<br>↪ Lisi passed.       |

This paper argues that the exceptive inference is an implicature while the additive one a presupposition. A unified proposal is sketched: *chule* (and arguably its cross-linguistic kin) encodes only subtraction [10, 16, 8]; it removes something from a Roberts-style Question under Discussion (QUD) [30]. Exceptive *chule* corresponds to removing an *individual* from the domain of the *wh* of the QUD, while additive *chule* removes a *proposition*. In the first case, reasoning about alternatives (quantificational statements with other domains) leads to the relevant implicature [10, 16, 8]. In the latter case, subtraction from a QUD is employed since the subtracted proposition is already known to be true, and not under discussion. Finally, the subtrahend QUD is indicated by focus in the matrix and the obligatory presence of additive particles in (2) is an *obligatory additive effect* [21, 32, 1, 3, 2].

\*Thanks to Dun Deng, Jianhua Hu, Haoze Li, Jo-wang Lin, Haihua Pan, Stanley Peters, Dag Westerståhl, Zhuang Wu, Beibei Xu, Mengxi Yuan, Linmin Zhang, and reviewers for AC 2019 for valuable discussion and comments. Support from the National Social Science Foundation of China (17CYY062) is acknowledged.

<sup>1</sup>In Mandarin, universally quantified NPs such as *suoyou.ren* ‘everyone’ in preverbal positions have to occur with the adverb *dou* ‘all/even’. See [22] for an explanation and [23] for a different perspective.



Intuitively, the proposal reflects two ways semantic subtraction could be useful in natural language: either to save the speaker from uttering a false quantificational statement, or to indicate the presuppositional status of a proposition within a larger QUD.

## 2 Exceptive Inference as Implicature

The mere fact that *chule* can express either exception or addition suggests that the exceptive inference might not be an inherent part of *chule*. In addition, while it has been debated whether the exceptive inference of English (connected) exceptives can be suspended/canceled (see for example distinct judgements for similar sentences in (3)), the exceptive inference of Mandarin *chule* can be easily suspended in ignorance contexts as in (4). There is even a grammaticalized way illustrated in (5) to do so, where *I don't know* immediately follows the *chule*-phrase. The strategy is fully general and systematically suspends *chule*'s exceptive inference.

- (3) a. Well, except for Dr. Samuels everybody has an alibi, inspector. Let's go see Dr. Samuels to find out if he's got one too. [17]  
 b. Well, we can't find Karl, but we've verified that everyone except Karl has an alibi, so let's find out whether he does too. [29]  
 c. #I don't know about Smith, but no corporate attorney but Smith wears cheap suits. [12]

- (4) *Wo bu zhidao Lisi guo.mei.guo;*  
 I NOT know Lisi pass.NOT.pass;  
*dan chule Lisi, suoyou.ren dou guo.le*  
 but CHULE Lisi, all.people all pass.ASP  
 'I don't know if Lisi has passed (the exam). But other than Lisi, everyone has passed.'

- (5) *Chule Lisi wo bu zhidao, suoyou.ren dou guo.le* ↗Lisi didn't pass.  
 CHULE Lisi I NEG know, all.people all pass.ASP  
 'I do not know if Lisi has passed, but everyone else has passed.'

The above suspension facts suggest the exceptive inference of *chule* is an implicature. However, different from ordinary conversational implicatures, it cannot be directly canceled, as (6) illustrates.

- (6) *Chule Lisi, suoyou.ren dou guo.le;*  
 CHULE Lisi, all.people all pass.ASP  
*# shijishang, Lisi ye guo.le.*  
 actually, Lisi also pass.ASP  
 'Except for Lisi, everyone passed; #actually, Lisi also passed.'

We suggest *chule*'s exceptive inference is an obligatory implicature [7, 24], similar to the plurality implicatures of plural NPs [34, 18]. The similarity is supported by the fact that the plurality inference (more than one), though not directly cancellable, can also be suspended in ignorance contexts, as illustrated in (7) [33].

- (7) a. [Context: You are inviting an old friend who you have not seen in years. you heard that he has a family now, but you have no idea how many children he has.]  
 You are welcome to bring your children. [33]  
 ↗ The speaker is certain that the addressee has more than one children.  
 b. #I just fed some cats. In fact, I fed only one. [6]

Based on the above facts, I conclude that the exceptive inference of *chule* is an implicature.

### 3 Additive Inference as Presupposition

Different from the exceptive inference, *chule*'s additive inference cannot be suspended. (8) shows that suspending the relevant additive inference in an ignorance context leads to contradiction, and (9) illustrates adding *I don't know* after an additive *chule*-phrase produces infelicity. The contrasts (4)/(8) and (5)/(9) clearly show while the exceptive inference is an implicature, the additive inference is not.

- (8) *Wo bu zhidao Lisi guo.mei.guo;*  
 I NOT know Lisi pass.NOT.pass;  
 # *chule Lisi, Zhangsan ye guo.le*  
 CHULE Lisi, Zhangsan also pass.ASP  
 'I don't know if Lisi has passed (the exam). #In addition to Lisi, Zhangsan also passed.'

- (9) # *Chule Lisi wo bu zhidao, Zhangsan ye guo.le*  
 CHULE Lisi I NEG know, Zhangsan also pass.ASP  
 'I do not know whether Lisi has passed, #Zhangsan also passed.'

Furthermore, the additive inference projects. (10) shows that it projects over polar questions, possibility modals, negation and conditional antecedents: all of the sentences in (10) imply that Lisi passed.

- (10) a. *Shi.bu.shi [chule Lisi, Zhangsan ye guo.le]?*  
 be.not.be CHULE Lisi Zhangsan also pass.ASP  
 'Is it the case that in addition to Lisi, Zhangsan also passed?'  $\rightsquigarrow$  Lisi passed.
- b. *Keneng [chule Lisi, Zhangsan ye guo.le].*  
 possibly CHULE Lisi Zhangsan also pass.ASP  
 'It is possible that in addition to Lisi, Zhangsan also passed.'  $\rightsquigarrow$  Lisi passed.
- c. *Bingfei [chule Lisi, Zhangsan ye guo.le].*  
 not CHULE Lisi Zhangsan also pass.ASP  
 'It is not the case that in addition to Lisi, Zhangsan also passed.'  $\rightsquigarrow$  Lisi passed.
- d. *Ruguo [chule Lisi, Zhangsan ye guo.le],*  
 if CHULE Lisi Zhangsan also pass.ASP  $\rightsquigarrow$  Lisi passed.  
*na women.ban jiu you liang.ge ren guo.le.*  
 then our.class then have two person pass.ASP  
 'If in addition to Lisi, Zhangsan also passed, then two persons from our class passed.'

The above facts suggest that *chule*'s additive inference is presuppositional. The claim is strengthened by the fact that *chule*'s additive inference cannot be used to answer questions. This is illustrated by (11A<sub>1</sub>): even though (11A<sub>1</sub>) implies that Lisi passed, the information cannot be used to answer *did Lisi pass?*, revealing its not-at-issue status. Crucially, (11A<sub>1</sub>) contrasts with (11A<sub>2</sub>) (a conjunction with *also* in the second conjunct), which *can* be used to answer (11Q). This is expected, since in (11A<sub>2</sub>) the presupposition triggered by *also* is filtered out by the first conjunct, and no longer a presupposition of the entire conjunction.

The contrast also suggests that the presuppositional status of the additive inference might not inherit directly from *ye* 'also', since that would indicate a parallel behavior between (11A<sub>1</sub>) and (11A<sub>2</sub>). The suggestion is confirmed by (11A<sub>3</sub>) in which the additive-*chule* embeds a

clause<sup>2</sup>. (11A<sub>3</sub>) cannot be used to answer (11Q) either, indicating the inference that Lisi passed is again presuppositional. Since the presupposition is not available without *chule* in (11A<sub>2</sub>), it must be triggered by *chule*.

(11) Q: Did Lisi pass?

A<sub>1</sub>: # *Chule Lisi, Zhangsan ye guo.le*  
 CHULE Lisi Zhangsan also pass.ASP  
 ‘In addition to Lisi, Zhangsan also passed.’  $\rightsquigarrow$  Lisi passed.

A<sub>2</sub>: *Lisi guo.le, Zhangsan ye guo.le*  
 Lisi pass.ASP, Zhangsan also pass.ASP  
 ‘Lisi passed; Zhangsan also passed.’  $\rightsquigarrow$  Lisi passed.

A<sub>3</sub>: # *Chule Lisi guo.le, Zhangsan ye guo.le*  
 CHULE Lisi pass.ASP, Zhangsan also pass.ASP  
 ‘In addition to Lisi, Zhangsan also passed.’  $\rightsquigarrow$  Lisi passed.

To summarize the empirical picture presented so far, the above facts show that while its exceptive inference is an implicature, *chule*’s additive inference is presuppositional. The rest of the paper offers an analysis where *chule* uniformly denotes subtraction from a QUD, and its two uses correspond to two ways (under different circumstances) subtraction from a QUD is employed by a speaker.

## 4 Exceptive-*chule* Subtracts Individuals

Adopting the QUD framework [30, 4], I assume that *chule*-sentences have *wh*-questions as their QUDs and exceptive-*chule* subtracts individuals from the domain of the *wh*-item. Recall that exceptive-*chule* co-occurs with a quantificational element (usually a universal<sup>3</sup>) in the matrix clause (1). The quantificational element inherits the domain argument, and has a subtracted set as its domain of quantification<sup>4</sup>. Next, *chule* requires the matrix quantifier to trigger alternatives, which then need exhaustification [7]. Finally, the extra exceptive inference is not inherent to *chule* but the result of exhaustification by an external operator EXH (12) (with the meaning of *only* as is standardly assumed in the literature) over the alternatives the matrix quantificational-sentence triggers [10, 16, 8]. The use of EXH puts the exceptive inference of *chule* into the same category as other types of implicatures, in particular obligatory implicatures [7, 24].

(12)  $\llbracket \text{EXH } S \rrbracket = 1$  iff  $\llbracket S \rrbracket = 1 \forall S' \in \text{ALT}(S) [\wedge \llbracket S' \rrbracket \not\subseteq \wedge \llbracket S \rrbracket \rightarrow \llbracket S' \rrbracket = 0]$   
 (Alternatives not entailed by *S* are false.)

To illustrate, the matrix universal in (13b) after *chule Lisi* quantifies over a subtracted set  $D \setminus \{\text{Lisi}\}$ , inherited from the subtracted domain of the *wh* in the QUD. Its alternatives are specified in (13d), and exhaustified by EXH as in (13c). The exhaustification delivers the exceptive inference: since everyone who is not Lisi in *D* passed and it is not the case that everyone in *D* passed, it must be the case that Lisi didn’t pass.

<sup>2</sup> *Chule* can systematically embed clauses on both its exceptive and additive uses. I remain neutral on whether there is a direct syntactic connection (such as ellipsis) between a clausal-*chule* sentence and its phrasal-*chule* counterpart (such as (11A<sub>3</sub>) and (1)), or *chule* is simply cross-categorical. See section 20 for some discussion.

<sup>3</sup> See [11] and especially [16] on how the co-occurrence restriction is captured in the alternatives-and-exhaustification framework adopted in the current section.

<sup>4</sup> What if the quantifier does not pick the subtracted set as its domain? Then exhaustification will be vacuous and this presumably violates the non-vacuity constraint that prohibits EXH from applying when it cannot negate any alternative [11, 16].

- (13) a. Who $^{D \setminus \{Lisi\}}$  passed? QUD subtracted by *chule*  
 b. *Chule Lisi, suoyou.ren* $^{D \setminus \{Lisi\}}$  *dou guo.le*  
 CHULE Lisi all.person all pass.ASP  
 ‘Except for Lisi, everyone passed.’  $\rightsquigarrow$  Lisi didn’t pass.  
 c. LF of the matrix: EXH(everyone $_{D \setminus \{Lisi\}}$  passed)  
 d. ALT of the matrix: {everyone $_{D \setminus \{Lisi\}}$  passed, everyone $_D$  passed}  
 e. Result of exh: ( $\llbracket \text{person} \rrbracket \cap D$ )  $\setminus$  {Lisi}  $\subseteq$   $\llbracket \text{passed} \rrbracket \wedge$  ( $\llbracket \text{person} \rrbracket \cap D$ )  $\not\subseteq$   $\llbracket \text{passed} \rrbracket$

The result is weak exception. Consider changing ‘*chule Lisi*’ into ‘*chule Lisi and Zhangsan*’; exhaustification over the alternatives {every $_{D \setminus \{Lisi, Zs\}}$  passed, every $_D$  passed} delivers ( $\llbracket \text{person} \rrbracket \cap D$ )  $\setminus$  {Lisi, Zs}  $\subseteq$   $\llbracket \text{passed} \rrbracket \wedge$  ( $\llbracket \text{person} \rrbracket \cap D$ )  $\not\subseteq$   $\llbracket \text{passed} \rrbracket$ , which is too weak to capture the intuition that both Lisi and Zhangsan failed: suppose everyone who is not Lisi passed while Lisi didn’t; then ( $\llbracket \text{person} \rrbracket \cap D$ )  $\setminus$  {Lisi, Zs}  $\subseteq$   $\llbracket \text{passed} \rrbracket \wedge$  ( $\llbracket \text{person} \rrbracket \cap D$ )  $\not\subseteq$   $\llbracket \text{passed} \rrbracket$  is true while the Mandarin sentence *Chule Lisi and Zhangsan, everyone passed* is intuitively false.

Weak Exception can be strengthened into Uniqueness of Exception by bringing in more alternatives [11, 16], such as the ones in {every $_{D \setminus \{Lisi, Zs\}}$  passed, every $_{D \setminus \{Lisi, John\}}$  passed, every $_D$  passed, ...} where different individuals are subtracted from  $D$ . It is easy to verify that exhaustification (negating the alternatives not weaker than the prejacent of EXH — ( $\llbracket \text{person} \rrbracket \cap D$ )  $\setminus$  {Lisi, Zs}  $\subseteq$   $\llbracket \text{passed} \rrbracket$ ) correctly strengthens the exceptive inference into *both Lisi and Zhangsan failed* (just consider the other two strictly stronger alternatives ( $\llbracket \text{person} \rrbracket \cap D$ )  $\setminus$  {Lisi}  $\subseteq$   $\llbracket \text{passed} \rrbracket$  and ( $\llbracket \text{person} \rrbracket \cap D$ )  $\setminus$  {Zs}  $\subseteq$   $\llbracket \text{passed} \rrbracket$ ). I speculate that the possibility of having different alternative sets (more alternatives  $\rightsquigarrow$  stronger exceptive inference) might explain variability of judgements in mixed scenarios with plural exceptive complements [27]<sup>5</sup>.

The proposal for exceptive-*chule* sketched above differs from most other theories of exceptives [17, 37, 26, 10, 16, 38, 28] in that *chule* does not directly subtract entities from the domain of the matrix quantifier. The departure is motivated by possible co-occurrences of *chule* with *qita* ‘other’ in the matrix<sup>6</sup>.

- (14) a. *chule Lisi*<sup>5</sup>, *qita*<sub>5</sub> *suoyou.ren dou guo.le*  
 CHULE Lisi<sup>5</sup>, other<sub>5</sub> all.person all pass.ASP  
 ‘Lisi<sup>5</sup> didn’t pass; all the other<sub>5</sub> people passed.’  
 b.  $\llbracket \text{qita}_i \rrbracket^g = \lambda P_{et} . \lambda x_e . \neg \text{OVERLAP}(g(i), x) \wedge P(x)$

(14a) is equivalent to (1) in meaning. With *qita* ‘other’ being a NP modifier removing individuals as in (14b) [19, 36] and anaphorically linked to Lisi in (14a), the restrictor of *all* in the matrix is  $\llbracket \text{person} \rrbracket \setminus \{Lisi\}$  and it is vacuous to subtract Lisi again from the set. This supports the proposal where the domain of the quantifier is *indirectly* modified via the domain of the *wh* of the QUD. This use of QUD furthermore connects the exceptive use of *chule* to its additive use, to which the next section turns.

<sup>5</sup>[27] reports that speakers tend (with variation among speakers) to judge the sentence ‘*no marble has a dot except/but the blue ones*’ true in a scenario with both red and blue marbles, where no red marble has a dot while only some blue marbles are dotless.

<sup>6</sup>English exceptives are reported to be compatible with *else* [15]: *Nobody (else) except/but Sarah is napping*. It is unclear how to compositionally analyze the combination of the two. See [35] for relevant discussion on *else* in questions.

## 5 Additive-*chule* Subtracts Propositions

I propose that additive-*chule* subtracts propositions, and this happens when a proposition within the QUD already belongs to the common ground (known to be true to the interlocutors), no longer under discussion and can be safely removed from the table [9]. In other words, the complement of *chule* ‘*Lisi*’ in (15) (repeated from (2)) actually stands for a proposition, just as the short answer ‘*John*’ stands for the proposition *John passed* when used as an answer to the question ‘*who passed?*’.

- (15) *Chule Lisi, Zhangsan<sub>F</sub> ye guo.le*  
 CHULE Lisi Zhangsan also pass.ASP  
 ‘In addition to Lisi, Zhangsan also passed.’  $\rightsquigarrow$  Lisi passed.

Next, to decide the polarity of the proposition that ‘*Lisi*’ stands for (*Lisi passed* instead of *Lisi didn’t pass*), we need to determine the shape of the Hamblin-Rooth QUD (*who passed?* vs. *who didn’t pass?*)<sup>7</sup>, which I suggest, assuming the matrix and the *chule*-phrase respond to the same QUD, is indicated by the focus structure of the matrix clause. Since the matrix clause is ‘*Zhangsan<sub>F</sub> also passed*’ instead of ‘*Zhangsan<sub>F</sub> also didn’t pass*’ in (15), the QUD must be *who passed?* as in (16), and thus the complement of *chule* in (15) ‘*Lisi*’ stands for the proposition *Lisi passed*.

- (16) [[*who passed?*]] = {that Lisi passed, that Zhangsan passed, that John passed, ... }

Finally, *chule* subtracts the proposition *Lisi passed* from the QUD, and indicates the proposition is already known to be true and can be safely removed from the table. We guarantee this by proposing the constraint in (17).

- (17) Constraint on Subtraction of Proposition from QUD  
 Subtraction of a proposition from a QUD happens only if the proposition is already known to be true by the interlocutors.

In this way, the analysis derives the correct additive inference of (15) and its presuppositional status ((15) can only be used in a context where the proposition *Lisi passed* is known to be true by the discourse participants).

The proposal predicts that *chule*’s additive inferences are sensitive to the focus structure of the matrix clause (see also [38]). This is indeed true, as is illustrated by (18): the two sentences carry different additive presuppositions, for their QUDs are different, which are further indicated by the focus markings in the matrix clauses. In (18a), ‘*Lisi*’ stands for the proposition *Mary introduced Lisi to John* since the QUD is *who did Mary introduce to John?*; In (18b), ‘*Lisi*’ stands for the proposition *Mary introduced Zhangsan to Lisi* since the QUD is *who did Mary introduce Zhangsan to?*

- (18) a. *Chule Lisi, Mali ye jieshao.le ZHANGSAN<sub>F</sub> gei Yuehan.*  
 CHULE Lisi Mary also introduce.ASP Zhangsan to John  
 ‘In addition to Lisi, Mary also introduced Zhangsan<sub>F</sub> to John.’  
 $\rightsquigarrow$  Mary introduced Lisi to John.

<sup>7</sup>The two questions denote different alternative sets in the Hamblin-Karttunen-Rooth framework [14, 20, 31], and this is a feature the current proposal makes use of to make sure that (15) has the correct additive presupposition. Note that this is impossible in a Groenendijk and Stokhof framework [13], where questions are treated as partitions over logical space, and the two questions (*who passed?* and *who didn’t pass?*) have the same denotation.

- b. *Chule Lisi, Mali ye jieshao.le Zhangsan gei YUEHAN<sub>F</sub>.*  
 CHULE Lisi Mary also introduce.ASP Zhangsan to John  
 ‘In addition to Lisi, Mary also introduced Zhangsan to John<sub>F</sub>.’  
 $\leadsto$  Mary introduced Zhangsan to Lisi.

Finally, additive particles such as *ye* ‘also’ are needed because of the independently attested *obligatory additive effect* which requires the presence of an additive particle whenever its presupposition is satisfied [21, 32, 5, 1, 3, 2], as illustrated in (19). Note that we do not need to decide between different explanations of the effect (*Maximize Presupposition* [5, 1] vs. *Obligatory Implicatures* [21, 32, 3]); both are compatible with the proposal of additive-*chule* sketched above.

(19) Dana went to a party. Lee went to a party(, #too).

For concreteness, I adopt the Maximize Presupposition approach: *too* is truth-conditionally vacuous but carries an additive presupposition that an alternative to its prejacent is also true; the presupposition is satisfied in its local context (the second clause in (19)), thus Maximize Presupposition favors *Lee went to a party too* over *Lee went to a party*, and *too* is obligatory. The same reasoning applies to (15): since the additive presupposition of *ye* ‘also’ is satisfied in its local context by the presupposition triggered by *chule*, it is obligatory by Maximize Presupposition.

Overall, the story captures the presuppositional status of the additive inference, and predicts that #*chule Lisi, Zhangsan passed* is bad, because (i) it cannot mean *Lisi didn’t pass, while Zhangsan passed* since the *chule*-phrase and the matrix correspond to different QUDs (*who didn’t pass?* for the former while *who passed?* for the latter), and (ii) it cannot express *Lisi passed, and Zhangsan passed* due to the absence of *also* and the resulting violation of Maximize Presupposition.

## 6 Phrasal vs. clausal *chule*

*Chule* can also take clauses as its complements (see also footnote 2), which our proposal is able to capture as well. In particular, the exceptive-*chule* in (20a) can still subtract individuals, if we assume *Lisi* is CT-marked as in [4] and subtraction from the domain of the *wh* in this case targets the referent of the contrastive topic. (20b) is bad, since subtracting *Lisi* from the domain of the matrix universal and subsequent exhaustification deliver *Lisi didn’t pass*, which contradicts the complement clause of *chule*. (20c) is good, for the complement of *chule* and the matrix respond to the same QUD *who passes?*, and the additive presupposition of *ye* is satisfied. (20d) on the other hand is bad, since (i) the complement of *chule* and the matrix don’t target the same QUD, and (ii) the presupposition of *ye* ‘also’ is not satisfied (just as ‘*John didn’t come. #Bill also came*’ is not acceptable).

- (20) a. *Chule Lisi<sub>CT</sub> mei guo, suoyou.ren dou guo.le*  
 CHULE Lisi not pass, all.person all pass.ASP  
 ‘Except Lisi didn’t pass, everyone passed.’  $\approx(1)$
- b. # *Chule Lisi<sub>CT</sub> guo.le, suoyou.ren<sub>D\{Lisi}</sub> dou guo.le*  
 CHULE Lisi pass.LE, all.person all pass.ASP  
 ‘#Except Lisi passed, everyone passed.’
- c. *Chule Lisi guo.le, Zhangsan ye guo.le*  
 CHULE Lisi pass.CL, Zhangsan also pass.ASP  
 ‘In addition to Lisi’s passing, Zhangsan also passed.’  $\approx(2)$

- d. # *Chule* *Lisi mei guo*, *Zhangsan ye guo.le*  
 CHULE Lisi not pass, Zhangsan also pass.ASP  
 ‘#In addition to Lisi’s not having passed, Zhangsan also passed.’

**Remaining issues** There are many unresolved issues, including a characterization and discussion of the *containment inference* of *chule* (in English, *every student except John came*  $\sim$  *John is a student*), and a proper formalization of the idea sketched in the paper. These are left for future work.

## References

- [1] Pascal Amsili and Claire Beyssade. Obligatory presuppositions in discourse. In P. Kuehnlein, A. Benz, and C. Sidner, editors, *Constraints in discourse 2: Pragmatics and Beyond*. John Benjamins, 2009.
- [2] Athulya Aravind and Martin Hackl. Against a unified treatment of obligatory presupposition trigger effects. In *Proceedings of Semantics and Linguistic Theory (SALT) 27*, 2017.
- [3] Nadine Bade. *Obligatory presupposition triggers in discourse-empirical foundations of the theories maximize presupposition and obligatory implicatures*. PhD thesis, Universitat Tubingen, 2016.
- [4] Daniel Büring. On d-trees, beans, and b-accent. *Linguistics and Philosophy*, 26(5):511–545, 2003.
- [5] Emmanuel Chemla. An epistemic step for antipresuppositions. *Journal of Semantics*, 25(2):141–173, 2008.
- [6] Gennaro Chierchia. Scalar implicatures and their interface with grammar. *Annual Review of Linguistics*, 2017.
- [7] Gennaro Chierchia, Danny Fox, and Benjamin Spector. Scalar implicature as a grammatical phenomenon. In Klaus von Stechow, Claudia Maienborn, and Paul Portner, editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 3 of *Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK)*, chapter 87, pages 2297–2331. de Gruyter, 2011.
- [8] Luka Crnić. A note on connected exceptives and approximatives. *Journal of Semantics*, 2018.
- [9] Donka F Farkas and Kim B Bruce. On reacting to assertions and polar questions. *Journal of Semantics*, 27(1):81–118, 2010.
- [10] Jon Gajewski. NPI *any* and connected exceptive phrases. *Natural Language Semantics*, 16:69–110, 2008.
- [11] Jon Gajewski. An analogy between a connected exceptive phrase and polarity items. In *Beyond ‘any’ and ‘ever’: New Explorations in Negative Polarity Sensitivity*. de Gruyter, 2013.
- [12] Ivan Garcia-Alvarez. *Generality and exception: a study in the semantics of exceptives*. PhD thesis, Stanford University, 2008.
- [13] Jeroen Groenendijk and Martin Stokhof. *Studies on the semantics of questions and the pragmatics of answers*. PhD thesis, University of Amsterdam, Amsterdam, Netherlands, 1984.
- [14] Charles J. Hamblin. Questions in Montague English. *Foundations of Language*, 10(1):41–53, 1973.
- [15] Jesse A. Harris. Who else but sarah? In *Connectedness: Papers in celebration of Sarah Van Wageningen*. UCLA Working Papers in Linguistics, 2014.
- [16] Aron Hirsch. An unexceptional semantics for expressions of exception. In *University of Pennsylvania Working Papers in Linguistics*, volume 22, 2016.
- [17] Jacob Hoeksema. The semantics of exception phrases. In Jaap van der Does and Jan van Eijk, editors, *Quantifiers, logic, and language*, pages 145–177. CSLI Publications, 1995.
- [18] Natalia Ivlieva. Obligatory implicatures and grammaticality. In *Logic, Language and Meaning: Revised Selected Papers of the 18th Amsterdam Colloquium*, 2011.

- [19] Hans Kamp. Presupposition computation and presupposition justification: One aspect of the interpretation of multi-sentence discourse. In *Meaning and the Dynamics of Interpretation*. Brill, 2013.
- [20] Lauri Karttunen. Syntax and semantics of questions. *Linguistics and Philosophy*, 1(1):3–44, 1977.
- [21] Manfred Krifka. Additive particles under stress. In *Semantics and Linguistic Theory (SALT) 8*, 1999.
- [22] Jo-Wang Lin. Distributivity in Chinese and its implications. *Natural Language Semantics*, 6(2):201–243, 1998.
- [23] Mingming Liu. Varieties of alternatives: Mandarin focus particles. *Linguistics and Philosophy*, 40(1):61–95, 2017.
- [24] Giorgio Magri. Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics*, 4:1–51, 2011.
- [25] Rolf Mayer. Anaphora and domain restriction. In *Text Understanding in LILOG*. Springer, Berlin, Heidelberg, 1991.
- [26] Friederike Moltmann. Exception sentences and polyadic quantification. *Linguistics and Philosophy*, 18(3):223–280, 1995.
- [27] Prerna Nadathur and Daniel Lassiter. An experimental look at the negative implications of exceptives. In *NELS 48*, 2017.
- [28] Stanley Peters and Dag Westerståhl. The semantics of exceptives. ms, 2019.
- [29] Christopher Potts. Introduction to semantics and pragmatics. Lecture notes, 2019.
- [30] Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69, 2012.
- [31] Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116, 1992.
- [32] Kjell Johan Saebo. Conversational contrast and conventional parallel: Topic implicatures and additive presuppositions. *Journal of Semantics*, 2004.
- [33] Uli Sauerland, Jan Anderssen, and Kazuko Yatsushiro. The plural is semantically unmarked. In Stephan Kepser and Marga Reis, editors, *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*, pages 413–434. de Gruyter, Berlin, Germany, 2005.
- [34] Benjamin Spector. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland and Penka Stateva, editors, *Presupposition and Implicature in Compositional Semantics*, pages 243–281. Palgrave, London, UK, 2007.
- [35] Nadine Theiler. When additive particles can associate with wh-phrases. In *Proceedings of Sinn und Bedeutung*, volume 23, pages 345–362, 2019.
- [36] Guillaume Thomas. Another additive particle. In *Semantics and Linguistic Theory (SALT) 21*, pages 634–651, 2011.
- [37] Kai von Stechow. Exceptive constructions. *Natural Language Semantics*, 1:123–148, 1993.
- [38] Ekaterina Vostrikova. *Phrasal and Clausal Exceptive-Additive Constructions Crosslinguistically*. PhD thesis, University of Massachusetts, 2019.



# Learnability as a window into universal constraints on person systems\*

Mora Maldonado and Jennifer Culbertson

Centre for Language Evolution  
University of Edinburgh, Edinburgh, UK  
mora.maldonado@ed.ac.uk

## Abstract

Zwicky (1977) made the following observation regarding the cross-linguistic distribution of person systems: Languages that do not have a dedicated form for the inclusive meaning (speaker+addressee) always assimilate it into the first person and never into the second or third. Different theories of person have been put forward to explain this asymmetry, positing different kinds of semantic constraints on the human capacity to categorize the person space. Here we use an Artificial Language Learning methodology to investigate whether learners are sensitive to these typological asymmetries. Our results reveal that (1) learners prefer person systems where there is homophony between inclusive and first person meanings; and (2) they find it easier to learn that the inclusive is a form of ‘you’ than a form of ‘them’. Given that second-inclusive and third-inclusive homophony patterns are both absent in the typology, our findings suggest that not all typological tendencies are built equal: while some might be seen as the result of strong, universal cognitive constraints on grammar, others should be modeled as weaker learning biases.

## 1 Introduction

In a classic paper from 1977, Zwicky made the following observation regarding the cross-linguistic distribution of person systems: Languages that do not have a dedicated phonological form for an inclusive person (speaker+addressee) always assimilate it into the first person and never into the second or third. That is: in these languages (e.g. English), the *you and us* inclusive meaning is expressed as a form of ‘us’, and never as a form of ‘you’ (or ‘them’).<sup>1</sup> The basic distinction between the typological patterns is illustrated in Table 1.

At first glance, Zwicky’s generalization is quite surprising. Most feature-based approaches to person systems (e.g. binary features account in Table 1, see [5] for discussion) assume that the inclusive person shares features with both the first exclusive (e.g. +speaker) and the second exclusive (e.g. +addressee). Indeed, a number of languages have inclusive pronouns that can be morphologically decomposed into first plus second forms (e.g., Bislama, 11), suggesting that there is (at least at some level) a semantic overlap between second and inclusive categories.<sup>2</sup> This leads naturally to the expectation that languages should be as likely to assimilate the

---

\*This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 757643).

<sup>1</sup>Zwicky’s generalization holds for languages taken as a system and not for individual paradigms within a language, which might show accidental homophony. For example, one can find languages in which some paradigms collapse inclusive and second persons to the exclusion of the first exclusive (and third); see [12] and examples therein. None of these languages, however, systematically group the inclusive and the second person in all its paradigms; instead, they have some paradigms in which the two differ. This is crucially different from *true* three-person systems like English, where first person collapses exclusive and inclusive in *all* its paradigms.

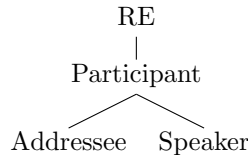
<sup>2</sup>Note that there are also inclusive languages where the verbal agreement of the inclusive form is shared with the second person and not with the first (e.g. Ojibwa pronouns, [13] from [20]).

1 <sub>ST</sub>	+speaker, -addressee													
INCL	+speaker, +addressee													
2 <sub>ND</sub>	-speaker, +addressee													
3 <sub>RD</sub> PL	-speaker, -addressee													
Binary-features account of person systems														
<table style="border-collapse: collapse; margin: auto;"> <tr><td style="padding: 2px 5px;">1<sub>ST</sub>PL</td></tr> <tr><td style="padding: 2px 5px;">INCL</td></tr> <tr><td style="padding: 2px 5px;">2<sub>ND</sub>PL</td></tr> <tr><td style="padding: 2px 5px;">3<sub>RD</sub>PL</td></tr> </table>	1 <sub>ST</sub> PL	INCL	2 <sub>ND</sub> PL	3 <sub>RD</sub> PL	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="padding: 2px 5px;">1<sub>ST</sub>PL</td></tr> <tr><td style="padding: 2px 5px;">INCL</td></tr> <tr><td style="padding: 2px 5px;">2<sub>ND</sub>PL</td></tr> <tr><td style="padding: 2px 5px;">3<sub>RD</sub>PL</td></tr> </table>	1 <sub>ST</sub> PL	INCL	2 <sub>ND</sub> PL	3 <sub>RD</sub> PL	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="padding: 2px 5px;">1<sub>ST</sub>PL</td></tr> <tr><td style="padding: 2px 5px;">INCL</td></tr> <tr><td style="padding: 2px 5px;">2<sub>ND</sub>PL</td></tr> <tr><td style="padding: 2px 5px;">3<sub>RD</sub>PL</td></tr> </table>	1 <sub>ST</sub> PL	INCL	2 <sub>ND</sub> PL	3 <sub>RD</sub> PL
1 <sub>ST</sub> PL														
INCL														
2 <sub>ND</sub> PL														
3 <sub>RD</sub> PL														
1 <sub>ST</sub> PL														
INCL														
2 <sub>ND</sub> PL														
3 <sub>RD</sub> PL														
1 <sub>ST</sub> PL														
INCL														
2 <sub>ND</sub> PL														
3 <sub>RD</sub> PL														
(a) First-Inclusive <b>attested</b>	(b) Second-Inclusive <b>unattested</b>	(c) Third-Inclusive <b>unattested</b>												

**Table 1:** Illustration of Zwicky’s observation. Grayed cells are mapping to a single pronominal form, white cells to different and distinct forms.

inclusive with the second person as they are to assimilate it with the first. In contrast, no theory would predict the inclusive meaning to be homophonous with the third person, as the inclusive and the third person do not have any features in common (although see discussion of Tupinamba in [18] for a potential exception).

A number of theories of person have been developed with an eye toward accounting for Zwicky’s observation. These theories take one of two general approaches: either maintain the traditional set of features (e.g. Table 1), but add default feature specifications, or posits a different set of features. For example, Harley and Ritter [13] take the first approach. Following ideas from phonology [7], they put forward a universal feature geometry for person, with primitives similar to the binary-features account but also hierarchical relations between them (see also 4, 15, 8 for similar approaches). In their system, the primitive features Speaker and Addressee are dependent nodes of the feature Participant. However, they specify the Speaker feature as less marked than the Addressee feature (see Figure 1). Consequently, in languages without an inclusive distinction, a preference for assimilating the inclusive meaning into the first person is expected, as they share the default feature. Defaults can be overridden, therefore the second-inclusive homophony pattern can still arise in contrast to a third-inclusive system which is predicted to be impossible.



**Figure 1:** Feature geometry account in Harley and Ritter (2002)

The second approach is adopted by Harbour [12], who posits a different set of binary features, ±author and ±participant. While the features themselves denote lattices of possible entities (speaker, addressee, etc.), the values of the features are modelled as complementary operations on lattices. Because features are similar to functions, languages can differ not only in which

features are active, but also in the order of feature composition.<sup>3</sup> Importantly, the absence of a  $\pm$ addressee feature in the system creates an inherent asymmetry between speaker and addressee discourse roles. This asymmetry is essential to derive Zwicky’s observation as a *strong* constraint on possible person systems: Systems showing systematic homophony between the first and inclusive person are derived, but homophony between inclusive and second or third person are both impossible. For example, a language which makes use of the  $\pm$ author feature only will have a bipartition of the person space in which first and inclusive persons are homophonous, and second and third are homophonous. A language in which both  $\pm$ author and  $\pm$ participant are active, with the  $\pm$ participant composing last, will have a first-inclusive tripartition (indeed this is the only tripartition of the space generated by this theory). A language where  $\pm$ participant feature composes first will have not a second-inclusive, but a quadripartition of the space (see [12] for details), where each category is mapped into a unique phonological form. Without a corresponding  $\pm$ addressee feature, there is no way to have any partition which picks out the set of categories including the addressee.<sup>4</sup>

The theories outlined above account for Zwicky’s observation, but differ critically in how second-inclusive and third-inclusive are treated. Under Harley and Ritter’s system [13], third-inclusive is singled out as underivable, while second-inclusive is possible but more marked than first-inclusive. By contrast, the system proposed by Harbour [12] takes as its starting point the idea that only first-inclusive can be generated by the grammar. Based on the typology alone, it is impossible to adjudicate between these theories: both second- and third-inclusive patterns are unattested. Further neither theory provides an explicit mechanism for linking the feature-based representations (and operations) they posit to typology. The implicit link is *learnability*; only a subset of possible person partitions are learnable by humans, or alternatively, some are learned readily while others are more difficult (e.g., can be learned but require substantially more evidence). However, there are many factors that can shape typological distributions—from learning biases to historical accidents, to genetic relations between languages, to pathways of diachronic change. In the domain of person, there is currently no evidence that learning is the crucial driving force.

The present paper follows a growing body of work using artificial language learning to provide new sources of behavioral data to test theoretical claims made on the basis of typological data (for reviews see [9, 10, 16]). Here, we investigate learners’ sensitivity to predicted asymmetries among non-inclusive paradigms. We train English-speaking learners on an artificially-constructed personal pronoun paradigm in which the inclusive is a form of ‘us’ (first-inclusive), a form of ‘you all’ (second-inclusive) or a form of ‘them’ (third-inclusive). Participants are taught one of these three paradigms and are tested on how accurately they are able to learn it (within a set number of trials).

Given that English features first-inclusive homophony *and* it is the only tripartition sys-

<sup>3</sup>For the sake of brevity, here we are simplifying Harbour’s account. More specifically, assume a person space that contains a speaker ( $i$ ), an addressee ( $u$ ) and an undefined number of others ( $o$ ) as basic entities in the ontology. Harbour’s proposed ontology consists on *egocentrically* nested subsets, such that the smallest subset in the ontology contains the speaker alone:  $\{i\} \subset \{i, iu, u\} \subset \{i, iu, u, o\}$ . The  $\pm$ author and  $\pm$ participant features will then denote part of this ontology by denoting the semi-lattices  $\{i\}$  and  $\{i, iu, u\}$  respectively. When the author feature is set up in its positive value, it will denote a semi-lattice containing only those elements containing  $i$  (i.e.  $\{i, iu\}$ ). Otherwise, it will denote those elements that *exclude* the speaker:  $\llbracket \text{-author} \rrbracket = \{i, iu, \}$

<sup>4</sup>An intermediate proposal can be found in Ackema and Neeleman’s system [3]. In their proposed feature structure, “there is no natural class (...) that comprises the first person inclusive and the second person, but not the first person exclusive” (p.910,[2]). However, second-inclusive patterns can still be obtained by incorporating an impoverishment rule in the system. This is not possible for inclusive-third homophony, creating an asymmetry between the two unattested patterns.

tematically attested in the typology, learners are predicted to prefer such paradigms over the alternatives. This mainly serves as a sanity check that participants understand the task and are able to learn a language that has the same structure as their own. Regarding second-inclusive and third-inclusive homophony patterns—both unattested in the typology—the accounts outlined make different predictions. If both of these patterns are directly ruled out by the grammar (*à la* Harbour), learners should be equally unlikely to learn either of them. By contrast, if learners are sensitive to the semantic commonalities between the inclusive and the second person (e.g. + addressee), a second-inclusive system should be easier to learn than a third-inclusive one [13, 15]. This pattern of results would moreover suggest that any apparent asymmetry between first-inclusive and second-inclusive languages should not be encoded as a hard constraint on person systems (*contra* Harbour).

## 2 Methods

This experiment, including all hypotheses, predictions, and analyses, was preregistered <https://osf.io/5h4m6>.

### 2.1 Design

Participants were randomly assigned to one of three possible conditions, which correspond to the (a-c) patterns in Table 1. Participants in all conditions were taught three pronominal forms mapped into four *plural* person categories (first exclusive, inclusive, second exclusive, and third). Each condition instantiated a different form-to-meaning mapping: the pronominal system could assimilate the inclusive meaning into the first plural person (First-Inclusive condition), into the second plural person (Second-Inclusive condition), or into the third plural person (Third-Inclusive condition).

Participants in all three conditions were also exposed to three additional distinct pronominal forms corresponding to the first, second, and third *singular* persons. Participants' learning of these forms was used as an exclusion criteria (see below).

#### 2.1.1 Materials

The language consisted of 6 different pronominal forms: 3 forms were used for the plural pronouns (critical categories), and 3 different forms were used for the singular pronouns (filler categories). For each participant, these 6 lexical items were randomly drawn from a list of 8 CVC non-words created following English phonotactics: 'kip', 'dool', 'heg', 'rib', 'bub', 'veek', 'tosh', 'lom'. Items were presented orthographically.

To express the pronoun meanings, we commissioned a cartoonist to draw scenarios involving a family of three sisters and their parents. Each family member has a clearly-defined role in the conversational context. The two older sisters are speech act participants (in all scenarios they are either speaker or addressee). The third (little) sister was spatially close, but never a speech act participant. The parents were seated in the background (serving as additional others).

Pronouns were used as one-word answers to questions like 'Who will be rich?'. Meanings were expressed by visually highlighting subsets of family-members, as in Table 2. In some cases, more than one pattern of visual highlighting could match the target meaning, options were then randomly selected. An example illustrating the 1<sup>st</sup> plural trial is provided in Figure 2. All questions were English interrogative sentences of the form 'Who will...?', which were randomly drawn from a list of 60 different tokens.

Category	Highlighted set
1 <sup>st</sup> <sub>SG</sub>	speaker
2 <sup>nd</sup> <sub>SG</sub>	addressee
3 <sup>rd</sup> <sub>SG</sub>	one other
1 <sup>st</sup> <sub>PL</sub>	speaker, other(s)
INCL	speaker, addressee (other(s))
2 <sup>nd</sup> <sub>PL</sub>	addressee, other(s)
3 <sup>rd</sup> <sub>PL</sub>	multiple others

**Table 2:** Highlighted family members for each person category. 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> plural categories randomly include one or two additional others; the inclusive category could refer to speaker and addressee alone or include as well one or two others.

### 2.1.2 Procedure

Participants were first introduced to the family, including the names of the sisters, and were told they were going to see the sisters playing with a hat that had two magical properties: whoever wore it could see the future but would also talk in a mysterious ancestral language. Participants were instructed to figure out the meanings of words in this new language. They were given a hint that the words were not names, but they always refer to people present in the scene. In addition, the speaker and addressee roles switched several times during the experiment to highlight that the words were dependent on contextually-determined speech-act role. This was induced by swapping who had the magical hat.

The experiment had two phases, each composed of exposure and testing blocks (e.g., ??). Exposure trials (e.g., Figure 2a) had two parts: a scene where a question was asked, and a scene where the question was answered with a pronoun form in the language. To check that participants were paying attention, they were then asked to select among three different alternatives the one corresponding to the pronominal form they had just seen. Testing trials (e.g., Figure 2b) consisted of a question, just as in exposure trials, followed by a scene where a set of individuals was highlighted (reference of the pronoun) without presenting the pronoun form. Participants had to pick the correct word for that meaning among three different options.

During the first phase, participants were trained and tested on the three singular pronouns. There were a total of 12 exposure and 12 testing trials (4 repetitions per form/meaning). Participants who responded accurately to at least 2/3 of the testing trials in this phase (corresponding to 8 correct responses) passed to the second critical phase. In the critical phase, participants were exposed to and tested on the mapping between three plural pronouns and four person meanings. This phase was comprised of two alternating exposure and testing blocks. There were a total of 24 exposure trials (6 repetitions per meaning) and 48 testing trials (12 repetitions per meaning).

The complete experimental session lasted approximately 20 minutes. The order of presentation of meanings was fully randomized within exposure and testing blocks for each participant. Participants were given a debrief at the end of the experiment to check how they interpreted the forms they were trained on. For example, participants in the Second-Inclusive condition would describe the critical form as 'me and you or you all' or as 'group containing Ann or Mary', whereas participants in the First-Inclusive condition would just use the pronoun 'us'.

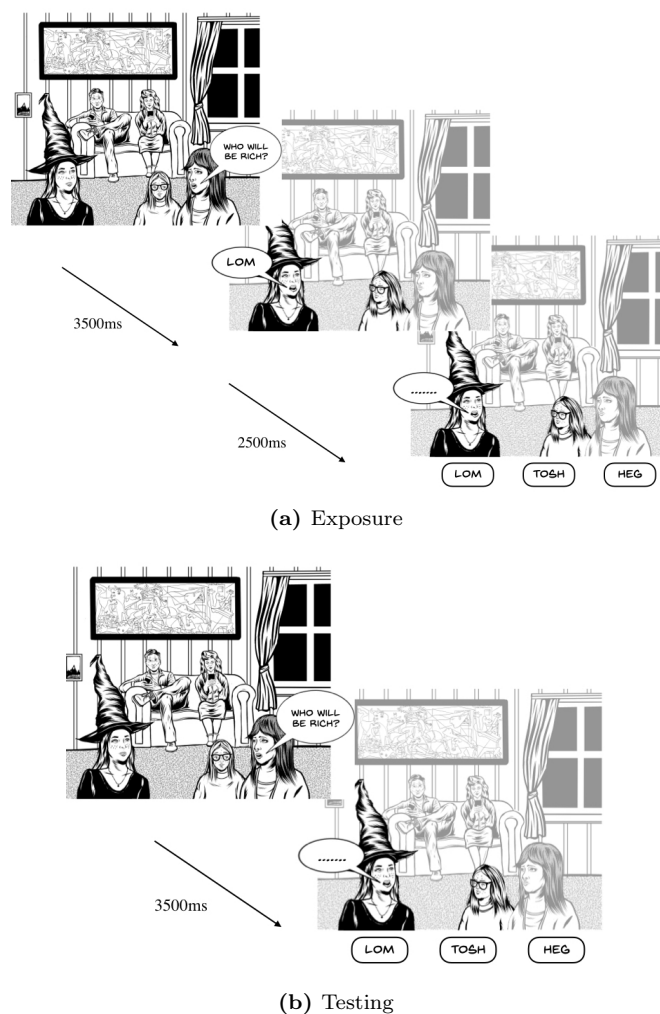


Figure 2: Illustration of Exposure and Testing trials for the 1<sup>st</sup> PL category.

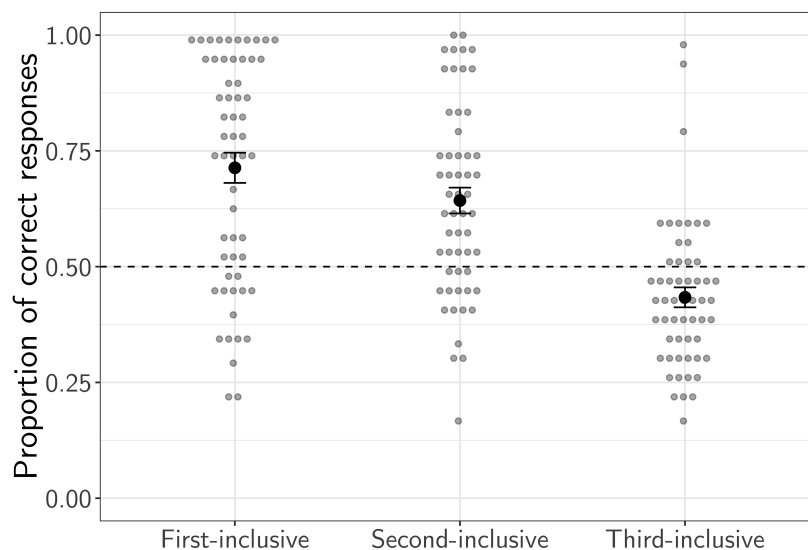
### 2.1.3 Participants

A total of 320 English-speaking adults were recruited via Amazon Mechanical Turk (First-inclusive group: 109, Second-inclusive group: 101, Third-inclusive: 110).<sup>5</sup> 167 participants responded accurately on more than 8 singular testing trials and were allowed to continue with the *critical* plural pronoun phase, according to our pre-registered plan (First-inclusive group: 57, Second-inclusive group: 55, Third-inclusive: 55). Participants who passed the singular test phase were paid 3.5 USD for their participation and 1 USD otherwise.

<sup>5</sup>The number of participants reported here does not include workers who were excluded for not being self-reported native speakers of English (10) and participants who fail to pass an attention check (AC) included at the very beginning of the experiment (35). This AC was added to exclude participants who had not read the instruction or were bots [19]. While these participants could in principle have been excluded based on their performance singular testing trials, the AC allowed us to filter them out in advance, distinguishing them from participants who just found the experiment hard.

### 3 Results

Mean accuracy rates on testing trials during the *critical* phase are given in Figure 3. The effect of Condition and Block on accuracy rates was analyzed using logit mixed-effect models in R [1].<sup>6</sup> The model included the maximal random effect structure, random intercepts per subject and slopes per block [following ?]. The standard alpha level of 0.05 was used to determine significance, and  $p$ -values were obtained based on asymptotic Wald tests.



**Figure 3:** Accuracy rates in *critical* testing trials by condition. Error bars represent standard error on by-participant means; gray dots represent individual participant means.

We first compared First-Inclusive with Second-Inclusive and Third-Inclusive (contrasts were treatment coded, with First-Inclusive and Block 2 as baselines). The model revealed that the proportion of accurate responses in the First-inclusive condition (Block 2) was significantly above chance (intercept:  $\beta = 1.59$ ;  $p < .001$ ). In addition, accuracy rates in the First-inclusive condition were significantly higher than in the Third-inclusive condition ( $\beta = -1.83$ ;  $p < .001$ ), and marginally different from accuracy rates in the Second-inclusive condition ( $\beta = -0.572$ ;  $p = .055$ ).

A second model was fitted to test the difference in accuracy between Second-inclusive and Third-inclusive conditions (contrasts were treatment coded, with Second-inclusive and Block 2 as baseline). Accuracy rates in the Second-inclusive condition (Block 2) were found to be significantly above chance (intercept:  $\beta = 0.97$ ;  $p < .001$ ) and higher than the ones in the Third-inclusive condition ( $\beta = -1.2$ ;  $p < .001$ ).<sup>7</sup>

<sup>6</sup>There were two testing blocks in the critical phase, each preceded by an exposure block. Participants were generally expected to improve with accumulated exposure, but this improvement could vary across conditions. Each model included the effect of Block on accuracy, as well as the interaction with Condition. However, we report here only simple effects regarding the second testing block. The complete model output can be found in the OSF repository.

<sup>7</sup>Following our preregistration, we ran a second version of each of these models, restricting the analysis to inclusive meanings. The idea behind this move was that our hypotheses target specifically the inclusive category



## 4 Discussion

Our findings first confirm that English-speaking participants are most successful at learning a new language that features native-like homophony between inclusive and first person meanings. This result confirms that participants understand the task and are able to learn a language that has the same structure as their own.

If, as predicted by Harbour [12], there is additionally a hard constraint on tripartitions which derives only first-inclusive homophony, then we might expect the preference for English-like systems to be very strong indeed. However, the difference in accuracy between First-inclusive and Second-inclusive conditions was only marginally significant. This result is consistent with the idea that learners are sensitive to the semantic overlap between inclusive and second person, as predicted by [5, 13, 15]. In line with these accounts, participants in the Second-Inclusive condition may treat these as a natural class, relying on a shared feature,  $\pm$ addressee, to learn this new language. This result supports a theory in which first- and second-inclusive patterns are both generated by the grammar, but the latter is just dispreferred (contra 12, and possibly 2, 3, see Footnote 4).

Importantly, we also found that learners have a bias against systems that assimilate the inclusive into the third person. This result reveals that second- and third-inclusive systems, despite being unattested in the typology, are not equal from a learnability perspective: there is a stronger pressure against third-inclusive than against second-inclusive homophony. This is again as predicted by theories like [5, 13, 15] (but arguably also by [3]) which posit that, unlike first and second person, third person differs from the inclusive in more than one feature and therefore does not form a natural class with it (cf., e.g., [18]). As a result, homophony between inclusive and third persons is not predicted to occur systematically, though it might arise accidentally.

Returning to Zwicky's observation, our findings suggest that the typological asymmetry between alternative non-inclusive systems may result from multiple interacting factors, not just learnability as tested in our experiment. In the typology first-inclusive systems are attested systematically, while second- and third-inclusive systems are not; in our experiment, third-inclusive systems were clearly dispreferred, but the advantage for first-inclusive over second was much weaker. While this is consistent with a theory positing weak learning biases (i.e., constraints that can be over-ridden given sufficient evidence) which penalize third-inclusive most, it still leaves the typological data partially unexplained.

One possibility is that there is an additional weak bias, not at play in our experiments, which further advantages first-inclusive systems. An obvious such candidate is a general *egocentric* bias, i.e., increased importance or salience of the speaker to him or herself. Indeed, a version of this bias has been argued to hold during early pronoun acquisition [6, 14, 17]. If individuals (and potentially children more so than adults) perceive the world as a function of their presence in it, then they may be more likely to adopt categorization systems which preserve this distinction. This would lead to a stronger asymmetry between first-inclusive and second-inclusive systems. There is good reason to suspect that any obvious egocentric bias is weakened in the context of our experiment, where participants are passive learners and do not themselves feature in the meanings they are learning—our learners are never themselves speakers. Future work will investigate whether a stronger egocentric bias, and therefore a stronger preference for first-inclusive paradigms, arises when the learner plays a more active role in the conversation.

Before concluding, let us entertain an (partial) alternative interpretation of our results.

---

(expressed by the ambiguous pronoun in each system). The output of these models follows the pattern of results described above.



Recall that Zwicky’s generalization holds specifically for languages as a system and not for particular paradigms. In our experiment, we teach participants a pronominal system, which in principle may be just one paradigm of many in the language. Arguably, our participants might be treating the pronominal system we teach them as an instance of accidental syncretism within an inclusive language (i.e., a language that makes an inclusive/exclusive distinction). If this were the case, theories like Harbour’s could still account for our results. However, given that our participants are speakers of a non-inclusive language, this seems very unlikely. Further, this would likely not predict the difference we find between second- and third-exclusive.

## 5 Conclusion

In this study, we used an artificial language learning paradigm to test the different predictions made by theories of person systems. These theories were designed to, among other things, explain Zwicky’s generalization: when a language does not have a dedicated pronominal form for the inclusive, it will assimilate the inclusive with the first person (not the second or the third). We targeted two influential approaches which differ in the set of universal person features they posit. According to Harley and Ritter [13], first-inclusive homophony is predicted to occur systematically, second-inclusive homophony is marked but possible, and third-inclusive homophony is impossible. For Harbour [12], only first-inclusive homophony can be generated by the grammar. Our results revealed that English-speaking learners have a slight preference for first-inclusive over second-inclusive paradigms, but clearly disprefer third-inclusive. Our findings support theories like Harley and Ritter’s, which posit semantic overlap between second and inclusive persons which makes homophony between these categories possible. Thus apparent regularities in typology are not necessarily the result of strong constraints on possible grammars (*à la* Harbour). Here we have argued that (morpho-)semantic universals such as Zwicky’s observation might be best explained in terms of weak learning biases (i.e., sensitive to cross-category similarity, and the importance of the ego), which can be overcome by learners given enough evidence.

## References

- [1] R: A language and environment for statistical computing. 2013.
- [2] Peter Ackema and Ad Neeleman. Person features and syncretism. *Natural Language & Linguistic Theory*, 31(4):901–950, November 2013.
- [3] Peter Ackema and Ad Neeleman. *Features of Person: From the Inventory of Persons to Their Morphological Realization*, volume 78. MIT Press, 2018.
- [4] Susana Bejar. *Phi-Syntax: A theory of agreement*. PhD thesis, University of Toronto, Toronto, 2003.
- [5] Jonathan David Bobaljik. Missing persons: A case study in morphological universals. *The Linguistic Review*, 25(1-2), January 2008.
- [6] Rosalind Charney. Speech roles and the development of personal pronouns. *Journal of child language*, 7(3):509–528, 1980.
- [7] G. N. Clements. The Geometry of Phonological Features. *Phonology Yearbook*, 2:225–252, 1985.

- [8] Elizabeth Cowper and Daniel Currie Hall. Argumenthood, pronouns, and nominal feature geometry. *Determiners: Universals and variation*, 147:97–120, 2009.
- [9] Jennifer Culbertson. Typological Universals as Reflections of Biased Learning: Evidence from Artificial Language Learning: Typological Universals as Reflections of Biased Learning. *Language and Linguistics Compass*, 6(5):310–329, May 2012.
- [10] Jennifer Culbertson and Kathryn Schuler. Artificial Language Learning in Children. *Annual Review of Linguistics*, 5(1):353–373, January 2019.
- [11] Daniel Harbour. Descriptive and explanatory markedness. *Morphology*, 21(2):223–245, May 2011.
- [12] Daniel Harbour. *Impossible Persons*. Linguistic Inquiry Monographs. The MIT Press, Cambridge, MA, 2016.
- [13] Heidi Harley and Elizabeth Ritter. Person and Number in Pronouns: A Feature-Geometric Analysis. *Language*, 78(3):482–526, 2002.
- [14] Katherine A. Loveland. Learning about points of view: Spatial perspective and the acquisition of ‘I/you’. *Journal of child language*, 11(3):535–556, 1984.
- [15] Martha McGinnis. On Markedness Asymmetries in Person and Number. *Language*, 81(3):699–718, 2005.
- [16] Elliott Moreton and Joe Pater. Structure and Substance in Artificial-phonology Learning, Part I: Structure. *Language and Linguistics Compass*, 6(11):686–701, 2012.
- [17] Morgan Moyer, Kaitlyn Harrigan, Valentine Hacquard, and Jeffrey Lidz. 2-year-olds’ comprehension of personal pronouns. In *Online Proceedings of the 29th Annual Boston University Conference on Language Development*, page 11, Boston, 2015.
- [18] Aryon D Rodrigues. You and I = Neither You nor I: The personal system of Tupinamba. page 13, 1990.
- [19] Steven V. Rouse. A reliability analysis of Mechanical Turk data. *Computers in Human Behavior*, 43:304–307, February 2015.
- [20] Linda J. Schwartz and Timothy Dunnigan. Pronouns and pronominal categories in Southwestern Ojibwe. *Pronominal systems*, pages 285–322, 1986.

# A Variable-Force Variable-Flavor Attitude Verb in Koryak\*

Maša Močnik and Rafael Abramovitz

Massachusetts Institute of Technology  
{mocnik, rafabr}@mit.edu

## Abstract

We enrich the typology of modal expressions with the attitude verb *ivək* from Koryak, which shows a wide range of flavors (doxastic, bouletic, assertive, directive) and is the first documented variable-force attitude verb. Variation in both domains goes against the universal that modal items can't vary in both force and flavor (Nauze 2008). We use the existential-universal doxastic-assertive variation to argue against this generalization. For the bouletic flavor, we show that it is triggered by the material in the embedded clause; we propose a new technical way of composing the bouletic flavor at LF.

## 1 Introduction

Research on understudied languages has uncovered modal systems that carve up the space of modal meaning differently from the English modals.<sup>1</sup> For example, St'át'imcets *k'a* has a fixed epistemic flavor but varies in force (possibility, necessity) (Rullmann et al. 2008, henceforth RMD). Nauze (2008) suggests that there is limited variability along these two dimensions<sup>2</sup>:

- (1) Modal elements [...] either vary on the [flavor] axis and thus are polyfunctional in the original sense of expressing different types of modality or they vary on the [force] axis and can express possibility and necessity, but they cannot vary on both axes. (p. 222)

English attitude verbs have traditionally been treated as modal items with a (lexically) fixed force and flavor (e.g. *think*: necessity force, doxastic flavor). Recent work on understudied languages has shown some variability within this class of expressions, too. Navajo *nízin* (Bogal-Allbritten 2016), for example, at least on the surface appears to vary in flavor: it has doxastic ('think') and bouletic ('want', 'hope') uses.

This paper puts forth a counterexample to (1) with an attitude verb from the Chawchiven dialect of Koryak<sup>3</sup>, henceforth 'Koryak'. *Ivək*, typically translated out of the blue as 'say' (**assertive**), is also used as a **doxastic** ('think', 'allow for the possibility'), **bouletic** ('hope', 'fear', 'wish'), and **directive** ('tell/order', 'propose/suggest') attitude. The directive flavor requires transitive agreement and we set it aside here.<sup>4</sup> Consider (2)–(3).

\*Special thanks to Kai von Fintel and Roger Schwarzschild for commenting on earlier drafts. Thanks also to Elizabeth Bogal-Allbritten, Jonathan Bobaljik, Luka Crnić, Ksenia Ershova, Sabine Iatridou, and the LFRG audience. We are deeply grateful to our Koryak teachers: O. K. Aleksejeva, L. A. Aslapova, L. J. Avilova, E. I. Dedyk, L. P. Kiseljova, N. S. Kuznetsova, S. N. Moisejeva, T. I. Nutelxut, A. E. Urkachan, G. N. Xarjutkina, and especially V. R. Dedyk. R. Abramovitz is partially supported by an NSF GRF under grant no. 1122374.

<sup>1</sup>Matthewson et al. 2005, Rullmann et al. 2008, Vander Klok 2008, Deal 2011, Bochnak 2015a,b, a.o.

<sup>2</sup>See Bochnak (2015a,b) for a Washo modal verb counterexample.

<sup>3</sup>Koryak is a highly endangered Chukotko-Kamchatkan language spoken in northern Kamchatka (Russia). Our transcription uses the IPA, except that we use *č* for the voiceless alveolo-palatal affricate.

<sup>4</sup>The directive flavor is essentially contributed by the obligatory infinitive or imperative.

(a) tikwi                    {jənnəmatək / qənnəmatyən}                    təllətəl  
*ivək*.1SG>2SG.PST {close.INF / close.2SG>3SG.IMP} door  
'I told you to close the door.' [translation from Russian to Koryak]

- (2) *meλλo kivəŋ, (əno) kumuqetəŋ*  
 Melljo *ivək*.3SG.PRS that rain.3SG.PRS  
 ‘Melljo {says, thinks, allows, hopes, fears, \*knows, \*imagines, \*wishes} that it’s raining.’
- (3) *meλλo kivəŋ, (iwke) nəʔəmuqetən*  
 Melljo *ivək*.3SG.PRS if.only rain.3SG.CF  
 ‘Melljo wishes it would rain.’

The assertive, doxastic, and some bouletic readings of *ivək* are available without special marking in the embedded clause, see (2). Example (2) also indicates that *ivək* is not a completely underspecified attitude verb – there are certain meanings, e.g. ‘imagine’, that are not available. For the bouletic ‘wish’, the embedded clause needs the counterfactual prefix *ʔ-*, see (3).<sup>5</sup>

We analyze *ivək* as an attitude verb whose domain of possible-worlds quantification is underspecified for the doxastic-assertive distinction (modeled with a modal-base-like variable) and whose quantificational force varies due to a restriction on an underlying universal quantifier (in the spirit of RMD). By contrast, we argue that the bouletic flavor is not one of the flavors of *ivək* and is instead triggered by the material in the embedded clause (overt for ‘wish’ and covert for ‘hope’ and ‘fear’). While this has been proposed for the Navajo *nizín* (Bogal-Allbritten 2016), our innovation is to show that the bouletic meaning can be split at LF into a matrix-clause *doxastic quantifier* and an embedded-clause preference item.

After a comment on methodology (§1.1), we turn to each of the three components in turn: the doxastic flavor and its force variability (§2), the assertive flavor and its underspecification with respect to the doxastic one (§3), and the bouletic flavor with its origins in the embedded clause (§4). Each section presents the empirical argument, followed by the analysis.

## 1.1 Elicitation Methodology

We employ an elicitation technique that we call a “matching task”.<sup>6</sup> We provide a context (typically in Russian) and two sentences: a Koryak one and a Russian one. The speakers are first asked to provide a contextual felicity judgment on the Koryak sentence, and then are asked whether it can express the same ‘thought’ (Rus. *mysl’*) as the Russian one in the given context.<sup>7</sup>

The motivation for using this new elicitation technique is that the existing ones have proven to be inadequate for eliciting our data in Koryak. We will briefly describe the issues we encountered and provide motivation for using this technique in fieldwork more generally.

The second author’s previous work with our Koryak consultants showed that they are prone to ignoring salient features of the context, whether it is provided verbally or pictorially. Therefore, eliciting a purely contextual felicity judgment is not sufficient. Asking the speaker to explain how they understood the Koryak sentence (in order to check to what extent the context was ignored) is also not sufficient because the speakers import features of Koryak into Russian, producing infelicitous Russian sentences.<sup>8</sup> For example, the speakers use ‘think if only’ to translate sentences like (3), even though this is not the locution for expressing wishes in Russian. By contrast, when asked to translate *želat’* (‘wish’) into Koryak, the speakers had no trouble using *ivək*. We thus found that the best elicitation technique was to use a matching

<sup>5</sup>Many speakers prefer to add *iwke* ‘if only’ to the embedded clause, whereas others merely tolerate it.

<sup>6</sup>Unless otherwise indicated, the examples provided in this paper were obtained in this manner.

<sup>7</sup>Using *mysl’* worked well, unlike ‘meaning’ (Rus. *značeniye*), which triggered a word-for-word translation.

<sup>8</sup>Interestingly, RMD (fn. 32) note a similar issue for “*k’a p & k’a -p*”, for which the consultant offered an infelicitous English translation.

task, where the context was aided by the Russian sentence. In this task we found consistent results. Importantly, this technique allowed us to also obtain infelicity judgments.

To illustrate the effect of the matching task, let us consider a concrete example. The speaker was provided with (4), where the context and the target sentences were all in Koryak. The speaker at first rejected it, expressing confusion as to how Hewngyto could think two incompatible things. When the speaker was explicitly asked whether *ivək* could mean *dopuskat* ('allow for the possibility') in this discourse, the speaker readily confirmed this and changed her judgment about the coherence of the discourse. We observed similar effects when the speakers had been previously exposed to *dopuskat* during the elicitation session. In the absence of previous exposure or a matching task, the judgments across speakers were not stable enough to warrant further inquiry, for example to test whether a piece of information can bias the weak meaning of *ivək*.

## 2 Variable-force doxastic flavor

Doxastic attitudes are attested not just in the necessity force (*think, believe*) but also in the possibility force (see Močnik (2019a,b) on Slovenian *dopusčati* 'allow for the possibility'). While the necessity force of *ivək* is the default (speakers have a preference for the 'say' translation), the possibility force reading is felicitous too, as shown in (4).

- (4) Hewngyto is walking down the street. Melljo sees him and asks: 'Where is your wife? Is she making jam at home?' He replies:

qoo. təkivəŋ əno kotavareŋjaŋəŋ jajak.  
dunno *ivək*.1SG.PRS that make.jam.3SG.PRS at.home

'I don't know. I allow for the possibility that she's making jam at home.'

He continues walking. Qechghylqot sees him and asks: 'Where is your wife? Is she picking berries in the forest?' Hewngyto replies:

qoo. təkivəŋ əno keluŋ umkək.  
dunno *ivək*.1SG.PRS that pick.berries.3SG.PRS in.forest

'I don't know. I allow for the possibility that she's in the forest picking berries.'

Importantly, force variability does not arise with all attitude verbs in Koryak. For example, *əmalavək* ('believe') is infelicitous on a possibility reading in (5), in contrast to *ivək* in (6). This argues against a general, covert variable-force item in the Koryak embedded clause.

- (5) #ʔewŋəto kolmalavəŋ əno kumuqetəŋ, ʔam ʔopta kolmalavəŋ əno uŋje emuqetke.  
H. believe.3SG.PRS that rain.PRS but also believe.3SG.PRS that NEG rain

'Hewngyto allows that it is raining but also allows that it is not raining.' (intended)

- (6) ʔewŋəto kivəŋ əno kumuqetəŋ, ʔam ʔopta kivəŋ əno uŋje emuqetke.  
H. *ivək*.3SG.PRS that rain.3SG.PRS but also *ivək*.3SG.PRS that NEG rain

'Hewngyto allows that it is raining but also allows that it is not raining.'

In downward entailing contexts, the strong assertive interpretation ('say') of *ivək* remains the default one. As shown in (7), the possibility reading is accepted in a matching task though.

- (7) We're walking down the street and there are many people with raincoats. Melljo says:  
əməŋ ʔujemtewilʔu mekiw ewləj əno jemuqetiki nejetən muqeicʔən  
all people who *ivək*.3PL.PRS that rain.FUT.IPFV bring.3PL>3SG.PST raincoat

‘Everybody who said that it will rain brought a raincoat.’ [volunteered]

‘Everybody who allowed that it will rain brought a raincoat.’ [accepted]

Under negation, there are two readings. Consider first a conjunction of two *ivək* statements without syntactic negation, as in (8a). The necessity interpretation of *ivək* (the speaker believes the ball to be both white and black) is infelicitous, while the possibility interpretation is acceptable (the ball could be white or black). The judgment for (8b) is that it expresses the same thought as (8a), felicitous in this context, but that it also expresses another thought that is infelicitous in this context (namely, the one where the ball is half white and half black).<sup>9</sup>

(8) Two balls are in a box: one white, one black. I pull out one and do not show it to you.

a. *təkivəŋ əno əŋŋin qapəl nilʏəqin to təkivəŋ əno əno luqin*  
*ivək.1SG.PRS that that ball white and ivək.1SG.PRS that it black*  
 ‘I allow that the ball is white and I allow that it is black.’

b. *ujje iwke təkಿತəŋ əno əŋŋin qapəl nilʏəqin to ujje iwke təkಿತəŋ əno*  
*NEG ivək AUX.1SG.PRS that that ball white and NEG ivək AUX.1SG.PRS that*  
*əno luqin*  
*it black*  
 ‘I don’t think that the ball is white and I don’t think that it is black.’

The felicitous interpretation of (8b) could be obtained via: (i) negated necessity or (ii) possibility with neg-raising. Similarly, the infelicitous (half-white-half-black) interpretation could be obtained through (i’) necessity with neg-raising or (ii’) negated possibility. Thus, the felicitous and the infelicitous reading of (8b) are consistent with assigning a single force to *ivək* and using neg-raising to obtain the other reading. The English translation, for example, seems to have the infelicitous reading too, though it is much dispreferred (to force it, it helps to insert an overt ‘only’ in the embedded clause). Notice that resorting to neg-raising would not be sufficient to account for the double reading in (7) since there is no negation in that example.<sup>10</sup>

We conclude the empirical portion of this section with a brief note on the role of epistemic adverbials. Recall that *ivək* needs no embedded material to be felicitous on most readings, as shown in (2). Nevertheless, the embedded clause may contain epistemic adverbials to facilitate the weaker reading. We illustrate with *amu* (‘might’; used to form *wh*-indefinites).

(9) a. *yəmmo təkivəŋ, amu jemuqejuʔəŋ*  
 I *ivək.1SG.PRS* might begin.to.rain.3SG.FUT

‘I allow for the possibility that it will rain.’ [trans to Koryak]

b. *ʔewŋəto kivəŋ, əno (amu) qəjəwjeŋpəlʔo vəʔəjək jəpkejʏəŋ ənəkənəmənəmetəŋ*  
 H. *ivək.3SG.PRS* that might herders soon arrive.3PL.FUT to.his.village

‘Hewngyto hopes that the reindeer herders will soon arrive to his village.’

Speakers have occasionally insisted on using *amu*, though no generalization as to when it is obligatory is forthcoming. We suspect that the weaker reading is simply harder to access, and that adverbs like *amu* facilitate it.<sup>11</sup>

<sup>9</sup>Note that the non-future negated form of the verb *ivək* is *iwke*, which is formed by combining the default agreement/negative circumfix *e-ke* with the root *iv-* and applying the normal phonological rules of the language. We take this form’s homophony with the particle *iwke* ‘if only’ (3) to be a synchronic accident.

<sup>10</sup>We have been unable to confirm the existence of existential versions of ‘say’, ‘wish’ and ‘hope/fear’.

<sup>11</sup>This may be related to RMD’s observation that St’át’imcets *sxək* (‘maybe’) is frequently used in clauses with variable-force modals when the possibility reading is intended.

RMD develop an analysis of the St'át'imcets epistemic modal *k'a* as a universal quantifier whose modal force is modulated by a “modal choice function”  $f$  in (10a) (cf. Lewis 1974, Stalnaker 1975). Since  $f$  in (10a) selects a set of possible worlds (rather than a world), we'll refer to it as a “subset selection function” (cf. von Stechow 1999b).

- (10) a.  $f_{(st)st}$  is a function s.t. for any non-empty set of worlds  $W$ :  $f(W) \subseteq W$  and  $f(W) \neq \emptyset$   
 b.  $[[\text{modal}]]^{c,w}$  is only defined if  $c$  provides a modal base  $B$ .  
 $[[\text{modal}]]^{c,w} = \lambda f_{(st)st} \lambda p_{st} . \forall w' [w' \in f(B(w)) \rightarrow p(w')] \quad (\text{RMD 2008, pp. 337–338})$

The authors provide the example *t'cum k'a kw s-John* (‘John must/may have won’) in a situation where the speaker knows that John played bingo and that he is spending a lot of money today. If  $f$  is the identity function (assumed to be the default, hence the preference for a necessity interpretation), there is no effect on the modal base, so the strong reading obtains (‘John must have won’). Alternatively,  $f$  can pick out a proper subset of the modal base, for instance, by restricting it to those worlds in which John is unemployed.<sup>12</sup> Proper subset selection yields the weak reading ‘John may have won’.<sup>13</sup>

Since *ivak* parallels *k'a* in its preference for a necessity interpretation, we adopt the idea that *ivak* denotes a universal quantifier whose force is restricted by a subset selection function. There are two differences between our data and those of RMD: (i) unlike the epistemic *k'a*, which always outscopes negation, *ivak* has two readings in the presence of a negation in the same clause, see (8b), and (ii) to ensure consistent results, *ivak* needed to be paired up with a translation in a matching task, cf. §1.1. This means that it would not have been reliable to test the felicity of *ivak* in a situation where a piece of evidence, such as John being unemployed, was salient and would therefore trigger the weak reading of *ivak*.

Instead of combining *ivak* with a subset selection function as a sister node at LF, as in (10b), we follow a suggestion by Roger Schwarzschild (p.c.) to encode existential quantification over the subset selection function in the denotation of *ivak* and restrict its choices to either the identity function (default) or to the space of all subset selection functions (that apply). We will incorporate this suggestion as in (11), with an object language variable that is presuppositionally constrained to these two options. We write the restrictions on  $C$  in set notation for simplicity.

- (11) Denotation of *ivak* (to be amended for flavor)  
 $[[ivak]]^{c,g,w} = \lambda C \lambda p \lambda x : C = \{f \mid f(\mathcal{B}_w^x) = \mathcal{B}_w^x\} \vee C = \{f \mid f(\mathcal{B}_w^x) \subseteq \mathcal{B}_w^x \wedge f(\mathcal{B}_w^x) \neq \emptyset, \exists f \in C [\forall w' \in f(\mathcal{B}_w^x) [p(w') = 1]]\}$ , where  $\mathcal{B}_w^x$  is the set of worlds compatible with  $x$ 's beliefs at  $w$  and  $C$  is a cover that limits the choice of  $f_{(st)st}$  (so that  $f$  is either the identity function or some subset selection function on  $\mathcal{B}_w^x$ )

Examples like (4) and (6) would be contradictory if the domain of quantification were constant across the two *ivak* statements (cf. universal quantification over non-disjoint sets). Take (6). Let's use  $\mathcal{C}_{id}$  (Option 1; assumed to be the default strategy) and  $\mathcal{C}_{all}$  (Option 2) for the two resolutions of  $C$  under  $g$ . Option 1 yields a contradiction, while Option 2 yields a felicitous interpretation (there needs to be one  $f$  for which it is raining in all the worlds in its output and one  $f$  for which it is not).<sup>14</sup>

- (12)  $[[ (6) ]]^{c,g,w} = 1$  iff  $\exists f \in g(C) [\forall w' \in f(\mathcal{B}_w^h) [r(w')]] \wedge \exists f \in g(C) [\forall w' \in f(\mathcal{B}_w^h) [\neg r(w')]]$   
 Option 1 (infelicitous):  $\forall w' \in \mathcal{B}_w^h [r(w')] \wedge \forall w' \in \mathcal{B}_w^h [\neg r(w')]$   
 Option 2 (felicitous):  $\exists f \in \mathcal{C}_{all} \forall w' \in f(\mathcal{B}_w^h) [r(w')] \wedge \exists f \in \mathcal{C}_{all} \forall w' \in f(\mathcal{B}_w^h) [\neg r(w')]$

<sup>12</sup>Note that the speaker cannot *know* that John is unemployed, otherwise ‘John is unemployed’ would be true throughout the modal base, so  $f$  would not be selecting a proper subset

<sup>13</sup>They set aside ordering sources and their relationship to the function in (10a).

<sup>14</sup>In RMD, see ex. (58)–(59), the function is simply existentially closed in each conjunct.



Consider now (8b). The felicitous reading is obtained via Option 1.<sup>15</sup>

- (13)  $[[\text{(8b)}]]^{c,g,w} = 1$  iff  $\neg\exists f \in g(C)[\forall w' \in f(\mathcal{B}_w^h)[b(w')]]$  and  $\neg\exists f \in g(C)[\forall w' \in f(\mathcal{B}_w^h)[\neg b(w')]]$   
 That is,  $\forall f \in g(C)[\exists w' \in f(\mathcal{B}_w^h)[\neg b(w')]]$  and  $\forall f \in g(C)[\exists w' \in f(\mathcal{B}_w^h)[b(w')]]$   
 Option 1 (felic):  $\exists w' \in \mathcal{B}_w^h[\neg b(w')] \wedge \exists w' \in \mathcal{B}_w^h[b(w')]$   
 Option 2 (infelic):  $\forall f \in \mathcal{C}_{all}[\exists w' \in f(\mathcal{B}_w^h)[\neg b(w')]] \wedge \forall f \in \mathcal{C}_{all}[\exists w' \in f(\mathcal{B}_w^h)[b(w')]]$

Option 2 is felicitous only if both balls are half white and half black ( $\mathcal{C}_{all}$  contains functions that yield singleton sets, for example).<sup>16</sup>

### 3 Underspecified assertive flavor

What is the relationship between the doxastic and the assertive flavor of *ivək*? By contrast to the bouletic flavor (see §4), we have no evidence to suppose that this distinction is triggered by something other than *ivək*. There are two basic options then: (i) there is a single lexical entry or (ii) there are two: *ivək*<sub>1</sub> ‘think’ and *ivək*<sub>2</sub> ‘say’ (polysemy). The generalization in (1) could in principle address a theoretical choice between these two.

The formal semantics tradition follows Kratzer (1977, 1981, 1991) in adopting a version of (i) for modal verbs, whereby a modal-base variable receives different values in different contexts. Arguments against this view and in favor of (ii) have been offered though, see Viebahn and Vetter (2016) in particular. Nauze (2008) also argues for a polysemous account of modals. We will take it then that the debate between (i) and (ii) is not what (1) is about.

For English, Kratzer (1991) has noted that adverbial expressions like *according to the law* can be used to *specify* the flavor. We observe that in Koryak adverbial markers of manner (‘openly’, ‘with words’, ‘secretly’, ‘to self’) seem to play a similar role, as shown in (14).

- (14) A teacher is always complaining to his wife about how bad his students are. One day, the principal asks him about his students, and he tells him that they are great.
- a. inenyəjulevəčʔən ivi                      əno əninew jejyučewŋəʔu metʔaŋ kojajyočawŋəʔaŋ ʔam  
 teacher                      *ivək*.3SG.PST that his students well study.3PL.PRS but  
 #(činin) ivi                      əno əčču qekwaŋ kojajyočawŋəʔaŋ.  
 self                      *ivək*.3SG.PST that they badly study.3PL.PRS  
 ‘The teacher said that his students studied well but thought to himself that they studied badly.’
- b. inenyəjulevəčʔən ivi                      əno əninew jejyučewŋəʔu qekwaŋ kojajyočawŋəʔaŋ ʔam  
 teacher                      *ivək*.3SG.PST that his students badly study.3PL.PRS but  
 #(?ojaŋ) ivi                      əno əčču metʔaŋ kojajyočawŋəʔaŋ  
 openly *ivək*.3SG.PST that they well study.3PL.PRS  
 ‘The teacher thought that his students studied badly but openly said that they studied well.’

Formally, we will model the flavor variation of *ivək* as a version of (i).<sup>17</sup> We use a free (modal-base-like) variable *i* that gives a set of possible words for an agent *x* and a world *w*, as

<sup>15</sup>RMD would derive this with a clause-level existential closure.

<sup>16</sup>In the system of RMD, Option 2 could be obtained by applying the existential closure below negation.

<sup>17</sup>Roger Schwarzschild (p.c.) points out that negation could be used to rule out a disjunctive fleshing-out of option (i). We indeed observe that the example below is infelicitous. If *ivək* denoted ‘The teacher thought or said that the students studied badly’, then negating the disjunction would rule out both thinking and saying, which would be inconsistent with the continuation that the teacher believed that the students studied badly.

(a) Context: The school principal goes into the classroom of a teacher whose students are doing poorly in



in (15).<sup>18</sup> The flavor restriction comes in as a presupposition on the variable, following RMD and others, so that for *ivək* it is either the set of *x*'s belief worlds at *w* ( $\mathcal{B}_w^x$ ) or the set of *x*'s sayings at *w* ( $\mathcal{S}_w^x$ ). The rest of the notation is as in (11).

- (15) Denotation of *ivək* (final):  $\llbracket ivək \rrbracket^{c,g,w} = \lambda i \lambda C \lambda p \lambda x: (i(x)(w) = \mathcal{B}_w^x \vee i(x)(w) = \mathcal{S}_w^x) \wedge (C = \{f \mid f(i(x)(w)) = i(x)(w)\} \vee C = \{f \mid f(i(x)(w)) \subseteq i(x)(w) \wedge f(i(x)(w)) \neq \emptyset\})$ .  
 $\exists f \in C [\forall w' \in f(i(x)(w)) [p(w') = 1]]$

## 4 Bouletic flavor via the embedded clause

This section makes an empirical and a theoretical point. Empirically, we show that the apparent bouletic flavor of *ivək* ('hope', 'fear', 'wish') is in fact due to the semantics of *ivək* in combination with material in the embedded clause. This has been previously argued also for the Navajo *nizín* (Bogal-Allbritten 2016), and it gives support for the idea (Heim 1992, von Fintel 1999a) that bouletic verbs like *hope* or *wish* encode a doxastic component, contra Anand and Hacquard (2013) for 'wish'. The technical contribution of this section is to show that we can arrive at the bouletic meaning at LF even if the doxastic quantifier (*ivək* on the doxastic interpretation) is not a clause-mate with the preference component. In this respect we differ from Bogal-Allbritten, who takes *nizín* to be a predicate over situations (and not a quantifier).

Consider the difference between *γajmatək* ('want/wish') and *ivək* in (16). The bouletic meaning of the former but not the latter is found in nominalizations.

- (16) a. *γajmat-γəjɣən* (*γajmat*-NMLZ.ABS.SG) 'wish/desire'  
 b. *ek-wəjɣən* (*ivək*-NMLZ.ABS.SG) 'utterance, thought, sth. allowed, \*hope, \*fear, \*wish'

Furthermore, when *ivək* embeds two clauses, as below, these two clauses can differ with respect to the presence or absence of the bouletic meaning, showing that the bouletic meaning is contributed to by the embedded clause – see Bogal-Allbritten (2016, pp. 149–151) for argumentation (including a discussion of gapping) and the use of this test with English.

- (17) *təkivəŋ* [əno tatjana kotvaŋ novosibirskək] to [əno ečyi kukəčviŋletəŋ]  
*ivək*.1SG.PRS that Tatiana be.3SG.PRS in.Novosibirsk and that today happy.3SG.PRS  
 'I think that Tatiana is in Novosibirsk and I hope that she is happy today.'
- (18) Hewngyto and Qechghylqot are competing in a race, and I want Hewngyto to win.  
*təkivəŋ* [əno ɣewŋəto jenalvatəŋ] ɣam [əno ewənčam qečɣəlqot jenalvatəŋ]  
*ivək*.1SG.PRS that H. win.3SG.FUT but that nonetheless Q. win.3SG.FUT  
 'I hope that Hewngyto will win, but I allow that nonetheless Qechghylqot will win.'
- (19) ɣewŋəto kinvəŋ [əno mełlo mitʔajin ] to [iwke nəʔəŋawtəŋəŋ ənək]  
 Hewngyto *ivək*.3SG.PRS that Melljo beautiful and if.only marry.3SG.CF her  
 'Hewngyto thinks that Melljo is beautiful and wishes he would marry her.'

class, and asks the teacher how the students are doing. The teacher doesn't want to disappoint him, so he says, 'The students are doing well'.

inenɣəjulevəčʔən uŋɣe iwke itti, əno əninew jeɣyucəwŋəlʔu qekwaŋ kojajyočawŋələŋ, ɣam ŋanko  
 teacher NEG *ivək* AUX.3SG.PST that his students badly study.3PL.PRS but then  
 ləmalave, əno qekwaŋ kojajyočawŋələŋ.  
 believe.3SG.PST that badly study.3PL.PRS

'The teacher did not say that his students studied badly, but he believed then that they studied badly.'

<sup>18</sup>Adopting this for attitude verbs is not new, see von Fintel (1999a).

For the doxastic-assertive reading of the conjuncts, however, preliminary results in (20) show the opposite behavior.<sup>19</sup> This is consistent with our proposal since we do not take the doxastic-assertive distinction to be contributed to by the embedded clause.

- (20) A principal enters the classroom of a teacher whose students are doing poorly in class and asks him how the students are doing. The teacher doesn't want to disappoint the principal, so he says 'The students are doing well'.

#inenyäjulevæŋæn ivi                      əno ənnew jejyufæwŋəŋu metʔaŋ kojajyufæwŋəŋaŋ ʔam əno  
 teacher                      *ivək*.3SG.PST that his            students            well    study.3PL.PRS    but that  
 qekwaŋ kojajyufæwŋəŋaŋ  
 badly    study.3PL.PRS  
 'The teacher said that his students are studying well but thought that they were studying badly.' (intended)

We do not have the space to discuss the distinction between 'hope' and 'fear'.<sup>20</sup> Note, however, that (21) shows that the orientation of the preference is not specified in the covert item that we end up placing in the embedded clause.<sup>21</sup> This test is also based on Bogal-Allbritten (2016).

- (21) Hewngyto and Vanja are in a race. Qotaw and I have bet money on the winner: I bet money on Hewngyto, and Qotaw bet on Vanja.

muji    qotaw    mətkivəŋ    [amu    ʔewŋəto    jenalvatəŋ    ʔijek]  
 we.two    Qotaw    *ivək*.1DU.PRS    might    Hewngyto    win.3SG.FUT    in.the.race  
 'I hope and Qotaw fears that Hewngyto will win the race.'

We propose that the bouletic meaning arises from the material in the embedded clause. For 'wish' we will simply encode it into the semantics of the counterfactual mood while for 'hope' we will attribute it to a covert item.<sup>22</sup>

Our technical contribution is to unpack the semantics many have proposed for verbs like *wish* and distribute it between *ivək*, since it can be a doxastic quantifier, and an item in the embedded clause that encodes a preference for the prejacent. The challenge is that it is standardly assumed that preferences are not evaluated in the belief worlds, see  $>_w^x$  in (22), which means that the item that we propose for the embedded clause needs access to the matrix world of evaluation.

- (22) If defined,  $\llbracket \text{wish} \rrbracket^{g,c}(\geq, p, x, w) = 1$  iff  $\forall w' \in \mathcal{B}_w^x : \text{SIM}(w', \text{rev}_p(\mathcal{B}_w^x) \cap p) >_w^x w'$  (Crnič 2011, p. 75) (based on Heim 1992, p. 204)

<sup>19</sup>This is perhaps unsurprising given that adverbial help was needed when we had two *ivək*'s, in (14).

<sup>20</sup>We also note briefly here that various epistemic adverbials can express a speaker-oriented desire reading in matrix clauses (e.g. 'I hope it'll be hot' is translated to Koryak as *amu jatyəlpəŋ* (might be.hot.3SG.FUT)).

<sup>21</sup>The example contains the adverb *amu*, which facilitates the possibility reading. There are also adverbs that specify the direction of the preference, e.g. *wəjɪnvan* 'fortunately', which is only acceptable with the 'hope' reading: (a) *təktivəŋ wəjɪnvan mitiw təjəpəŋətəkuŋəŋ wuččɪn vetyəjɪŋəŋ* (*ivək*.1SG.PRS fortunately tomorrow finish.1SG>3SG.FUT this work) 'I hope that I will finish this work tomorrow' versus (b) *təktivəŋ (#wəjɪnvan) mitiw qəjəŋ məpəŋətəkuŋəŋ wuččɪn vetyəjɪŋəŋ* (*ivək*.1SG.PRS fortunately tomorrow NEG.FUT finish.1SG>3SG this work) 'I fear that I will not finish this work tomorrow' (intended).

<sup>22</sup>A more precise analysis would build the meaning of 'wish' from the counterfactual mood and 'want', see Iatridou (2000). Since *ivək* does not mean 'want', we will not encode a connection between the 'hope' reading and the 'wish' reading, but will provide two separate lexical entries. Furthermore, we have not yet uncovered a fear-version of 'wish' (possibly due to the lack of a suitable matching item in Russian, see §1.1), where 'I *ivək* that I was happy' would presuppose that I am not happy and assert that being happy is dispreferable.

To solve this, we build on Yalcin’s (2007) idea that the index of evaluation contains (in addition to the world of evaluation) an information state (a set of worlds) and that attitude verbs shift this parameter (e.g. to the set of belief worlds). Instead of  $\llbracket \cdot \rrbracket^{g,w}$  (where  $g$  is the assignment function), Yalcin proposes  $\llbracket \cdot \rrbracket^{g,\langle w,s \rangle}$  where  $s$  is an information state. The information state is contextually determined in the matrix and is shifted by a verb like *believe* to, say,  $\mathcal{B}_w^x$  (the set of  $x$ ’s beliefs at  $w$ ). The modification we propose is to replace this notion of an information state (a set of worlds) with the triple that produces it  $\llbracket \cdot \rrbracket^{g,\langle w,\langle a,v,I \rangle \rangle}$ : the information state holder ( $a$ ), the world from which the state is generated ( $v$ ), and the way in which it is generated ( $I$  of type *esst*). That is, replace  $\mathcal{B}_w^x$  with  $\langle x,w,\mathcal{B} \rangle$ . Since we use subset selection functions like  $f$ , we replace  $f(\mathcal{B}_w^x)$  with  $\langle x,w,\lambda y\lambda v.f(\mathcal{B}_v^y) \rangle$ . This allows the counterfactual mood to access the matrix world of evaluation, as in (23). The denotation of *ivək*, in (24), changes only in that the information state gets updated.<sup>23</sup>

$$(23) \quad \text{If defined, } \llbracket \text{CF} \rrbracket^{c,g,\langle w',\langle a,v,I \rangle \rangle}(p) = 1 \text{ iff } \text{SIM}(w', \text{rev}_p(\mathcal{I}_v^a) \cap p) >_v^a w'$$

$$(24) \quad \llbracket \text{ivək} \rrbracket^{c,g,\langle w',\langle a,v,I \rangle \rangle}(i)(C)(p)(x) \text{ is defined only if } i(x)(w) = \mathcal{B}_w^x \text{ or } i(x)(w) = \mathcal{S}_w^x, \text{ and } \\ C = \{f \mid f(i(x)(w)) = i(x)(w)\} \text{ or } C = \{f \mid f(i(x)(w)) \subseteq i(x)(w) \wedge f(i(x)(w)) \neq \emptyset\} \text{ and,} \\ \text{if defined, is true iff } \exists f \in C [\forall w' \in f(i(x)(w)) [p(\langle w', \langle x,w,\lambda y\lambda v''.f(i(y)(w'')) \rangle \rangle) = 1]]$$

By contrast, the ‘hope’ and ‘fear’ readings of *ivək* have no obligatory overt reflex in the embedded clause, recall (2). We postulate a covert item DES in the embedded clause that plays the analogous role to the counterfactual mood (cf. fn. 22).

$$(25) \quad \text{If defined, } \llbracket \text{hope} \rrbracket^{g,c}(\geq, p, x, w) = 1 \text{ iff } \forall w' \in \mathcal{B}_w^x : \text{SIM}(w', \mathcal{B}_w^x \cap p) >_w^x \text{SIM}(w', \mathcal{B}_w^x \setminus p) \\ (\text{Crnić 2011, p. 76}); \text{ (roughly, for any belief world } w': x \text{ prefers } p \text{ at } w')$$

$$(26) \quad \text{If defined, } \llbracket \text{DES} \rrbracket^{c,g,\langle w',\langle a,v,I \rangle \rangle}(p) = 1 \text{ iff} \\ [\text{SIM}(w', \mathcal{I}_v^a \cap p) >_v^a \text{SIM}(w', \mathcal{I}_v^a \setminus p)] \vee [\text{SIM}(w', \mathcal{I}_v^a \setminus p) >_v^a \text{SIM}(w', \mathcal{I}_v^a \cap p)] \\ \text{(roughly, } a \text{ prefers } p \text{ at } w' \text{ or } a \text{ disprefers } p \text{ at } w', \text{ where } a \text{'s preference is set at } v)$$

The disjunction in (26) is one way of encoding the idea that *ivək* seems to be unspecified for the direction of the preference (that is, between ‘hope’ vs ‘fear’), see (21). We do not know whether *ivək* also has a ‘mixed-feelings’ reading. This reading would be one where in some some doxastic worlds  $p$  is preferred and in others  $\neg p$  is preferred. If this reading turns out to not be available, one can place a homogeneity condition (as a definedness condition) on the preference ( $>$ ) relation.

## 5 Conclusion

We have documented a variable-force (existential, universal) variable-flavor (doxastic, assertive) attitude verb *ivək* from Koryak. *Ivək* joins the Washo variable-force variable-flavor modal *e?* (Bochnak 2015a) in counterexemplifying a proposed universal on modal items that bans variation in both force and flavor. What is new is that we have shown that this generalization does not hold for attitude verbs. *Ivək* also joins the Navajo *nízin* (Bogal-Allbritten 2016) in

<sup>23</sup>We use intensional FA. Example: If defined,  $\llbracket \text{John} [\llbracket \text{ivək } i \rrbracket C] [\text{that CF it's raining}] \rrbracket^{c,g,w,\langle a,v,I \rangle} = 1$  iff  $\llbracket \text{ivək} \rrbracket^{c,g,w,\langle a,v,I \rangle}(g(i))(g(C))(\llbracket \text{that CF it's raining} \rrbracket_e^{c,g})(\text{John}) = 1$  iff  $\exists f \in g(C) \forall w' \in f(g(i)(\text{John})(w)) : \llbracket \text{that CF it's raining} \rrbracket_e^{c,g}(w', \langle \text{John}, w, \lambda x\lambda v.f((g(i)(x)(v))) \rangle) = 1$  iff  $\exists f \in g(C) \forall w' \in f(g(i)(\text{John})(w)) : \llbracket \text{that CF it's raining} \rrbracket^{c,g,\langle w',\langle \text{John},w,\lambda x\lambda v.g(C)((g(i)(x)(v))) \rangle \rangle} = 1$  iff  $\exists f \in g(C) \forall w' \in f(g(i)(\text{John})(w)) : \text{SIM}(w', \text{rev}_\varphi(f(g(i)(\text{John})(w))) \cap \varphi) >_w^{\text{John}} w'$  ( $\varphi$  is  $\lambda w$ .it’s raining at  $w$ ) (If  $g(C) = C_{id}$ ,  $g(i) = \mathcal{B}$ , it amounts to:  $\forall w' \in \mathcal{B}(\text{John})(w) : \text{SIM}(w', \text{rev}_\varphi(\mathcal{B}(\text{John})(w))) \cap \varphi >_w^{\text{John}} w'$ )

showing that there are languages that split the bouletic flavor at LF, with a separate doxastic component (this gives support to the analyzes that base the bouletic meaning on a doxastic component). What is new is that we proposed a very different technical implementation of this interaction since, as we argued, the quantificational force comes from *ivək*.

## References

- Pranav Anand and Valentine Hacquard. Epistemics and attitudes. *Semantics and Pragmatics*, 6:1–59, 2013.
- Ryan Bochnak. Variable force modality in Washo. In Thuy Bui and Deniz Özyıldız, editors, *Proceedings of NELS 45*, pages 105–114, 2015a.
- Ryan Bochnak. Underspecified modality in Washo. In Natalie Weber and Sihwei Chen, editors, *Proceedings of Workshop on the Structure and Constituency of Languages of the Americas (WSCLA) 18*, pages 3–17, 2015b.
- Elizabeth Bogal-Allbritten. *Building Meaning in Navajo*. PhD thesis, UMass Amherst, 2016.
- Luka Crnić. *Getting Even*. PhD thesis, Massachusetts Institute of Technology, 2011.
- Amy Rose Deal. Modals without scales. *Language*, 87:559–585, 2011.
- Kai von Fintel. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, 16:97–148, 1999a.
- Kai von Fintel. Quantifier domain selection and pseudo-scope. Handout from a talk at the Cornell Conference on Theories of Context Dependency (<http://mit.edu/fintel/fintel-1999-cornell-context.pdf>), 1999b.
- Irene Heim. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9(3):183–221, 1992.
- Sabine Iatridou. The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31(2):231–270, 2000.
- Angelika Kratzer. What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy*, 1:337–355, 1977.
- Angelika Kratzer. The notional category of modality. In Hans-Jürgen Eikmeyer and Hannes Rieser, editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter, Berlin, 1981.
- Angelika Kratzer. Modality. In Arnim von Stechow and Dieter Wunderlich, editors, *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, pages 639–650. de Gruyter, Berlin, 1991.
- David Lewis. Semantic analyses for dyadic deontic logic. In Sören Stenlund, editor, *Logical Theory and Semantic Analysis: Essays Dedicated to Stig Kanger on His Fiftieth Birthday*, pages 1–14. Reidel, 1974.
- Lisa Matthewson, Hotze Rullmann, and Henry Davis. Modality in St’át’imcets. In *40th international conference on Salish and neighbouring languages*, pages 93–112, 2005.
- Maša Močnik. Slovenian ‘dopuščati’ and the semantics of epistemic modals. In *Formal Approaches to Slavic Linguistics (FASL) 27*, Ann Arbor, 2019a. Michigan Slavic Publications.
- Maša Močnik. Embedded epistemic modals pragmatically. In M.Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, and Cristina Real-Puigdollers, editors, *Proceedings of Sinn und Bedeutung 23*, volume 2, pages 197–206, Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), 2019b.
- Fabrice Nauze. *Modality in typological perspective*. PhD thesis, University of Amsterdam, 2008.
- Hotze Rullmann, Lisa Matthewson, and Henry Davis. Modals as distributive indefinites. *Natural Language Semantics*, 16:317–357, 2008.
- Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975.
- Jozina Vander Klok. Javanese modals. In *Proceedings of the 2008 Annual Conference of the CLA*, 2008.
- Emanuel Viebahn and Barbara Vetter. How many meanings for ‘may’? The case for modal polysemy. *Philosophers’ Imprint*, 16(10), 2016.
- Seth Yalcin. Epistemic modals. *Mind*, 116:983–1026, 2007.

# Are conservative quantifiers easier to learn? Evidence from novel quantifier experiments \*

Jennifer Spenader<sup>1</sup> and Jill de Villiers<sup>2</sup>

<sup>1</sup> University of Groningen, The Netherlands

[j.k.spenader@rug.nl](mailto:j.k.spenader@rug.nl)

<sup>2</sup> Smith College, Massachusetts, U.S.A.

[jdevilli@smith.edu](mailto:jdevilli@smith.edu)

## Abstract

Natural language quantificational determiners seem to always be conservative. Some researchers suggest that one explanation for this link between syntactic form and semantic interpretation may be due to conservative quantifiers being easier to learn. Previous experimental research showed that children learned a conservative novel quantifier (i.e. *not all*) more easily than a novel non-conservative quantifier (i.e. *not only*) [Hunter and Lidz, 2013].

In a series of four experiments we further investigated this learnability claim. Experiment 1 replicated the original study, but with adult participants. We found no evidence that the conservative novel quantifier was easier to learn. Experiment 2 replicated and extended the original study, testing children. Again, we found no difference between the quantifiers. Experiments 3 and 4 introduce a new training and testing method, situation verification with correction. In Experiment 3, we tested this new method with known quantifier meanings, conservative *all* and non-conservative *only*, using a novel quantifier expression. Children were highly successful, and there was no difference between the two quantifier meanings. In Experiment 4, we used the same method to teach children the novel conservative and non-conservative meanings *not all* and *not only*. Most children were unable to learn either quantifier, and we again found no difference between the conservative and non-conservative quantifier in terms of learnability. In conclusion, in four experiments, we failed to replicate findings that conservative quantifiers are easier to learn. These results suggest that a learnability advantage is not a plausible explanation for conservativity's universality.

## 1 Introduction

It is a linguistic universal that quantificational determiners like *all* and *some* are always conservative (Keenan and Stavi, 1986; Barwise and Cooper, 1981). This means interpreting (1) or (2) only requires examining the restrictor set (pirates) and the intersection between the restrictor and scopal set (pirates with treasure chests). Other individuals with treasure chests are irrelevant. Syntactically, quantificational determiners also form a syntactic constituent with their restrictor set. In contrast, non-conservative quantifiers, like *always* and *only* (3), are adverbs, and do not form a syntactic constituent with their restrictor. To verify (3), we must first identify relevant non-pirates and then check that none of them possess treasure chests.

- (1) All pirates have treasure chests.
- (2) Some pirates have treasure chests.

---

\*Thanks to Anna de Koster and the members of the Language Acquisition Lab in Groningen for useful comments on the poster associated with this paper.

- (3) Only pirates have treasure chests.

Verifying the non-conservative quantifier may be more complex because syntactic information does not help in identifying the restrictor set, and the interpretation requires looking for the complement of the mentioned set, rather than the explicitly mentioned set. Since languages include both conservative and non-conservative quantifiers, where does this syntactic restriction that quantificational determiners are always conservative come from? Its typological universality has led to speculation that conservativity makes quantificational determiners easier to learn.

Hunter and Lidz [2013] investigated this claim experimentally by teaching young children novel quantifier meanings. To create two comparable novel quantifiers, they used the negations of two known quantifier meanings, the novel conservative *not all* and the novel non-conservative quantifier *not only*. Using a picture presentation method for training and testing, children were taught that a novel quantifier expression, *gleeb*, had one of these meanings. They found better performance for the conservative quantifier than for the non-conservative quantifier, suggesting that conservative quantifiers are easier to learn.

These experimental results suggest that conservative quantifiers are easier to learn. However, using Recurrent Neural Network computational models and semantic features as input, Steinert-Threlkeld and Szymanik [2019] were unable to find a learnability advantage for conservative quantifiers. They point out that this is unsurprising; without syntactic information there was actually no difference between the conservative/non-conservative quantifiers. They note that their results are consistent with a proposal by Romoli [2015] that determiner conservativity is a consequence of the syntax-semantic interface. Simplified, if quantifier raising is obligatory, the way in which sentences are constructed naturally leads to non-conservative quantifier meanings that, regardless of definition, will either be contradictory, tautological, or completely equivalent to existing quantifier meanings. This work suggests that conservativity is not universal because of a conceptual difference in inherent complexity, and therefore learnability, but because the syntax-semantics interface works similarly across languages in a way that makes non-conservative determiner meanings trivial.

We focus on experimentally investigating the learnability issue in four experiments. In Experiment 1, we replicated Hunter and Lidz [2013] study with adults. In Experiment 2, we replicated and extended it with children. Based on these results, we conclude that the method used in these and the original experiment were unlikely to have taught quantifier meanings, which means the results cannot support the conclusion that conservative quantifiers are easier to learn. We then present a new training and testing method, situation verification with correction, which solves a number of problems with the picture-presentation method. In Experiment 3, we show that this new method succeeds in teaching known quantifier meanings with novel expressions with just ten training items. This experiment serves as a control for teaching children novel quantifiers. Experiment 4 then used the new method to teach novel quantifier meanings with novel expressions. In both Experiment 3 and Experiment 4 we found no difference between the quantifiers.

## 2 Previous experiments teaching conservative and non-conservative quantifiers

Novel word learning experiments have been long used as a method to show that adults and children are sensitive to certain syntactic features when learning novel words (Naigles and Hoff-Ginsberg, 1995, Fisher et al., 2010), so to show a learnability difference between conservative



and non-conservative quantifiers, a novel word task is an obvious choice. Hunter and Lidz [2013] choose *not all* and *not only* as their novel quantifier meanings. Neither meaning exists as a single word quantifier in English or Dutch,<sup>1</sup> and while negation often adds complexity to interpretation, since they both incorporate negation they are still comparable. The set-theoretic definition that they use for the interpretation of the novel quantifiers is presented below:

- (4) a. conservative: gleeb (A,B) is true iff  $A \not\subseteq B$ .  
 b. non-conservative: gleeb' (A,B) is true iff  $B \not\subseteq A$ .

Hunter and Lidz [2013] found better performance for the conservative quantifier than for the non-conservative quantifier. After five training instances, the mean accuracy for the conservative determiner on five test items was 4.1, significantly better than chance ( $\chi^2 = 74.160$ ,  $df=5$ ,  $p < 0.0001$ ) while for the non-conservative determiner the mean accuracy was only 3.1, not significantly different from chance ( $\chi^2 = 6.640$ ,  $df=5$ ,  $p > 0.25$ ). Even stronger evidence comes from their finding that half the children exposed to the conservative quantifier had perfect responses to the five test instances. For the non-conservative quantifier, only one child gave perfect responses.

The original study done by Hunter and Lidz [2013] aimed to show that children will be more easily able to learn a conservative quantifier, either because it is easier to verify or because it is conceptually easier. They chose to work with children between the ages of 4-5. At this age, children are still in the process of acquiring the full meaning of existing natural language quantifiers, but should be able to easily interpret the basic meaning of the non-negated counterparts *all* and *only*.

They used a method that teaches a contrast in an indirect way, the “picky puppet” paradigm. This method links a contrast to what a puppet likes and does not like. We outline the method here since we also use it in both our Experiment 1 and 2. Large cards, printed in color, displayed various scenes on a beach and a grassy area next to the beach, with boys and girls (See Figures 1 and 2). A target sentence is either a correct or incorrect description of the cards. The same target sentence was used for all items, (5). Children were introduced to a puppet who likes some cards but dislikes others. The puppet is shy, and only whispers their likes and dislikes to the experimenter. For each card the children were told whether the puppet likes the card or not, with an explicit statement that was framed around the target sentences, e.g. ((6-a) or (6-b)):

- (5) Gleeb girls are on the beach.  
 (6) a. Puppet told me that she likes this card because gleeb girls are on the beach.  
 b. Puppet told me that she doesn't like this card because it is not the case that gleeb girls are on the beach.

Liked cards were placed in a pile under a green smiley face, and disliked cards were placed under a red smiley face. After this training, participants were presented with five new cards to sort. All training and testing cards were presented in the same fixed order.

With conservative *gleeb* meaning *not all*, sentence (5) presented with Figure 1 will be true, because some of the girls are on the grassy area, but the same sentences will be false with Figure 2, because all the girls are on the beach. With non-conservative *gleeb* meaning *not only*, sentence (5) will be false with Figure 1, because only girls are on the beach, but it will be true with Figure 2 because some of the boys are also on the beach.

<sup>1</sup>Or any other language, as far as we know

**Figure 1:**

Gleeb=*Not all*: True  
 Gleeb=*Not only*: False

**Figure 2:**

Gleeb=*Not all*: False  
 Gleeb=*Not only*: True

### 3 Experiment 1: Replication with Adults

18 English speaking adults participated in Experiment 1 ( $M_{age}=21;1$ , Range=18;6-26;4, 17 female). In a between subjects design, 9 were taught conservative *not all*, and 9 were taught non-conservative *not only*, using the same materials, method and order of presentation used in Hunter and Lidz [2013].<sup>2</sup>

#### 3.1 Results and Discussion Experiment 1

Unlike the children in the experiment by Hunter and Lidz [2013], our adults were not very successful at learning novel quantifier meanings with this method, with a mean accuracy of 56% with the conservative quantifier and 69% with the non-conservative. A mixed-effects model analysis also found no significant difference between the quantifier types ( $p>0.11$ ).

We did however find a difference between the quantifiers in terms of the number of subjects that showed perfect accuracy. 4 subjects in the non-conservative condition (taught *not only*) were perfect (and one had 80% correct), but only 1 participant in the conservative condition had a perfect score. In debriefing the participants that gave perfect responses to the *not only* condition stated that they had noticed that if there are boys on the beach, the card was ‘liked’ by the puppet. They then simply sorted cards with boys on the beach as ‘liked’. Following this strategy doesn’t require any reference to the target sentences. Another explanation given by some participants in the *not only* condition was that *gleeb* meant *non-*. These participants then also followed a strategy of looking checking if *gleeb girls=non-girls* were on the beach, which is in essence the same strategy.

There is also a potential simple non-linguistic strategy for *not all*, which is to look for girls on the grassy area. However, no participants reported using such a strategy, perhaps in part because few participants did well with the conservative quantifier.

The strategies mentioned by successful participants in the *not only* condition highlights how correct responses can be achieved for this task without any reference to language or without interpreting *gleeb* as a quantifier. These strategies are in part possible because all items used the same target sentence. Additional target sentences that refer to another character, e.g. *Gleeb boys are on the beach* would make the first strategy impossible.

Our results do not support the conclusion that conservative quantifiers are easier to learn, and in fact seems to support the opposite conclusion. But the adults are still much worse than the children. It is certainly the case that sometimes children are able to show performance

<sup>2</sup>Participants were not permitted to examine already sorted cards to make their decision, though many asked if they could.





**Figure 3:** Experiment 2, Test item 11. This item was designed to distinguish a strategy of just looking for boys on the beach from a quantifier interpretation for *not only* items.

# of cards	Quantifier meaning	# sorted correctly
5	Conservative	3.0
	Non-conservative	3.4
10	Conservative	5.6
	Non-conservative	6.7
11	Conservative	6.0
	Non-conservative	7.2

**Figure 4:** Experiment 2 results: Number of cards correctly sorted for each quantifier, out of 45, 10 and 11 test items. No differences significantly better than chance ( $\chi^2$  tests.)

that adults are not able to achieve, for various other reasons (e.g. overthinking). Given the abstract nature of the task, which we replicated exactly the original study used with children, may have encouraged adults to look for a non-linguistic strategy. Remember, the task simply stated that the puppet liked or did not like certain situations because of a linguistically quite complex reason (e.g. see (6-a) and (6-b)). This makes the target sentence easy to ignore. Is it possible that the children were attending to the linguistic information but that adults instead focused on non-linguistic strategies because the training method de-emphasized meaning? To investigate this further we replicated and extended the original experiment by Hunter and Lidz [2013] in Experiment 2

## 4 Experiment 2: Replication and Extension with Children

In Experiment 2 we had two goals, first to confirm the original results, and second, to make sure that responses given in the testing are reliable. The original study had 5 training items and 5 test items. In order to get a stronger confirmation of children's knowledge, we extended the testing part of the task to include 6 new test items. Each of these 5 test items was a minimal variation on the original 5 test items. The 6th new item differed by adding monkeys on the beach (and boys on the grassy area, see Figure 3). This picture should be true with *not only*, but if children followed the successful strategy of some of the adults in Experiment 1 and simply looked for boys on the beach, then they will incorrectly reject this item.

Using a between subjects design, 21 English-speaking children ( $M_{age}=5;0$ , 7 girls) participated. 10 children were taught *gleeb* with the novel meaning *not all* and 10 children were taught *gleeb* with the novel meaning *not only*. Experiment 2 duplicated the materials, procedure and order of presentation of Hunter and Lidz's original work and Experiment 1 for the training items and the first five testing items. Immediately following the five original test items, testing items 6-11 were presented, in a different random order for each participant.

### 4.1 Results and Discussion Experiment 2

Children were equally bad at either quantifier. For the first 5 test items which were an exact replica of the Hunter and Lidz [2013] study, participants had an accuracy with conservative

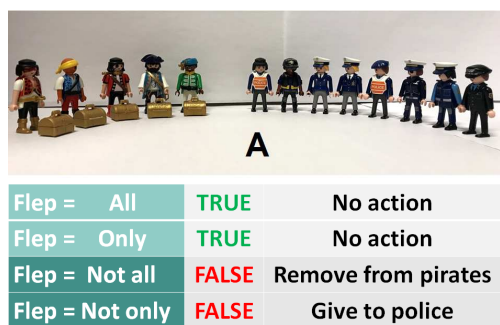


Figure 5: Type A Situation

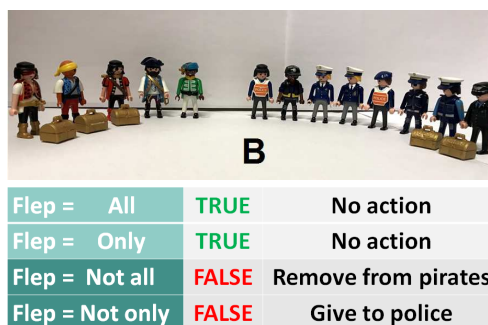


Figure 6: Type B Situation

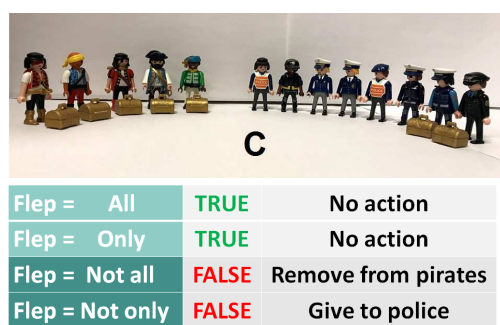


Figure 7: Type C Situation

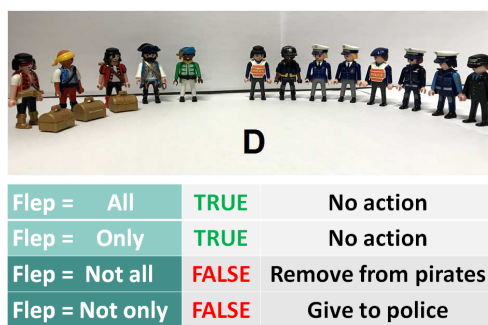


Figure 8: Type D Situation

*gleeb* of 60% and for non-conservative *gleeb* 68%, sorting 3.0 and 3.4 cards correctly out of five respectively. A mixed-effects model analysis found no significant difference between the quantifiers ( $p > 0.43$ ). The number of cards correctly sorted out of 10 and 11 are also given in Table 4. Also using the analysis from Hunter and Lidz [2013] comparing the cards sorted to chance with  $\chi^2$ , there was no significant difference between the quantifiers.

Subjects were barely above chance in both Experiment 1 and 2, suggesting the picture presentation method fails to teach quantifier meanings. Instead, we developed an act-out alternative where true and false situations were presented for judgment to a puppet. Crucially, false descriptions were corrected by the puppet by manipulating objects, allowing children to immediately compare true and false situations.

### 5 Experiment 3: Known quantifier meanings using Situation Verification with Correction

To be sure that participants in testing actually understand quantifier meaning, we needed more information than simple picture verification can provide. For this reason we developed a novel training and testing paradigm: situation verification with correction.

In Experiment 3 we used this new method for training and testing with the known quantifiers *all* and *only*, presented as the nonce word *flep*<sup>3</sup> using five pirates and eight police officer figurines.

<sup>3</sup>We replaced *gleeb* with *flep*, as a more natural novel word for Dutch.

- (7) a. Flep piraten hebben een schatkist. Klopt dat?  
 Flep pirates have a treasure chest. Is that true?  
 b. Hebben flep piraten een schatkist?  
 Have flep pirates a treasure chest? (*Do flep pirates have a treasure chest?*)

The training used a puppet and proceeded as follows. First, the experimenter distributed objects to the two different sets of figurines (*police and pirates*). Then the experimenter asked the puppet about how the situation related to the meaning of *flep*, using *flep* in a statement and in a question, e.g. (7-a), (7-b). The puppet judged the descriptive accuracy. If the description was incorrect, the experimenter asked the puppet to correct the situation by either adding or removing objects. When *flep* was used to mean *all*, the puppet gave an object to any mentioned characters that did not have one. When *flep* was used to mean *only* the puppet removed any objects that the non-mentioned characters had. There was always a surplus of objects for giving to figurines or for parking removed objects, as necessary, placed near the figurines for easy access.

16 Dutch native speaking children participated in Experiment 3. Seven children were taught *all* and nine children were taught *only*. Participants were first trained by being asked to observe the interaction between the experimenter and the puppet with 10 training items. Four different situations types were used, see Figures 3-6. Participants were presented with 3 Type A items, 3 Type B items, 2 Type C items and 2 Type D item. This resulted in 5 true items and 5 false items. After training, participants were presented with 10 test items, in the same distribution of types as the training items. Training and testing took approximately 10 minutes per child.

### 5.1 Results and Discussion Experiment 3

The children were highly successful in learning both *all* (98.5% correct) and *only* (96.6% correct), with most children giving perfect responses for all 10 test items. Many spontaneously used *flep* correctly even before the training was completed. There was no difference between the conservative and non-conservative quantifier.

These results showed that the new method, situation verification with correction, was successful in teaching known quantifier meanings with a novel quantifier expression with only 10 training items. It also showed that using 8 and 5 figurines was sufficient for encourage quantifier interpretations. Using this same method, we next attempted to teach children novel quantifier meanings with a novel quantifier expression in Experiment 4.

## 6 Experiment 4: Novel Quantifier Meaning using Situation Verification with Correction

In Experiment 4 we applied our new training method to the novel quantifiers *not all* and *not only* with the novel expression *flep*.

37 Dutch children ( $M_{age}=5;6$ , Range =5;0-6;6, 15 girls) participated in a between subjects design. 18 were taught *not all* and 19 were taught *not only* using the novel expression *flep* and 10 training and 10 test items. The method was the same as Experiment 3 except for the quantifier meaning.

When *flep* was used to mean *not all*, the puppet removed one object from the mentioned character. When *flep* was used to mean *not only* the puppet corrected false situations by giving objects to the non-mentioned character. The same four situation types were used as in Experiment 3. Participants were presented with 3 Type A items, 3 Type B items, 2 Type C

items and 2 Type D items for a total of 10 items each in training and testing. After training, children were asked to take over the role of the puppet, first verifying whether or not the situation presented was true or false, given the sentence and question, and then, if the sentence was false, children were asked to change the situation by manipulating the objects to make it true. Training and testing took approximately 10 minutes per child.

## 6.1 Results and Discussion Experiment 4

Few children were able to learn the novel quantifiers. The mean accuracy for *flep* with the novel quantifier meaning *not all* was 63% while the mean accuracy with the meaning *not only* was 54%. However, these means are not really informative. Examining individual accuracy, for children taught that *flep* meant *not all*, only five children out of 18 were successful, but these children actually had nearly perfect scores. For *flep* meaning *not only*, four children out of 19 were successful, again with nearly perfect scores. Similar to the results of Experiments 1 and 2, there was no difference between quantifiers.

Looking more closely at the children who made errors, a large subset of children picked up on the manipulation necessary for correcting false items, e.g. whether or not objects needed to be removed or added. But these children applied this action indiscriminately, showing they did not understand the intended meaning and also leading to low accuracy scores. The remaining children showed no identifiable pattern.

## 7 General Discussion

In four experiments, we failed to replicate findings that conservative quantifiers are easier to learn, instead finding no difference between the novel quantifiers and children and adult's success in the experiments. While our experiments were designed to test the learnability explanation directly, our results may also be relevant for evaluating other explanations for conservativity's universality.

If conservativity is instead a consequence of the way the syntax-semantic interface constructs meaning, then children's success in recognizing the non-conservative determiner meaning *only* distributionally presented as a determiner in Experiment 3 is surprising. These results show that children could comprehend and produce, even after only a handful of examples, a meaning with a syntactically dubious derivation. This suggests that at least for the ages and tasks tested, syntactic information was not necessary to learn the quantifier meaning. More research, is needed.

Why were children not able to learn the novel quantifier meanings *not all* and *not only* even with the new training method? This was unexpected given the ease with which the children were able to learn the known quantifier meanings in Experiment 3. One possible reason is that the negation makes the quantifier meaning much more complex. It is also well-known that children can have difficulties processing negation.

Our own impression is that the children who learned the quantifiers were children who were particularly good at the kind of puzzle the task represents. If this is the case, then the difficulty most children showed in the current study can be attributed to the conceptual novelty of the the quantifier meanings being taught. If we provide children with more training items, we might see a greater proportion of successful children. In that case, any difference found in the success rates between the two quantifiers would be more meaningful. We plan to investigate this further.

## 8 Conclusion

In four experiments we investigated the claim that natural language quantificational determiners are universally semantically conservative because conservativity makes them easier to learn. Contrary to previous experimental findings, we were unable to find a difference in ease of learnability between conservative and non-conservative quantifiers, both in experiments with adults and with children. We conclude that there is reason to believe that a learnability advantage is not a plausible explanation for conservativity's universality.

## References

- Jon Barwise and Robin Cooper. Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence*, pages 241–301. Springer, 1981.
- Cynthia Fisher, Yael Gertner, Rose M Scott, and Sylvia Yuan. Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(2):143–149, 2010.
- Tim Hunter and Jeffrey Lidz. Conservativity and learnability of determiners. *Journal of Semantics*, 30(3):315–334, 2013.
- Edward L Keenan and Jonathan Stavi. A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9(3):253–326, 1986.
- Letitia R Naigles and Erika Hoff-Ginsberg. Input to verb learning: Evidence for the plausibility of syntactic bootstrapping. *Developmental Psychology*, 31(5):827, 1995.
- Jacopo Romoli. A structural account of conservativity. *Semantics-Syntax Interface*, 2(1):28–57, 2015.
- Shane Steinert-Threlkeld and Jakub Szymanik. Learnability and semantic universals. *Semantics and Pragmatics*, 12:4, 2019.

# Quantifiers in natural language optimize the simplicity/informativeness trade-off\*

Shane Steinert-Threlkeld

Department of Linguistics  
University of Washington  
shanest@uw.edu

## Abstract

While the languages of the world vary greatly, linguists have discovered many restrictions on possible variation. Semantic universals are restrictions on the range of variation in meaning across languages. Recently, in several domains—e.g. kinship terms, color terms—such universals have been argued to arise from a trade-off between simplicity and informativeness.

In this paper, we apply this method to a prominent domain of function words, showing that the *quantifiers* in natural language also appear to be optimized for this trade-off. We do this by using an evolutionary algorithm to estimate the optimal languages, systematically manipulating the *degree of naturalness* of languages, and showing that languages become closer to optimal as they become more natural.

Our results suggest that very general communicative and cognitive pressures may shape the lexica of natural languages across both content and function words.

## 1 Introduction

While the languages of the world vary greatly, linguists have discovered many restrictions on possible variation (Croft, 1990; Hyman, 2008; von Stechow and Matthewson, 2008). *Semantic universals* are restrictions on the range of variation in meaning across languages. Recently, in several domains—kinship terms, color terms—such universals have been argued to arise from a trade-off between simplicity and informativeness (Kemp and Regier, 2012; Kemp et al., 2018). Roughly: a language cannot be both maximally simple (in terms of, e.g. cognitive load) and at the same time maximally informative. Intuitively, a maximally simple language would have a single term, which could not be used to convey significant information. A maximally informative language, on the other hand, would contain individual expressions for every possible thought to be expressed; such a language is highly complex, relying on significant memorization. The general claim: the semantic systems of the world’s languages optimally balance these two competing pressures.

While the aforementioned case studies apply to domains of content words, the historically most prominent domain of semantic universals has been from a domain of function words, namely determiners (Barwise and Cooper, 1981; Peters and Westerstahl, 2006). In particular, the quantifiers expressed by determiners have been argued to have properties like monotonicity, quantitativity, and conservativity. Recent work has offered a different explanation for these

---

\*Many thanks to Milica Denić, Iris van de Pol, and Jakub Szymanik for helpful discussions and to three anonymous reviewers for the Amsterdam Colloquium for helpful comments. Special thanks are due to Wouter Postdijk; the results presented here originate in his ILLC masters thesis, which I had the pleasure of supervising. He wished to not be part of publishing these results in order to focus on an electronic music career. The author has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP/20072013)/ERC Grant Agreement n. STG 716230 CoSaQ.

universals in quantifiers, namely that they arise from a pressure of learnability: quantifiers satisfying the universals are easier to learn than those that do not, and therefore get lexicalized (Steinert-Threlkeld and Szymanik, 2019, 2020). This argument, however, does not rule out the possibility of explaining these universals in terms of the aforementioned trade-off.

In this paper, we argue that the semantic universals for quantifiers can in fact also be seen as arising from the trade-off between simplicity and informativeness. In the next section, we develop methods to make this argument, introducing measures for simplicity and informativeness that can apply to quantifier systems, a method for defining a degree of naturalness for artificial languages, and a novel technique for estimating the set of optimal languages (i.e. Pareto frontier) via an evolutionary algorithm. This new methodology allows us to perform statistical tests on the factors influencing the optimality of quantifier systems. Section 3 presents a regression showing that as languages become closer to being a natural language, they also become closer to being optimal. We then conclude with a discussion of the consequences of the main result and future directions.

## 2 Methods

In order to argue that more natural languages are more optimized for the trade-off, we need methods for generating artificial languages of varying degrees of naturalness and measuring how optimal they are. After presenting some preliminary background about quantifiers and precisely defining the measures of simplicity and informativeness, Sections 2.3 and 2.4 turn to those two tasks. Code for reproducing all of the results in this paper and carrying out further experiments may be found at [https://github.com/shanest/SimInf\\_Quantifiers](https://github.com/shanest/SimInf_Quantifiers).

We are measuring a *degree* of naturalness, instead of just comparing actual natural languages to artificial languages, for two reasons. On the one hand, in the case of quantifiers, unlike the cases of kinship (Kemp and Regier, 2012) and color (Regier et al., 2015), there is not yet an existing catalog of the quantifier systems of any significant number of natural languages. On the other hand, even in the cases where such data does exist, finding a correlation between a degree of naturalness and optimality provides more information and may strengthen the argument by linking to language change/evolution. As languages have changed over time, they became ‘closer’ to our current natural languages; a correlation result of the kind we are developing shows that during this process, they also become closer to optimally trading off the two competing pressures.

### 2.1 Preliminaries

We represent quantifiers—the denotations of determiners like *all*, *some*, *most*, etc.—as sets of models  $\langle M, A, B \rangle$  containing a domain of discourse and two distinguished subsets for the restrictor and nuclear scope (equivalently: a binary relation between those sets) (Barwise and Cooper, 1981; Peters and Westerståhl, 2006; Szymanik, 2016). For example:

$$\begin{aligned} \llbracket \text{every} \rrbracket &= \{ \langle M, A, B \rangle : A \subseteq B \} \\ \llbracket \text{at\_most\_3} \rrbracket &= \{ \langle M, A, B \rangle : |A \cap B| \leq 3 \} \\ \llbracket \text{most} \rrbracket &= \{ \langle M, A, B \rangle : |A \cap B| > |A \setminus B| \} \end{aligned}$$

In what follows, a *language* is a set of quantifiers. For computational reasons—namely because of the exponential growth of the space of possible models—we restrict ourselves to all models up to size 10 in the remainder of this paper.



## 2.2 Measuring Simplicity and Informativeness

Our measure of cognitive simplicity relies on representing quantifiers in a Language of Thought (Feldman, 2000; Piantadosi et al., 2016), i.e. using formulas in a logical language containing operations for set union, intersection, and complementation, as well as for measuring cardinalities and comparing, multiplying, and dividing them. Table 1 shows the entire set of operators used in this paper.

Boolean	Set-Theoretic	Numeric
$\wedge, \vee, \neg$	$\cap, \cup, \subset,   \cdot  $	$/, +, -, >, =, \%$

Table 1: The operators in the grammar for generating quantifiers.

The *complexity* of a quantifier is the length of the *shortest* formula in this language that denotes the quantifier. We found the shortest such formula by exhaustively enumerating all formulas with up to 12 operations and comparing the truth-values across all models up to size 10.<sup>1,2</sup> The complexity of a language is the sum of the complexities of the quantifiers in it. We specify an upper bound on the number of possible quantifiers in a language (10 in our experiments) and divide the sum by this number.

Our measure of informativeness stems from notions of communicative success: a speaker has an intended model that they want to communicate to a listener using the quantifiers in their language (Skyrms, 2010; Kemp et al., 2018). This is captured by the following:

$$I(L) := \sum_{\mathbb{M}} P(\mathbb{M}) \sum_{Q \in L} P(Q|\mathbb{M}) \sum_{\mathbb{M}' \in Q} P(\mathbb{M}'|Q) \cdot u(\mathbb{M}', \mathbb{M})$$

The prior over models, as well as the conditional distributions, are assumed to be uniform where defined (e.g.  $P(Q|\mathbb{M}) = 1/n$  if  $\mathbb{M} \in Q$ , 0 otherwise, where  $n = |\{Q \in L : \mathbb{M} \in Q\}|$  is the number of quantifiers in  $L$  containing  $\mathbb{M}$ ).

This measure captures the following communicative scenario: a speaker has a model ( $\mathbb{M}$ ) in mind, that it wishes to communicate to a listener. To do so, they can use the quantifiers in the language  $L$ . The speaker’s behavior is captured by  $P(Q|\mathbb{M})$ . The listener then guesses a model ( $\mathbb{M}'$ ) that the speaker has in mind, with probability  $P(\mathbb{M}'|Q)$ .

The utility  $u(\mathbb{M}', \mathbb{M})$  measures how good it is for the listener to guess  $\mathbb{M}'$  when the speaker had in mind  $\mathbb{M}$ . We base this on a measure of the distance between models, capturing the notion that non-exact matches can still be better or worse (Jäger, 2007; O’Connor, 2014). More precisely:

$$u(\mathbb{M}', \mathbb{M}) = \frac{1}{1 + d(\mathbb{M}', \mathbb{M})} \quad \text{where} \quad d(\mathbb{M}', \mathbb{M}) = \sum_{X \in A \setminus B, A \cap B, B \setminus A, M \setminus (A \cup B)} \max\{0, |X| - |X'|\}$$

Intuitively, this measure is inversely proportional to how many elements one has to move to transform the listener’s guessed model into the sender’s model (by summing this value across the four ‘zones’ in a model of the form  $\langle M, A, B \rangle$ ).<sup>3</sup> For example, suppose  $\mathbb{M}$  has 3, 4, 2, and 1 elements in  $A \setminus B, A \cap B, B \setminus A, M \setminus (A \cup B)$  respectively, and  $\mathbb{M}'$  has 2, 4, 3, and 1 elements

<sup>1</sup>For memory reasons, we collapse isomorphic models, representing a model  $\langle M, A, B \rangle$  by the cardinalities of the four sets  $A \cap B, A \setminus B, B \setminus A, M \setminus (A \cup B)$ . This prevents us from capturing quantifiers like *the first three* which do not satisfy the universal known as QUANTITY (Steinert-Threlkeld and Szymanik, 2019). Future work will explore methods that relax this assumption while simultaneously addressing the resulting combinatorial explosion.

<sup>2</sup>Using length only is equivalent to using the probability of generating an expression with a PCFG that assigns equal weight to all productions from the same non-terminal.

<sup>3</sup>The addition of 1 in the denominator both prevents division by zero and makes distance-0 models have maximal utility of 1.



in the same zones. We then have that  $d(\mathbb{M}, \mathbb{M}') = 1$ , since moving one element from  $A \setminus B$  to  $B \setminus A$  will make the four zones have the same size in the two models.

## 2.3 Sampling Languages

To answer the question of whether natural languages optimize the trade-off between these two measures, we need to (i) define artificial languages, (ii) identify the natural languages, and (iii) compare how well each does at this optimization. For (i) and (ii), we systematically control ‘how natural’ a language is by biased sampling. A completely random language can be generated by randomly sampling a specified number of quantifiers from the space of all quantifiers generated by our grammar.

While there is no existing dataset of quantifiers across a large set of natural languages, a major cross-linguistic study (Keenan and Paperno, 2012; Paperno and Keenan, 2017) found that all natural language quantifiers belonged to three classes:

- Generalized existential: depending only on  $|A \cap B|$ .  
For example:  $\llbracket \text{at least three} \rrbracket = \{\langle M, A, B \rangle : |A \cap B| \geq 3\}$ .
- Generalized intersective: depending only on  $|A \setminus B|$ .  
For example:  $\llbracket \text{every} \rrbracket = \{\langle M, A, B \rangle : |A \setminus B| = 0\}$ .
- Proportional: comparing  $|A \cap B|$  and  $|A \setminus B|$ .  
For example:  $\llbracket \text{most} \rrbracket = \{\langle M, A, B \rangle : |A \cap B| > |A \setminus B|\}$ .

We call a quantifier *quasi-natural* if it can be expressed in one of the three forms above. And a language will be considered natural if it contains only quasi-natural quantifiers.

Our complete sampling procedure, then, worked as follows: for each number of words between 1 and 10, we generated 8000 languages. Each language of size  $n$  was chosen to have  $m \leq n$  quasi-natural quantifiers, with  $m$  chosen uniformly from  $\{0, \dots, n\}$ . All remaining quantifiers were chosen randomly from the set of all quantifiers whose minimal formula has 12 or fewer operators. We refer to  $m/n$  as the *degree of naturalness* of a language. Thus, a language that has only quasi-natural quantifiers will have a degree of naturalness of 1 (and a language that has no quasi-natural quantifiers will have a degree of naturalness of 0).

## 2.4 Measuring Optimality

For (iii), we need a measure of optimality for a language, to see how it relates to the degree of naturalness. To do this, we measure how close a language is to the *Pareto frontier*, the set of languages which are not dominated (i.e. which have no language both simpler and more informative). The Pareto frontier contains the fully optimal languages: they cannot be made less complex or more informative without becoming worse on the other dimension. Writing  $P$  for the Pareto frontier, we define the optimality of a language as

$$\text{optimality}(L) := 1 - \min_{L' \in P} d(L, L')$$

where  $d$  is the Euclidean distance between points in the plane. This measure takes the *closest* point on the Pareto frontier to a given language. If a language is on the frontier, i.e. is optimal, that minimum distance will be 0, and so the degree of optimality will be 1. Because both communicative cost and complexity range from 0 to 1, the theoretically largest value for the

minimum distance is 1, and so **optimality** also ranges from 0 to 1. To summarize: the degree of optimality of a language increases as it gets closer to the Pareto frontier, the set of optimal languages.

A complication arises when trying to apply this measure: because the space of possible languages is enormous, we cannot exhaustively enumerate it and thereby uncover the *true* Pareto frontier. Moreover, our sampling procedures from the previous section are not guaranteed to uncover the Pareto frontier. Because of this, we need a method to estimate the Pareto frontier without being able to calculate it directly.

To estimate the true Pareto frontier, we used an evolutionary algorithm (Coello et al., 2007). Such algorithms take inspiration from evolutionary processes in that points in a space change over a sequence of generations, with ‘children’ arising via ‘mutation’ from previous points. More importantly, such algorithms are explicitly designed to solve *multi-objective optimization* problems. Since the Pareto frontier can be seen as the set of solutions to the problem of simultaneously optimizing multiple objectives (simplicity and informativeness), these algorithms are well-suited to estimating it.

Our algorithm—provided in full detail in Algorithm 1 in the Appendix—can be intuitively described as follows. We start with an initial seed of randomly generated languages. For some specified number of ‘generations’ we select the *dominant* languages among the current set of languages. Each language then has an equal number of ‘children’ languages (enough to maintain the size of the pool of languages). A child arises from a parent language by some small sequence of ‘mutations’. In our case, this was between 1 and 3 mutations, where a mutation could be: (i) deleting a quantifier from the parent language, (ii) adding a quantifier to the parent language, or (iii) swapping a quantifier in the parent language (i.e. deleting one and adding a new one).

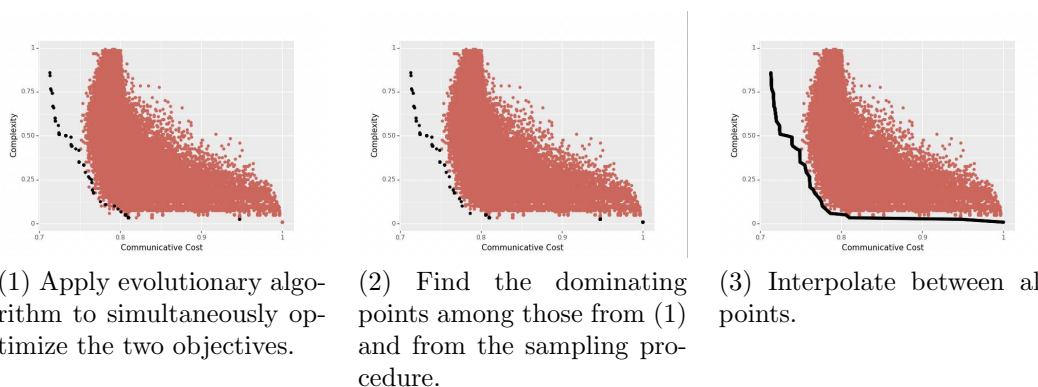


Figure 1: The overall Pareto frontier estimation algorithm, in three steps. The red points are the languages sampled as described in Section 2.3. The black points in panel (3) constitute the final estimate of the true Pareto frontier.

After running the above algorithm for some specified number of generations, we then take the dominant languages from the pool together with the languages we previously sampled, and then linearly interpolate between all of the points to form a smooth and dense frontier.<sup>4</sup> This entire process is depicted in Figure 1.

<sup>4</sup>More sophisticated evolutionary algorithms specify a convergence criterion. We leave the explorations of these refinements to future work.

### 3 Results

The main results can be seen in Figure 2. In this figure, the  $x$ -axis is communicative cost, which is  $1 - I(L)$ , and the  $y$ -axis is complexity. Each point represents a possible language, with the color of a point corresponding to degree of naturalness. The black line is the estimated Pareto frontier, i.e. the set of languages that optimally trade-off between these two factors.

A few things can be observed right away. All of the sampled points that were found to lie on the estimated Pareto frontier (i.e. which dominate all languages both sampled and discovered by the evolutionary algorithm) appear to have a very high degree of naturalness. These are the yellow points on the black frontier, where no brown or blue points (less natural languages) are to be found. Moreover, this seems to be a general trend: it appears that languages with a high degree of naturalness tend to be closer to the Pareto frontier than those with low degrees of naturalness.<sup>5</sup>

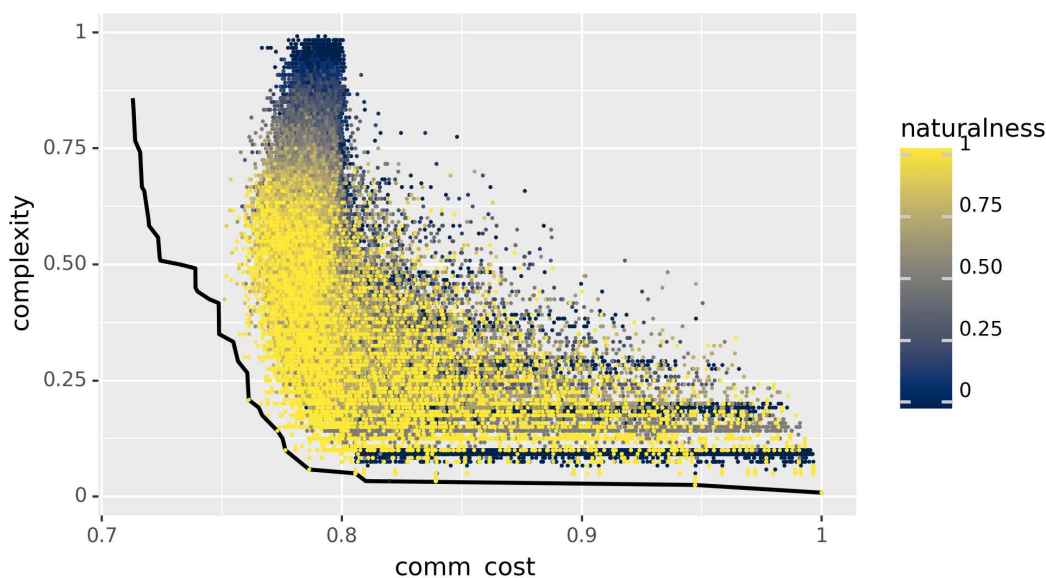


Figure 2: Languages in the space of communicative cost and complexity, colored by their degree of naturalness. Languages with more quasi-natural quantifiers appear to be closer to optimal, as measured by closeness to the (estimated) Pareto frontier, depicted in black.

In virtue of the methods described in the previous section, we can test this appearance statistically: a regression reveals that there is a significant positive correlation between the degree of optimality and degree of naturalness of a language ( $\beta = 0.30$ ,  $t = 88.95$ ,  $p \approx 0$ , 95% CI: [0.293, 0.307]). In other words, as languages become more similar to natural languages with respect to their quantifiers, they come closer to optimally trading off between the competing pressures of cost and complexity.

<sup>5</sup>A possible exception to the general trend lies in the bottom-right corner. Exploring the properties of those languages remains for future work. Thanks to an anonymous referee for this suggestion.

## 4 Discussion

The languages of the world do not express all logically possible meanings, but only a restricted subset thereof. The results of this paper show that in the domain of quantifiers, natural languages appear to be optimizing a trade-off between simplicity and informativeness. We demonstrated this by estimating the Pareto frontier and finding a significant correlation between the degree of naturalness of a language and its closeness to the frontier. This suggests that very general communicative and cognitive pressures may shape the lexica of natural languages across both content and function words (Chemla et al., 2019).

Much work remains to be done. Directly concerning the results in the present paper, a few natural directions emerge. Firstly, while we argued that the statistical analysis presented here provides more information than simply comparing natural and artificial languages, more detailed and systematic documentation of the quantifier systems of the actual languages of the world would provide a stronger empirical foundation for the theory. Secondly, at several junctures, we made concrete modeling choices (e.g. what operators to use in our grammar, the size of models and of languages, the parameters of the evolutionary algorithm). While on quick inspection our main result seems robust, a more detailed analysis of sensitivity to those parameters would be welcome. Thirdly, alternative methods for measuring naturalness can be developed. For instance, Carcassi et al. (2019) provide an information-theoretic measure of the degree of monotonicity of a quantifier, which extends naturally to other universals as well. This raises the question: are more monotone quantifiers closer to the Pareto frontier than less monotone ones? Finally, alternative methods of sampling languages should be explored. Figure 2 shows that our evolutionary algorithm explores a region of the space that the naive sampling procedures do not discover (e.g. in the top-left corner). Because exhaustive enumeration is not feasible, more sophisticated sampling methods that sufficiently explore low-density regions of the space<sup>6</sup> will help make the results more robust.

Most generally, this work raises a number of interesting questions about how to distinguish between alternative explanations for semantic universals. The present results contribute to a body of literature arguing that semantic variation can be explained in terms of the simplicity/informativeness trade-off, across a range of domains (Kemp and Regier, 2012; Regier et al., 2015; Kemp et al., 2018; Gibson et al., 2019). At the same time, a growing body of literature argues that semantic typology—often in the same domains (e.g. quantifiers, color terms)—reflects relative ease of learning (Steinert-Threlkeld and Szymanik, 2019, 2020; Steinert-Threlkeld, 2019; van de Pol et al., 2019; Saratsli et al., 2019; Maldonado and Culbertson, 2019). Future studies can and should be developed to probe whether these explanations are in conflict and, if so, which is to be preferred (while remaining open to the possibility that different factors may explain the typological facts in different domains). One promising avenue for this work comes from integration with explicit models of language change and evolution. If, for example, as languages change, they appear to be regularly and continuously *optimizing* one of the above factors but not the other, that would provide evidence in its favor.

## A Estimating the Pareto Frontier

The complete algorithm for estimating the Pareto frontier, described in Section 2.4, appears as Algorithm 1. It is a simplified version of the non-dominated genetic sorting algorithm

---

<sup>6</sup>For example: while executing the evolutionary algorithm, we can store all of the languages that are generated, in addition to the pool that constitutes the current generation.

---

**Algorithm 1** Estimating the Pareto Frontier

---

*Parameters:* num\_generations, num\_langs*Inputs:* set of languages  $L$ , Pareto dominance method FIND\_DOMINANT, INTERPOLATE method

```

function GENETIC_ESTIMATE(num_generations, num_langs)
  languages  $\leftarrow$  SAMPLE_RANDOM_LANGUAGES(num_langs)
  for  $i = 1, \dots, \text{num\_generations}$  do
    dominant_languages  $\leftarrow$  FIND_DOMINANT(languages)
    languages  $\leftarrow$  SAMPLE_MUTATED(dominant_languages, num_langs)
  end for
  return languages
end function

function SAMPLE_MUTATED(languages, amount)
  amount_per_lang, amount_random  $\leftarrow$  amount / |languages|
  mutated_languages  $\leftarrow$  []
  for language  $\in$  languages do
    for  $i = 1, \dots, \text{amount\_per\_lang}$  do
      Add MUTATE(language) to mutated_languages
    end for
  end for
  for  $i = 1, \dots, \text{amount\_random}$  do
    language  $\leftarrow$  RANDOM_CHOICE(languages)
    Add MUTATE(language) to mutated_languages
  end for
  return mutated_languages
end function

function MUTATE(language)
  mutated_language  $\leftarrow$  language
  num_mutations  $\leftarrow$  RANDOM_CHOICE([1, 2, 3])
  for  $i = 1, \dots, \text{num\_mutations}$  do
    mutation  $\leftarrow$  RANDOM_CHOICE(
      {ADD_QUANTIFIER, REMOVE_QUANTIFIER, SWAP_QUANTIFIER})
    mutated_language  $\leftarrow$  MUTATION(language)
  end for
  return mutated_language
end function

estimate  $\leftarrow$  GENETIC_ESTIMATE(num_generations, num_langs)
pareto_frontier  $\leftarrow$  FIND_DOMINANT(estimate  $\cup$   $L$ )
pareto_frontier  $\leftarrow$  INTERPOLATE(pareto_frontier)

```

---

(Srinivas and Deb, 1994). There are two main parameters: how many generations to run the algorithm for (num\_generations), and how large a pool of languages to maintain at each generation (num\_langs). For the experiments in this paper, we set num\_generations to 100 and num\_langs to 2000. The set of languages  $L$  is the full set that we sampled according to the procedures described in Section 2.3.

The three final lines in the algorithm correspond to the three steps described in Figure 1 above. The method `GENETIC_ESTIMATE` contains the basic loop over generations. For finding dominant languages, we used the `pygmo` library's `non_dominated_front_2d` method (Biscani and Izzo, 2019). `SAMPLE_MUTATED` generates the new population at each generation by giving each dominant language its offspring. The function `MUTATE` performs the mutation of a single language, by choosing a number of mutations to apply and then randomly choosing from the available mutations.

## References

- Jon Barwise and Robin Cooper. Generalized Quantifiers and Natural Language. *Linguistics and Philosophy*, 4(2):159–219, 1981. doi: 10.1007/BF00350139.
- Francesco Biscani and Dario Izzo. `pagmo/pygmo2`. 2019. doi: 10.5281/zenodo.1045336. URL <https://esa.github.io/pagmo2/>.
- Fausto Carcassi, Shane Steinert-Threlkeld, and Jakub Szymanik. The emergence of monotone quantifiers via iterated learning. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)*, pages 190–196, 2019. URL <https://psyarxiv.com/8swtd>.
- Emmanuel Chemla, Brian Buccola, and Isabelle Dautriche. Connecting Content and Logical Words. *Journal of Semantics*, 2019. doi: 10.1093/jos/ffz001.
- Carlos A Coello Coello, Gary B Lamont, and David A van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, 2007. doi: 10.1007/978-0-387-36797-2.
- William Croft. *Typology and Universals*. Cambridge University Press, Cambridge, 1990.
- Jacob Feldman. Minimization of Boolean Complexity in Human Concept Learning. *Nature*, 407:630–633, 2000. doi: 10.1038/35036586.
- Edward Gibson, Richard Futrell, Steven T Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy. How Efficiency Shapes Human Language. *Trends in Cognitive Sciences*, 2019. doi: 10.1016/j.tics.2019.02.003.
- Larry M. Hyman. Universals in phonology. *The Linguistic Review*, 25(1-2):83–137, 2008. doi: 10.1515/TLIR.2008.003.
- Gerhard Jäger. The evolution of convex categories. *Linguistics and Philosophy*, 30(5):551–564, 2007. doi: 10.1007/s10988-008-9024-3.
- Edward L Keenan and Denis Paperno, editors. *Handbook of Quantifiers in Natural Language*, volume 90 of *Studies in Linguistics and Philosophy*. Springer, Dordrecht, 2012. doi: 10.1007/978-94-007-2681-9.
- Charles Kemp and Terry Regier. Kinship categories across languages reflect general communicative principles. *Science*, 336(6084):1049–1054, 2012. doi: 10.1126/science.1218811.
- Charles Kemp, Yang Xu, and Terry Regier. Semantic typology and efficient communication. *Annual Review of Linguistics*, pages 1–23, 2018. doi: 10.1146/annurev-linguistics-011817-045406.

- Mora Maldonado and Jennifer Culbertson. Something about *us*: Learning first person pronoun systems. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)*, 2019.
- Cailin O'Connor. The Evolution of Vagueness. *Erkenntnis*, pages 1–21, 2014. doi: 10.1007/s10670-013-9463-2.
- Denis Paperno and Edward L Keenan, editors. *Handbook of Quantifiers in Natural Language: Volume II*, volume 97 of *Studies in Linguistics and Philosophy*. Springer, 2017. doi: 10.1007/978-3-319-44330-0.
- Stanley Peters and Dag Westerståhl. *Quantifiers in Language and Logic*. Clarendon Press, Oxford, 2006.
- Steven T Piantadosi, Joshua B Tenenbaum, and Noah D Goodman. The logical primitives of thought: Empirical foundations for compositional cognitive models. *Psychological Review*, 123(4):392–424, 2016. doi: 10.1037/a0039980.
- Terry Regier, Charles Kemp, and Paul Kay. Word meanings across languages support efficient communication Informativeness and simplicity as competing principles. In Brian MacWhinney and William O'Grady, editors, *The Handbook of Language Emergence*, pages 237–263. Wiley, 2015.
- Dionysia Saratsli, Stefan Bartell, and Anna Papafragou. Are Cross-Linguistically Frequent Semantic Systems Easier to Learn? The Case of Evidentiality. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)*, 2019.
- Brian Skyrms. *Signals: Evolution, Learning, and Information*. Oxford University Press, 2010.
- N. Srinivas and Kalyanmoy Deb. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, 2(3):221–248, 1994. doi: 10.1162/evco.1994.2.3.221.
- Shane Steinert-Threlkeld. An Explanation of the Veridical Uniformity Universal. *Journal of Semantics*, 2019. doi: 10.1093/jos/ffz019. URL <https://semanticsarchive.net/Archive/DI5ZTNmN/UniversalResponsiveVerbs.pdf>.
- Shane Steinert-Threlkeld and Jakub Szymanik. Learnability and Semantic Universals. *Semantics & Pragmatics*, 12(4), 2019. doi: 10.3765/sp.12.4.
- Shane Steinert-Threlkeld and Jakub Szymanik. Ease of Learning Explains Semantic Universals. *Cognition*, 195, 2020. doi: 10.1016/j.cognition.2019.104076.
- Jakub Szymanik. *Quantifiers and Cognition: Logical and Computational Perspectives*, volume 96 of *Studies in Linguistics and Philosophy*. Springer, 2016. doi: 10.1007/978-3-319-28749-2.
- Iris van de Pol, Shane Steinert-Threlkeld, and Jakub Szymanik. Complexity and learnability in the explanation of semantic universals of quantifiers. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)*, pages 3015–3021, 2019. URL <https://psyarxiv.com/f8dbp/>.
- Kai von Stechow and Lisa Matthewson. Universals in semantics. *The Linguistic Review*, 25: 139–201, 2008. doi: 10.1515/TLIR.2008.004.



# Equating by degrees or state-kinds, or both

Yenan Sun

The University of Chicago  
yenansun@uchicago.edu

## Abstract

This paper provides novel data of a cross-categorial equation construction in Mandarin Chinese to support the cross-linguistic connection between kinds, manners, and degrees. I further identify two modes of scalar equatives in Mandarin, showing that equating by state-kinds represents a distinct equation mode than equating by degree objects. The syntactic and semantic differences between the two modes motivate us to preserve both degree state-kinds and degree objects in our ontology.

## 1 Introduction

Anderson and Morzycki [1] (A&M henceforth) observe that a variety of languages exhibit a systematic connection between kinds, manners, and degrees. English uses the same preposition *as* to introduce arguments for kind modifiers, manner modifiers, and degree modifiers:

- (1) a. Such dog as Toby is adorable. (KIND)  
b. Fiddo barked as Toby did. (MANNER)  
c. Fiddo is as tall as Toby is. (DEGREE)

In Polish, *tak* ‘such’ can be anaphoric to kinds (2-a), manners (2-b), and degrees (2-c); a single *wh*-word *jak*, is used to question those three categories (3).

- (2) a. **taki** pies (KIND)  
such-MASC dog  
‘such a dog’  
b. **tak** się zachowywać (MANNER)  
such REFL behave  
‘behave that way’  
c. **tak** wysoki (DEGREE)  
such tall  
‘that tall’
- (3) a. **jaki** pies (KIND)  
WH-MASC dog  
‘what kind of dog’  
b. **Jak** się zachowywał (MANNER)  
WH REFL behave-3MASC  
‘How did he behave?’  
c. **Jak** wysoki jest Toby (DEGREE)  
WH tall is Toby  
‘How tall is Toby?’

Motivated by those facts, A&M propose that manners and degrees are both Chierchia-style kinds of eventualities. Based on Chierchia [6], a kind is the plurality of all of its possible instances. For instance, the kind DOG is the plurality of all the dogs in every possible world. Turning to manners and degrees, they can be viewed as kinds of events and kinds of states respectively: all possible events performed fast constitute the event-kind FAST, and all possible states that have six feet as their measure along the spatial dimension constitute the degree-kind SIX-FEET. The introduction of kinds over eventualities renders possible a uniform analysis of Polish *tak* in various equation constructions.

One issue raised by their paper is, if degree can be represented as a particular state-kind, whether we still need degree objects<sup>1</sup> in ontology. Though a dual analysis to keep both auto-

<sup>1</sup>I will use ‘degree objects’ henceforth to refer to the objects in the model and ‘degree’ for the abstract concept.



matically achieves a wider empirical coverage as admitted by A&M<sup>2</sup>, the question remains why languages will have these two systems co-exist if much of the work they do overlap.

This paper argues that both degree state-kinds and degree objects should be kept in our ontology based on the data from Mandarin Chinese. First of all, just like Polish *tak* and many other languages, Mandarin has the cross-categorial *na-yang* ‘such’ which can express equation across kinds, manners, and degrees. This supports a cross-linguistic connection between those three categories (Section 2.1). Moreover, Mandarin has an equation construction involving the (stressed) *yi-yang*<sub>EMP</sub>, which is specific to equating degrees thus NOT cross-categorial. I show that there is a correlation for an equation construction in Chinese between being cross-categorial and triggering an obligatory positive inference of the standard (Section 2.2). I argue that this correlation points to an important difference between equating by state-kinds and equating by degrees for a scalar property, which shows both are needed in our model (Section 3). Section 4 concludes.

## 2 Equation constructions in Mandarin

This section first talks about a cross-categorial equation construction in Mandarin to support A&M’s claim (Section 2.1) and then turns to an equation construction that is specific to degree equation (not cross-categorial) to identify two modes of scalar equatives (Section 2.2).

### 2.1 Mandarin cross-categorial *na-yang*

Mandarin exhibits a striking parallel to Polish: a single morpheme *-yang*, which literally means ‘sort, kind’ (Liao and Wang [10]), is used to build the anaphoric form with the demonstrative *na* ‘that’ for kinds, manners, and degrees, as in (4).<sup>3</sup>

- (4) a. **na-yang** de shu zhide du  
that-kind MOD book worth read  
‘That kind of book worths reading.’ (KIND)
- b. Yuehan hui **na-yang** tiaowu.  
John will that-kind dance  
‘John will dance that way.’ (MANNER)
- c. Yuehan shi **na-yang** gao  
John be that-kind tall  
‘John is that tall.’ (DEGREE)

That same morpheme can be used to form the *wh*-word to question those three categories, as shown in (5).

- (5) a. ni xihuan **zen-yang** de shu?  
you like WH-kind MOD book  
‘What kind of book do you like?’ (KIND)
- b. ni hui **zen-yang** tiaowu?  
you will WH-kind dance  
‘How will you dance?’ (MANNER)

<sup>2</sup>For instance, A&M mention that the standard degree approaches can better account for phenomena like differential comparatives and factor phrases (*two times taller*).

<sup>3</sup>Abbreviations in this paper: CL=classifier, PERF=perfective marker, MOD=modification marker, YNQ=yes-or-no question marker.

- c. wulun Yuehan **zen-yang** congming, ta haishi yinggai shang ke  
 no.matter John WH-kind clever he still should take class  
 ‘No matter how clever John is, he still should take classes’ (DEGREE)

We will focus on the data of *na-yang* in this paper. Notice (4) involve the anaphoric use of *na-yang*, namely there is a salient kind in the context feeding the interpretation of those modifiers. For instance, if (4-a) is uttered in a discourse like (6), the referent introduced by the previous discourse will highlight a relevant kind like CLASSIC (which constitute all possible classic books). Following the comparative literature we call such a relevant kind the *standard* of equation.

- (6) a. A: shangzhou wo du-le Zhanzhengyuheping  
 last.week I read-PERF War.and.Peace  
 ‘Last week I read *War and Peace*’  
 b. B: hao! ni yinggai duo du **na-yang** de shu  
 good you should more read that-kind MOD book  
 ‘Great! You should read that kind (=CLASSIC) of books more.’

Alternatively, such a standard can be overtly introduced by the preposition *xiang* ‘like’ as in (7), which are often called non-anaphoric use.

- (7) a. xiang Zhanzheng.yu.heping **na-yang** de shu...  
 like War.and.Peace that-kind MOD book  
 ‘such book like *War and Peace*...’  
 b. Yuehan hui xiang Bier **na-yang** tiaowu  
 John will like John that-kind dance  
 ‘John will dance in such way like Bill does.’  
 c. Yuehan xiang Bier **na-yang** gao  
 John like Bill that-kind tall  
 ‘John is such tall like Bill.’

This cross-categorial equation construction involving *na-yang* provides support for A&M’s proposal to render manners and degrees as kinds of eventualities.

## 2.2 Non-cross-categorial *yiyang*<sub>EMP</sub> and two modes of equatives

In Mandarin there is another equation construction that can only equate degrees but not kinds and manners, which is featured by *yi-yang*<sub>EMP</sub> (8).

- (8) a. ?\*gen Zhanzheng.yu.heping **yi-yang**<sub>EMP</sub> de shu  
 as War.and.Peace one-kind MOD book  
 Int: ‘such book like *War and Peace*...’ (KIND)  
 b. ?\*Yuehan hui gen Bier **yi-yang**<sub>EMP</sub> tiaowu  
 John will like John one-kind dance  
 Int: ‘John will dance in such way like Bill does’ (MANNER)  
 c. Yuehan gen Bier **yi-yang**<sub>EMP</sub> gao  
 John as Bier one-kind tall  
 ‘John is as tall as Bill.’ (DEGREE)

The subscript ‘EMP’ indicates it bears some salient intonation in the sentence, in order to distinguish it from another *yi-yang*<sub>LIGHT</sub>, which does not bear stress and behaves essentially like

the cross-categorical *na-yang*. We will turn back to this *yi-yang*<sub>LIGHT</sub> later. Comparing the scalar equative involving the non-cross-categorical *yi-yang*<sub>EMP</sub> and the scalar equative involving the cross-categorical *na-yang*, we find they differ both in syntax and semantics.

Firstly, they select different prepositions, either *gen* or *xiang*, to introduce the standard in the equation as in (9). Switching the prepositions leads to ungrammaticality:

- (9) a. Yuehan {gen, \*xiang} Bier **yi-yang**<sub>EMP</sub> gao  
 John as like Bier one-kind tall  
 ‘John is as tall as Bill.’  
 b. Yuehan {\*gen, xiang} Bier **na-yang** gao  
 John as like Bier that-kind tall  
 ‘John is such tall like Bill.’

Secondly, the *na-yang* equative invokes a positive interpretation of the standard in the equation. Take (9) for instance, when uttering (9-b), the speaker not only commits to the equation statement that John’s height equals Bill’s height, but also a positive interpretation of Bill’s tallness, namely Bill’s height must exceed the contextually relevant standard of being tall. For (9-a), only the equation of John’s height and Bill’s height is asserted and no such positive interpretation is enforced. This contrast can be shown in (10) by having a denial of the positive inference as the follow-up.

- (10) a. Yuehan gen Bier **yi-yang**<sub>EMP</sub> gao. suiran Bier hen ai.  
 John as Bier one-kind tall though Bill very short  
 ‘John is as tall as Bill. Though Bill is quite short.’  
 b. Yuehan xiang Bier **na-yang** gao. #suiran Bier hen ai.  
 John like Bier that-kind tall though Bill very short  
 ‘John is such tall like Bill. #Though Bill is quite short.’

Testing with this extra inference from the *na-yang* equative, we further find it to be a presupposition which survives under negation or interrogative operators, as shown in (11).

- (11) a. Yuehan bu xiang Bier **na-yang** gao. #suiran Bier yijing gou ai le.  
 John NEG like Bill that-kind tall though Bill already enough short SFP  
 ‘John isn’t so tall like Bill. #Though Bill is short enough.’  
 b. Yuehan xiang Bier **na-yang** gao ma? #suiran Bier yijing gou ai le.  
 John like Bill that-kind tall YNQ though Bill already enough short SFP  
 ‘Is John so tall like Bill? #Though Bill is short enough.’

The fact that the cross-categorical equation construction has this positive inference in the case of scalar equatives is not random – the unstressed *yi-yang*<sub>LIGHT</sub> is another way of equating things in Mandarin and it shares almost identical properties with *na-yang*:

- (12) a. {??gen, xiang} Zhanzheng.yu.heping **yi-yang**<sub>LIGHT</sub> de shu  
 as like War.and.Peace one-kind MOD book  
 ‘such book like *War and Peace*...’ (KIND)  
 b. Yuehan hui {??gen, xiang} houzi **yi-yang**<sub>LIGHT</sub> tiaowu  
 John will as like monkey one-kind dance  
 Int: ‘John will dance in such way like monkeys’ (MANNER)  
 c. Yuehan {??gen, xiang} Bier **yi-yang**<sub>LIGHT</sub> gao  
 John as like Bill one-kind tall  
 ‘John is such tall like Bill.’ (DEGREE)

Crucially, under this use, the standard must be (or at least strongly preferred to be) introduced by the preposition *xiang* and the positive inference of the standard is presupposed as well, shown by the impossibility of denying Bill is tall when uttering (12-c):

- (13) Yuehan xiang Bier **yi-yang**<sub>LIGHT</sub> gaO. #suiran Bier hen ai.  
 John like Bill one-kind tall though Bill very short  
 ‘John is such tall like Bill. #Though Bill is quite short.’
- (14) Yuehan bu xiang Bier **yi-yang**<sub>LIGHT</sub> gaO. #suiran Bier yijing gou ai le.  
 John not like Bill one-kind tall though Bill already enough short SFP  
 ‘John isn’t’s such tall like Bill. #Though Bill is short enough.’

In short, we identify two modes of scalar equatives, one represented by *yi-yang*<sub>EMP</sub> and another represented by *na-yang* and *yi-yang*<sub>LIGHT</sub>. Their properties are summarized in Table 1.

	Non-cross-categorial	Cross-categorial
The preposition used to introduce the standard	<i>gen</i> ‘as’	<i>xiang</i> ‘like’
The positive interpretation of the standard	NA	presupposed

Table 1: Two modes of equatives

We are particularly interested in the correlation between the existence of a positive interpretation of the standard in a certain equative, and its ability to equate manners and kinds besides degrees. The next section proposes an analysis to capture such a correlation.

### 3 The analysis

I propose that the two modes of scalar equatives differ in whether state-kinds or (sets of) degree objects are equated. The equative involving *na-yang* (or *yi-yang*<sub>LIGHT</sub>) equates state-kinds, while the equative involving *yi-yang*<sub>EMP</sub> equates sets of degree objects. In other words, both state-kinds and degree objects are required in our ontology. Based on A&M [1] and other standard models which include degree objects, I assume the following subsets of the domain  $D$ :

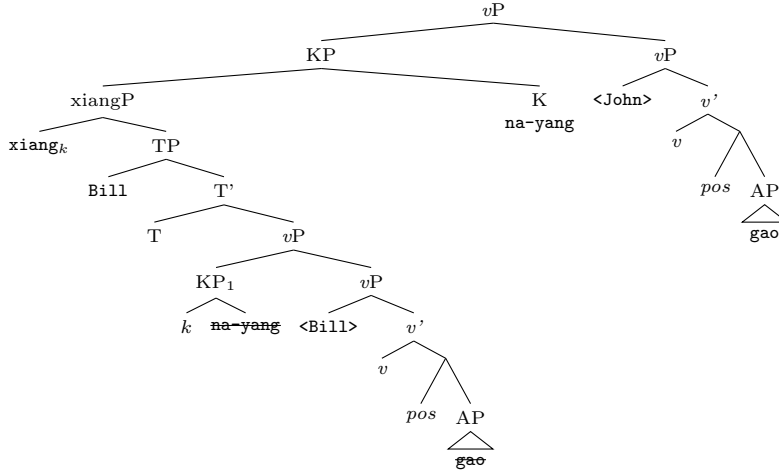
- (15) a.  $D_k$  is a set of kind objects in  $D$  (represented by  $k, k', \dots$ )  
 b.  $D_o$  is a set of non-kind objects in  $D$ :  
 $D_e$  is a set of non-kind individuals in  $D_o$  ( $x, y, z, \dots$ )  
 $D_v$  is a set of non-kind events in  $D_o$  ( $e, e', \dots$ )  
 $D_s$  is a set of non-kind states in  $D_o$  ( $s, s', \dots$ )  
 $D_d$  is a set of degree objects in  $D_o$  ( $d, d', \dots$ )

I further assume Kratzer [9]’s version of event semantics. For instance, an intransitive verb like *dance* denotes a set of dancing events, and the agent is introduced by  $v$  (voice head) via Event Identification, as in (16).

- (16)
- $$\begin{array}{c}
 \text{vP} \\
 \lambda e. \mathbf{dance}(e) \& \mathbf{ag}(e, j) \\
 \text{John} \quad \text{v}' \\
 \lambda x \lambda e. \mathbf{dance}(e) \& \mathbf{ag}(e, x) \\
 \text{v} \quad \text{VP} \\
 \lambda x \lambda e. \mathbf{ag}(e, x) \quad \lambda e. \mathbf{dance}(e) \\
 \mathbf{dance}
 \end{array}$$



(20)



The standard clause (i.e. *xiangP*) is elided and contains a free kind variable which is abstracted over by the preposition *xiang*. The resulted denotation is shown in (21):

- (21) a.  $\llbracket [{}_{vP} \text{ Bill pos gao}] \rrbracket = \lambda s. \mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b)$   
 b.  $\llbracket [{}_{vP} [{}_{KP1} k \text{ na-yang}] [{}_{vP} \text{ Bill pos gao}]] \rrbracket = \exists s [\mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b) \& {}^{\cup}k(s)]$   
 c.  $\llbracket [\text{xiangP}] \rrbracket = \lambda k. \exists s [\mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b) \& {}^{\cup}k(s)]$

Since *na-yang* does not head a DegP over AP, it does not saturate the degree argument of the predicate *gao* 'tall'. Thus the predicate *gao* goes through its regular composition with *pos*.

Following A&M, the type mismatch between  $\llbracket [\text{xiangP}] \rrbracket$  and *na-yang* triggers Iota type shift (Caponigro [5]), since there is precisely one degree state-kind that a state instantiates:

$$(22) \quad \llbracket [\text{SHIFT xiangP}] \rrbracket = \iota k [\exists s [\mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b) \& {}^{\cup}k(s)]]$$

This unique kind saturates the first argument of *na-yang* in the matrix clause and results in a property of states:

$$(23) \quad \llbracket [{}_{KP} [\text{SHIFT xiangP}] \text{ na-yang}] \rrbracket = \lambda o. {}^{\cup} \iota k [\exists s [\mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b) \& {}^{\cup}k(s)]](o)$$

This property intersects with the property contributed from the matrix clause:

- (24) a.  $\llbracket [\text{John pos gao}] \rrbracket = \lambda s'. \mathbf{tall}_{pos}(s') \& \mathbf{holder}(s', j)$   
 b.  $\llbracket [{}_{KP} [\text{SHIFT xiangP}] \text{ na-yang}] [\text{John pos gao}] \rrbracket = \lambda s'. {}^{\cup} \iota k [\exists s [\mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b) \& {}^{\cup}k(s)]](s') \& \mathbf{tall}_{pos}(s') \& \mathbf{holder}(s', j)$

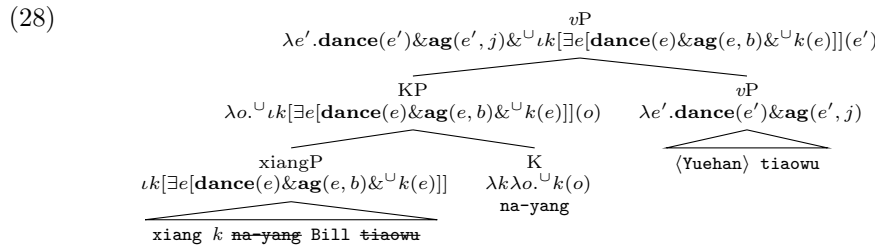
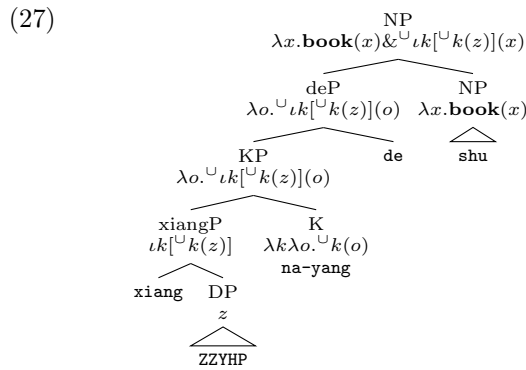
After the existential closure of the state variable, we derive the semantics for the scalar equative (18-a) in (25): there is a state of having a positive (i.e. exceeding the contextual threshold of being tall) height which is held by John such that it instantiates a unique degree state-kind which is instantiated by a state of having a positive height whose holder is Bill.

$$(25) \quad \llbracket [(18-a)] \rrbracket = \exists s' [{}^{\cup} \iota k [\exists s [\mathbf{tall}_{pos}(s) \& \mathbf{holder}(s, b) \& {}^{\cup}k(s)]](s') \& \mathbf{tall}_{pos}(s') \& \mathbf{holder}(s', j)]$$

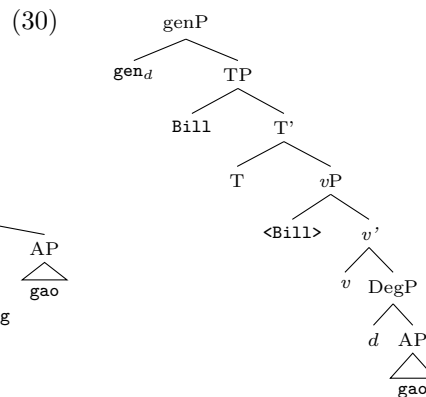
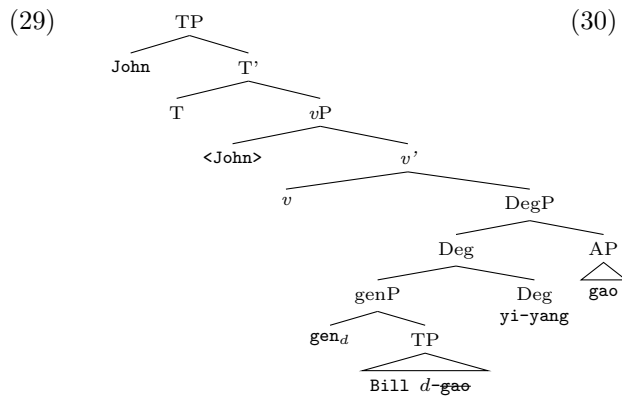
This is equivalent to saying John's state of having some positive height has the same measure along the spatial dimension as Bill's state of having some positive height. We further capture the presupposition that the standard (Bill) must be considered as tall by the  $\iota$ -closure in the formula: since a unique degree state-kind which is instantiated by Bill's state of having some

positive height is presupposed (under the scope of the  $\iota$  operator), we expect the inference that Bill is tall can project. Finally, since *na-yang* equates general kinds instead of specific degrees, it is expected that it is cross-categorical. We briefly present how the current proposal extends to the equation of nominal kinds and manners (26) in (27)-(28).

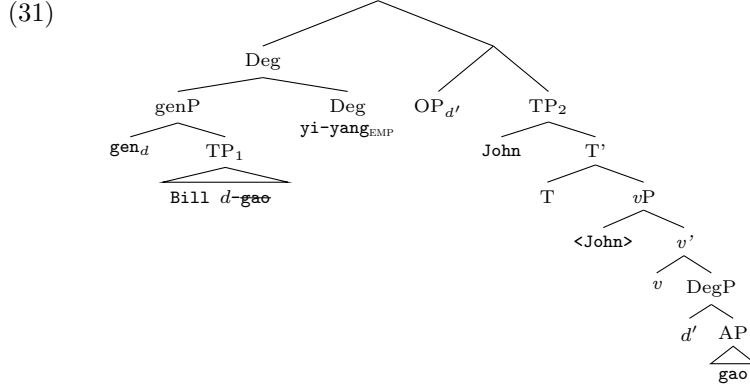
- (26) a. xiang Zhanzheng.yu.heping **na-yang** de shu...  
 like War.and.Peace that-kind MOD book  
 ‘such book like *War and Peace*(ZZYHP)...’  
 b. Yuehan hui xiang Bier **na-yang** tiaowu  
 John will like John that-kind dance  
 ‘John will dance in such way like Bill does.’



Turning to the non-cross-categorical *yi-yang*<sub>EMP</sub> in (18-b), I assume it is a regular Deg head which selects an elided clause introduced by the preposition *gen* and AP, as in (29). The syntax of the elided clausal standard is elaborated in (30).



Following a common approach to comparatives (Bhatt and Takahashi [3], Rett [11], a.o.), the Deg head scopes out together with genP to form the LF in (31) and some operator (a null one in matrix clause and the preposition *gen* in the elided clause, following Rett [11]) abstracts over the occurrences of free degree variable to form a set of degrees.



The corresponding semantic composition is as follows:

$$(32) \quad \llbracket \text{yi-yang}_{\text{EMP}} \rrbracket = \lambda D_{dt} \lambda D'_{dt} . D = D'$$

$$(33) \quad \begin{aligned} \text{a.} \quad & \llbracket d' \text{ gao} \rrbracket = \lambda s . \mathbf{tall}(s, d') \\ \text{b.} \quad & \llbracket v\text{P} \rrbracket = \exists s [\mathbf{tall}(s, d) \& \mathbf{holder}(s, j)] \\ \text{c.} \quad & \llbracket \text{OP}_{d'} \text{ TP}_2 \rrbracket = \lambda d' . \exists s [\mathbf{tall}(s, d) \& \mathbf{holder}(s, j)] \end{aligned}$$

$$(34) \quad \llbracket \text{genP} \rrbracket = \lambda d . \exists s [\mathbf{tall}(s, d) \& \mathbf{holder}(s, b)]$$

$$(35) \quad \llbracket (18\text{-b}) \rrbracket = \llbracket [\text{genP } \text{yi-yang}_{\text{EMP}}] [\text{OP}_{d'} \text{ TP}_2] \rrbracket = \\ (\lambda d . \exists s [\mathbf{tall}(s, d) \& \mathbf{holder}(s, b)]) = (\lambda d' . \exists s [\mathbf{tall}(s, d') \& \mathbf{holder}(s, j)])$$

Since *yi-yang<sub>EMP</sub>* is a degree quantifier, it cannot equate the nominal kinds or manners. With a standard treatment for the scalar equative involving *yi-yang<sub>EMP</sub>*, we also capture the fact that it does not presuppose Bill counts as tall in (18-b).

In sum, I extend A&M's proposal of Polish *tak*, with some revisions, to capture that equatives involving *na-yang* are (i) cross-categorial between degrees, kinds, and manners; and (ii) presuppose the positive interpretation of the standard wrt the relevant scalar property. I also show that degree objects should still be kept in our ontology since equatives with *yi-yang<sub>EMP</sub>* indeed equate degree objects, and thus is non-cross-categorial and lacks such presupposition.

## 4 Concluding remarks

This paper investigates the equation constructions in Mandarin Chinese and especially how they contribute to Anderson and Morzycki [1]'s proposal that degrees and manners are kinds of eventualities. First, the existence of the cross-categorial *na-yang* 'such', which just like Polish *tak* can express the equation of nominal kinds, manners, and degrees, provides support for their proposal. Second, Mandarin has a non-cross-categorial equation construction featured by *yi-yang<sub>EMP</sub>*, which is restricted to equation of degrees. Those two different equation constructions exhibit differences in syntax and semantics when expressing the equation of degrees. Crucially, there is a correlation for an equation construction between being cross-categorial and presupposing the positive interpretation of the standard for a relevant scalar property. The current



analysis accounts for the correlation by proposing that the cross-categorial *na-yang* equates state-kinds (and also heads a syntactically higher projection as the modifier of *vP*), whereas the non-cross-categorial *yi-yang*<sub>EMP</sub> equates sets of degree objects (as a standard degree quantifier). Such a proposal relies on the co-existence of state-kinds and degree objects in our ontology. While the distinction between two modes of equatives in a single language like Mandarin might not be enough to fully answer A&M's question that whether state-kinds should be the only representation of degree in our model, I believe it is an interesting starting point, and it is worthwhile to look into more languages to see if it is a cross-linguistic tendency for a language to have certain means to distinguish between equating by state-kinds and by degree objects.

## References

- [1] Curt Anderson and Marcin Morzycki. Degrees as kinds. *Natural Language & Linguistic Theory*, 33(3):791–828, 2015.
- [2] Rebekah Baglini. Stative predication and semantic ontology: A crosslinguistic study. *Chicago IL: University of Chicago PhD dissertation*, 2015.
- [3] Rajesh Bhatt and Shoichi Takahashi. Reduced and unreduced phrasal comparatives. *Natural Language & Linguistic Theory*, 29(3):581–620, 2011.
- [4] Manfred Bierwisch. The semantics of gradation. *Dimensional adjectives*, 71:261, 1989.
- [5] Ivano Caponigro. The semantic contributions of wh-words and type shifts: Evidence from free relatives crosslinguistically. In *Semantics and linguistic theory*, volume 14, pages 38–55, 2004.
- [6] Gennaro Chierchia. Reference to kinds across language. *Natural Language Semantics*, 6(4):339–405, 1998.
- [7] Max J Cresswell. The semantics of degree. In B. Partee, editor, *Montague grammar*, pages 261–292. New York: Academic Press, 1976.
- [8] Christopher Kennedy. Projecting the adjective: The syntax and semantics of gradability and comparison: Uc santa cruz dissertation. *Published in*, 1999.
- [9] Angelika Kratzer. Severing the external argument from its verb. In *Phrase structure and the lexicon*, pages 109–137. Springer, 1996.
- [10] Wei-wen Roger Liao and Yuyun Wang. The same difference: Comparative syntax-semantics of english same and chinese tong/xiang-tong. *Chinese Syntax in a Cross-Linguistic Perspective*, 128, 2014.
- [11] Jessica Rett. Similitives and the argument structure of verbs. *Natural Language & Linguistic Theory*, 31(4):1101–1137, 2013.
- [12] Alexis Wellwood. States and events for s-level gradable adjectives. In *Semantics and linguistic theory*, volume 26, pages 166–184, 2016.

# Animal linguistics and the puzzle of titi monkeys alarm sequences

Mélissa Berthet<sup>1</sup>, Juan Benjumea<sup>2</sup>, Juliette Millet<sup>3,4,5</sup>, Cristiane Cäsar<sup>6</sup>,  
Klaus Zuberbühler<sup>7,8</sup> and Ewan Dunbar<sup>3,5</sup>

<sup>1</sup>Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL  
Research University, France

<sup>2</sup>LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University,  
France

<sup>3</sup>Laboratoire de Linguistique Formelle, Université de Paris, CNRS, Paris, France

<sup>4</sup>CRI, Département Frontières du vivant et de l'apprendre, IIFR, Université de Paris,  
France

<sup>5</sup>Cognitive Machine Learning, ENS, EHESS, CNRS, INRIA, PSL Research University,  
France

<sup>6</sup>Vale S.A., Brazil

<sup>7</sup>Department of Comparative Cognition, Institute of Biology, University of Neuchâtel,  
Switzerland

<sup>8</sup>School of Psychology and Neurosciences, University of St Andrews, UK  
melissa.berthet.ac@gmail.com

## Abstract

This article aims to illustrate how animal linguistics can increase our understanding of animal communication using the example of the alarm system of Black-fronted titi monkeys. Titi monkeys produce sequences composed of A- and B-calls in alarm situations. Previous biological and linguistic works investigated the semantics of these utterances but relied on a preliminary dataset. We followed up on this work by carrying out an extensive analysis of meaning at the call and sequence level, based on 18 months of field observations and experiments followed by acoustical and linguistic analyses. Our results suggest that alarm calls refer to the emotional state of the caller at the time of emission of the call. Listeners most likely attend to the proportion of BB-grams in each sequence, which provides them with information about the caller's emotional state, to infer what the predator is (aerial or terrestrial) and where it is (ground or canopy) using their world knowledge. Overall, this work suggests that the emotional states of animals can convey complex and reliable information to listeners. Our study illustrates that combining field data with linguistic analyses provides a powerful new approach to better understand animal communication.

## 1 Introduction

Animal linguistics, or the application of linguistic methods to the study of nonhuman animal data, is a recent approach that aims to deepen our understanding of nonhuman communication, from bird songs (e.g. Yip, 2006) to primate vocalizations (e.g. Heesen et al., 2019) and gestures (Graham et al., 2018). When exploring semantics and syntax in primate vocalizations, Schlenker et al. (2016a) proposed an approach that relies on three main pillars derived from formal linguistics: i) calls are combined into sequences that are subject to syntactic rules and can be studied under a formal syntax approach that distinguishes

possible and impossible sequences, ii) there is a correlation between a call type and the context(s) in which it occurs, thus calls have a semantics and can be studied using truth-conditional semantics that establishes whether a call is true (or appropriate) in a given context, and iii) there are rules of call competition that determine which call must be chosen over another one. Overall, their approach aims to explore “the division of labor between syntax, semantics, pragmatics, and properties of the environmental context (‘world’ knowledge and context change)” (Schlenker et al., 2016a, p. 20).

Black-fronted titi monkeys *Callicebus nigrifrons* are good candidates to apply this approach and demonstrates its efficiency, since titi monkeys’ vocal communication has been described as “the maximal elaboration that can be attained by species-specific language” (Moynihan, 1966, p. 77) and that several successive theories have tried to decode its puzzling alarm system. Thus, the current work aims to illustrate how animal linguistics can increase our understanding and unveil unexpected mechanisms in a nonhuman communication system.

It should be noted that the notion of *meaning* we use here is similar to that of *natural meaning*, for which information is transmitted when there is a statistical correlation between a signal and the context (like *smoke* means fire). Thus, this work does not aim to explore *Gricean meaning* (in the sense of *intentional* communication, see for example Townsend et al., 2017) in titi monkey calls, but rather to explore the formal properties of their communication system while remaining agnostic about the underlying cognitive mechanisms.

## 2 First theories

Titi monkeys produce sequences composed of A-calls and B-calls in alarm situations. Cäsar et al. (Cäsar, 2011; Cäsar et al., 2012b, 2013) conducted extensive field observations followed by experiments such as predator presentations (raptor or felid models placed either on the ground or in the canopy; snake, tayra and deer models on the ground) and playbacks of titi monkey alarm sequences. They analysed the 30 first calls emitted and described the titi monkey alarm system as simplified in (1).

### (1) Cäsar et al.’s description of the different alarm sequences and their context of emission

- a. Terrestrial predator on the ground: Sequences of B-calls (B+)
- b. Terrestrial non-predator animal on the ground: Sequences of B-calls (B+)
- c. Caller foraging near the ground in absence of predator: Sequences of B-calls (B+)
- d. Raptor in the canopy: Sequences of A-calls (A+)
- e. Capuchins in the canopy: Sequences of A-calls (A+)
- f. Felid in the canopy: Sequences of one single A-call followed by B-calls only (AB+)
- g. Raptor on the ground: Sequences of A-calls with interspersed B-calls (A+B+)

Cäsar et al. (2012a) conducted playback experiments, in which they broadcasted A+ and B+ sequences to titi monkeys to assess the reaction of the listeners to the utterances. They obtained the results presented in (2).

### (2) Results of Cäsar et al.’s playback experiments

- a. A+ sequences make the listener look upward
- b. B+ sequences make the listener look toward the speaker, as if searching for additional cues to determine the eliciting event

Using the results presented in (1) and (2), Cäsar et al. drew the conclusions detailed in (3).

**(3) Cäsar et al.'s theory**

- a. A-calls refer to danger within the canopy
- b. B-calls are general alarm calls, elicited by terrestrial disturbances
- c. Sequences encode type (aerial vs terrestrial) and location (ground vs canopy) of the predator using a sophisticated syntactic/semantic interface: unusual locations are encoded by a modified predator-specific sequence (namely, addition of a A-call at the beginning of the B+ sequence when the terrestrial predator is in the canopy and interspersed B-calls into the A+ sequence when the aerial predator is on the ground)

The possibility that this species exhibits a complex syntactic/semantics interface drove linguists (Schlenker et al., 2017) to analyse this communication system using the formal linguistic approach mentioned above. The authors rejected Cäsar et al.'s theory. First, they showed that the semantics of A-and B-calls proposed by Cäsar et al. does not hold: the data exposed in (1) do not support the idea that A-calls refer to danger within the canopy, since A-calls are given to predators in the canopy but also on the ground, to aerial predators and to terrestrial predators. Similarly, B-calls can not refer to terrestrial disturbances, since they are emitted in contexts in which the threat is on the ground as well as in the canopy, and to terrestrial and aerial predators. Schlenker et al. also rejected the claim that titi monkeys exhibit a complex syntactic/semantic interface. Instead, Schlenker et al. (2017) analysed the calls using scalar implicatures, as stated in (4) and proposed the conclusions detailed in (5).

**(4) Schlenker et al. (2016a)'s Informativity Principle**

If a sentence S was uttered and if S' is (i) an alternative to S, and (ii) strictly more informative than S (i.e. asymmetrically entails S), infer that S' is false

**(5) Schlenker et al.'s theory**

- a. A-calls do not refer to the type or location of the predator, but rather provide information about the appropriate reaction to the threat: A-calls refer to serious non-ground threats
- b. B-calls are general alarm calls, referring to noteworthy events. According to the Informativity Principle, the strengthened meaning of the B-call is applicable in case there is a noteworthy event but no serious non-ground threat (or else the A-call would have been emitted)
- c. Calls are individual units that reflect the state of the environment at the time of utterance of the call
- d. There is no syntax: the meaning of the sequence is a conjunction of claims made by individual calls

The generalizations that fall out of Schlenker et al. (2017)'s theory are detailed in (6).

**(6) Generalizations from Schlenker et al. (2017)'s theory**

- a. Terrestrial predator on the ground: noteworthy event but not serious non-ground threat, B+
- b. Terrestrial non-predator animal on the ground: noteworthy event but not serious non-ground threat, B+
- c. Caller foraging near the ground without the presence of predator. noteworthy event but not serious non-ground threat, B+
- d. Raptor in the canopy: serious non-ground threat, A+
- e. Capuchins in the canopy: serious non-ground threat, A+
- f. Felid in the canopy: this is a serious non-ground threat ("non-ground" because the threat is in the canopy) so one A-call is emitted. But felid predators give up hunting after detection by their preys (Zuberbühler et al., 1999), so the predatory

threat becomes non-serious after the A-call is emitted, thus only B-calls are emitted afterwards: AB+

- g. Raptor on the ground: this is a serious non-ground threat (“non-ground” because the bird attacks by flying) so A-calls are emitted. After a while, the caller realizes that the bird is not in a hunting position, because it is on the ground, thus it becomes non-serious, and the caller emits B-calls: A+B+

Cäsar et al.’s and Schlenker et al.’s conclusions were drawn using one preliminary dataset, for which the number of observations was small, and crucial experiments (namely, playbacks of AB+ and A+B+ sequences to titi monkeys) were lacking. In order to complement Cäsar et al.’s dataset, we conducted an extensive study involving 18 months of field observations and experiments on a wild, observer-habituated population of titi monkeys at the Santuário do Caraça, Brazil. We then conducted acoustic and linguistic analyses, and revised the conclusions proposed by Cäsar et al. and Schlenker et al., both at the call and sequence level.

### 3 Meaning from the caller’s perspective

We recorded spontaneous alarm sequences, given to natural predator encounters and when individuals were descending near the ground to forage in the absence of predators, but also conducted predator presentation experiments. We presented 6 groups of monkeys with aerial (raptor *Caracara plancus*) or terrestrial (oncilla *Leopardus guttulus* and tayra *Eira barbara*) stuffed predators, placed either on the ground or in the canopy. We recorded the vocal reactions of the monkeys (10 first calls) and completed our dataset with alarm sequences from Cäsar (Cäsar, 2011; Cäsar et al., 2013). Description of this new dataset (N=74 sequences) is presented in (7). Our observations are mostly congruent with those of Cäsar’s et al., except for the “felid in the canopy” condition (7f) for which we did not systematically observe the AB+ pattern described in (1f).

#### (7) Updated description of the different alarm sequences and their context of emission

- a. Terrestrial predator on the ground: Sequences of B-calls (B+)
- b. Terrestrial non-predator animal on the ground: Sequences of B-calls (B+)
- c. Caller foraging near the ground in absence of predator: Sequences of B-calls (B+)
- d. Raptor in the canopy: Sequences of A-calls (A+)
- e. Capuchins in the canopy: Sequences of A-calls (A+)
- f. Felid in the canopy: Sequences of B-calls; A-call(s) interspersed in the sequence at irregular places in 7/14 sequences (B+(A))
- g. Raptor on the ground: Sequences of A-calls with interspersed B-calls (A+B+)

As presented in (1), B-calls are found in a large set of contexts, including predatory and non-predatory situations, which have nothing in common from a human point of view. These puzzling observations led both Cäsar et al. and Schlenker et al. to propose a general meaning for B-calls, as presented in (3) and (5). These claims were built on the assumption that B-calls emitted in predatory and non-predatory contexts are similar, but acoustic analyses were sorely lacking to verify and thus validate these assumptions. To investigate this possibility, we first conducted acoustic analyses on B-calls spontaneously given when the caller is foraging in low strata (“ground B-calls”) and experimentally induced during terrestrial model presentations (“terrestrial predator B-calls” collected from AB+ and B+ sequences as described in (1)). We found that B-calls exhibit a different acoustic structure depending on their eliciting context, and that terrestrial predator B-calls are higher-pitched than ground B-calls (Berthet et al., 2018). These results are presented in (8).

B-calls are also emitted in response to aerial predators on the ground (A+B+ sequences). To fully grasp the meaning of the B-calls, we investigated whether these B-calls are similar to ground or terrestrial predator B-calls. We measured 45 B-calls given to aerial predators on the ground (“aerial B-calls”) and compared them to 192 ground B-calls and 113 terrestrial B-calls by classifying them using a linear discriminant analysis on mel filterbank. Our analysis confirmed that ground and terrestrial predator B-calls are distinguishable based on spectral properties. Importantly, the classification of aerial B-calls was unambiguously bimodal, with 28 aerial B-calls classified as clear terrestrial predator B-calls, and 17 classified as clear ground B-calls, confirming that aerial B-calls do not form a separate class of B-calls. Rather, aerial B-calls are a mix of lower-pitched ground B-calls, and higher-pitched terrestrial B-calls. Results are also presented in (8).

**(8) Results of our successive acoustic analyses on B-calls**

- a. Terrestrial predator B-calls are higher-pitched than ground B-calls
- b. Aerial B-calls are a mix of terrestrial predator and ground B-calls

As presented in (8b), ground B-calls and terrestrial predator B-calls are given to diverse situations that, again, have little in common from an observer’s point of view. Indeed, ground B-calls are given when the caller is descending near the ground and when presented with an aerial predator on the ground, while terrestrial predator B-calls are given to terrestrial predators but also when presented with an aerial predator on the ground. The fact that one context can elicit different B-call types is puzzling, and suggest that either B-calls exhibit two acoustic variants that both refer to general alerts – but this system does not seem of great relevance – or that B-calls do not function as labels as previously thought. In line with the second idea, we suggest that B-calls actually refer to the emotional state of the caller. Moreover, as presented in (8a), the frequency of the B-calls varies with the eliciting situation. Yet, the frequency of a call is highly dependent on the emotional state of the caller (Briefer, 2012), with higher arousal levels eliciting higher frequencies. If the acoustic variation of the B-call was linked to a variation in the emotional state of the caller, our data would support the hypothesis that the caller is more aroused in presence of a terrestrial predator than when it is foraging in a vulnerable position (near the ground, where terrestrial predators are likely to be) but in the absence of immediate threats. We can thus stipulate that B-calls reflect the emotional state of the caller.

When considering the case of A-calls, the picture is clearer. As presented earlier in (3), Schlenker et al. stated that A-calls refer to serious non-ground threats. However, Cäsar (2011) argued that aerial predators represent the major danger to titi monkeys, which suggests that the proposed meaning “serious non-ground threat” is redundant. We thus argued that the “non-ground” specification can be eliminated from the meaning of A-call, resulting in a simplification such that A-calls refer to serious threats (Commier & Berthet, 2019).

However, results presented in (8) suggest that B-calls are linked to the emotional state of the caller. Since A-calls are very high-pitched calls, and because they are given when the caller is facing a serious threat, we emit the hypothesis that A-calls are also related to the emotional state of the caller, and are emitted when the caller is highly aroused. Since B-calls are emitted in less dangerous situations, and because B-calls exhibit two acoustic variants, we conclude that B-calls reflect either low or medium arousal levels in the caller. These hypotheses are synthesized in (9).

**(9) Meaning at the call level: our hypothesis**

- a. Alarm calls of titi monkeys reflect the emotional state of the caller
- b. A-calls reflect high emotional states of the caller
- c. B-calls reflect lower emotional states of the caller

- d. Modular theory: Acoustic structure of the B-calls indicates whether the level of stress experienced by the caller is low (low pitched B-calls) or medium (high pitched B-calls)

Since alarm calls are a reflection of the emotional state of the caller, and in line with Schlenker et al.'s theory described in (5d), we propose that the calls reflect the emotional state of the caller at the precise time at which they are uttered. If this is true, then the sequence reflects the temporal evolution of the emotional state of the caller, as detailed in (10).

**(10) Sequence encoding: the caller emits calls that reflects its emotional state at the time of emission**

- a. Caller descending near the ground: vulnerability to potential terrestrial predator, low arousal: low-pitched B+
- b. Terrestrial predator on the ground: noteworthy threat that is easy to escape from, medium arousal: high-pitched B+
- c. Aerial predator in the canopy: serious threat, high arousal: A+
- d. Terrestrial predator in the canopy: noteworthy threat that is easy to escape, medium arousal with anecdotal peaks of excitement: B+(A)
- e. Aerial predator on the ground: serious threat at first, the caller then realizes that it is not that serious (because not in a hunting position or because it is a fake model that does not move): A+B+ (high- and low-pitched B-calls)

Contrary to the previous proposals of Cäsar et al. and Schlenker et al., this theory can also account for irregularities of the calling patterns, like the one presented in (10d): these irregularities can be due to sudden changes in the arousal, which can be explained by the experimental paradigm (e.g. previous experience of the caller with this model, the fact that the model does not move), the presence of human observer and their distance to the caller, the presence or absence of conspecifics and the changes in the audience, the movement of the caller toward or away from the model, etc.

## 4 Meaning from the listener's perspective

We ran an extensive field study to investigate what information was extracted by listeners at the sequence-level, and what mechanisms they used to decode information (Berthet et al., 2019). We conducted predator presentation as detailed above, recorded the vocal reaction, completed this dataset with that of Cäsar et al. and extracted the first ten calls of each sequence. Each of the N=50 sequences was then characterized by 15 metrics referring to the composition, ordering, temporal structure and complexity of the sequence<sup>1</sup>. We conducted a multimodel inference to investigate whether each metric conveyed information about predator type and/or location. In order to identify the mechanisms used by listeners to extract information from the vocalizations, we conducted playback experiments. We broadcasted the recorded alarm sequences to other individuals, coded the gaze reaction of the subjects and conducted multimodel inference to investigate whether gaze direction of listeners was influenced by i) the origin of the sequence (i.e. the location and type of the predator that elicited the broadcasted sequence) and ii) the metrics characterizing the sequences used as playback stimuli. We present the main results of this study in (11) and our resulting conclusions, in (12).

**(11) Results from Berthet et al., 2019**

- a. Several sequence metrics encode for predator type or predator type and location

---

<sup>1</sup>It must be noted that this analysis was conducted before we investigated the acoustic variants of B-calls. Thus, this study did not consider whether B-calls of the sequences were high- or low-pitched B-calls.

- b. Listeners look more upward and less toward the speaker when listening to sequences recorded from encounters with an aerial predator/a predator in the canopy
- c. Listeners strongly react to the proportion of BB-grams (i.e. the proportion of combinations of 2 B-calls) of the sequence: they look more upward and less toward the speaker when listening to sequences containing a smaller proportion of BB-grams
- d. Sequences recorded from encounters with an aerial predator/a predator in the canopy contain a smaller proportion of BB-grams than sequences recorded from encounters with a terrestrial predator/a predator on the ground

**(12) Conclusions from Berthet et al., 2019**

- a. Sequences encode predator type and location
- b. Listeners can extract information about predator type and location from the sequences
- c. Information is transferred through the proportion of BB-grams of the sequence

But if alarm sequences are mere reflections of the temporal evolution of the arousal state of the caller, how can we explain that alarm sequences convey information about the predator type and location of the predator? As shown in (12), sequences are decoded by listeners using the proportion of BB-grams. This suggests that listeners do not process the alarm sequence call by call, but rather synthesize information by using the proportion of BB-grams. We thus postulate that listeners extract the overall level of stress experienced by the caller during the situation. We did not have the opportunity to conduct playback experiments to determine whether low- and high-pitched B-calls were discriminated by listeners in titi monkeys yet. However, several studies suggest that, in nonhuman animals, listeners can extract emotional information from the acoustic structure of the calls (e.g. Manser et al., 2002). Thus, we postulate that, in addition to relying on the proportion of BB-grams, listeners also process the acoustic structure of the calls to assess the overall emotional state of the caller – a statement that remains to be tested in the field. Listeners could then infer the eliciting situation using their world knowledge, as detailed in (13).

**(13) Sequence decoding: the listener extracts the proportion of BB-grams and uses world knowledge**

- a. Low-pitched B+: maximum proportion of BB-grams, reflecting a low arousal level of the caller. The listener infers that the caller is near the ground
- b. High-pitched B+: maximum proportion of BB-grams, reflecting a medium arousal level of the caller. The listener infers that there is a terrestrial predator
- c. A+: minimum proportion of BB-grams, reflecting a high arousal level. The caller infers that there is an aerial predator in the canopy
- d. B+(A): high proportion of BB-grams, reflecting a medium arousal level with some peaks of excitement. The listener infers that there is a terrestrial predator in the canopy
- e. A+B+: low proportion of BB-grams, reflecting a high arousal that slowly decreases over time but still exhibit some sudden peaks of excitements. The listener infers that there is an aerial predator on the ground

## 5 Conclusion

The titi monkey system appears to rely on a division of labour between semantics (with the natural meaning of A- and B-calls being linked to the emotional state of the caller) and on



pragmatics (with the listener extracting information about the type and location of the predator from the emotion-related sequences, by using its world knowledge). Although based on several observational and experimental datasets, and inferred from rigorous analyses, these conclusions remain to be tested in the field, through playbacks (e.g. by broadcasting artificial sequences with varying proportion of BB-grams) or other innovative measurements, such as the use of thermography to track emotional changes in the caller (e.g. Ermatinger et al., 2019).

Overall, our work suggests that an emotional reaction from an individual can convey reliable and complex information to listeners about the external world, provided that they use their world knowledge to process it. This confirms that the traditional dichotomy between referential signals (i.e. signals that convey information about the external world) and affective signals (i.e. signals that are linked to the emotional state of the caller) is wrong: the affective property of a signal depends on mechanisms of call production in the caller, while the referential property of a signal depends on the listener's ability to extract information from events (Seyfarth & Cheney, 2003).

Here, we remained agnostic about the cognitive processes that may be used by titi monkeys when communicating about alarm situations. First, we investigated meaning as a correlation between a signal and an event, regardless of the intention of the caller. Second, our work suggests that the Informativity Principle of Schlenker et al. (2016a) is not required to explain the vocal behaviour of titi monkeys, limiting debates on the cognitive capacities needed by the caller to apply this concept (Jäger, 2016; Schlenker et al., 2016b). Third, the world knowledge used by titi monkeys to extract information from the sequences does not require any high cognitive process, such as theory of mind or empathy: a simple associative learning between the alarm sequences and the eliciting event can account for it. Of course, we do not exclude the possibility that titi monkeys possess high cognitive processes, but the current data do not need such claims to be explained.

Overall, this work demonstrates that methods from animal linguistics can be highly beneficial to the study of animal communication. Indeed, by applying methods and concepts from ethology and linguistics on several field datasets, we built clear hypotheses that i) clarified a puzzling behaviour, ii) could not have been reached with linguistic or biological analysis only, iii) do not make high claims about the cognitive capacities of the species and iv) highlight new communication mechanisms in a nonhuman species. Animal linguistics represents a powerful new tool to study animal communication and future research should benefit from this approach, to explore the diversity of communicative systems, investigate the similarities and differences between human language and animal communication, and eventually understand better how communication evolved over time.

## Acknowledgements

This work was funded by the Fyssen Foundation, the European Research Council under the European Union's Seventh Framework Program (FP7/2007–2013)/ERC grant agreement No. 283871 (PI: Zuberbühler K.). No. 324115–FRONTSEM (PI: Schlenker), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788077, Orisem, PI: Schlenker). Research was conducted at

Institut d'Études Cognitives, Ecole Normale Supérieure - PSL Research University. Institut d'Études Cognitives is supported by grants ANR-10-IDEX-0001-02 and FrontCog ANR-17-EURE-0017. We acknowledge further fundings from the Swiss National Science Foundation and the University of Neuchâtel. We thank N. Buffenoir, A. Colliot, G. Duvot, G. Mesbahi, C. Ludcher, F. Müschenich, C. Rostan, A. Pajot and A. Pessato for help with data collection, and G. Mesbahi, G. Duvot E. Chemla, and P. Schlenker for helpful discussions and comments on the early drafts of this manuscript. We acknowledge logistic support from Rogério Grassetto Teixeira da Cunha (UNIFAL), the Santuário do Caraça and the Natural Sciences Museum of the Pontifícia Universidade Católica de Minas Gerais.

## References

- Berthet, M., Mesbahi, G., Pajot, A., Cäsar, C., Neumann, C., & Zuberbühler, K. (2019). Titi monkeys combine alarm calls to create probabilistic meaning. *Science Advances*, *5*(5), eaav3991. <https://doi.org/10.1126/sciadv.aav3991>
- Berthet, M., Neumann, C., Mesbahi, G., Cäsar, C., & Zuberbühler, K. (2018). Contextual encoding in titi monkey alarm call sequences. *Behavioral Ecology and Sociobiology*, *72*(8). <https://doi.org/10.1007/s00265-017-2424-z>
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, *288*(1), 1-20. <https://doi.org/10.1111/j.1469-7998.2012.00920.x>
- Cäsar, C. (2011). *Anti-predator behaviour of black-fronted titi monkeys (Callicebus nigrifrons)*. University of St Andrew.
- Cäsar, C., Byrne, R. W., Hoppitt, W., Young, R. J., & Zuberbühler, K. (2012a). Evidence for semantic communication in titi monkey alarm calls. *Animal Behaviour*, *84*(2), 405-411. <https://doi.org/10.1016/j.anbehav.2012.05.010>
- Cäsar, C., Byrne, R., Young, R. J., & Zuberbühler, K. (2012b). The alarm call system of wild black-fronted titi monkeys, *Callicebus nigrifrons*. *Behavioral Ecology and Sociobiology*, *66*(5), 653-667. <https://doi.org/10.1007/s00265-011-1313-0>
- Cäsar, C., Zuberbühler, K., Young, R. W., & Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. *Biology Letters*, *9*(5), 20130535-20130535. <https://doi.org/10.1098/rsbl.2013.0535>
- Commier, S., & Berthet, M. (2019). Commentary: Titi semantics: Context and meaning in Titi monkey call sequences. *Frontiers in Psychology*, *10*, 512. <https://doi.org/10.3389/fpsyg.2019.00512>
- Ermatinger, F. A., Brügger, R. K., & Burkart, J. M. (2019). The use of infrared thermography to investigate emotions in common marmosets. *Physiology & Behavior*, *211*, 112672. <https://doi.org/10.1016/j.physbeh.2019.112672>
- Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T., & Byrne, R. W. (2018). Bonobo and chimpanzee gestures overlap extensively in meaning. *PLOS Biology*, *16*(2), e2004825. <https://doi.org/10.1371/journal.pbio.2004825>
- Heesen, R., Hobaiter, C., Ferrer-I-Cancho, R., & Semple, S. (2019). Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society B: Biological Sciences*, *286*, 20182900. <https://doi.org/10.1098/rspb.2018.2900>
- Jäger, G. (2016). Grice, Occam, Darwin. *Theoretical Linguistics*, *42*(1-2). <https://doi.org/10.1515/tl-2016-0004>
- Manser, M. B., Seyfarth, R. M., & Cheney, D. L. (2002). Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences*, *6*(2), 55-57. [https://doi.org/10.1016/S1364-6613\(00\)01840-4](https://doi.org/10.1016/S1364-6613(00)01840-4)
- Moynihan, M. (1966). Communication in the Titi monkey, *Callicebus*. *Journal of Zoology*, *150*, 77-127. <https://doi.org/10.1111/j.1469-7998.1966.tb02999.x>

- Schlenker, P., Chemla, E., Cäsar, C., Ryder, R., & Zuberbühler, K. (2017). Titi semantics: Context and meaning in Titi monkey call sequences. *Natural Language and Linguistic Theory*, 35(1), 271-298.
- Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., ... others. (2016a). Formal monkey linguistics. *Theoretical Linguistics*, 42(1-2), 1-90.
- Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., ... Zuberbühler, K. (2016b). Formal monkey linguistics: The debate. *Theoretical Linguistics*, 42(1-2), 173-201.
- Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, 54(1), 145-173. <https://doi.org/10.1146/annurev.psych.54.101601.145121>
- Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., ... Manser, M. B. (2017). Exorcising Grice's ghost: An empirical approach to studying intentional communication in animals. *Biological Reviews*, 92(3), 1427-1433. <https://doi.org/10.1111/brv.12289>
- Yip, M. J. (2006). The search for phonology in other species. *Trends in Cognitive Sciences*, 10(10), 442-446. <https://doi.org/10.1016/j.tics.2006.08.001>
- Zuberbühler, K., David, J., & Bshary, R. (1999). The predator deterrence function of primate alarm calls. *Ethology*, 105(6), 477-490.

# Situating Rich Demonstrations in Discourse\*

Christian De Leon

University of California, Los Angeles  
cdeleon@humnet.ucla.edu

## Abstract

I discuss a phenomenon I call *rich demonstration*, wherein a demonstrative gesture is seemingly used to communicate an entire thought. I present an analysis based in discourse coherence theory on which rich demonstrations have their functions in virtue of the interpretive effects of discourse-structuring mechanisms.

## 1 Introduction

Standard cases of demonstrative gestures in conversation are deictic—they serve to fix the reference of some word or phrase (McNeill, 1992; Kendon, 2004). (1) is an example of such a case; the speaker employs a co-speech demonstrative gesture to fix the reference of “this”. (2), by contrast, is a case in which the speaker uses a slightly different combination of speech and gesture to accomplish the same communicative goal as in (1) without any reference-fixing.

- (1)
  - a. X: [This]<sub>b</sub> is going to make my parents furious.
  - b. [X turns their arm, showing a fresh tattoo]
  - c. Y: But they have tattoos themselves!
- (2)
  - a. X: My parents are going to be furious.
  - b. [After speaking, X turns their arm, showing a fresh tattoo]
  - c. Y: But they have tattoos themselves!

I call gestures like the one in (2b) *rich demonstrations*. This is because they seem to be the causal source of crucial discourse interpretation—in this case, a) that there is some reason for the speaker’s parents to be upset, and b) that the reason is the speaker’s new tattoo. What is puzzling about rich demonstrations is that they are formally identical to simple demonstrations—those gestures that merely supply an object for reference, e.g. (1b)—yet a more complex discourse function is attributable to them. *Prima facie*, the discourse function of rich demonstration is like that of assertion. That is, rich demonstrations seem to be intentionally used to inform, and the information they communicate can be responded to by an interlocutor, typically in the same way.

The purpose for this paper is to investigate the discourse function of rich demonstration. I begin by filling in details of the phenomenon and arguing that general models of discourse focused on common ground (Stalnaker, 1978) fail to capture its features. I present an analysis based in discourse coherence theory (Hobbs, 1979; Kehler, 2002; Asher and Lascarides, 2003) on which a rich demonstration’s apparent complex discourse contribution is actually attributable to the semantic effect of the coherence relation structuring the discourse. On my view, there is no need to assign specific propositional content to perceptual scenes or to rich demonstrations themselves. This contrasts with the analysis given by Hunter et al. (2018) who conceive of the

---

\*I’m grateful to Josh Armstrong, Sam Cumming, Gabe Greenberg, Jessica Rett, and Matthew Stone for discussion and encouragement. Thanks also to the Amsterdam Colloquium’s anonymous reviewers for helpful feedback.

mechanisms of discourse coherence as guiding the ascription of content to gestures and to visual scenes more generally.

## 2 Puzzle

### 2.1 More Cases and Details

Grice (1957) famously took the case of Herod's presenting the head of John the Baptist to Salome to not be a case of non-natural meaning (or, for the purposes here, of communication). He (1969) also discussed the related case of a squash player communicating specifically that they are unable to play but not that their leg is bandaged. An adapted version is given in (3):

- (3) a. A: Do you want to go play squash?  
 b. [*B displays their bandaged leg*]

It was important to Grice that one cannot communicate what is perceptually apparent. He allowed for, but did not focus on, the possibility that one can use what is perceptually apparent to communicate information that is not. When this is done via a demonstrative gesture, it is a case of rich demonstration. To further illustrate the phenomenon, I will quickly discuss some other cases.

- (4) a. I moved the table into the living room this morning.  
 b. [*The speaker nods toward some scratches on the wall*]  
 c. I had to buy some new paint.<sup>1</sup>

The speaker of (4) does not nod to merely inform the hearer that the wall is scratched. Instead, they intentionally use the nod to inform the hearer of what happened while moving the table—the speaker scratched the wall with it. That information is necessary for understanding the point of (4c) and how the need to buy paint arose from the events earlier in the day.

- (5) a. [I went to the grocery store earlier]<sub>b</sub>  
 b. [*While talking, the speaker motions to an unopened box of cookies with an open hand*]  
 c. ... if you want any.

In (5), the speaker does not merely present two situations, the shopping and the presence of the cookies. Instead, they successfully communicate that they bought that box of cookies while on their recent shopping trip. Recognizing this is necessary for properly interpreting the biscuit conditional in (5c).

- (6) *Maria is helping Carl hunt for a new apartment. While walking up to a new complex:*  
 a. [*C taps M's shoulder, then points to the building's pool*]  
 b. M: You would get such a bad sunburn.

Maria comments on what would happen in a very specific hypothetical scenario, one in which Carl rents the apartment and decides to frequent the complex's pool. She can only felicitously make that comment given that she understands Carl to be raising that as a live option with his gesture.

---

<sup>1</sup>From Hunter et al. (2018)

These cases illustrate that rich demonstrations do not correspond to any particular formal properties of gestures. Rich demonstrations can take the form of pointing with the index finger, an open hand, nodding the head, and so on. What matters is that the speaker directs the hearer’s attention to something in the conversational vicinity. Rich demonstrations can, but need not, be co-speech, and can even be discourse-initial.

These cases also help illustrate two related reasons for believing that rich demonstrations are functionally assertion-like. First, they exhibit hallmark properties of assertions—they are informative, can be followed up, responded to, etc. Second, as [Hunter et al. \(2018\)](#) observe, rich demonstrations can frequently be replaced by actual assertions without issue. For example, (2b) could be replaced by “I got a new tattoo”, (3b) by “I hurt my leg”, (4b) by “. . . and scratched the wall”, etc., and the same communicative goals would be accomplished. If this is right, an account is needed of how a simple demonstrative gesture can fill the same role as (or a similar role to) assertion.

## 2.2 Beyond Noticing

Here I want to quickly elaborate on why rich demonstrations are puzzling from the perspective of traditional approaches to the semantics and pragmatics of discourse. It is important to first emphasize that it is not puzzling how the demonstrated eventuality comes to be mutually known in the conversation. In the tradition of [Stalnaker \(1978\)](#), we can conceive of the semantic content of a sentence  $s$  as (the characteristic function of) a set of worlds—those worlds in which  $s$  is true. Discourse is modeled as a sequence of updates to the Context Set ( $CS$ ), which represents the body of information compatible with what the discourse participants mutually believe. For now I will represent a discourse as a double consisting in a list of moves and the  $CS$ . Assertions update the conversation by being added to the list of moves and by restricting what is mutually believed.<sup>2</sup>

- (7) For a sentence  $s$ , a proposition  $\llbracket s \rrbracket = \{w : s = 1 \text{ in } w\}$
- (8) A conversation state  $K = \langle M, CS \rangle$  :
- a.  $M = \langle m_1 \dots m_n \rangle \approx$  an ordered list of moves in the discourse
  - b.  $CS = \bigcap \{\llbracket s \rrbracket : \llbracket s \rrbracket \text{ is mutually believed by interlocutors}\}$
- (9) Assertion update:  $K_1; [A(s)] = K_2$  :
- a.  $M_2 = M_1 \cup \{A(s)\} \approx$  add the assertion to the list of moves
  - b.  $CS_2 = CS_1 \cap \llbracket s \rrbracket$

Importantly, given its definition in (8b), the  $CS$  can be updated by non-linguistic events. [Stalnaker \(2002, 2014\)](#) calls such events *manifest events*; they are events that happen in the conversational vicinity that all parties are mutually aware of. Because they are mutually aware of them, the fact that they happened becomes mutually believed, and so the  $CS$  adjusts accordingly. To use his famous example, were a goat to walk suddenly into the room, we would be able to use anaphoric expressions to refer to it, presuppose its presence, and so on (*how did that get in here?*). This means that part of what is informative about a rich demonstration can already be accounted for. That the demonstrated scene is the way it is can be incorporated into the discourse via the same mechanism as all manifest events—via update to mutual belief.

<sup>2</sup>Technically for Stalnaker an assertion  $A(s)$  only directly updates the common ground, which models mutual belief,  $CG = \{\llbracket s \rrbracket : \llbracket s \rrbracket \text{ is mutually believed}\}$ ,  $CG_1[A(s)] = CG_2 = CG_1 \cup \{\llbracket s \rrbracket\}$ . The  $CS$  is then defined in terms of the  $CG$ ,  $CS = \bigcap CG$ . So assertions indirectly update the  $CS$ . But I follow most in focusing on the  $CS$ , and so fold the definition of the  $CG$  into the  $CS$ , per (8b).

What cannot be accounted for by this mechanism is the communicative purpose a speaker has in richly demonstrating. In (2), that the speaker has a fresh tattoo can be handled via traditional means because that fact is mutually salient to the interlocutors (because the speaker made it salient via their gesture). What is missing is the fact that the tattoo constitutes the reason why the parents will be upset. It is not perceptually available, of course, that the tattoo is the reason, and the speaker failed to explicitly mention that there exists any reason at all. Similarly, (3)-(6) all involve interpretations above and beyond what is manifestly available, and those interpretations are crucial for understanding what has been communicated in each discourse.

### 3 Coherence and Demonstration

What is missing from a purely *CS*-focused approach to rich demonstration is the conceptual connection between the demonstrated scene and whatever else the speaker is talking about. For this reason, discourse coherence theory (Hobbs, 1979; Kehler, 2002; Asher and Lascarides, 2003) is well-suited for explaining the phenomenon. In pursuing this approach, I follow others in explaining phenomena involving situated discourse by appeal to coherence mechanisms (Lascarides and Stone, 2009a,b; Stojnić et al., 2013; Stone and Stojnić, 2015). Before discussing the details of my proposal, though, I should first discuss an analysis of rich demonstration from Hunter et al. (2018) that is similar in spirit to mine, but diverges at key points.

The project undertaken by Hunter et al. is specifically aimed at modelling how non-linguistic events in a *Settler's of Catan* game can come to be semantically relevant in a conversation among the game's players. But more generally they are motivated by the same kind of data that motivate me here. They present an expansion of Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003) that allows for the assignment of content to non-linguistic events on the basis of reasoning about coherence relations. That content can interact with standard linguistic SDRT content upon incorporation into a discourse.

What is distinctive about their analysis is that it represents two related conceptual shifts, one from coherence theory specifically and the other from standard discourse modelling generally. With respect to coherence theory, Hunter et al. shift from conceiving of coherence relations as relating discourse segments and affecting discourse on that basis to conceiving of them as holding between events generally, some of which are discourse events, some of which are not. Conceiving of coherence relations in this way leads naturally to the second shift: from taking the mechanisms of discourse interpretation to model how communicative actions have the contents they do to taking them to model how inferences are made about the world generally, even when no communication is involved. These two shifts are made clear when they say,

In our view, semantic structures composed entirely of what are traditionally classified as discourse moves (including, perhaps, discourse dependent nonlinguistic moves) are just a subclass of the kinds of structures that we can use such moves to build. In fact, we think that the kinds of semantic structures built up from coherence relations need not involve any discourse moves at all. Suppose Peter looks out into the garden and sees his cat, Lupin, staring at a pile of leaves. The leaves suddenly move, and Lupin pounces. Peter goes to investigate and finds a baby whipsnake. He now understands why Lupin was staring at the leaves and why the leaves rustled; he also understands that Lupin's pounce was a result of the leaf movement. Yet, neither the snake nor the cat intended to communicate anything, and certainly the snake didn't intend its presence to explain Lupin's behaviour and Lupin didn't

intend his pounce to be a result of the leaf movement. Nevertheless, both the result and meta-level explanation are inferred. (Hunter et al., 2018, pp. 18-19)

The authors take this position regarding discourse for two broad reasons. The first is that we rely on conceiving of external events in particular ways in planning and interpreting discourse events. The second is that specific data (some of which is rich demonstration data) seems difficult to explain without explicitly representing external events as having structured content.

Regarding the first broad reason, while I am generally friendly to the kinds of motivating factors for Hunter et al., I do not think that they necessitate a total reconceptualization of the theoretical role of discourse modelling. But whether this is correct is far beyond the scope of this paper. Here, I merely hope to touch on the second broad reason. In the next section, I present my analysis of rich demonstration and argue that the data can be explained with minimal assumptions from coherence theory and *CS*-focused approaches without the need to model the psychological conceptualization of external events.

## 4 Analysis

The core of my analysis is simple. Per usual in coherence theory, there are some coherence relations such that when they relate two discourse moves, there is an additional semantic effect on the discourse. Rich demonstrations should be considered full, complete, discourse moves, so this kind of semantic effect is operative in cases of rich demonstration. That is what accounts for their assertion-like quality. So rich demonstrations are functionally like assertions in that they feed some object or scene to be utilized by the semantic effect of the operative coherence relation. The intuition, then, is that the richness of rich demonstrations is accounted for by the very mechanisms that account for the non-explicit content in (10), from Kehler (2002).<sup>3</sup>

- (10) a. John took a train from Paris to Istanbul.  
b. He has family there.

Following Kehler (2002), the two sentences in (10) are related by the *Explanation* relation. When *Explanation* holds, the discourse is updated with the further information that the semantic content of the former obtains because the semantic content of the latter obtains.

### 4.1 Adding Coherence

To reflect this formally, I update the definitions of *conversation state* and *update* from (8)-(9):

- (11) A conversation state  $K = \langle M, R, CS \rangle$  :
- $M = \langle m_1 \dots m_n \rangle \approx$  an ordered list of moves in the discourse
  - $R = \{ \langle r, m_i, m_j \rangle \}, \forall m_i \in M, \exists m_j \in M, \exists r \in \mathcal{R} \approx$  a record of coherence relations between moves
  - $CS = \{ w \dots w_n \}$
- (12) Assertion update:  $K_1; [A(s)] = K_2$  :
- $M_2 = M_1 \cup \{ A(s) \} \approx$  add the assertion to the list of moves
  - $R_2 = R_1 \cup \{ \langle r, A(s), m \rangle \}, \exists r \in \mathcal{R}, \exists m \in M_1 \approx$  add the new move's (contextually-determined) coherence relation with some prior move to the record

<sup>3</sup>In Kehler (2002), the case is meant to contrast with the classic case of incoherence from Hobbs (1979): "John too a train from Paris to Istanbul. #He likes spinach."



$$c. \quad CS_2 = CS_1 \cap \llbracket s \rrbracket \cap \llbracket \{ \llbracket r(A(s), m) \rrbracket : \langle r, A(s), m \rangle \in R_2 \} \rrbracket$$

$M$ , the record of moves, and how it's updated remain unchanged. I add to  $K$  a record of coherence relations between moves  $R$  that requires that each move in  $M$  coherently relate to some other move(s).<sup>4</sup> The specific relation  $r$  is pulled from  $\mathcal{R}$ , the set of all coherence relations.<sup>5</sup> The determination of which  $r \in \mathcal{R}$  obtains is, per usual, a matter of the morpho-syntactic features of the utterance and surrounding discourse, as well as general pragmatic reasoning on the basis of world knowledge. The relations that matter for me here will have semantic constraints, reflected as a set of worlds  $\llbracket r(m_n, m_m) \rrbracket$ .<sup>6</sup> Those semantic constraints restrict what the new  $CS$  can be in order to capture content that is not explicit in any assertions. The LF of (10) is in (13).<sup>7</sup>

$$(13) \quad [A(\langle 10a \rangle = \text{took-train}(\text{john}))]; [A(\langle 10b \rangle = \text{lives}(\text{john's-fam}, \text{istanbul}))]$$

To see how this captures the basic intuition that the speaker of (10) communicates that John went to Istanbul in order to visit family even though that information is not explicit in the discourse, we can use the following simple model:

$$\begin{aligned} \llbracket \text{took-train}(\text{john}) \rrbracket &= \{w_1, w_2, w_3\} & \llbracket \text{lives}(\text{john's-fam}, \text{istanbul}) \rrbracket &= \{w_1, w_2, w_4\} \\ \llbracket \text{visit}(\text{john}, \text{john's-fam}) \rrbracket &= \{w_2, w_5\} \end{aligned}$$

The first update adds the assertion to the list of moves  $M$  and requires that  $CS \subseteq \llbracket \text{took-train}(\text{john}) \rrbracket$ .  $R = \emptyset$  because  $A(\text{took-train}(\text{john}))$  is discourse-initial, and so cannot coherently relate to any prior discourse moves. The second adds the new assertion to  $M$ , and adds to  $R$  that the new assertion relates to the first via *Explanation*. Following Kehler (2002), the reason for explanation involves the tense/aspect of the two sentences as well as the general world knowledge that a normal explanation for taking trips is to visit family. The  $CS$  is updated not only with the information that  $\text{lives}(\text{john's-fam}, \text{istanbul})$  but also that  $\text{visit}(\text{john}, \text{john's-fam})$ , which falls out of the semantics of *Explanation*. The evolution of the discourse can be seen in the following representations of  $K$ :

$K_1 = [A(\text{took-train}(\text{john}))]$			$K_2 = K_1; [A(\text{lives}(\text{john's-fam}, \text{istanbul}))]$		
$M_1$	$R_1$	$CS_1$	$M_2$	$R_2$	$CS_2$
$\langle A(\langle 10a \rangle) \rangle$	$\{\}$	$\{w_1, w_2, w_3\}$	$\langle A(\langle 10a \rangle), A(\langle 10b \rangle) \rangle$	$\{\langle \text{Exp}, \langle 10b \rangle, \langle 10a \rangle \rangle\}$	$\{w_2\}$

## 4.2 Adding Demonstration

So far the simple formalism is able to handle conversation update with the semantic effects of operative coherence relations. What is still missing is a way for demonstrated scenes to factor in. To account for this, I make one final revision to the definition of discourse update from (12):

$$(14) \quad \text{Discourse update: } K_1; [\Phi] = K_2 : \\ a. \quad M_2 = M_1 \cup \{\Phi\} \approx \text{add the move to the list of moves}$$

<sup>4</sup>I gloss over discourse-initial assertions for simplicity. The definition also allows for the possibility that new moves coherently relate to multiple prior moves.

<sup>5</sup>I remain mostly agnostic about the composition of  $\mathcal{R}$  since my analysis does not essentially depend on which or how many relations exist. I intend this sketch to be compatible with approaches in the style of Hobbs/Kehler, SDRT, RST (Mann and Thompson, 1988), etc.

<sup>6</sup>For purely structural relations like *Parallell*, we can simply assume their semantic value is identical to the set of all worlds  $\mathcal{W}$ .

<sup>7</sup>I am ignoring many factors in the representation of (10), e.g. pronoun resolution, that this system could easily be adapted to explain using coherence mechanisms (cf. Stojnić et al., 2017).

- b.  $R_2 = R_1 \cup \{\langle r, \Phi, i \rangle\}$ ,  $\exists r \in \mathcal{R}, \exists m \in M_1 \approx$  add the new move's (contextually-determined) coherence relation with some prior move to the record
- c.  $\Phi = A(s) \Rightarrow CS_2 = CS_1 \cap \llbracket s \rrbracket \cap \bigcap \{\llbracket r(A(s), m) \rrbracket : \langle r, A(s), m \rangle \in R_2\}$
- d.  $\Phi = D(\delta) \Rightarrow CS_2 = CS_1 \cap \bigcap \{\llbracket r(D(\delta), m) \rrbracket : \langle r, D(\delta), m \rangle \in R_2\}$

I distinguish between two kinds of update, assertion  $A(s)$  and demonstration  $D(\delta)$ , where  $\delta$  is a perceptual scene. Any discourse move  $\Phi$  updates  $M$  and  $R$  in the same way—by being added to  $M$  and by resolving to an appropriate coherence relation  $r$ . Assertions and demonstrations update the  $CS$  in slightly different ways. Assertions restrict the  $CS$  to the content of the asserted sentence as well as the content of the operative coherence relation(s). Demonstrations restrict the  $CS$  only with respect to the operative coherence relation(s). The important result of this is that the semantic content of coherence relations accounts for all of the communicative informational updates triggered by the gesture. That is, I do not employ coherence mechanisms to interpret or determine the content of the gesture or visual scene, as do Hunter et al.. Instead, I employ them to account for the general communicative strategy used by the speaker. To see how this explains the intuitions surrounding rich demonstration, I will use a simple model paired with the LF of (2) in (15).

$$(15) \quad [A(\langle 2a \rangle = \text{furious}(\text{parents}))]; [D(\langle 2b \rangle = \text{TAT})]; [A(\langle 2c \rangle = \text{has-tattoo}(\text{parents}))]$$

$$\begin{aligned} \llbracket \text{furious}(\text{parents}) \rrbracket &= \{w_1, w_2, w_3, w_4\} & \llbracket \text{cause}(\text{tattoo}, \text{furious}) \rrbracket &= \{w_2, w_3, w_4\} \\ \llbracket \text{has-tattoo}(\text{speaker}) \rrbracket &= \{w_1, w_2, w_3\} & \llbracket \text{has-tattoo}(\text{parents}) \rrbracket &= \{w_2\} \end{aligned}$$

$K_1 = [A(\text{furious}(\text{parents}))]$			$K_2 = K_1; [D(\text{TAT})]$		
$M_1$	$R_1$	$CS_1$	$M_2$	$R_2$	$CS_2$
$\langle A(\langle 2a \rangle) \rangle$	$\{\}$	$\{w_1, w_2, w_3, w_4\}$	$\langle A(\langle 2a \rangle), D(\langle 2b \rangle) \rangle$	$\{\langle \text{Exp}, \langle 2b \rangle, \langle 2a \rangle \rangle\}$	$\{w_2, w_3\}$
$K_3 = K_2; [A(\text{has-tattoo}(\text{parents}))]$					
$M_3$		$R_3$		$CS_3$	
$\langle A(\langle 2a \rangle), D(\langle 2b \rangle), A(\langle 2c \rangle) \rangle$		$\{\langle \text{Vio}, \langle 2c \rangle, \langle 2b \rangle \rangle, \langle \text{Exp}, \langle 2b \rangle, \langle 2a \rangle \rangle\}$		$\{w_2\}$	

$K_1$  is the result of the first update, the assertion of `furious(parents)`. The assertion is added to  $M$  and the  $CS$  shrinks such that  $CS \subseteq \llbracket \text{furious}(\text{parents}) \rrbracket$ . The demonstration in (2b) is then added to  $M$  and recorded in  $R$  as coherently relating to the prior assertion via *Explanation*. I elaborate on what goes into the selection of the appropriate coherence relation with respect to rich demonstrations in §4.3, but for now the fact that the assertion of (2a) heavily biases *Explanation* should suffice.<sup>8</sup> The  $CS$  is updated so that it is a subset of  $\llbracket \text{cause}(\text{tattoo}, \text{furious}) \rrbracket$ , which is the semantic fallout of  $\text{Exp}(A(\text{furious}(\text{parents})), D(\text{TAT}))$ . But that by itself would merely result in  $CS = \{w_2, w_3, w_4\}$ . The reason that  $w_4 \notin CS_2$  is that as a byproduct of the gesture, it is manifestly obvious to all interlocutors that `has-tattoo(speaker)`. That is, typical Stalnakerian mechanisms for  $CS$ -updates are perfectly capable of reflecting what is mutually known in a conversation, as discussed in §2.2. Finally, the third discourse move  $A(\text{has-tattoo}(\text{parents}))$  updates  $M$  in the normal way, records in  $R$  that the assertion stands in the *Violated Expectation* relation with  $D(\text{TAT})$ , and shrinks the  $CS$  in the appropriate way.

I have so far said little about the visual scene  $\delta$  (in this case,  $\delta = \text{TAT}$ ) that is the argument to the demonstration  $D(\cdot)$ . One might worry that without an account of how  $\delta$  is conceptualized in the minds of the speaker and hearer, the general account of rich demonstrations is lacking. After all, it will matter for the discourse whether  $\delta$  is the tattoo on the speaker's arm, the

<sup>8</sup>This is a large part of why *Explanation* is operative in (1), which involves the same discourse-initial assertion, followed by another assertion instead of by a rich demonstration.

image it depicts, or something behind the arm entirely. This is the kind of consideration that leads Hunter *et al.* to discuss visual interpretation beyond discourse. I choose to remain mostly agnostic about the nature of  $\delta$ 's content for two reasons.

First, while it is undoubtedly important to investigate how visual signals are interpreted, I consider such an investigation peripheral to the question of the discourse function of rich demonstration. So long as  $\delta$  is capable of encoding information of the same type as is normally used in coherence-based explanations, the account here captures how and why rich demonstrations fulfil the task they intuitively do. I take the question of  $\delta$ 's resolution to be analogous to the question of how phonetic items are resolved to particular semantic items; such questions are deeply important, but their answers are not answers to questions about how semantic items themselves serve the communicative functions they do. The analysis presented here is intended to (begin to) answer questions regarding the semantic and pragmatic functions of rich demonstration.

Second, filling in the content of  $\delta$  on the basis of reasoning about what the communicative goals of the speaker are risks limiting the discourse representation unnecessarily, and thereby wrongly predicting that some followups are unavailable. For example, it may be genuinely unclear which aspect of the tattoo will upset the parents. My intuition is that what the speaker should be taken to say is, roughly, nothing more than that their parents will be upset because of the tattoo. Were the interlocutor to ask why the parents would be upset, they should be taken to ask a question about the psychology and motivations of the parents, not a metalinguistic question about what the speaker in fact said. This means that leaving the details of  $\delta$ 's content to the perceptual faculties and not to the mechanisms of discourse interpretation will suffice for investigating how rich demonstrations accomplish their tasks.

This analysis has consequences for the theoretical question about the meanings of gestures. My preferred interpretation of this analysis is that rich demonstrations are themselves no more meaningful than are simple demonstrations (i.e. gestures that resolve referential expressions, as in (1b)); they merely designate a perceptual scene. What is interesting about our gestural conventions is that such a designation can suffice for accomplishing a complex communicative task—the same task accomplished by assertion. The reason this is possible has its roots in discourse structure. On a coherence-theoretic approach, explicit moves must relate in various ways to one another, and the specific ways they relate result in informational updates above and beyond the updates mandated by the meanings of the moves. That rich demonstrations exist shows *a)* what kind of actions can be counted as full discourse moves in their own right, and *b)* that some gestures, *qua* full discourse moves, need not have propositional meanings in order to function. Assertions are of propositional contents and also cause further propositional updates; rich demonstrations can cause those same further updates despite not being meaningful in the way assertions are.

### 4.3 Constraints on Coherence

The analysis I have presented relies on demonstrative discourse moves being related for appropriate coherence relations. I would like to conclude by briefly discussing the constraints on the selection of available coherence relations for connecting rich demonstrations to surrounding discourse moves. The determination of the operative coherence relation connecting two assertions is a matter of pragmatic reasoning on the basis of the grammatical features of each assertion, and the same is true when connecting an assertion with a rich demonstration.

First, the preceding assertion will bias or require resolution to some coherence relations, per usual. This will be a function of the particular phrases used and the sentence's tense/aspect.

I claim further that rich demonstrations require (or at least strongly bias) subordinating coherence relations to prior discourse. That is, a rich demonstration must maintain the current discourse topic or goal. Particularly common relations, as is obvious at this point, are *Explanation* and *Elaboration*. The former is operative in cases such as (2), the latter in cases like (4). It might seem at first glance that some coordinating relations, namely *Result* (which can be considered the dual of *Explanation*), are possible candidate relations. I contend, however, that *Result* readings are strange at best. Contrast the following cases:

- (16) *Explanation*:
- a. [The speaker gestures toward a broken vase]
  - b. Alex threw a football earlier.
- (17) *Result*:
- a. Alex threw a football earlier.
  - b. ?[The speaker gestures toward a broken vase]

The intuitions about these cases are admittedly subtle, but instructive. It is clear what the speaker intends to convey in (17), but it is difficult to force the gesture to come after the completion of the utterance. It is much more natural for the gesture to occur during the utterance or before it, as in (16). But such an occurrence would permit *Explanation*.<sup>9</sup> The connection between throwing a football and breaking a vase is easy enough to make, but it feels like the speaker would more naturally point as a followup, not to introduce a new event to be discussed. That (16) is arbitrarily more natural than (17) is explained by the fact that (17) only makes sense with a coordinating relation, which rich demonstrations resist. Nevertheless, there are variants where the gesture's being after the utterance is perfectly natural:

- (18) *Elaboration*:
- a. Alex was throwing a football around,
  - b. [The speaker gestures toward a broken vase]

Changing the aspect of the sentence to the progressive makes the trailing gesture more natural. The reason for this is that the progressive biases *Elaboration*. The eventuality described in (16) and (17) involves one throwing event, but the eventuality described in (18) involves many throwing events. Expanding the eventuality to Alex's general activity means that the broken vase eventuality can be contained within it. That is, more details can be given about the general activity, which the speaker does via the rich demonstration. This is, of course, a first gloss. More research should be done into the constraints gestures impose on discourse structure.

## 5 Conclusion

Broadly, the purpose for this paper was to expand on current research into gestural communication and situated conversation more generally within a coherence-theoretic framework. Specifically, I have focused on the phenomenon of rich demonstration and argued for an analysis on which its complex discourse function is accounted for by being a discourse move in a full sense, and thereby being capable of relating to surrounding discourse via coherence relations.

---

<sup>9</sup>Note, however, that my claim is that rich demonstrations ought to relate to *prior* discourse via subordinating relations. I make no claim about posterior discourse. So (16) is not an instance of my prediction; its purpose is merely to provide a more natural way of using a gesture to communicate the intended thought in (17).

This analysis makes clear several paths for further research into multimodal communication. First, other phenomena should be investigated and situated with respect to their expressive content and their discourse-structural function. Other gestures, facial expressions, gaze, and more all have deep communicative effects to be studied. Second, further research can be done with respect to the nature of coherence relations specifically and discourse interpretation generally. As I touched on in §3, I have approached this topic with an eye toward maintaining the distinction between discourse interpretation and general reasoning about world events. I have not argued for the distinction, but take myself to have provided an analysis of a phenomenon that is compatible with it.

## References

- Nicholas Asher and Alex Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.
- Herbert Paul Grice. Meaning. *The Philosophical Review*, 66(3):377–388, 1957.
- Herbert Paul Grice. Utterer’s meaning and intention. *The Philosophical Review*, 78(2):147–177, 1969.
- Jerry R. Hobbs. Coherence and Coreference. *Cognitive Science*, 3(1):67–90, 1979.
- Julie Hunter, Nicholas Asher, and Alex Lascarides. A formal semantics for situated conversation. *Semantics and Pragmatics*, 11(10), 2018.
- Andrew Kehler. *Coherence, Reference, and the Theory of Grammar*. Center for the Study of Language and Information, 2002.
- Adam Kendon. *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- Alex Lascarides and Matthew Stone. Discourse coherence and gesture interpretation. *Gesture*, 9(2):147–180, 2009a.
- Alex Lascarides and Matthew Stone. A Formal Semantic Analysis of Gesture. *Journal of Semantics*, 26(4):393–449, 2009b.
- William C. Mann and Sandra A. Thompson. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 1988.
- David McNeill. *Hand and Mind*. The University of Chicago Press, 1992.
- Robert Stalnaker. Assertion. In Peter Cole, editor, *Pragmatics*, pages 35–332. Academic Press, 1978.
- Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25(5/6):701–721, 2002.
- Robert Stalnaker. *Context*. Oxford University Press, 2014.
- Una Stojnić, Matthew Stone, and Ernie Lepore. Deixis (Even Without Pointing). *Philosophical Perspectives*, 27:502–525, 2013.
- Una Stojnić, Matthew Stone, and Ernie Lepore. Discourse and Logical form: Pronouns, Attention, and Coherence. *Linguistics and Philosophy*, 40:519–547, 2017.
- Matthew Stone and Una Stojnić. Meaning and Demonstration. *Review of Philosophy and Psychology*, 6(1):69–97, 2015.

# Towards a uniform super-linguistic theory of projection\*

Maria Esipova


Princeton University, Princeton, NJ, USA  
mesipova@princeton.edu

## Abstract

Formal semantics/pragmatics has recently seen an upsurge in studying secondary modality content like gestures and facial expressions. Much of this work has focused on how such content projects from under semantic operators, but most of it has assumed that projection of secondary modality content is governed by independent rules. I argue that projection behavior of compositionally integrated content is guided by the same linguistic principles in all modalities, but to see that we need to treat secondary modality morphemes as linguistic objects across the board. I apply this approach to two novel case studies: conventionalized co-speech gestures and degree modifiers in various modalities.

## 1 Introduction

The question of how gestures accompanying speech contribute meaning has lately been gaining traction in formal semantics/pragmatics (Lascarides & Stone 2009; Ebert & Ebert 2014; Hunter 2018; Schlenker 2018a, a.o.). This line of inquiry has been extended to other types of secondary modality content, e.g., facial expressions and suprasegmental morphemes, in Schlenker 2018b. Much of this work has focused on how contributions of non-conventionalized, iconic gestures, such as in (1), project from under various semantic operators, such as negation, *if*, modals, etc.<sup>1</sup>

- (1) a. Lea might bring  her dog<sub>LARGE</sub>.  
→ Lea's dog is large.
- b. Zoe might  shoot at the target<sub>LONGBOW</sub>.  
→ If Zoe shoots, she'll shoot a longbow.

Most of this literature has assumed that projection patterns of secondary modality content are determined by modality-specific rules. This view is most prominent in Schlenker's (2018b) typology of "iconic enrichments", which aims to predict if/how a piece of secondary modality content  $X$  projects, based on: (i) whether  $X$  has its own time slot or co-occurs with something in a more primary modality, and (ii) whether  $X$  is an "internal" or "external" enrichment, i.e., whether it's "syntactically eliminable". Thus, *co-speech gestures* like in (1) co-occur with speech, a more primary modality, and are "external", as they can be omitted. Such content is predicted by Schlenker's typology to trigger "cosuppositions" (assertion-dependent presuppositions).

---

\*I thank the committee of my dissertation, where I first made some of the observations discussed here: Lucas Champollion, Ailís Cournane, Kathryn Davidson, Stephanie Harves, and Philippe Schlenker; members of the Super Linguistics Research Group at the University of Oslo—with special thanks to Patrick Grosz for the very insightful interactions; members of the MorphBeer reading group at NYU for a recent inspiring discussion about degrees; and four anonymous reviewers for the Super Linguistics workshop at the Amsterdam Colloquium.

<sup>1</sup>I write gesture labels in ALL CAPS. Underlining approximately indicates temporal alignment of the gesture without making any syntactic claims. Illustrations are sometimes added at the approximate onset of the gesture.

While Schlenker’s (2018b) paper contains some extremely valuable empirical observations, the typology proposed there is not based on linguistically meaningful notions or principles. For one thing, when discussing the various “enrichments”, the paper rarely, if at all, makes explicit assumptions about their lexical semantics or how they compose, if at all, with the expressions they are “enriching”.

Next, the “internal” vs. “external” distinction is not defined in a theoretically meaningful way. The decisions about what’s “internal” and what’s “external” in speech and sign are made in a seemingly ad hoc fashion, and no coherent view of the architecture of grammar in which this distinction could be grounded is presented. The notion of “eliminability” is applied to at least two completely different distinctions: syntactic adjuncts vs. non-adjuncts (with the latter comprising both verbs and their arguments) and segmental vs. suprasegmental morphemes (with the latter exemplified by vowel lengthening, discussed in subsection 4.3).

While the distinction between content that has its own time slot and content that doesn’t is potentially useful, it is not well-defined in Schlenker 2018b either, and prosodic phrasing is ignored, even though it can be indicative of syntactic structure (Price et al. 1991 et seq.) and is especially crucial in distinguishing between appositive and non-appositive content, which is central to Schlenker’s notion of “post-speech” gestures.

All this makes Schlenker’s typology hard to apply to new empirical phenomena. However, when we do manage to do so, we start seeing this typology make the wrong predictions, as I will show in this paper. I will also show that approaching secondary modality content as bona fide linguistic objects at all levels of representation reveals that there is no need for a typology like Schlenker’s, since secondary modality expressions fit into the same typology of projection patterns as primary modality expressions. I will ground this super-*linguistic* approach in the theoretical assumptions from Esipova 2019a,b, summarized in section 2. I will then discuss two new case studies supporting this approach over Schlenker’s: conventionalized gestures (section 3) and degree modifiers in various modalities (section 4). Section 5 concludes. Audios and videos of selected examples from this paper are available at: <https://osf.io/5gu4d/>.

## 2 Background

One important distinction I’ll be making is between MODIFIERS and SUPPLEMENTS. Modifiers, exemplified by attributive adjectives in (2), are, roughly, expressions of type  $\langle \tau, \tau \rangle$ . SUBJECTIVE MODIFIERS exhibit SUBJECTIVE ENTAILMENT, i.e., they combine with an expression  $\alpha$  yielding  $\beta$  that entails  $\alpha$  via generalized entailment. For example, in (2a), *blond* and *skillful* are subjective modifiers. NON-SUBJECTIVE MODIFIERS, like *alleged* in (2b), don’t exhibit subjective entailment.

- (2) a. Zoe is a {blond, skillful} stuntwoman.  
       → Zoe is a stuntwoman.  
     b. Daisy is an alleged criminal.  
       ↯ Daisy is a criminal.

Supplements, exemplified in (3) by appositives (an appositive relative clause and a nominal appositive), combine with an expression, pass it unchanged for subsequent composition, but also yield a proposition of a special kind about it. I’ll remain agnostic about how exactly the special status of this proposition is operationalized (see, e.g., Potts 2005; Koev 2013).

- (3) I will invite Zoe, (who is) a stuntwoman.

Specific instances of subjective modifiers can be RESTRICTING, when they are used to pick out

a smaller part of the denotation of the expression they are combining with, as in (4a), or NON-RESTRICTING. The latter are truth-conditionally vacuous modifiers that are used to add some extra information about the denotation of the expression they are combining with, as in (4b).<sup>2</sup>

- (4) a. I should be eating less **saturated** fats and more **healthy** fats.  
 ↗ I should be eating less fats and more fats.  
 ↗ All fats are {saturated, healthy}.
- b. *Context: The speaker believes that processed meat causes cancer.*  
 I shouldn't be eating so many deadly hot dogs.  
 → I shouldn't be eating so many hot dogs.  
 → All hot dogs are deadly.

Non-restricting modifiers give rise to inferences about the truth-conditional equivalence of the expression they modify and the result of modification; in (4b), it is the inference that being a hot dog is equivalent to being a deadly hot dog. Crucially, as noted in Esipova 2019a,b as a refinement of the original definition of non-restricting modifiers in Leffel 2014, this equivalence is assessed relative to the local context of the modified expression, as shown in (5).

- (5) *Context: The speaker just read an article saying that processed meat might be causing cancer, but they are not ready to embrace it as a fact just yet.*  
 If processed meat causes cancer, I shouldn't be eating so many deadly hot dogs.  
 ↗ All hot dogs are deadly.  
 → If processed meat causes cancer, all hot dogs are deadly.

In Esipova 2019a,b, I argue that these non-restricting modifier inferences arise pragmatically when the addressee recognizes the speaker's intention to use a modifier as a truth-conditionally vacuous one. Various pragmatic factors affect whether a given modifier instance is restricting (e.g., adjectives with an attitudinal meaning component like *lovely* are often non-restricting).

Whichever specific analysis one assumes for supplements, it's an empirical fact that, in contrast to modifiers, supplements can never be restricting, as shown in (6), and project very strongly, as shown in (7), in a way that suggests conventional rather than pragmatic triggering.

- (6) #<sub>(IP I invited Zoe)</sub>, (<sub>IP (who is) a **stuntwoman**</sub>), (<sub>IP not Zoe</sub>), (<sub>IP (who is) a **politician**</sub>).<sup>3</sup>
- (7) #I don't know if Zoe is a stuntwoman, but if you invite Zoe, (who is) a stuntwoman, you should show her your muscle car.  
 Intended: '...if (you invite Zoe and she is a stuntwoman)...'

The exceptional cases in which appositives do interact with semantic operators, noted first in Schlenker 2013, are subject to strict discourse constraints (Jasinskaja & Poschmann 2018), which is distinct from the default availability of restricting uses for subsecutive modifiers.

For a more detailed and technical discussion of modifiers and supplements, the viewer is referred to Esipova 2019a,b. For the purposes of this paper, the discussion above will suffice.

Next, I assume an inverted Y model of grammar whereby the syntax creates hierarchical, non-linearized structures of abstract objects (labels, features, etc.), which get shipped off to the compositional semantics to be interpreted and to the phonology to eventually create pronounceable sequences of (possibly overlapping) elements. Vocabulary insertion and linearization

<sup>2</sup>I use **bold** to indicate prosodic (contrastive) focus marking; for English spoken material, this is usually an (L+)H\* pitch accent on the stressed syllable of the bolded word in MAE-ToBI terms (Beckman & Ayers 1997).

<sup>3</sup>IP stands for 'Intonational Phrase' in MAE-ToBI terms; the appositives in (6) are not to be confused with prosodically integrated nominal modifiers as in (*IP I invited Zoe the stuntwoman*).



happen after the split. This view entails that at the level of syntax and compositional semantics no distinctions can be made between primary and secondary modality exponents, nor can there be distinctions made at this level based on the linearization properties of a given exponent. Any modality- or linearization-specific effects emerge in the phonology broadly construed (i.e., during linearization, prosodic phrasing, articulation, etc.) or in the pragmatics, i.e., during post-compositional reasoning about the speaker's beliefs and intentions.

The modifier vs. supplement distinction is first and foremost a compositional one, and it affects which projection mechanism is made available (in the case of modifiers) or enforced (in the case of supplements) for a given expression. Therefore, under the architectural assumptions above, it is impossible to maintain a view whereby gestures that are linearized in a certain way uniformly project as supplements or as non-restricting modifiers, which unsubstantiates the earlier debate about gesture projection in Ebert & Ebert 2014 vs. Schlenker 2018a.<sup>4</sup>

In Esipova 2019a,b, I showed that non-conventionalized gestures like those in (1) should be analyzed under the assumptions above. Thus, *LARGE* can be construed of as iconically representing the property of being large and compose as a modifier (like the adjective *large*), giving rise to a projecting inference only if non-restricting. Or it can be construed of as iconically representing a large object, composing and projecting as a supplement (like the nominal appositive *a large animal*). I furthermore maintain that co-speech modifier gestures prefer to be non-restricting for pragmatic and possibly prosodic reasons, although restricting interpretations, as in (8), are in principle available for them, which is supported by experimental results.



- (8) *Context: We are going on a group tour and need to rent a van. Stephanie, who has a Pug and a Great Dane, is planning to bring one of her dogs with her.*  
 ?%Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing her her **dog**<sub>SMALL</sub>, we'll be fine, but if she's bringing her her **dog**<sub>LARGE</sub>, we should get a bigger van. (video available)

However, non-conventionalized gestures like in (1) are, on the one hand, too unconstrained, because their semantic type is determined on the spot based on how their iconic content is interpreted. On the other hand, they are too constrained, because the types of meaning that can be encoded with purely iconic means are limited. Studying non-conventionalized gestures only is, thus, potentially misleading, since if most non-conventionalized gestures can be construed of as modifiers, and independent post-syntactic considerations make it hard for modifier gestures to have their own time slot in English (as I argue is the case in Esipova 2019b), it is easy to convince oneself that gestures project in a uniform way depending on their linearization. Once, however, we start expanding our empirical scope to include more conventionalized secondary modality content, we start seeing even more clearly that the projection behavior of a given piece of content can't possibly be determined by the parameters assumed in Schlenker's typology.

### 3 Conventionalized gestures

Some conventionalized gestures can be substantive modifiers, too, and, thus, exhibit the projection behavior pattern proper to modifiers. For instance, in both examples in (9), the contribution of the gesture does seem to project by default, even though a restricting interpretation of the gesture would be perfectly reasonable and can be obtained under pressure, as shown in (10). *DRINK* is a Russian conventionalized gesture that means 'drink (alcohol)' and involves flicking one's finger on one's neck (or tapping one's neck with the back of one's hand).

<sup>4</sup>As argued in Esipova 2019a,b, Schlenker's cosuppositions are non-restricting modifier inferences in disguise.


- (9) a. If you bring  a semanticist<sub>CRAZY</sub> to my talk, I'll likely fight with them.  
 → If you bring a semanticist to my talk, I'll likely fight with them.  
 → All semanticists are crazy.
- b. If we wanna  celebrate<sub>DRINK</sub> my defense, we better go to a store now.  
 → If we wanna celebrate my defense, we better go to a store now.  
 → If we celebrate my defense, we'll do so by drinking alcohol.
- (10) a. ?If Kim brings her brother<sub>CRAZY</sub>, I'll fight with him, but if she brings her **normal** brother, that's OK.<sup>5</sup>  
 ≈ If Kim brings her **crazy** brother...
- b. ?If we wanna celebrate<sub>DRINK</sub> my defense, we better go to a store now, but, of course, we can also celebrate **without** alcohol.  
 ≈ If we wanna celebrate my defense by drinking **alcohol**...

Schlenker's typology could treat these gestures like those in (1) (with the same objections as those raised in Esipova 2019a,b). However, other conventionalized gestures seem to be non-subjective modifiers, e.g., *AIR-QUOTES* in (11a) is akin to *so-called* or *quote-unquote*.<sup>6</sup>

- (11) a. Kim is bringing her  friend<sub>AIR-QUOTES</sub> to the party.  
 b. Kim is bringing her {so-called, quote-unquote} friend to the party.

However we go about describing and analyzing projection behavior of non-subjective modifiers, Schlenker's cosuppositions (which are, once again, just non-restricting modifier inferences) aren't the way to go, because if *friend*<sub>AIR-QUOTES</sub> doesn't even entail *friend*, the two certainly can't be truth-conditionally equivalent.

Finally, some conventionalized gestures can only be supplements, e.g., *FINGERS-CROSSED*, which is akin to sentence-level optative adverbs like *hopefully* or a parenthetical *fingers crossed*.<sup>7</sup>

- (12) a. If  a friend of mine wins the race<sub>FINGERS-CROSSED</sub>, I'll buy them a drink.  
 b. If a friend of mine, {hopefully, fingers crossed}, wins the race, I'll buy them a drink.  
 ✓ → I want a friend of mine to win.  
 ✗ ≈ If (a friend of mine wins and I wanted them to)...

Again, Schlenker's typology makes the wrong predictions for *FINGERS-CROSSED* in (12): it's a co-speech gesture, but doesn't trigger a cosupposition; instead, it projects conventionally, like a supplement. But if we exclude all conventionalized gestures from a typology of projection

<sup>5</sup>I apologize for the potentially sanist nature of this example.

<sup>6</sup>I thank Patrick Grosz (p.c.) for drawing my attention to both *AIR-QUOTES* and *FINGERS-CROSSED*.

<sup>7</sup>Such optatives don't only adjoin to clauses, actually, at least not on the surface. Here's a naturally occurring example of a DP-adjoining *FINGERS-CROSSED* (cf. DP-adjoining epistemic adverbs, conveniently also illustrated in (i), discussed in Ernst 1984 et seq.); a spoken parenthetical *fingers crossed* would work here, too:

(i) That was Clive's friend in Vice, who I've enlisted in my search for that Beanpole Bob guy and potentially a zombie cure<sub>FINGERS-CROSSED</sub>. ('iZombie', S05E08, video available)

patterns that includes gestures, we would be missing the parallel between gestures like *LARGE* and gestures like *CRAZY*. Also, this would require one to take a stance on what's conventionalized and what isn't, treating it as a categorical distinction, which seems implausible. It would be impossible to exclude some, but not other conventionalized gestures without making reference to their syntactic and/or lexical properties, which would require assuming a properly linguistic approach to gestures. But assuming said approach reveals that a typology like Schlenker's isn't necessary, as gestures fit into the same typology of projection patterns as spoken content.

## 4 Degree modifiers cross-modally

### 4.1 Degree modifiers in the primary modality

Open-scale<sup>8</sup> degree modifiers are persistently truth-conditionally non-vacuous (i.e., restricting) by default. This is true not only for run-of-the-mill degree modifiers like *very*, *extremely*, *truly*, etc., but also for degree modifiers with an attitudinal component like *surprisingly* and even for expressives used as degree modifiers, which are never restricting in their purely expressive uses:

- (13) a. If the movie is {very, extremely, truly, surprisingly, fucking, bloody, damn} good, I'll stay till the end of the credits.  
 ↗ If the movie is good, I'll stay till the end of the credits.  
 b. A: Which of her dogs is Lea bringing?  
 B: The {lovely, disgusting, #fucking, #bloody, #damn} one.

Two further relevant cases of degree modification in the primary modality are (i) modifier repetition, as in (14), which has a gradient iconic effect, and (ii) so-called "contrastive reduplication", whereby a string of the form *x x* gets interpreted along the lines of '{true, proper, prototypical} *x*', as in (15), which also has iconic roots. Both are truth-conditionally non-vacuous as well.

- (14) a. You're really, really sick. (Ghomeshi et al. 2004, fn. 3)  
 b. You are a sick, sick man. (Ghomeshi et al. 2004, fn. 3)  
 c. If the movie is very, very, very good, I'll stay till the end of the credits.  
 ↗ If the movie is good, I'll stay till the end of the credits.
- (15) a. **Like'em-like'em?** Or, I'd-like-to-get-store-credit-for-that-**amount** like'em?  
 (corpus example from Ghomeshi et al. 2004, (1b))  
 b. I'm up, I'm just not **up-up**. (corpus example from Ghomeshi et al. 2004, (1d))  
 c. Lea doesn't have a chihuahua, she has a **dog-dog**.

It is not obvious to me how exactly the gradient mapping between the form and meaning should be operationalized in cases of modifier repetition like in (14). It is possible, however, that each occurrence of the modifier is interpreted independently, as a further restriction on the input. Compare the effect obtained in (14) to that of multiple degree modifiers:

- (16) Zoe is an [incredibly, [extremely, [amazingly talented]]] stuntwoman.

An extra component would need to be introduced, assuring that this step-wise restriction always restricts towards the higher degree of the scale. This can be done in many ways, but I will leave it at that and will focus on a perhaps more straight-forward phenomenon of contrastive reduplication. In line with my architectural assumptions, I maintain that contrastive

<sup>8</sup>I'm making this caveat here, because the maximal-degree modifier 'completely' has been argued to project by default in spoken Italian and Italian Sign Language in Aristodemo 2017.

reduplication expones a DegP head, CRED, whose (coarse-grained) semantics is given in (17). I mostly assume the degree semantics from Kennedy & McNally 2005, except, following Esipova 2019b, I separate out the existential closure that they build into their lexical entries for degree modifiers so as to make sure that degree modifiers are in fact modifiers (i.e., expressions of type  $\langle \tau, \tau \rangle$ ) and to allow for multiple degree modification like in (16). The **prototype** relation in (17) requires that the degree  $d$  is within the prototypical range for the input expression.

$$(17) \quad \llbracket \text{CRED} \rrbracket(\llbracket \alpha_{\langle d, \langle \tau_1 \dots \tau_n, st \rangle} \rrbracket) = \lambda d \lambda X_{\tau_1}^1 \dots X_{\tau_n}^n \lambda w. \llbracket \alpha \rrbracket(d)(X^1) \dots (X^n)(w) \wedge \text{prototype}_{\llbracket \alpha \rrbracket}(d)(w)$$

Inherently scalar predicates like *sick* combine with CRED directly, as in (18).

$$(18) \quad \begin{array}{l} \text{a. } \llbracket \textit{sick} \rrbracket = \lambda d \lambda x \lambda w. \textit{sick}(x)(w) = d \\ \text{b. } \llbracket \text{CRED} \rrbracket(\llbracket \textit{sick} \rrbracket) = \lambda d \lambda x \lambda w. \textit{sick}(x)(w) = d \wedge \text{prototype}_{\textit{sick}}(d)(w) \end{array}$$

Predicates born non-scalar first type-shift into scalar versions of themselves, e.g., *dog* type-shifts from a property of individuals into a function that maps individuals onto the scale of dogness:

$$(19) \quad \begin{array}{l} \text{a. } \llbracket \text{SCALAR} \rrbracket(\llbracket \alpha_{\langle \tau_1 \dots \tau_n, st \rangle} \rrbracket) = \lambda d \lambda X_{\tau_1}^1 \dots X_{\tau_n}^n \lambda w. \text{scale}_{\llbracket \alpha \rrbracket}(X^1) \dots (X^n)(w) = d \\ \text{b. } \llbracket \textit{dog} \rrbracket = \lambda x \lambda w. \textit{dog}(x)(w) \\ \text{c. } \llbracket \text{SCALAR} \rrbracket(\llbracket \textit{dog} \rrbracket) = \lambda d \lambda x \lambda w. \text{scale}_{\textit{dog}}(x)(w) = d \end{array}$$


The  $d$  variable is existentially closed off after all degree modifiers have composed.

With these tools at our disposal, let us look at secondary modality degree modifiers.

## 4.2 Degree modification via facial expressions

Like spoken mirative adverbs, the mirative facial expression whose most salient feature is eyes wide open, *OO*, can act as a proposition-contributing supplement or as a degree modifier:<sup>9</sup>

- (20) a. Yesterday there was a party, and, [*surprisingly, impressively*], [Mia got drunk].  
 $\approx$  It is  $\langle$ surprising, impressive $\rangle$  that Mia got drunk.  
 b. Yesterday there was a party, and Mia got [*surprisingly, impressively*] [drunk].  
 $\approx$  Mia got drunk to a(n)  $\langle$ surprising, impressive $\rangle$  extent.

- (21) a. Yesterday there was a party,  and  $\overline{\llbracket \text{Mia got drunk} \rrbracket}^{\text{OO}}$ .  
 $\approx$  It is surprising that Mia got drunk.  
 b. Yesterday, there was a party, and Mia got  $\overline{\llbracket \text{drunk} \rrbracket}^{\text{OO}}$ .  
 $\approx$  Mia got drunk to a {surprising, high} extent.

Now, more needs to be said eventually about the interaction of temporal alignment of *OO* and its syntactic construals. As pointed out by Patrick Grosz (p.c.), in its supplement use, *OO* is focus-sensitive, just like *surprisingly* and its kin, which, I believe, does affect the docking of *OO* (see some preliminary observations in Esipova 2019c). For the purposes of this paper, however, what matters is that both interpretations are in principle available for a co-speech *OO*.

The degree modifying *OO* often comes with an intonational morpheme on the modified predicate, discussed in the next subsection. However, *OO* can make this contribution independently, as shown in (22), where *OO* co-occurs with the Russian conventionalized gesture *DRINK* from the previous section (which here is a predicative adjective morphosyntactically, showing that the gesture itself is a root).

<sup>9</sup>Overlining approximately indicates temporal alignment for facial expressions.

- (22) a. Yesterday, there was a party, and  $\overline{[[\text{Mia got DRUNK}]^{\text{O-O}]}$ . (video available)  
 $\approx$  It is surprising that Mia got drunk.  
 b. Yesterday there was a party, and Mia got  $\overline{[[\text{DRUNK}]^{\text{O-O}]}$ . (video available)  
 $\approx$  Mia got drunk to a {surprising, high} extent.

Aptly, *OO* projects like a supplement when it is interpreted as one, and is restricting by default when it is used as a degree modifier (just like *surprisingly* and the like):<sup>10</sup>

- (23) a.  $\overline{[[\text{a friend of mine gets DRUNK}]^{\text{O-O}]}$ , I sometimes comment on that.  
 $\rightarrow$  When a friend of mine gets drunk, that is surprising.  
 b. When, a friend of mine gets  $\overline{[[\text{DRUNK}]^{\text{O-O}]}$ , I sometimes comment on that.  
 $\not\rightarrow$  When a friend of mine gets drunk, they do so to a {surprising, high} extent.  
 c. #When a friend of mine gets DRUNK, I don't say anything,  
 but when  $\overline{[[\text{a friend of mine gets DRUNK}]^{\text{O-O}]}$ , I sometimes comment on that.  
 Intended: '...when (a friend of mine gets drunk and I am surprised by that)...'.

Again, Schlenker's typology can't predict this variable projection behavior for co-speech/gesture facial expressions.<sup>11</sup> But an approach that treats facial expressions as linguistic objects and starts by asking the question about their lexical semantics and compositional integration, once again, reveals that they fit neatly into a cross-modal typology of projection patterns: a propositional *OO* patterns with propositional spoken adverbs, and a degree modifier *OO* patterns with other degree modifiers, whatever the reason for their persistently restricting nature.

### 4.3 Suprasegmental degree modification

The aforementioned intonational degree modifier *DEG-INT*, written as a subscript on the word whose stressed syllable it docks to, can make its contribution independently, too. Its precise phonetic and phonological properties remain to be established,<sup>12</sup> but it seems that it can involve an L\*+H pitch accent, creaky phonation, lengthening of the accented syllable (not just the vowel, the onset lengthens, too, if it can), and higher intensity (audio available).

Like CRED, *DEG-INT* is morphosyntactically promiscuous and restricting by default:

- (24) a. If the movie is good<sub>DEG-INT</sub>, I'll stay till the end of the credits.  $\approx$  *very good*  
 $\not\rightarrow$  If the movie is good, I'll stay till the end of the credits.  
 b. Lea has a dog<sub>DEG-INT</sub>.  $\approx$  *big dog* or *proper dog*  
 c. When I saw that snake, I ran<sub>DEG-INT</sub>.  $\approx$  *ran fast* or *ran properly*

*DEG-INT* can be given the same semantics as CRED or a version thereof. I think there are some discrepancies between the two, but leave a proper exploration of these for future research.

Schlenker (2018b) discusses what he calls "iconic vowel lengthening" in (25).

- (25) If the talk is loooong, I'll leave before the end.  
 $\not\rightarrow$  If the talk is long, I'll leave before the end.

<sup>10</sup>A reviewer asks if the attitudinal component of a degree modifier *OO* can project. The judgements are subtle, but, if need be, we can separate the 'high degree' and 'that degree is surprising' components for *OO* or its spoken kin and have the latter project (cf. items like *fucking*, which can simultaneously make a truth-conditional contribution as degree modifiers and perform their expressive function as emotional outlets; Esipova 2019c).

<sup>11</sup>Schlenker (2018b) claims that co-speech facial expressions trigger cosuppositions. He only looks at the "disgusted" facial expression, which I believe should be fit into a typology of attitudinal expressions along with expressives like *fucking* and evaluative adjectives like *disgusting*, whose beginnings can be seen in Esipova 2019c.

<sup>12</sup>I am currently working with Zoe Kahana and Reis White on establishing these properties and whether they are distinct from those of the modified focus marking accompanying purely mirative, non-degree-modifier *OO*.

Schlenker maintains that vowel lengthening is an “internal enrichment” without its own time slot, and those can, but don’t have to be at-issue. It’s not entirely clear to me in what sense such suprasegmental events are “internal”, but that aside, this approach clearly misses the generalization about degree modifiers cross-modally.

More needs to be said about the role of iconicity, however. I do agree that examples like (25) can contain extra segment lengthening beyond that associated with *DEG-INT*. However, this extra lengthening in (25) isn’t obligatory to convey the high degree reading, and a more general phenomenon is at play, as shown in (24). Schlenker also notes that having the same extra lengthening in *shooort* is weird. This seems correct, although I don’t think the same holds about the regular *DEG-INT* on *short*. Similarly, *DEG-INT* seems to be equally good with *slow* and *fast*, but only the former allows for the extra lengthening.

(26) It was {slow<sub>DEG-INT</sub>, slooow<sub>DEG-INT</sub>, fast<sub>DEG-INT</sub>, #faaast<sub>DEG-INT</sub>}.

There might still be some remnant effects of iconicity in *DEG-INT*, however, due to the low pitch (coming from the L\* target, which tends to be extra-low in *DEG-INT*) and the overall syllable lengthening, constraining the distribution of this morpheme. How we should operationalize all these iconicity effects architecturally is an extremely interesting question, which I leave for future research. What is clear, however, is that this problem needs a properly linguistic approach.

Another question that I will leave for future research is how to treat simultaneous combinations of various degree modifiers in a principled way. For instance, (27) is a naturally occurring example in which the image of a larger, fuller chest is simultaneously conveyed by *OO*, *DEG-INT*, and a co-speech gesture. The compositional possibilities here are numerous, and it is not clear if and how we should distinguish among them.

(27) *Context: The speaker is giving advice on how to build pec muscles.*  
 You won’t have a chest that looks like this<sub>TOUCHES-ANOTHER’S-CHEST</sub>, you’ll actually  
 have a chest<sub>DEG-INT</sub><sup>OO</sup><sub>LARGE-ROUND</sub>. (‘Athlean-X’ YouTube channel, video available)

One final direction for future research that I will mention relates to the claims in the literature that some languages either don’t have degree variables at all or can’t bind them, because they lack spoken lexical items that would operate on degree variables, such as degree modifiers or comparatives (e.g., Motu in Beck et al. 2009, Washo in Bochnak 2015).<sup>13</sup> It would be worthwhile investigating whether speakers of such languages can use and perceive secondary modality degree modification. This could shed light on how deep this cross-linguistic asymmetry lies.

## 5 Conclusion

In this paper, I hoped to show that by putting *linguistics* into *super linguistics*, i.e., by treating secondary modality content such as gestures, facial expressions, and intonational morphemes as linguistic objects at all levels of representation, we gain predictive and explanatory power. In particular, this approach reveals parallels in projection behavior between primary and secondary modality content that void the need for modality-specific typologies of projection patterns. I believe other classifications of meaning-bearing expressions should be cross-modal, too. For example, I develop the beginnings of a cross-modal typology of attitudinal and expressive content in Esipova 2019c and establish the need for a cross-modal typology of expressions that integrate with the host utterance at some level(s) of representation, but not compositionally (e.g., pure expressives or some morphemes with purely social functions) in Esipova To appear.

<sup>13</sup>I thank Yining Nie (p.c.) for pointing out this literature to me.

## References

- Aristodemo, Vita Maria Valentina. 2017. *Gradable constructions in Italian Sign Language*: Institut Jean-Nicod dissertation.
- Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Stefan Hofstetter, Christiane Savelsberg, John Vanderelst & Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9(1). 1–66. doi:10.1075/livy.9.01bec.
- Beckman, Mary E. & Gayle Ayers. 1997. Guidelines for ToBI labelling. Version 3.0. The Ohio State University Research Foundation.
- Bochnak, M. Ryan. 2015. The degree semantics parameter and cross-linguistic variation. *Semantics and Pragmatics* 8(6). 1–48. doi:10.3765/sp.8.6.
- Ebert, Cornelia & Christian Ebert. 2014. Gestures, demonstratives, and the attributive/referential distinction. Talk given at *Semantics and Philosophy in Europe (SPE 7)*, Berlin.
- Ernst, Thomas Boyden. 1984. *Towards an integrated theory of adverb position in English*: Indiana University dissertation.
- Esipova, Maria. 2019a. Composition and projection in co-speech gestures. To appear in *Proceedings of Semantics and Linguistic Theory (SALT) 29*.
- Esipova, Maria. 2019b. *Composition and projection in speech and gesture*: New York University dissertation.
- Esipova, Maria. 2019c. Emotion and composition. Talk given at *SURGE*, Rutgers University.
- Esipova, Maria. To appear. Polar responses in Russian across modalities and across interfaces. *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 28*.
- Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory* 22(2). 307–357. doi:10.1023/B:NALA.0000015789.98638.f9.
- Hunter, Julie. 2018. Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics and Philosophy* 1–16. doi:10.1007/s10988-018-9244-0.
- Jasinskaja, Katja & Claudia Poschmann. 2018. Attachment in syntax and discourse: towards an explanation for the variable scope of non-restrictive relative clauses. In Sireemas Maspong, Brynhildur Stefánsdóttir, Katherine Blake & Forrest Davis (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 28*, 433–453. doi:10.3765/salt.v28i0.4438.
- Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381. doi:10.1353/lan.2005.0071.
- Koev, Todor. 2013. *Apposition and the structure of discourse*: Rutgers University dissertation.
- Lascarides, Alex & Matthew Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26(4). 393–449. doi:10.1093/jos/ffp004.
- Leffel, Timothy. 2014. *The semantics of modification: Adjectives, nouns, and order*: New York University dissertation.
- Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199273829.001.0001.
- Price, Patti J, Mari Ostendorf, Stefanie Shattuck-Hufnagel & Cynthia Fong. 1991. The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America* 90(6). 2956–2970. doi:10.1121/1.401770.
- Schlenker, Philippe. 2013. Supplements without bidimensionality. Ms., Institut Jean-Nicod and New York University.
- Schlenker, Philippe. 2018a. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. doi:10.1007/s10988-017-9225-8.
- Schlenker, Philippe. 2018b. Iconic pragmatics. *Natural Language & Linguistic Theory* 36(3). 877–936. doi:10.1007/s11049-017-9392-x.

# Pronominal Typology and Reference to the External World\*

Patrick Georg Grosz  
University of Oslo, Oslo, Norway  
p.g.grosz@iln.uio.no

## Abstract

Recent years have seen a surge in research on speech-accompanying gestures, often focusing on iconic gestures. This paper explores pointing gestures and their interaction with different types of 3<sup>rd</sup> person pronouns (e.g., anaphoric *vs.* deictic pronouns). While pointing gestures are often mentioned in the formal linguistic literature on deictic expressions, such as demonstratives, these gestures are less frequently investigated for their own sake. I aim to contribute to our understanding of pointing gestures by mapping the relationships that pronominal expressions can have with referents present in the utterance situation, and by investigating how they interact with pointing. I argue that pointing gestures are attention-directing devices that freely co-occur with anaphoric expressions; by contrast, the introduction of a new referent through deixis requires contrastive focus in addition to pointing.

## 1 Anaphora, Exophora, Deixis, and Pointing

In recent years, speech-accompanying gestures (McNeill 1992, Kendon 2004) have attracted a fair share of attention in the formal linguistics literature (e.g., Lascarides & Stone 2009, Ebert & Ebert 2014, Schlenker 2018a, Esipova 2019). This paper explores interactions between pointing gestures, pronominal expressions, and referents in naturalistic discourse situations (as opposed to written text). My core questions are: [i.] what kinds of DPs *can* co-occur with a pointing gesture (and to what effect), and [ii.] when does pointing become *necessary*?<sup>1</sup> As of now, such questions have largely been backgrounded, often with an implicit consensus that Lücking (2018:256) illustrates with examples such as (1a) and (1b), and which he describes as follows: “a demonstration [i.e., pointing gesture] cannot be associated with an endophoric [i.e., anaphoric or cataphoric] demonstrative, while it is obligatorily connected to [deictic]<sup>2</sup> uses”. The relevant intuition is that the phrase *that donkey* cannot be accompanied by a pointing gesture, [☞], if it is used anaphorically (here: as a donkey-pronoun), (1a). By contrast, a pointing gesture is required to introduce a new referent into the discourse based on the individuals that are present in the utterance situation, (1b).<sup>3</sup> In what follows, I approach (1a) and (1b) by pursuing a strategy of focusing

---

<sup>1</sup>For feedback and comments, I thank Masha Esipova, Pritty Patel-Grosz, Philippe Schlenker, Sarah Zobel, the poster session audience at the 3rd Crete Summer School of Linguistics, and five anonymous reviewers for the *Super Linguistics* workshop at the Amsterdam Colloquium, whose input was very valuable and has been incorporated wherever possible.

<sup>2</sup>Here, the term ‘kinds of DPs’ subsumes both semantic types (e.g., definite expressions that are anaphoric *vs.* deictic) and morphosyntactic types (e.g., pronominal expressions that are clitic *vs.* non-clitic).

<sup>3</sup>The original statement in Lücking (2018:256) uses *exophoric* (i.e., “... obligatorily connected to exophoric uses”). In this paper, I follow Cornish (1999; 2010) in treating *exophora* as a type of *anaphora*, which contrasts with *deixis*.

<sup>4</sup>I use bold type in examples for emphasis only, not for stress. Stress is marked by capitalization. Lücking (2018:257) does not provide stress marking for (1a)-(1b); however, intuitively, *that* must be stressed in (1b). This will be crucial for my discussion, and I come back to this in section 3.



on more basic questions, such as, with regards to (1a): “Is there an entity in the external world that can felicitously be pointed at?” – or, with regards to (1b): “Is there a potential discourse referent in the external world that is salient enough to be picked up without pointing?”

- (1) *Context: k is a non-salient individual in the utterance context* (Lücking 2018:256, adapted)
- a. Every farmer who owns a donkey<sub>i</sub> beats that donkey<sub>i/\*k</sub>      *covarying / anaphoric*  
 b. Every farmer who owns a donkey<sub>i</sub> beats [ $\square$ ]**that** donkey<sub>k/\*i</sub>      *fixed / deictic*

To begin with, let me state the somewhat trivial fact that there are two relationships that a speaker and hearer can have to a discourse referent that is associated with a definite description such as *that/the donkey*, or with a pronoun such as *it*. The referent can either be present in the external world in the utterance situation, or not. In (2), the external world (utterance situation) does not contain a referent for the pronoun *it*. The interpretation of *it* covaries with the indefinite *a leprechaun* across situations that are quantified over, and (2) might well be true if there are no leprechauns that exist in the whole world of evaluation. This use of *it* is prototypically *anaphoric*, as the pronoun is referentially dependent on its antecedent (*a leprechaun*); see Evans (1980:358). Adding a pointing gesture to (2) is non-trivial, since there is no entity in the external world in the utterance situation that could easily be pointed at.<sup>4</sup>

- (2) Whenever I see **a leprechaun**, I will do my best to avoid **it**.

In contrast to (2), the pronoun *it* in (3) is usually classified as an *exophoric* pronoun (e.g., by Cornish 1999:112; see also Halliday & Hasan 1976:32-36), which picks up a referent present in the external world (a dog that both hearer and speaker can see in the utterance situation). I will adopt the terminology from Cornish (2010:220), who contrasts the exophoric pronoun *it* in (3) with the deictic expression *that* in (4), and argues that exophoric uses are a sub-type of anaphoric uses.<sup>5</sup> He treats both exophora and non-exophoric anaphora as different from *deixis*, defined as follows: “Deixis serves prototypically to orientate the addressee’s attention focus towards a *new discourse entity* – or to a *new aspect* of an already-existing discourse referent” (Cornish 2010:218, italics added). Exophoric pronouns, (3), thus presuppose their referent’s existence and salience, whereas deictic expressions, (4), assert the existence and salience of their referent.

Note that (3) seems to be most natural when accompanied by a type of pointing where the speaker slightly tosses or tilts her head, but, crucially, (3) is acceptable without any pointing. Zooming out, we must bear in mind that the boundaries between exophora, (3), and deixis, (4), are necessarily fluent. If, in the context in (4), A can assume that B has already noticed the strange bird, then *that* may well assume an exophoric (referent-presupposing) use rather than a deictic (referent-asserting) use, thus turning (4) into another example of what (3) illustrates.

- (3) [A and B turn a corner on the pavement, and suddenly find themselves face to face with a rather large dog]  
 A to B: Do you think **it**’s friendly?  
 (Cornish 1999:112, who adapts it from Yule 1979)

<sup>4</sup>However, see (19) for a strategy that involves establishing schematic positions in gesture space, so-called *gestural loci* (coined after the loci of sign language; see Schlenker, to appear). This strategy is available for both (1a) and (2).

<sup>5</sup>There is no single standard use for the labels *anaphoric* vs. *exophoric*. While Cornish (1999; 2010) argues that exophoric uses are a sub-type of anaphoric uses, Halliday & Hasan (1976:33) treat *anaphora* vs. *exophora* as two distinct ways of referring. Correspondingly, Cornish groups anaphora/exophora together, set apart from deixis, whereas many scholars group exophora/deixis together, set apart from anaphora. I adopt Cornish’s use of the terminology.

- (4) A to B: Hey, look at [☞] **that!** (uttered with a pointing gesture towards a strange bird perched on the branch of a tree near the interlocutors)  
(Cornish 2010:219, adapted)

Having illustrated three types of relationships between pronominal expressions and the external world, we can now operationalize our definitions. *Non-exophoric anaphoric expressions* (short: *anaphoric expressions*), (2), are referentially dependent on an overt linguistic antecedent. By contrast, *exophoric anaphoric expressions* (short: *exophoric expressions*), (3), are referentially dependent on a presupposed non-linguistic ‘antecedent’<sup>6</sup> that is present in the external world in the utterance situation. Finally, *deictic expressions*, (4), select a referent that is externally present in the utterance situation, and introduce it into the conversation as a new salient discourse entity.

Having defined *anaphoric* and *exophoric* in this way, a non-trivial issue emerges from examples (5a) and (5b), which a reader should keep in mind. In (5a), the 3<sup>rd</sup> person singular bound pronoun *it* has an overt linguistic antecedent (the distributive quantifier *each of them*), yet the three individuals that are quantified over are also all present in the external world. In spite of these individuals being present, *it* is clearly anaphoric in (5a), as it cannot be referentially dependent on any one of the three individuals in the external world. The same conclusion does not carry over to the plural pronoun *they* in (5b), which refers to the whole group of three dogs, and seems to give rise to the familiar *bound vs coreferential* ambiguity from Reinhart (1983). At least in principle, we expect that (5b) has a (*bound*) anaphoric reading where *they* referentially depends on the DP *these dogs* and is syntactically bound by it. In addition, we expect (5b) to have a (*coreferential*) exophoric reading where *they* referentially depends on the actual group of dogs that are present in the external world and is not syntactically/semantically bound by *these dogs*. I mark these two readings of (5b) with a subscript *i* for the bound reading and *x* for the exophoric reading. Note that exophoric reading may be blocked by a preference of binding over coreference.

- (5) [A and B turn a corner, and find themselves face to face with three extremely large dogs]  
a. A to B: Do you think [**each of them**]<sub>i</sub> knows that **it**<sub>i</sub> could easily overpower one of us?  
b. A to B: [**These dogs**]<sub>i</sub> look like **they**<sub>i/x</sub>’re angry.

Having thus established a number of factors that need to be taken into consideration in a discussion of pointing, I proceed by outlining the core data set that I analyze in this paper.

## 2 Core Data: Pointing with Anaphoric and Deictic Pronouns

As discussed, my core question is when pointing is possible and/or necessary for different pronoun types. Quite generally, the interaction between pointing gestures and non-demonstrative pronouns (such as *it*, *they*, or *him*) remains underexplored. Cardinaletti & Starke (1999:153), in their paper on pronominal typology, discuss how pointing relates to the distinction between full pronouns (*elle* ‘her’) and clitic pronouns (*l’/la* ‘her’) in French. Examples (6a) and (6b) presumably involve the introduction of a new referent into the

<sup>6</sup>Here, *antecedent* (in single quotation marks) refers to the actual individual in the external world. Cornish (1999:ch.4, 2010) treats exophoric pronouns as *pragmatically controlled anaphoric pronouns*. Many scholars working on such expressions raise the question of whether anaphoric pronouns like *it* in (3) should be called “antecedentless”. Yule (1979:127) writes, on this matter: “The antecedent, if that is at all an appropriate term, is in the context or the situation.”

discourse by means of deixis (as defined in section 1). In this case, a pointing gesture can combine with the non-clitic pronoun *elle* ‘her’ in (6a), but not with the clitic pronoun *la* ‘her’ in (6b). However, pointing is not limited to deixis; it is also possible in anaphoric contexts like (6c), where pointing accompanies the clitic *la* ‘it.FEM’, whose referent has already been introduced into the discourse by virtue of the antecedent *cette maison* ‘this house’ in the previous sentence. Note that (6c) is parallel to (5b) in that *la* ‘it.FEM’ may referentially depend on *cette maison* ‘this house’ (in which case it would be non-exophoric) or on the actual house in the utterance situation (in which case it would be exophoric).

- (6) a. J’ai vu Marie puis j’ai vu [☞]elle. *French*  
 I=have seen Marie then I=have seen her  
 ‘I have seen Marie and then I have seen [☞]her(≠ Marie).’  
 b. \*J’ai vu Marie puis je [☞]l’ai vue.  
 I=have seen Marie then I her=have seen  
 c. Mets-toi ici et regarde **cette maison**. Tu [☞]la vois bien maintenant?  
 come here and look.at this house you it see well now  
 ‘Come here and look at **this house**. Do you see [☞]it(=the house) well now?’

Clitics are an optimal case study to begin with, since they typically cannot be used in deictic contexts, as in (6b); they are restricted to anaphoric contexts. Notably, they can also be used in exophoric anaphoric contexts, as illustrated for English *’im* in (7), modeled after (3).

- (7) [In an old castle, A sees a man’s ghost come towards A and B]  
 A to B: Do you see **’im**?

However, French clitics sometimes exhibit hallmark signs of non-clitic pronouns; for instance, the subject clitic *il* ‘he’ can be contrastively focused (Cardinaletti & Starke 1999:218-219) and occur in the disjunction *il ou elle* ‘he or she’ (Sportiche 1998:311), both of which are usually not possible for clitics. I thus proceed with data from Colloquial Viennese German, another language with a full vs. clitic distinction, but which disallows focused or coordinated clitics.

Recall my initial questions, stated above: [i.] what kinds of DPs *can* co-occur with a pointing gesture (and to what effect), and [ii.] when does pointing become *necessary*? In the remainder of this section, I answer these questions as follows: first, both anaphoric/exophoric pronouns and deictic pronouns can co-occur with a pointing gesture, the effect of which is explored in sections 3 and 4; second, pointing is required for deixis in scenarios with more than one possible referent.

Let us start by setting up the context in (8). This is a context that introduces a salient and familiar individual into the discourse, which is also present in the utterance situation, namely Peter. Moreover, the context explicitly contains several other possible discourse referents. We can now introduce a pronoun in the next sentence to see if it permits an anaphoric/exophoric reading (referring to Peter) or a deictic reading (selecting a new referent from the people in the room).

- (8) *Context*: at a party with fancy costumes, we are observing how people are dressed.  
**Der Peter** ist angekommen.  
 the Peter is arrived  
 ‘Peter has arrived.’ [followed by a short pause]

An anaphoric reading, where the pronoun refers to Peter, can be forced by using the clitic pronoun *’en* ‘him’. We can now check whether such an anaphoric pronoun can combine with a pointing gesture. As shown in (9), the answer is affirmative. Crucially, (9) can only have the anaphoric reading in (9a) (which may be exophoric or not, in line with (5b) and (6c)), and it cannot have a (new-discourse-entity-introducing) deictic reading, as shown in (9b).

(9) clitic pronoun with pointing gesture (anaphoric reading only)

Hast'en<sup>[☞]</sup> schon gsehn?  
 have.you=him already seen

- a. OK: 'Have you already seen **Peter**, who, by the way, is over there?' (<sup>OK</sup> *anaphoric*)  
 b. #: 'Have you noticed **the new guy** over there already?' (# *deictic*)

Intuitively, in order to introduce a new discourse referent, the pointing gesture has to co-occur with contrastive focus (see also Ahn 2017:37), which is only possible for non-clitic pronouns, as shown in (10). In the context in (8), the combination of pointing and contrastive focus appear to be necessary for introducing a new discourse referent, (10b) *vs.* (9b), but it is not sufficient, as shown by (10a); the anaphoric reading remains available. In the anaphoric reading, contrastive stress presumably targets Peter as the current aboutness topic, giving rise to a contrastive topic interpretation in Krifka's (2008) sense; i.e., *HIM* is a focused topic constituent, which serves to highlight that the speaker is answering sub-questions of *Who have you seen?* (e.g., {*Have you seen Peter?*, *Have you seen Sue?*, *Have you seen Ann?*, ...}); see also Beaver et al. (2017). An open question with regards to (10a) *vs.* (10b) is whether the pointing gesture itself differs across contexts, e.g., the pointing gesture may be less pronounced and emphatic in (10a) than in (10b).<sup>7</sup>

(10) stressed non-clitic pronoun with pointing gesture (ambiguous)

Hast **IHN**<sup>[☞]</sup> schon gsehn?  
 have.you HIM already seen

- a. OK: 'Have you already seen **PETER**, who, by the way, is over there?' (<sup>OK</sup> *anaphoric*)  
 b. OK: 'Have you noticed **the new guy** over there already?' (<sup>OK</sup> *deictic*)

Note that it is the combination of these two ingredients (pointing and contrastive focus) that seems to give rise to the introduction of a new (3<sup>rd</sup> person) discourse referent.<sup>8</sup> In the context in (8), contrastive focus alone is not sufficient to introduce a new discourse referent, (11b).<sup>9</sup>

(11) stressed non-clitic pronoun without pointing gesture (anaphoric reading only)

Hast **IHN** schon gsehn?  
 have.you HIM already seen

- a. OK: 'Have you already seen **PETER?**' (<sup>OK</sup> *anaphoric*)  
 b. #: 'Have you noticed **the new guy** already?' (# *deictic*)

To summarize the observations in (9)-(11), we can conclude that there are at least two ways in which pointing gestures interact with pronominal expressions whose referent is present in the utterance situation. They can co-occur with anaphoric pronouns, (9a) and (10a), in which case they add surplus information on where the intended referent is located. Alternatively, they can co-occur with non-anaphoric deictic pronouns, (10b), as part of introducing a new discourse referent. Importantly, the second case also requires contrastive stress on the pronoun in the context in (8).

Having established that pointing can occur both with anaphoric/exophoric pronouns (with or without contrastive stress) and with deictic pronouns (with contrastive stress), we can ask what the pointing gesture contributes. In sections 3 and 4, I argue that, as proposed by Lücking (2018), pointing is merely an attention-directing device. In anaphoric/exophoric

<sup>7</sup>I thank Masha Esipova (p.c.) for pointing out this possible difference in the pointing gesture itself.

<sup>8</sup>As pointed out by an anonymous reviewer, 1<sup>st</sup> and 2<sup>nd</sup> person pronouns behave differently with regards to some of the properties that I discuss, which is presumably related to the fact that their intended referent is, by default, self-evident.

<sup>9</sup>One may wonder whether pointing would still be necessary for (10b) in a context where only one other person is in the room; crucially, this is the type of scenario, briefly discussed for (3) *vs.* (4), where the lines between an exophoric/anaphoric and a deictic reading are blurred, as speaker and hearer may both already be aware of the single other person.

contexts, it signals the location of the intended referent in the utterance situation. In deictic contexts, contrastive focus introduces a set of alternative individuals, and signals that the intended antecedent is not the most prominent one (i.e., it is a non-default antecedent); pointing is merely the auxiliary attention-directing device that enables the hearer to identify the intended referent by virtue of signaling the referent's location. Crucially, pointing serves as an attention-directing device across all of the constructions that I discuss; it never introduces discourse referents by virtue of its own semantics.

### 3 Analysis, Part I: Contrastive Focus + Pointing = Deixis

Let us start with the observation that clitic pronouns, as opposed to stressed full pronouns, can never be used to deictically introduce a new referent, as shown in (12) (repeated from (9b)/(10b)).

- (12) Hast **IHN**<sup>[☐]</sup> schon gesehn? / #Hast'en<sup>[☐]</sup> schon gesehn?  
 have.you HIM already seen have.you=im already seen  
*intended*: 'Have you already seen **the new guy** over there?' (*deictic new referent*)

For reasons of space, I only discuss 3<sup>rd</sup> person pronouns in this paper; building on Patel-Grosz & Grosz's (2017) proposal of structural asymmetry between weak/clitic and strong pronouns (based on Schwarz 2009), I propose that the clitic *vs.* full pronoun distinction in Colloquial Viennese German reflects a different internal structure. This is given in (13a) *vs.* (13b) for the personal pronouns in (12);  $[[NP_n]]^g$  represents the denotation of a salient null NP<sub>n</sub>; the clitic '(e)n involves a weak article *the<sub>weak</sub>* (which only conveys uniqueness), whereas the strong *ihn* involves a strong article *the<sub>strong</sub>*, and a referential index (here: 1), by virtue of which it also marks familiarity. Since (13ab) are adopted directly from Patel-Grosz & Grosz (2017), I refer the interested reader to this paper with regards to the details of the implementation that are orthogonal to the current aims.<sup>10</sup>

- (13) a.  $[[ '(e)n ]^g ] = [[ [the_{weak} s_r] NP_n ] ]^g = \iota x [[ [NP_n ]^g (x)(g(s_r)) ] ]$   
 b.  $[[ ihn ]^g ] = [[ [1 [the_{strong} s_r] NP_n ] ] ]^g = \iota x [[ [NP_n ]^g (x)(g(s_r)) \& x = g(1) ] ]$

In Colloquial Viennese German, non-clitic 3<sup>rd</sup> person pronouns of the (13b) type can be contrastively stressed with a deictic use, as in (12), whereas clitic pronouns of the (13a) type cannot be stressed or used in a deictic way. Based on this observation, I propose that stress on *ihn* 'him' in (12) is contrastive focus on the index, (14ab), which flags alternative individuals in the resource situation  $s_r$  that fulfil the NP property. For reasons of simplicity, I assume that the NP property here is resolved towards  $[\lambda x. \lambda s. x \text{ is a man in } s]$ . The set of alternatives given in (14b) thus does not pick out alternatives for the unique individual who is a man in the situation of evaluation, but it picks out index-based alternatives, given that, even in a situation that contains more than one man, there will only be one man who is associated with the index 1, 2, 3, etc.

- (14) a.  $[[ [1_F [the_{strong} s_r] man ] ] ]_o^g = \iota x [x \text{ is a man in } g(s_r) \& x = g(1)]$   
 b.  $[[ [1_F [the_{strong} s_r] man ] ] ]_{Alt}^g = \{ \iota x [x \text{ is a man in } g(s_r) \& x = g(i)] \mid i \in \mathbb{N} \}$

<sup>10</sup>Specifically, I remain agnostic as to whether the silent NP is provided by means of NP deletion (see Elbourne 2013) or as a contextually supplied variable of type  $\langle e, \langle s, t \rangle \rangle$ ; for bound uses of both (13a) and (13b) (which involves situation binding), see Patel-Grosz & Grosz (2017:278-279) and Elbourne (2013:196). Note that I differ from Patel-Grosz & Grosz (2017) in that they propose the structural asymmetry in (13ab) for German personal pronouns *vs.* demonstrative pronouns; by contrast, I here apply this analysis to the clitic *vs.* full pronoun distinction in Colloquial Viennese German.

I propose the following analysis, illustrated for (15a), a variant of (10b). First, contrastive focus on *ihn* gives rise to the alternatives in (15b), which spell out (14b). Second, contrastive focus on *ihn* also triggers an inference, in (15c), that the intended antecedent is not the default antecedent. In other words, in the deictic reading of (15a), contrastive stress signals a non-default reference continuation, i.e., reference to a non-discourse-prominent individual, which triggers the search for a suitable discourse referent. Such a discourse referent is not provided in the relevant context (originally provided in (8)), which only mentions Peter. Therefore, the pointing gesture is needed to direct the hearer's attention towards a possible (and intended) referent, (15d). We thus derive that pointing in combination with contrastive focus gives rise to the deictic introduction of new referents, while pointing itself purely amounts to an attention-directing device.

(15)a. *Context*: at a party with fancy costumes, we are observing how people are dressed.

Ich schau mir grad **IHN**<sub>[☞]</sub> an.

I look.at me right.now HIM V.PRT

<sup>OK</sup> 'I'm just looking at **the new guy** over there.' (<sup>OK</sup> *deictic*)

b. *alternatives activated by contrastive focus on 'IHN'*

$\llbracket \text{IHN} \rrbracket_{\text{Alt}}^g = \{ \iota x [x \text{ is a man in } g(s_r) \& x = g(\mathbf{1})], \iota y [y \text{ is a man in } g(s_r) \& y = g(\mathbf{2})], \dots \}$

c. *inference triggered by contrastive focus*

$\rightsquigarrow \llbracket \text{IHN} \rrbracket_o^g$  is not the most prominent discourse referent

d. *non-at-issue inference (presupposition) triggered by the pointing gesture*

$\rightsquigarrow \llbracket \text{IHN} \rrbracket_o^g$  is at the position the speaker is pointing to

In contrast to approaches that intimately tie pointing gestures to deictic reference, I have thus proposed that pointing gestures merely locate a referent, (15d), while (contrastive) focus triggers alternatives (Krifka 2008), (15b), and a shift towards a non-default antecedent, (15c). All of these components are necessary conditions that license, but do not entail, the deictic introduction of a new discourse referent. Section 4 will elaborate on pointing outside of deixis contexts.

As for (15c), contrastive stress has previously been observed to trigger a non-default interpretation with pronouns (e.g., in Larson & Luján 1989), as illustrated in (16), from Givón (2001:419). Givón (2001) observes that the stressed pronoun in (16b) gives rise to an *object-to-subject switch* (i.e., the preference is for *HE* to refer to Joe), whereas (16a) is typically understood to involve *subject continuity* (i.e., the preference is for *he* to refer to Bill). Since subjects (here: *Bill*) are typically more prominent than objects (here: *to Joe*) and thus qualify as the default antecedents for subsequent pronouns (*he*), contrastive stress can be said to trigger an inference towards a non-default antecedent (namely *Joe*). As of now, it is an open question how this non-default-antecedence inference comes about, but this could plausibly be an instantiation of emphatic focus (as described, e.g., in Eckardt 2001). Note, also, that the triggering of (15c) seems to be more of a preference rather than an obligatory inference, as (16a) retains a reading where Joe is the antecedent, and (16b) has a reading with Bill as the antecedent. (Similarly, the availability of anaphoric readings, (10a)/(11a), entails that this inference does not always arise.)

(16)a. Bill talked to Joe and then **he** left. (*preference for Bill as antecedent*)

b. Bill talked to Joe and then **HE** left. (*preference for Joe as antecedent*)

Revisiting Lücking's (2018:257) example (1b) from this perspective, we notice that a similar role of contrastive focus can be observed in (17a) vs. (17b); in order to trigger a deictic reference to the individual in the utterance context, contrastive focus (not marked by Lücking) is required.

(17) *Context:  $k$  is a non-salient individual in the utterance context*

- a. Every farmer who owns a donkey <sub>$i$</sub>  beats [ $\mathbb{F}$ ] **THAT** donkey <sub>$k/i$</sub>   
 b. <sup>?</sup>\* Every farmer who owns a donkey <sub>$i$</sub>  **BEATS** [ $\mathbb{F}$ ] **that** donkey <sub>$k/i$</sub>

Section 4 proceeds by exploring the semantic contribution of the pointing gesture itself, which, in (15d), I sketched as follows:  *$x$  is at the position the speaker is pointing to.*

## 4 Analysis, Part II: Pointing with Anaphoric Expressions

To capture the core data in section 2, I propose three components, which I will discuss in turn: first, all co-speech pointing can be analyzed as pointing at a situation rather than at an individual (a revision of (15d)); second, co-speech pointing is purely an “attention directing device” (Lücking 2018:277); and, third, a pointing gesture is interpreted like any other co-speech gesture.

Let me start with the first component, i.e., the idea that all co-speech pointing can be analyzed as pointing at a situation (i.e., *situation-denoting pointing*) and may not involve pointing at an individual (i.e., *individual-denoting pointing*). If pointing is situation-pointing, individuals in the respective situations can be picked out via situational uniqueness (see Elbourne 2013:193). So, in a case like (18a), repeated from (12), a speaker points at a situation  $s$ , contained in the utterance situation, which contains a unique referent for the definite expression in (18b). An illusion of individual-denoting pointing arises when the situation that is being pointed at is small enough to only contain one potential referent (e.g., a *man-sized* situation, in line with Elbourne 2013:198).

- (18)a. Hast            **IHN**[ $\mathbb{F}$ ]    schon        gsehn?  
           have.you    **HIM**        already    seen  
 b.  $[[\text{ihn}_i]]^g = \iota x[x \text{ is a man in } g(s_i) \& x=g(1)]$

As of now, I am not aware of any evidence to the end that co-speech pointing ever involves pointing at individuals; by contrast, Schlenker (2018b) argues that pointing in sign language can target individuals, situations or situation stages of individuals. We can set up a tentative empirical argument in favor of situation-pointing with co-speech gestures, which is modeled after Schlenker (2018b). First, we can establish that co-speech pointing gestures in spoken languages can mimic the establishment of *loci* (positions in signing space) in sign languages (compare Schlenker, to appear). An illustrative example is sketched in (19), where subscripted indices are associated with distinct positions in the gesturing space in front of the speaker. In this case, the pointing gestures are presumably used ‘metaphorically’, i.e., they may still trigger the inference ‘ $x$  is at the position the speaker is pointing to’; however, this is not a real-world position, but a position in a schema.

- (19) Whenever [ $\mathbb{F}_1$ ][**a baritone**] tells [ $\mathbb{F}_2$ ][**another baritone**] gossip about [ $\mathbb{F}_3$ ][**a third baritone**], [ $\mathbb{F}_2$ ][**he**] ends up telling [ $\mathbb{F}_3$ ][**him**] about it and [ $\mathbb{F}_1$ ][**he**] gets into trouble.

Having established that gestural pointing can use gestural loci, (20) shows the following;<sup>11</sup> the speaker establishes a position  $k$  in gesture space, as the locus of New York. After stating that the hearer will be transferred to New York, it becomes possible to point at the New-York-locus  $k$  in connection with the pronoun *you*. As argued in Schlenker (2018b), “locative shifting” the second person pronoun in this way is not possible for individual-denoting loci in sign language, so (20) should be infelicitous if it involved pointing at individuals; this

<sup>11</sup>Thanks to Philippe Schlenker (p.c.) for suggesting this example.

suggests that the gestural pointing in (20), in fact, involves pointing at situations, as proposed.

(20) Since [<sub>hearer</sub>]**you** can't seem to work with [<sub>n</sub>]**John**, I'll have you transferred to [<sub>k</sub>]**New York**. And if later I need to downsize, I'll fire [<sub>k</sub>]**you**.

On a related note, non-deictic pointing gestures improve in connection with covarying DPs if the pointing gesture is attached to a 'kind-denoting' referent (here: a picture of a donkey), as in (21). The pointing gesture in (21) does not serve to establish a discourse referent for *that donkey*; what it seems to convey is that the *donkey*-property is illustrated at the indicated location.

(21) *Context: there are three pictures on the table, depicting a sheep, a horse and a donkey*  
Every farmer who owns a donkey<sub>i</sub> BEATS that [<sub>i</sub>]**donkey<sub>i</sub>** (*points at donkey picture*)

A question that emerges from the discussion of situation-denoting pointing (*vs.* individual-denoting pointing) is whether pointing gestures ever truly attach to pronouns (or, more generally, DPs), or rather to larger constituents, such as the VP. This, in turn, raises questions concerning the semantic type and lexical semantics of pointing gestures; see Esipova (2019).<sup>12</sup> It also ties in with research such as Ebert et al. (2011) on how gestures relate to information structure and, more specifically, focus exponents. Quite generally, we face the question of where in a clause such attention-directing pointing gestures can (or must) be anchored. A reviewer points out that a treatment of pointing gestures as attention-directing devices would predict that (22a) is acceptable (contrary to fact) as long as the presupposition in (22b) is satisfied.<sup>13</sup>

(22) a. ?\* [<sub>i</sub>]**[No student]<sub>i</sub>** believes that [<sub>i</sub>]**he<sub>i</sub>** is a genius.

b. *intended*:  $\rightsquigarrow$  For every student  $x$ ,  $x$  is at the position the speaker is pointing to.

Crucially, the pointing gesture appears to be well-formed when accompanying the quantifier rather than the bound pronoun, as shown in (23a), though it is unclear whether the correct rendering of the gesture's contribution should quantify over atomic variables, (23b), or make a claim about a group, (23c) (which may be related to pointing with 'kind-denoting' reference, as in (21)). In any case, there appear to be rules for gesture anchoring that favor (23a) over (22a) regardless of its interpretation; crucially, focus does not seem to matter in this case, as the gesture placement in (23a) seems to be invariant and does not track focus stress, e.g. on *she*.

(23) a. [<sub>i</sub>]**[No student]<sub>i</sub>** believes that **he/she<sub>i</sub>** is a genius.

(*pointing, e.g., at student dorms*)

b. *option 1*:  $\rightsquigarrow$  For every student  $x$ ,  $x$  is at the position the speaker is pointing to.

c. *option 2*:  $\rightsquigarrow$  The group of all students is at the position the speaker is pointing to.

As the second component of my proposal, I follow Lücking (2018:277-278) in treating pointing as a pure "attention directing device", which does not directly identify a referent. Initial evidence for such a view stems from the fact that pointing can be systematically connected to DPs that refer on their own without any need for pointing whatsoever. This isn't only true for weak (clitic) pronouns, but also for proper names, as shown in (24).

<sup>12</sup>As pointed out to me by Masha Esipova (p.c.), one possible approach to many instances of co-speech pointing would be to treat them on analogy with VP-level modifiers such as appositional *over there*, which in turn may involve pointing connected to *there*. This is illustrated by (i), which is interpreted very much in parallel to examples like (24).

i. Have you already seen Peter, **over** [<sub>i</sub>]**there**?

<sup>13</sup>Note that (22a) improves if, in line with (19), there are pointing gestures both on *no student* and on *he*, which would then serve to indicate loci in gesture space.



- (24) Hast [den [☞] Peter] schon gesehn?  
 have.you the Peter already seen  
 ‘Have you already seen **Peter**, who, by the way, is over there?’

As the third and final component, I argue that pointing gestures are interpreted just like any other co-speech gestures (see, in particular, Schlenker 2018a). As shown in (25a) and (25b), pointing gestures trigger a non-at-issue inference, even when the existence of a referent for an anaphoric pronoun is challenged, a typical feature of co-speech gestures. In both (25a) and (25b), the pointing gesture targets an empty part of the stage, which is the presumed position of individuals who are absent.

- (25) *We’re seated at a pop concert; we don’t know if the singer will bring special guests*  
 a. If there are **special guests** who join her today, you might see **em**[☞] from here.  
 ↪ If there are special guests, they will be at the location that I am pointing to.  
 b. There are **no special guests** in this concert. Or can you see **em**[☞]?  
 ↪ If there were special guests, they would be at the location that I am pointing to.

These non-at-issue inferences may be related to the *conditionalized presuppositions* (or *cosuppositions*) that Schlenker (2018a) assumes for iconic co-speech gestures. This connection can be strengthened by looking at examples like (26a), where the pointing gesture gives rise to the inference in (26b). In this case, *it* must be a non-exophoric anaphoric pronoun, as it is bound by the quantifier *every mole in our garden*, and there are no moles present in the utterance situation. Nevertheless, the pointing gesture can indicate the position of potential referents.

- (26) *The speaker is standing at the entrance to a large garden.*  
 a. **Every mole in our garden** must realize that we can sometimes see [☞]**it**.  
 (*pointing at an opening in the ground where moles have been spotted, though none is visible right now*)  
 b. ↪ If we can see any mole, then we can see it at the location that I am pointing to.

## 5 Conclusion

In this paper, I argued that a connection between pointing gestures and referential expressions is, by and large indirect; anaphoric expressions that are exophoric can be freely accompanied by pointing, which indicates the position of the respective discourse referents. Non-anaphoric deictic expressions, which introduce a new referent, involve exactly the same meaning for the pointing gesture. In such deictic examples, contrastive focus introduces a set of alternative individuals in a given situation, and signals that the intended antecedent is not the most prominent one (i.e., it is a non-default antecedent). Pointing is merely the auxiliary attention-directing device that enables the hearer to identify the intended referent by virtue of its usual contribution. Crucially, the upshot is that, even in deictic cases, the pointing gestures only serve to pick out a situation in the external context of the utterance, and they do not introduce new discourse referents *per se*.

## References

- Ahn, D. (2017). Definite and Demonstrative Descriptions: A Micro-typology. *Proceedings of GLOW in Asia XI*, (pp. 33-48).
- Beaver, D., Roberts, C., Simons, M., & Tonhauser, J. (2017). Questions Under Discussion: Where Information Structure Meets Projective Content. *Annual Review of Linguistics* 3, 265-284.
- Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency: A case study of the three classes of pronouns. In H. van Riemsdijk, *Clitics in the languages of Europe* (pp. 145-234). Berlin: Walter de Gruyter.
- Cornish, F. (1999). *Anaphora, discourse and understanding. Evidence from English and French*. Oxford: Clarendon Press.
- Cornish, F. (2010). Anaphora: Text-based or discourse-dependent? Functionalist vs. formalist accounts. *Functions of Language* 17, 207-241.
- Ebert, C., & Ebert, C. (2014). Gestures, demonstratives, and the attributive/referential distinction. Slides from a talk given at Semantics and Philosophy in Europe (SPE 7), ZAS, Berlin, June 2014. Retrieved from <https://semanticsarchive.net/Archive/GJjYzkwN>
- Ebert, C., Evert, S., & Wilmes, K. (2011). Focus Marking via Gestures. *Proceedings of Sinn & Bedeutung* 15.
- Eckardt, R. (2001). Reanalysing selbst. *Natural Language Semantics* 9, 371-412.
- Elbourne, P. (2013). *Definite descriptions*. Oxford: Oxford University Press.
- Esipova, M. (2019). Composition and projection in speech and gesture. PhD dissertation, New York University.
- Evans, G. (1980). Pronouns. *Linguistic Inquiry* 11, 337-362.
- Givón, T. (2001). *Syntax: An Introduction, Volume 1*. Amsterdam: John Benjamins.
- Halliday, M. A., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Krifka, M. (2008). Basic Notions of Information Structure. *Acta Linguistica Hungarica* 55, 243-276.
- Larson, R., & Luján, M. (1989). Emphatic pronouns. Manuscript, Stony Brook University, New York, and University of Texas, Austin. Retrieved from <http://semlab5.sbs.sunysb.edu/~rlarson/emphpro.pdf>
- Lascarides, A., & Stone, M. (2009). A formal semantic analysis of gesture. *Journal of Semantics* 26, 393-449.
- Lücking, A. (2018). Witness-loaded and witness-free demonstratives. In M. Coniglio, A. Murphy, E. Schlachter, & T. Veenstra, *Atypical Demonstratives* (pp. 255-284). Berlin: de Gruyter.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Patel-Grosz, P., & Grosz, P. G. (2017). Revisiting Pronominal Typology. *Linguistic Inquiry* 48, 259-297.
- Reinhart, T. (1983). Coreference and Bound Anaphora: A Restatement of the Anaphora Questions. *Linguistics and Philosophy* 6, 47-88.
- Schlenker, P. (2018a). Gesture projection and cosuppositions. *Linguistics and Philosophy* 41, 295-365.
- Schlenker, P. (2018b). Locative Shift. *Glossa: a journal of general linguistics* 3, 115.1-46.
- Schlenker, P. (to appear). Gestural Grammar. *Natural Language and Linguistic Theory*.
- Schwarz, F. (2009). Two types of definites in natural language. Doctoral dissertation, University of Massachusetts, Amherst.
- Sportiche, D. (1998). *Partitions and Atoms of Clause Structure. Subjects, agreement, case and clitics*. London: Routledge.
- Yule, G. (1979). Pragmatically Controlled Anaphora. *Lingua* 49, 127-135.

# Gestures as markers on non-canonical questions\*

Michela Ippolito

University of Toronto, Toronto, Ontario, Canada  
michela.ippolito@utoronto.ca

## Abstract

In this paper I argue that both the co-speech and pro-speech symbolic gesture MAT (*mano a tulipano*) used by native speakers of Italian characterizes non-canonical *wh* questions. MAT can with a fast tempo contour and a slow tempo contour. Tempo is semantically distinctive: descriptively, a fast tempo characterizes a biased but information-seeking non-canonical question; a slow tempo characterizes a rhetorical non-canonical question. I will argue that the fast contour is the default tempo of MAT and that it brings about a ‘speaker bias’ interpretation. Slowing down the movement occurs when the feature [slow] is added: the semantic contribution of this feature is to generalize the bias introduced by MAT to all discourse participants (all discourse participants agree about the answer to the question). The “doxastic harmony” imposed by [slow] is the source of the rhetorical interpretation of the question. I speculate that tempo plays a similar role in the interpretation of a second symbolic gesture used in Italian, i.e. *mani giunte* (MG).

## 1 Introduction

This study is a preliminary investigation of a symbolic gesture used by native speakers of Italian: the *Mano a Tulipano* gesture (MAT), sometimes also labeled *grappolo* or *carciofo*. MAT can be used as co-speech and pro-speech gesture. Briefly described, MAT involves a path movement by which the speaker positions her hand at the level of her torso and during which the hand achieves the “tulip” configurations, i.e. all fingers tips touch, as shown in Figure 1. This path movement is then followed by a local movement generated at the wrist. Borrowing terminology from [16], I will call this local movement a “trill”: the hand moves repeatedly up and inward towards the speaker.



Figure 1: MAT

---

\*This work has been supported by a Jackman Humanities Research Fellowship awarded to the author by the Jackman Humanities Institute at the University of Toronto, and the SSHRC Research Grant #496697. I would like to thank the Jackman Humanities Institute, as well as the audience at the Università degli Studi di Milano, Bicocca, where this work was presented in the summer 2019. Special thanks for interesting discussions and feedback go to Francesca Panzeri, Donka Farkas, and Francesca Foppolo.

During our investigation of MAT, we will also consider a second symbolic gesture used by native speakers of Italian, i.e. the *Mani Giunte* gesture (MG). In MG, shown in Figure 2, both hands come together in the “prayer” position at the level of the torso and then move repeatedly up and down in place.



Figure 2: MG

Previous descriptive literature on gestures ([12], [15], [5], among others) has pointed out that both co-speech MAT and pro-speech MAT seem to have an interrogative component and are found in pragmatically marked constituent questions. For example, in his study of gestures used by speakers of a Southern Italian dialect spoken in the city of Naples, [12] suggests that a speaker uses MAT when confronted with something that undermines her expectations and therefore demands an explanation. More recently, [8] and [9] found occurrences of co-speech gestures MAT (which they call *carciofo*) and MG accompanying what they label *surprise/disapproval* questions (questions also marked by a special, non standard, intonation, and often introduced by the adversative coordinating conjunction *ma*, ‘but’).

With respect to the first observation (i.e. that MAT occurs in constituent questions), in line with the authors cited above, I am going to assume that MAT marks a *wh*-operator in a constituent question. As for the second observation made by the previous literature (that MAT is used to express surprise or disapproval), I will rephrase it by saying that MAT characterizes *non-canonical questions* as defined for example by [4], [7], and others. Non-canonical questions are questions where some of the assumptions which characterize standard (canonical questions) have been dropped. For example, canonical questions are characterized at the very least by two assumptions: an assumption that the speaker is ignorant about the answer to the question that is being asked, and an assumption that the addressee is competent about the answer to such a question. Unlike canonical questions, non-canonical questions are somehow marked: for example, they might have a non-standard syntactic form, or they might have a non-standard intonation. In the course of this paper, I will argue that the kind of questions marked by MAT are characterized by the absence of the speaker’s ignorance assumption. More specifically, I will argue that the *tempo* of MAT correlates with the kind of non-canonical question that is asked: a *fast* tempo MAT marks a non-canonical *biased* question, whereas a *slow* tempo MAT marks a non-canonical *rhetorical* question. I take a question to be biased if it signals that the speaker has a non-neutral doxastic attitude towards the possible answers to the questions and that there is a situation of doxastic *conflict* in the context: in this case, the question is still information-seeking and the goal of the speech act is ultimately to resolve the doxastic conflict. On the other hand, I take a question to be rhetorical if it is “doubly-biased”, that is, if it signals that both speaker and addressee have the same bias towards the answer to the question and that there is a situation of doxastic *harmony* in the context. So far, researchers

have focussed on the way that different kinds of declaratives and interrogatives sentences are characterized by different kinds of intonational contours ([11], [13], [14], [10], among others), where each intonational contour is identified as having a particular meaning (in the broad sense of the word). Borrowing the terminology from the literature on prosody, our hypothesis is that there are two semantically distinct *tempo contours* for MAT (and MG as well): the *fast contour* (FC), which characterizes a speaker-biased question, and the *slow contour*, which characterizes a rhetorical question.

Previous research and observations on these gestures are consistent with my hypothesis that both MAT and MG when used as co-speech gestures mark non-canonical questions. The research that I am reporting in this paper builds on this idea and expands it in two ways: (i) by looking at two different kinds of non-canonical questions that MAT and MG can mark and the specific feature of the gesture that characterizes each kind; (ii) by looking at *pro-speech* MAT and MG in order to establish whether the relation between the gesture and the meaning that is conveyed is direct and not mediated by the spoken utterance.

## 2 Hypothesis: tempo is significant

There are at least two kinds of non-canonical questions:<sup>1</sup> biased questions and rhetorical questions.<sup>2</sup> As we pointed out in the introduction, just like canonical questions, biased questions are information-seeking questions; however, unlike canonical questions, biased questions indicate that the speaker is not doxastically neutral towards the answer to the question. In uttering a biased question the speaker communicates that she has a bias towards one of the possible answers to the questions. Asking a question while at the same time communicating that the speaker has a bias towards an answer to the question is appropriate in a situation in which there is doxastic conflict, that is, in a situation where the speaker has reason to believe that the addressee might provide an answer to the question different from the one the speaker expects. In other words, speaker bias and the information-seeking nature of a question are compatible in a situation in which there is doxastic conflict with respect to the question under discussion. Rhetorical questions, on the other hand, are so to speak “doubly-biased” questions: the speaker and the hearer share the same bias. Since a rhetorical question conveys that both the speaker and the addressee have the same bias towards the answer to the question, in uttering a rhetorical question the speaker cannot be seeking information. Based on linguistic judgments of native speakers of Italian, I formulated the hypothesis that the type of non-canonical question marked by MAT depends on the tempo contour of the gesture: a fast contour (FC) marks an information-seeking biased question; a slow contour (SC) marks a rhetorical question. The same hypothesis holds for MG. The goal of the research reported in this paper was to test this hypothesis and to provide a semantic analysis consistent with the findings. This preliminary study is divided in two parts: a preference task and a forced choice task.

### 2.1 Part I: preference

In this part of the study, I recruited 13 native speakers of Italian, all undergraduate students at the University of Milan, Bicocca. 12 videos were prepared, each targeting a particular gesture: 4 out of these 12 videos were constructed to test our hypothesis concerning the semantic difference

<sup>1</sup>There are more kinds of non-canonical questions but the two I mentioned in the text are the ones that are relevant to the current project.

<sup>2</sup>In this paper, I will use the term ‘questions’ and ‘interrogatives’ interchangeably since here I am only looking at interrogative sentences.

between MAT-FC/MG-FC on the one hand and MAT-SC/MG-SC on the other; the other 8 videos targeted unrelated gestures. For each video, each participant (i) read a short paragraph on the screen describing a scenario ending with a character saying something; (ii) heard a sentence pronounced by this character, while the screen was dark; (iii) was shown four gestures made by the character who pronounced the sentence they heard. The face of the person making the gesture in the video could not be seen. Each participant was asked to rate each gesture on a scale from 1 to 7 with respect to how appropriate the gesture was, given what they heard the character say (1 = least appropriate, 7 = most appropriate). For the 4 relevant videos, the sentence that the participant heard was pronounced with either a biased/information-seeking intonation or with a rhetorical intonation. The target gesture was the one hypothesized to match the intonation of the utterance: the FC gesture for the biased-information-seeking intonation and the SC gesture for the rhetorical intonation. Whenever a MAT-FC/MG-FC was the target gesture, the video showed the competitor MAT-SC/MG-SC as well as two unrelated gestures. Similarly, when a MAT-SC/MG-SC was the target gesture, the video showed the competitor MAT-FC/MG-FC as well as two unrelated gestures. The participants were divided in two groups differing only with respect to the stories testing MAT and MG. Each participant saw 10 videos. In addition to the 8 videos targeting unrelated gestures, the first group was shown (i) one video where the target was MAT-FC and in which they also saw MAT-SC and (ii) one video where the target was MG-SC and in which they also saw the competitor MG-FC. The second group was shown (i) one video where the target was MAT-SC and in which they also saw MAT-FC and (ii) one video where the target was MG-FC and in which they also saw the competitor MG-SC. Thus, each participant saw all four relevant gestures: MAT-FC, MAT-SC, MG-FC, MG-SC. Below, I provide the content of two of the relevant videos.

- (1) Target gesture: MAT-FC
  - a. Text: Maria was invited to play at a friend's house. The friend's family knows that Maria is allergic to chocolate. However, Maria's mother sees her come back holding a chocolate bar in her hand and with chocolate stains around her mouth. Maria's mother is mad and asks:
  - b. Ma chi ti ha dato quella barretta? (biased utterance)  
But who has given you that bar  
Who gave you that chocolate bar?
  - c. Sequence of 4 gestures:  
1: *Basta* ('enough') gesture (unrelated); 2: MAT-SC (competitor); 3: *Smamma* ('go away') gesture (unrelated); 4: MAT-FC (target)
- (2) Target gesture: MAT-SC
  - a. Text: Elisa is talking with Anna about Elisa's chances of landing an academic job immediately after completing her Ph.D. Elisa is having a lot of trouble publishing. Anna tells her that someone will hire her.
  - b. Ma chi mi assumerà? (rhetorical utterance)  
But who me will-hire  
Who will hire me?
  - c. Sequence of 4 gestures:  
1: MAT-SC (target); 2: *parere personale* ('it's just my personal opinion') gesture (unrelated); 3: MAT-FC (competitor); 4: *basta* ('enough') gesture (unrelated)

The results are shown in Figure 4 for MAT and Figure 5 for MG.

Both Figure 4 and Figure 5 combine (a) both groups and (b) both biased and rhetorical

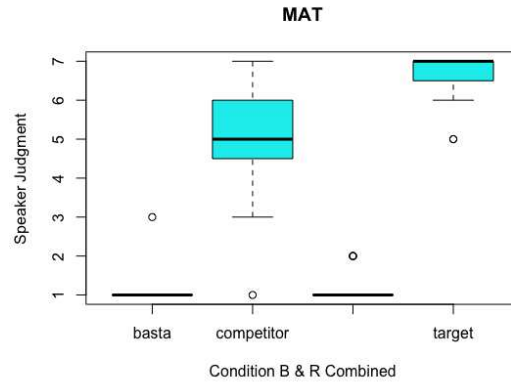


Figure 3: Results for MAT, both groups combined

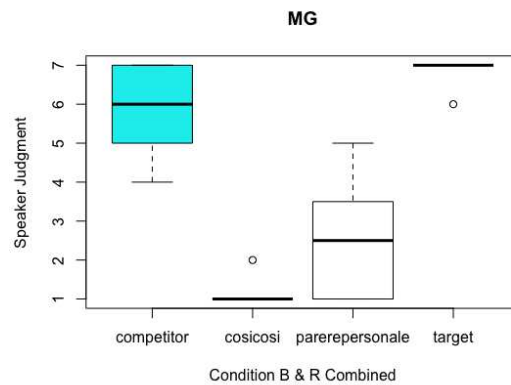


Figure 4: Results for MG, both groups combined

condition.<sup>3</sup> Figure 4 shows that the target MAT was preferred to its competitor (MAT with a different tempo contour); the unrelated gestures were considerably lower than both target and competitor. For example, in information-seeking/biased contexts, the target MAT-FC was preferred to its competitor MAT-SC; in rhetorical contexts, MAT-SC (target) was preferred to MAT-FC (competitor). Figure 5 shows the same results for MG.

## 2.2 Part II: forced choice

We collected the judgments of a large body of participants (98 undergraduate students at the University of Milan, Bicocca), divided in two groups. Participants read a text ending with a character making a gesture, and subsequently saw two videos, one showing the character performing a slow contour pro-speech gesture, and the other showing the character performing the same gesture but with a fast contour. Participants had to choose the one gesture that they

<sup>3</sup>In Figure 3, the column corresponding to the second unrelated gesture has not been labeled because, as the reader can see in (1) and (2), this gesture was different in the two videos (*smamma* and *parere personale*).

thought was more natural in the given context. Each participant saw two contexts: one for MAT and one for MG. In the end each participant saw all four relevant gestures once: MAT-SC, MAT-FC, MG-SC, MG-FC. The participants were divided in two groups: group A had 56 participants and group B had 42 participants. Group A was shown (i) a biased context with a choice between MAT-FC and MAT-SC, and (ii) a rhetorical context with a choice between MG-FC and MG-SC. Group B was shown (i) a rhetorical context with a choice between MAT-FC and MAT-SC, and (ii) a biased context with a choice between MG-FC and MG-SC. The rhetorical and biased contexts for the two groups were different. The results are shown in Figure 5.

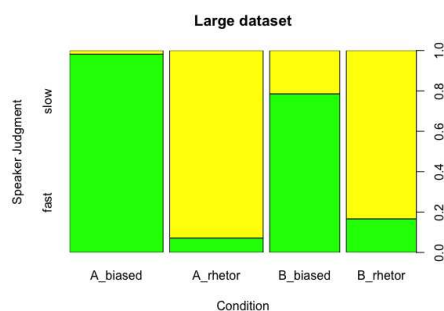


Figure 5: Results for Forced choice task, groups A and B

Group A's accuracy was 98% in the biased context and 92% in the rhetorical context. Group's B accuracy was 78% in the biased context and 83% in the rhetorical context. These results too support our hypothesis: the fast contour is strongly preferred in biased/information-seeking contexts, whereas slow contour is strongly preferred in rhetorical contexts.<sup>4</sup>

### 3 Semantics

In this section I focus on the semantics of MAT. In proposing a semantics for MAT in biased and rhetorical questions, there are at least three components of the gesture that might in principle contribute to its meaning: static MAT (that hand in the tulip configuration), movement, and tempo. The goal of the analysis I will sketch in this section is to capture the following points. First, MAT is a *wh*-operator: if MAT is a co-speech gesture, then there is also an overt constituent question (with an overt *wh*-word); if MAT is a pro-speech gesture, then the remainder of the question is covert. Second, movement indicates the divergence from some of the assumptions that characterize canonical, standard questions. In particular, movement with a fast (or, maybe more accurately, non-slow) tempo indicates *speaker's* bias towards the answer to the question: i.e. it indicates that the speaker believes that no positive answer to the

<sup>4</sup>There was an asymmetry between group A and group B in the way these judgments were collected. The participants in group A saw the MAT and MG contexts as part of a larger forced-choice task about gestures (which was part of a different study); the participants in the B group only saw the MAT and MG contexts I designed to check my hypothesis. Since the accuracy scores for the B group are lower than the scores for the A group, one might hypothesize that the higher accuracy in the A group is the result of a priming effect. Alternatively, the two groups read different scenarios and chose the gesture they thought was appropriate in the scenarios they saw. It is conceivable that this different might at least have contributed to the different degrees of accuracy we saw.



question is true. Movement with a slow tempo, on the other hand, indicates both *speaker and addressee*'s bias with respect to the question: more specifically, it indicates that the speaker believes that speaker and addressee share the same bias towards the answer to the question.

There are different ways to formalize these ideas. For example, with respect to the bias, one can formalize this idea more traditionally as a presupposition and as such closely linked to the compositional semantics of the question, or one can treat it as a conventional discourse effect – along the lines proposed by [7] and [6] – and as such as independent of the semantics. For convenience, here I will choose the presupposition analysis. A second issue that becomes important in formalizing these ideas is how one should understand the relation between MAT and movement. One possibility is that what we have described above is a complex structure formed by a static gesture (the hand in the tulip configuration) together with movement, which itself can be fast or slow. A different possibility is that there is no static gesture and that the basic non-decomposable unit is given by the moving MAT. In this scenario, the two different variants are given simply by slowing down the default tempo of the gesture: in this scenario, what we are calling the ‘fast contour’ is simply the default tempo contour of the gesture. If one represents tempo as a feature, then the question is whether we should formalize these two tempos as two independent features [fast] and [slow] that can combine with a static MAT, or whether we should only have the [slow] feature and represent the ‘fast’ tempo as the absence of [slow], i.e. as a kind of default tempo. Since I do not have the space here to argue for these points, I will proceed by making some assumptions with the caveat that, as long as they capture the main ideas that I am proposing, different formalizations can be considered, explored, and possibly adopted.

In order to simplify exposition, I will make the following assumptions. First, the gesture MAT (the hand in the tulip configuration) is essentially characterized by movement and it denotes a non-canonical *wh*-operator, i.e. a *wh*-operator carrying the speaker's presupposition (bias) that the speaker believes that no proposition in the denotation of the question is true. I will refer to this gesture as  $\text{MAT}_{\text{MVT}}$ .<sup>5</sup> Second, a feature [slow] can combine with  $\text{MAT}_{\text{MVT}}$ . If [slow] is not there, then we have what we called the ‘fast contour’, which now we understand to be a default/non-slow contour: this has the meaning of the biased MAT question. If, on the other hand, the feature [slow] is there, then we have what we called the ‘slow contour’: [slow] adds the presupposition that the speaker believes that the *addressee* shares the speaker's bias with respect to the answer to the question. A  $\text{MAT}_{\text{MVT}}$  marked with [slow] is interpreted as a rhetorical MAT question. Note that in this picture, what makes MAT a rhetorical question is not that the answer is part of the common ground but that it presupposes that both speaker and addressees agree on the answer.<sup>6</sup> I will assume that the denotation of a question is the set of possible answers and I will make the simplifying assumption that the *wh* operator combines with the open proposition  $P$  (of type  $\langle \text{et} \rangle$ ) in its complement and creates a set of propositions  $p$  where  $p$  is obtained by combining  $P$  with  $x$ , for every  $x$  in a relevant set of entities. The contextual parameter  $c$  includes a set  $A$  of salient entities, which are crucial in the construction of the propositions in the denotation of the question, and the set of discourse participants  $DP$ , which includes speaker  $s$  and addressee  $a$ .

$$(3) \quad \begin{aligned} & \llbracket [\text{MAT}_{\text{MVT}} P] \rrbracket^c \text{ defined if for every } x \in A_c, \text{DOX}_{s \in DP_c} \cap P(x) = \emptyset; \text{ if defined, } \llbracket [\text{MAT}_{\text{MVT}} \\ & P] \rrbracket^c = \{p : p = P(x) | x \in A_c\} \end{aligned}$$

<sup>5</sup>Our main reason for not having an interrogative *static* MAT is that the latter does not seem to be a possible gesture for Italian speakers. There is evidence that MAT occurs in LIS (Lingua Italiana dei Segni, ‘Italian sign language’) as well, as discussed in [2] and that too seems to be characterized by movement. An investigation of the features and semantics of MAT in LIS is lacking, a gap to be addressed in the future.

<sup>6</sup>This makes the current proposal closer to [1] than [3].

This speaker’s presupposition is carried by  $\text{MAT}_{\text{MVT}}$  itself. [slow] introduces the additional presupposition that all discourse participants in DP *agree* which respect to the answer to the question that is being asked, where I assume that to agree with respect to a proposition is to have the same doxastic attitude towards that proposition. Thus, I take it that to agree with respect to the answer to a question means to have the same doxastic attitude towards all the propositions in the denotation of the question. Compositionally, I take [slow] to be an identity function applying to  $\text{MAT}_{\text{MVT}}$  and adding the presupposition that for every discourse participant  $d$  and  $d'$ ,  $d$  has the same doxastic state as  $d'$  with respect to each proposition in the denotation of the question; that is,  $d$  and  $d'$  agree with respect to the answer to the question. This is shown in (4).

- (4)  $[[\text{slow}]^c = \lambda Q_{\langle st,t \rangle} : \text{for every } p \in Q \text{ and for every } d, d' \in DP_c, d \text{ has the same doxastic attitude as } d' \text{ with respect to } p. Q$

When applied to a MAT question, we obtain the semantics in (5).

- (5)  $[[[\text{slow}] \text{MAT}_{\text{MVT}} P]^c$  defined if for every  $p \in [[\text{MAT}_{\text{MVT}} P]^c$  and for every  $d, d' \in DP_c$ :  $d$  has the same doxastic attitude at  $d'$  towards  $p$ ; if defined,  $[[[\text{slow}]\text{MAT}_{\text{MVT}} P]^c = \{p : p = P(x)|x \in A_c\}$

Since [slow] applies to a  $\text{MAT}_{\text{MVT}}$ -question, requiring that speaker and addressee agree with respect to the answer to the question amounts to requiring that the speaker and addressee share the same bias towards the answer to the question.

As an illustration, consider the interrogative utterance in (6) accompanied by co-speech MAT-FC. The first line approximates the timing of gesture execution: the formation of the tulip hand overlaps with the utterance of the interrogative pronoun *chi*, ‘who’; the succession of wedges ( $\wedge$ ) is used to describe the trill, which begins immediately after the tulip is formed and continues until the end of the utterance.

- (6) MAT-FC  $\wedge \wedge \wedge \wedge \wedge \wedge$   
**Chi** ti aiuterà?  
 Who will help you?

In the following diagram, the elements that contribute to the semantics of the gesture in relation to the utterance are hierarchically arranged.  $\text{MAT}_{\text{MVT}}$  duplicates the *wh*-word and adds the speaker’s presupposition in (3).



The dialogue in (8) illustrates an occurrence of pro-speech MAT.

- (8) A: Qualcuno mi aiuterà.  
 Someone me will-help  
 Someone will help me.  
 B: **MAT-FC**  $\wedge \wedge \wedge \wedge \wedge \wedge$

In this case, part of the question is covert: the interrogative pronoun is contributed by  $\text{MAT}_{\text{MVT}}$  and the open proposition  $[\lambda x. x \text{ will help you}]$  is made salient by the previous utterance. More

specifically, in this case  $P$  in (7) is a covert variable indexed in such a way as to be assigned as its value the predicate  $[\lambda x. x \text{ will help you}]$  made salient by the previous utterance. The complete meaning of B’s gesture is equivalent to the meaning of (6).

To recap our proposal so far: (i)  $\text{MAT}_{\text{MVT}}$  has the semantics of a non-canonical *wh* interrogative pronoun; (ii)  $\text{MAT}_{\text{MVT}}$  is non-canonical in that it introduces the bias that the speaker believes all answers to the question to be false; (iii) speaker’s bias was modelled as a presupposition; (iv) the bias of the speaker (only) is characterized by what we labelled the ‘fast’ contour. A MAT-FC question is felicitous in a situation in which there is doxastic conflict between the discourse participants: in (6) and (8), for example, the speaker is at the same time asking a question whether someone will help A and presupposing that she believes that nobody will help A. This must be because the speaker has been given (e.g. by the addressee’s asserting that someone will help her) some evidence against what she believes is the answer to the question, and is now trying to resolve this situation of doxastic conflict.

The diagram in (9) shows the structure when movement is marked by the feature [slow].



Given the meaning for [slow] we established in (4), the presupposition introduced by  $\text{MAT}_{\text{MVT}}$  is attributed to all the discourse participants. For example, in the slow contour version of (8) – shown below in (10) – every discourse participant is required to believe that nobody will help A.

- (10) A: Qualcuno mi aiuterà.  
 Someone me will-help  
 Someone will help me.  
 B: MAT-SC/~~~~~

Unlike the biased case, the speaker is now presupposing that there is ‘doxastic harmony’ between her and the addressee in that they both share the speaker’s bias about the answer to the question. This doxastic harmony with respect to a bias about the answer to a question is, I claim, the feature of a rhetorical question.

## 4 Conclusion

MAT can occur as a co-speech and pro-speech gesture. Previous descriptive literature has observed that it occurs in non-canonical interrogative sentences, i.e. interrogatives denoting questions where one of the canonical assumptions about questions has been abandoned. In this paper I have argued that both co-speech and pro-speech MAT have the semantics of a non-canonical *wh*.  $\text{MAT}_{\text{MVT}}$  by itself brings about a ‘speaker bias’ interpretation – this is what we labeled the ‘fast contour’, which we now understand to be the unmarked tempo of the gesture. When the feature [slow] is added, the bias introduced by  $\text{MAT}_{\text{MVT}}$  is generalized to all discourse participant: the doxastic harmony imposed by [slow] is the source of the rhetorical interpretation of the question (all discourse participants agree about the answer to the question). I did not have space to analyze the semantics of MG in this paper. However, given the results of the tests I have reported above, my hypothesis is that  $\text{MG}_{\text{MVT}}$  has a non-canonical interrogative meaning and carries a ‘speaker bias’ presupposition of the kind introduced above; modification

by [slow] in this case too will introduce a ‘doxastic harmony’ that will give rise to the rhetorical interpretation.

## References

- [1] Maria Biezma and Kyle Rawlins. Rhetorical questions: Severing questioning from asking. In Dan Burgdorf, Jacob Collard, Sireemas Maspong, and Brynhildur Stefánsdóttir, editors, *Proceedings of Semantics and Linguistic Theory 27*, pages 302–322, 2017.
- [2] Chiara Branchini, Cardinaletti Anna, Carlo Cecchetto, Caterina Donati, and Carlo Geraci. Wh-duplication in italian sign language (LIS). *Sign Language and Linguistics*, 16(2):157–188, 2013.
- [3] Ivano Caponigro and Jon Sprouse. Rhetorical questions as questions. In Estela Puig-Waldmueller, editor, *Proceedings of Sinn und Bedeutung 11*, pages 121–133, 2007.
- [4] Veneeta Dayal. *Questions*. Oxford Surveys in Semantics and Pragmatics. Oxford University Press, 2016.
- [5] Perangela Diadori. *Senza parole - 100 gesti degli italiani*. Bonacci, 4th updated edition (first published in 1990) edition, 2003.
- [6] Donka Farkas. Non-intrusive questions. Handout, November 9, 2018.
- [7] Donka Farkas and Floris Roelofsen. Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics*, 34(2):237–289, 2017.
- [8] Alessandra Giorgi. On surprise questions and surprise-disapproval exclamatives introduced by adversative *ma* (but): syntax, prosody, and gesture. Ca’ Foscari University, Venice, 2018.
- [9] Alessandra Giorgi and Chiara Dal Farra. On the syntax/pragmatics interface: Expressing surprise and disapproval. Ca’ Foscari University of Venice, 2018.
- [10] Dan Goodhue and Michael Wagner. Toward an intonational bestiary. <http://prosodylab.org/data/bestiary/toronto/#1>, 2019.
- [11] Christine Gunlogson. *True to form: Rising and falling declaratives as questions in English*. PhD thesis, UCSC, 2001.
- [12] Adam Kendon. *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- [13] Mark Liberman and Ivan Sag. Prosodic form and discourse function. In *Proceedings of the Chicago Linguistics Society 10*. University of Chicago, 1974.
- [14] Janet Pierrehumbert and Julia Hirschber. The meaning of intonational contours in the interpretation of discourse. In Jerry Cohen, Philip & Morgan, editor, *Intentions in communication*, pages 271–311. Cambridge, MA: MIT Press, 1990.
- [15] Isabella Poggi. Gesti. *Enciclopedia Treccani dell’Italiano (Encyclopedia of Italian)*, 2010.
- [16] Wendy Sandler. The phonology of movement in sign language. In Marc van Oostendorp, Marc Ewen, Keren Rice, and Elizabeth Hume, editors, *Blackwell Companion to Phonology*, pages 577–603. Wiley-Blackwell, 2010.

# Picturing words: the semantics of speech balloons\*

Emar Maier

Theoretical Philosophy/MILLC, University of Groningen  
e.maier@rug.nl

Semantics traditionally focuses on linguistic meaning. In recent years, the Super Linguistics movement has tried to broaden the scope of inquiry in various directions, including an extension of semantics to talk about the meaning of pictures. There are close similarities between the interpretation of language and of pictures. Most fundamentally, pictures, like utterances, can be either true or false of a given state of affairs, and hence both express propositions (Zimmermann, 2016; Greenberg, 2013; Abusch, 2015). Moreover, sequences of pictures, like sequences of utterances, can be used to tell stories. Wordless picture books, comics, and film are cases in point. In this paper I pick up the project of providing a dynamic semantic account of pictorial story-telling, started by Abusch (2012) and continued by Abusch & Rooth (2017); Maier & Bimpikou (2019); Fernando (2020).

More specifically, I propose here a semantics of speech and thought bubbles by adding event reference and event modification to PicDRT. To get there I first review the projection-based semantics for pictures (section 1), noting the fundamental distinction between symbolic and iconic meaning that makes speech bubbles especially interesting (section 2). I then review the dynamic PicDRT framework for pictorial narratives (section 3), add events (section 4), and propose an account of speech bubbles as quotational event modification (section 5). I end with a brief look at other conventional event modifiers in comics such as motion lines (section 6).

## 1 Picture semantics

As a first pass at bringing our formal semantics toolkit to the study of pictures, we might say a picture expresses the proposition consisting of the set of all those worlds that ‘look like’ the picture. However, just as sentences with indexical elements express propositions only relative to a context of utterance (Kaplan, 1989), so pictures are usually assumed to express propositions relative to a viewpoint (formally, a unit vector located somewhere in space time, specifying from where, when, and in what direction we’re observing/picturing the world). With this notion of a viewpoint we can introduce the notion of a projection function, which will then replace the vague (and problematic, Goodman 1976; Greenberg 2013) notion of resemblance in our first rough characterization. In short, a projection is a recipe for turning a 3D scene (part of a world seen from a viewpoint) into a 2D representation. For example, using linear perspective digital photography projection, as implemented in my phone, we can project the actual world  $w$ , from a viewpoint  $v$  facing the Keizersgracht from the bridge onto a digital photo, (1a). With a different projection algorithm,  $\pi_2$ , treating projection lines coming from edges and surfaces somewhat differently, we can map the same world, from more or less the same viewpoint, onto a 2D drawing:<sup>1</sup>

---

\*Many thanks to Sofia Bimpikou and the anonymous AC reviewers, and to the workshop organizers. This research is supported by NWO Vidi Grant 276-80-004 (FICTION).

<sup>1</sup>Photo from <https://pixabay.com/photos/amsterdam-keizersgracht-netherlands-686460/>, drawing from <https://pixabay.com/illustrations/amsterdam-keizersgracht-netherlands-3609378/>.

$$(1) \quad \text{a. } \pi_1(w, v) = \img alt="A photograph of a canal scene in Amsterdam, showing buildings, a bridge, and a street lamp." data-bbox="305 158 448 225"/> \quad \text{b. } \pi_2(w, v') = \img alt="A painting of the same canal scene in Amsterdam, showing buildings, a bridge, and a street lamp." data-bbox="575 158 718 225"/>$$

We can now more precisely define the proposition expressed by a picture as the inverse of the relevant projection, say  $\pi_2$  for the interpretation of the drawing (Abusch, 2015).

$$(2) \quad \left[ \left[ \img alt="A painting of a canal scene in Amsterdam." data-bbox="208 281 335 332" \right] \right]^{v'} = \left\{ w \mid \pi_2(w, v') = \img alt="A painting of a canal scene in Amsterdam." data-bbox="498 281 611 332} \right\}$$

In other words, the meaning of a picture, relative to a fixed viewpoint  $v$  and projection function  $\pi$  provided by the model, is the set of worlds  $w$  that could be projected onto that picture from viewpoint  $v$ . However, when we interpret a picture we often don't have any prior, independent access to the spatio-temporal viewpoint the artist used to create her projection (this is especially, but not exclusively, true for fictional depictions). Instead, we may infer certain properties of the viewpoint from the picture itself, just like in interpretation we try to infer properties of the world depicted. To model this we use a more abstract, two-dimensional, or centered, notion of content, akin to a Kaplanian character, or its diagonal. Generally, for any picture  $\alpha$ , interpreted in a model providing a projection function  $\pi$  alongside a set of worlds and a set of viewpoints (Rooth & Abusch, 2017):

$$(3) \quad \llbracket \alpha \rrbracket = \{ \langle w, v \rangle \mid \pi(w, v) = \alpha \}$$

## 2 Iconic vs. symbolic meaning

The key differences between pictorial and linguistic meaning are in *how* they express propositions. Linguistic meaning is considered mostly *symbolic* (i.e., depending on an arbitrary, purely conventional lexicon), and *compositional* (i.e., depending on a grammar specifying how meanings of complex expressions are built up out of constituent meanings). Pictorial meaning by contrast is mostly *iconic*, i.e. representing by virtue of some more or less natural, structure preserving transformation relation (like our projection above), and *holistic*, i.e. independent of grammatical constituent structure.

However, on closer inspection, many aspects of language turn out to be more or less iconic (onomatopoeia (Henderson, 2016), co-speech gestures (Schlenker, 2018), sign language classifiers (Davidson, 2015), etc.). Quotation and quotatives, especially in spoken language, are often analyzed as iconic as well: in quoting, a reporter 'demonstrates' a previous speech act by producing something 'similar' (Clark & Gerrig, 1990; Recanati, 2001; Davidson, 2015).

On the other side, the degree to which certain drawing or even photography and film styles are purely iconic has also been debated. Linear perspective drawing, for instance, is arguably not so 'natural' that early medievals were able to apply it. Moreover, it's not clear that the visual system itself follows linear perspective projection in the interpretation of our surroundings (Giardino & Greenberg, 2015; Greenberg, 2013).

A rather different instance of symbolic/compositional meaning intruding in the pictorial domain comes (again, surprisingly) from quotation. If in a drawing we want to depict someone saying something, we often add writing. A common convention, especially in comics and car-

toons, is to use the speech balloon symbol to show who said what (4b). Similarly, a thought bubble shows what a character is thinking (4a).<sup>2</sup>



Consider also fight clouds (5a), sound effect descriptions (5b), explosion stars, motion lines/reduplication/blurring (5c), and what Cohn (2013) calls ‘affixes’, like the idea-lightbulb, angry-thundercloud, stars circling over a passed out character’s head (5d), or the bulging-vein anger symbol (5e). All of these symbols serve as conventional indicators of certain types of events or (mental) states, significantly enriching the purely projective, pictorial meaning of the panels in which they occur.



The goal of this paper is to give a compositional semantics of speech and thought balloons, and integrate that into a projection-based dynamic semantic account for pictorial narratives. In the final section I share some preliminary thoughts about some of the other symbolic enrichments in (5).

### 3 PicDRT

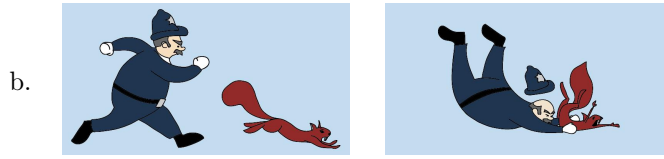
Following Abusch (2012) and Abusch & Rooth (2017), Maier & Bimpikou (2019) introduce PicDRT, a rather minimal extension of standard DRT (Kamp, 1981) for wordless comics. In a nutshell: the PicDRT construction algorithm adds the current panel in a new, labeled ‘picture condition’, identifying the ‘regions of interest’ (corresponding to salient individuals), and tagging these regions with fresh discourse referents. A postsemantic, pragmatic enrichment module takes care of linking some of the newly introduced discourse referents up with already established ones.

To illustrate, I reuse a minimal example from Maier & Bimpikou (2019), the visual analogue of a two-sentence discourse from a well-known introduction to DRT (Geurts, 1999):

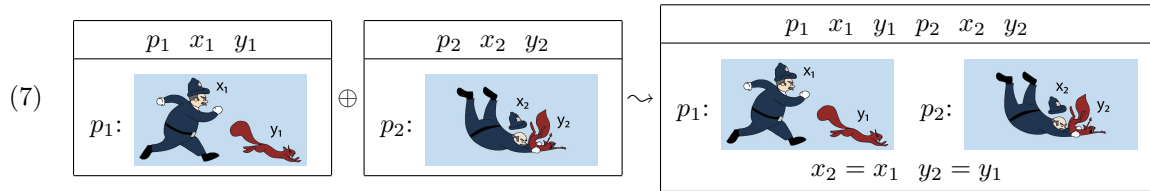
- (6) a. A policeman was chasing a squirrel. He caught it.

<sup>2</sup>Obviously, none of these or any other pictures in the paper are mine. They are panel fragments taken from googled images. These two *Donald Duck* panels are from an image published at <https://www.nrc.nl/nieuws/2015/12/10/donald-duck-verrijkt-de-taal-1565843-a1298767>.





The first picture is interpreted as adding three discourse referents:  $p_1$  representing the viewpoint of the entire picture; and  $x_1$  and  $y_1$  representing the individuals corresponding to the two salient picture regions (which we assume can be identified by the preliminary DRS construction algorithm, i.e. at the level of syntax/semantics, before pragmatic reasoning). The second picture is treated similarly, introducing three new discourse referents. After merging the two PicDRSs we pragmatically enrich the representation by inferring that the individual depicted in region  $x_1$  is (probably) the same as that depicted in  $x_2$ , on the basis of certain plausible assumptions about similarity, world-knowledge, temporal progression between panels and other aspects of coherent story-telling.



Pragmatic enrichment need not stop here. Typical readers will view the picture sequence not just as depicting two states of affairs in which two entities reoccur, but add that the one is a policeman, and the second a squirrel, etc. We can add these defeasible pragmatic inferences as additional DRS conditions on  $x_1$  and  $y_1$  in the PicDRS. We return the descriptive enrichment below, after we've introduced events.<sup>3</sup>

To interpret PicDRS boxes in a model we use an individual assignment function  $f$  to verify the regular DRS conditions, and a viewpoint assignment  $v$  to verify the pictorial conditions.







- (8)
- a. model:  $M = \langle D, W, I, V, \pi \rangle$  (where  $D$  is domain of individuals,  $W$  set of worlds,  $I$  and interpretation function,  $V$  a set of viewpoints, and  $\pi$  a projection function)
  - b. viewpoint assignment:  $v : \subset \{p_1, p_2, \dots\} \rightarrow V$
  - c. individual assignment:  $f : \subset \{x_1, x_2, y_1, \dots\} \rightarrow D$

The semantic definition of truth for PicDRSs combines a standard DRT semantics for descriptive conditions involving individual discourse referents, with a projective picture semantics for the picture conditions. Since the individual discourse referents thus refer to individuals via  $f$ , but also correspond to picture regions that refer to space-time regions via  $v$  (given  $\pi$ ), we must make sure that  $f$  and  $v$  are properly 'aligned'. Technically, this last step requires that we extend the projection function  $\pi$  to map not only worlds (seen from a viewpoint) to pictures, but also individuals in those worlds (seen from a viewpoint) to picture regions. For concreteness we apply here the semantics to the final, pragmatically enriched example DRS in (7).

- (9)
- a.  $\llbracket (7) \rrbracket^{w,v,f} = 1$  iff there is a verifying embedding  $f' \supset f$  with  $Dom(f') = \{x_1, y_1, x_2, y_2\}$  and a viewpoint assignment  $v' \supset v$  with  $Dom(v') = \{p_1, p_2\}$  such that:
    - a.  $f'$  verifies the descriptive conditions:

<sup>3</sup>Cf. Wildfeuer (2019) for a different take on the DRT modeling of pictorial narratives, without projective picture conditions and without a distinction between semantic DRS construction and post-semantic enrichment.



- (i)  $f'(x_2) = f'(x_1)$   
(ii)  $f'(y_2) = f'(y_1)$
- b.  $v'$  verifies the pictorial conditions:
- (i)  $\pi(w, v'(p_1)) =$  
- (ii)  $\pi(w, v'(p_2)) =$  
- c.  $f'$  and  $v'$  are aligned:
- (i)  $\pi(f'(x_1), w, v'(p_1)) =$    
(i.e. the policeman-region in the picture is a projective depiction of the policeman-entity  $f'(x_1)$  in  $w$ , from the viewpoint associated with the picture)
- (ii)  $\pi(f'(y_1), w, v'(p_1)) =$  
- (iii)  $\pi(f'(x_2), w, v'(p_2)) =$  
- (iv)  $\pi(f'(y_2), w, v'(p_2)) =$  

## 4 Picturing events

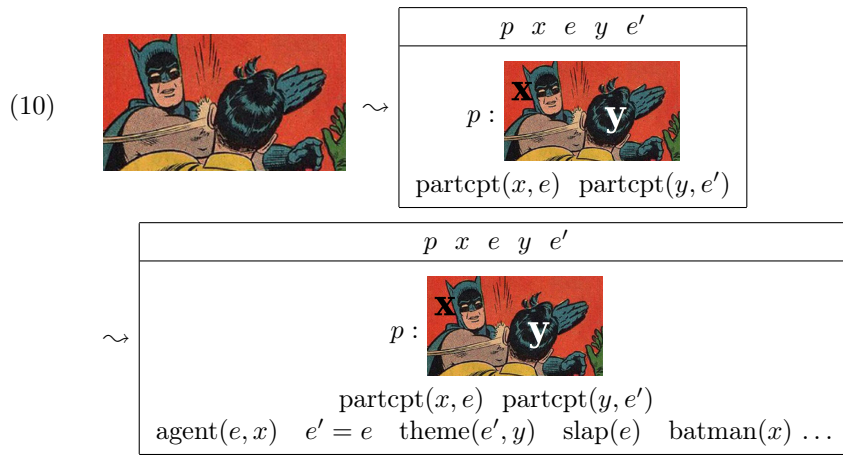
Speech and thought bubbles are readily thought of as the visual language analogue of quotation marks in written language (Saraceni, 2003; Cohn, 2013). To make this intuition precise I propose to extend PicDRT with some Davidsonian event semantics and then apply the powerful event-modification semantics of quotation (and reporting constructions more generally, see Kratzer 2006; Davidson 2015; Maier 2017).

Intuitively, we tend to read sequences of panels as representing sequences of events – in fact, that is what makes it a narrative. Panels should thus be understood as depicting (and introducing appropriate discourse referents for) not just individuals, but also events. Panels thereby become more like full-fledged utterances, which likewise introduce event discourse referents, though there it happens compositionally, through the lexical semantics of verbs. One further benefit of introducing events in visual narrative panels is that it's a first step toward applying a more general theory of discourse structure like SDRT, but for reasons of space I leave that general extension for another occasion.

It is tempting to simply treat each panel as introducing a single (main) event discourse referent. But it's not hard to find counterexamples: panel sequences depicting stages of a single event, or depicting stative scenes (without temporal progression), or single panels depicting multiple (non-simultaneous) events (McCloud, 1993). If we look at comics with lots of dialogue and/or action we quickly find single panels with multiple speech and thought balloons attaching to different characters (who may also be moving around or experiencing emotions at the same time, perhaps indicated by further event modifiers as in (5)). One drastic option would be to leave the introduction of event discourse referents entirely up to pragmatic enrichment. Semantically speaking, the picture depicts a scene in a world, with a number of salient individuals, and on the basis of world knowledge we may infer that some particular type of event is probably happening. I opt for a middle position between requiring a general semantic one-event-per-panel rule and a leave-it-to-pragmatics strategy.

I propose to link the introduction of event discourse referents to the introduction of individ-

uals. Each salient individual, as identified by the DRS construction algorithm, by stipulation, participates in some eventuality. Concretely, let's say that the DRS construction algorithm adds with each individual discourse referent  $x$  a new eventuality discourse referent  $e$ , and a condition stating that  $x$  participates in  $e$  ( $\text{partcpt}(x, e)$ ). On the basis of the picture, world-knowledge, context, and rationality assumptions we may then post-semantically infer exactly what thematic roles and what (kinds of) eventualities are depicted ( $\text{agent}(e, x)$ ,  $\text{chase}(e)$ ,  $\text{slap}(e), \dots$ ), at the same time as we're adding descriptive properties of the individuals ( $\text{policeman}(x_1)$ ,  $\text{batman}(x)$ ), and adding anaphoric relations between old and new discourse referents in the case of a multipanel narrative ( $x_2 = x_1$ ,  $e' \prec e, \dots$ ). In the example below we have tow characters, which leads the construction algorithm to introduce two associated event discourse referents. In this case, post-semantic pragmatic reasoning might equate the two events, leaving just one event, a slapping, with two individuals fulfilling different thematic roles (Batman the agent, Robin the patient).



## 5 Bubbles

To deal with speech bubbles (and other symbolic enrichments exemplified in (5)) we'll need to decompose the panel into its (syntactic) constituents: the picture itself<sup>4</sup> (to be interpreted projectively, as part of a picture condition) and the bubble (to be interpreted decriptively, as a linguistic quotation). To preserve the connection between the speech bubble and the individual it 'points to' in the picture we use the discourse referent that the construction algorithm also generates for salient picture regions. From here on let's denote the first phase of interpretation, the generation of a DRS representation of a panel (or meaningful sign more generally), with  $\langle \rangle$  (because we like to reserve  $\llbracket \rrbracket$  for the second phase, i.e. the modeltheoretic interpretation of DRSs).<sup>5</sup>

<sup>4</sup>Note that once we disentangle the speech balloon and the picture we are left with a hole in the picture. The projective semantics needs a bit of adjustment to properly interpret holes as underspecified regions.

<sup>5</sup>The example is a fragment of a panel from <https://www.smbc-comics.com/comic/complex>

$$(11) \quad \langle\langle \text{GOD I HOPE NOT} \rangle\rangle = \begin{array}{|c|} \hline p \ x \ e \\ \hline p: \text{ } \langle\langle \text{GOD I HOPE NOT} \rangle\rangle \\ \hline \text{partcpt}(x, e) \\ \hline \end{array} \oplus \langle\langle \text{GOD I HOPE NOT} \rangle\rangle(x)$$

We analyze bubbles semantically, in the construction algorithm, as quotations, i.e. operators that take a linguistic input, a string of letters (written inside the bubble, as a kind of infix notation), to yield a property of individuals, viz. saying something of that form. This gives it the right type to apply to an individual discourse referent, as needed in (11).<sup>6</sup>

$$(12) \quad \begin{array}{l} \text{a. } \langle\langle \text{ } \rangle\rangle = \lambda s \lambda y. \begin{array}{|c|} \hline e' \\ \hline \text{say}(e') \ \text{agent}(e', y) \\ \hline \text{form}(e', s) \\ \hline \end{array} \\ \\ \text{b. } \langle\langle \text{GOD I HOPE NOT} \rangle\rangle = \langle\langle \text{ } \rangle\rangle(\ulcorner \text{God, I hope not} \urcorner) = \lambda y. \begin{array}{|c|} \hline e' \\ \hline \text{say}(e') \ \text{agent}(e', y) \\ \hline \text{form}(e', \ulcorner \text{God, I hope not} \urcorner) \\ \hline \end{array} \end{array}$$

We can now plug (12) into (11), to get the output of the construction algorithm:

$$(13) \quad \begin{array}{|c|} \hline p \ x \ e \ e' \\ \hline p: \text{ } \langle\langle \text{GOD I HOPE NOT} \rangle\rangle \\ \hline \text{partcpt}(x, e) \\ \text{say}(e') \ \text{agent}(e', x) \\ \text{form}(e', \ulcorner \text{God, I hope not} \urcorner) \\ \hline \end{array}$$

In words, the picture depicts an individual  $x$  who participates in some event  $e$ , and who is also the agent of a speech event  $e'$ , which exhibits the linguistic form “God, I hope not”.

Note that I’ve now analyzed the speech balloon as introducing a new event discourse referent  $e'$  of its own, rather than automatically linking to the default event associated with the individual (stipulated in 4). Of course, this leaves open the possibility of equating the two event variables post-semantically, pragmatically simplifying the semantic representation generated by the construction algorithm into one where what we see is an individual who is the agent of a particular speech event. The advantage of the current proposal is that it straightforwardly allows one panel to depict a single individual with various speech balloons originating from them, who is also participating in some salient non-linguistic activity.

Other bubble shapes conventionally encode other event properties. Cloud-like contours indicate thoughts, (replace ‘say( $e$ )’ with ‘think( $e$ )’ in (12a), see Maier (2017) on thought quotation more generally), bolded spiky contours indicate shouting, etc.

<sup>6</sup>For readability I freely apply lambda conversions in DRS construction derivations.

## 6 Beyond bubbles

The basic analysis of verbal bubbles provided here can be further extended in various directions. For instance, Cohn's (2013) non-verbal affixes, like the aforementioned motion lines, idea-lightbulbs, head-circling stars, or bulging vein anger symbols in (5), can be analyzed rather similarly. In the DRS construction phase, like bubbles they are isolated from the picture to be interpreted symbolically, as conventionally (not projectively) denoting properties of individuals, and then they are applied to the discourse referent associated with the region they were visually attached to.

Some of the symbolic embellishments in (5) are not obviously attached to a specific individual. Perhaps some sound effect descriptions are better interpreted as modifiers of events, and thus semantically applied to the event discourse referents, that are in turn introduced by individual depictions or other symbols. Action stars and fight clouds may just introduce events, only pragmatically linked to entities represented in the surrounding narrative (Cohn & Wittenberg, 2015). Further research is required to properly analyze the various ways in which pictorial representations interact with symbolic enhancements.

As a final remark on thought bubbles, comics allow for the use of pictures inside bubbles to represent an agent's thoughts or other mental states iconically. We can model this with our event-based semantics by generalizing the predicate 'form( $e, s$ )' to take a picture as its second argument ( $s$ ). The idea is that thought events, imaginative projects, and similar mental processes can be either linguistic or visual. If  $s$  is a picture and  $e$  a thought event, 'form( $e, s$ )' means that the thought is essentially a visual experience of the worlds that are projective mapped to the given thought bubble picture. Interestingly, the viewpoint of the thought bubble picture need not coincide with the *de se* center of the experience, leading to distinction parallel to that between free perception panels (Abusch & Rooth, 2017) and blended perception panels (Maier & Bimpikou, 2019).

## References

- Abusch, Dorit. 2012. Applying Discourse Semantics and Pragmatics to Co-reference in Picture Sequences. *Proceedings of Sinn und Bedeutung* 17.
- Abusch, Dorit. 2015. Possible worlds semantics for pictures. In *The Blackwell Companion to Semantics*, .
- Abusch, Dorit & Mats Rooth. 2017. The formal semantics of free perception in pictorial narratives. *Proceeding of the 21st Amsterdam Colloquium* .
- Clark, Herbert & Richard Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805.
- Cohn, Neil. 2013. *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. A&C Black.
- Cohn, Neil & Eva Wittenberg. 2015. Action starring narratives and events: Structure and inference in visual narrative comprehension. *Journal of Cognitive Psychology* 27(7). 812–828. doi:10.1080/20445911.2015.1051535.
- Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38(6). 477–520. doi:10.1007/s10988-015-9180-1.

- Fernando, Tim. 2020. Pictorial narratives and temporal refinement. *Semantics and Linguistic Theory (SALT)* 29.
- Geurts, Bart. 1999. *Presuppositions and Pronouns*. Amsterdam: Elsevier.
- Giardino, Valeria & Gabriel Greenberg. 2015. Introduction: Varieties of Iconicity. *Review of Philosophy and Psychology* 6(1). 1–25. doi:10.1007/s13164-014-0210-7.
- Goodman, Nelson. 1976. *Languages of Art: An Approach to a Theory of Symbols*. Hackett Publishing.
- Greenberg, Gabriel. 2013. Beyond Resemblance. *Philosophical Review* 122(2). 215–287. doi:10.1215/00318108-1963716.
- Henderson, Robert. 2016. A demonstration-based account of (pluractional) ideophones. *Semantics and Linguistic Theory* 26(0). 664–683. doi:10.3765/salt.v26i0.3786.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal Methods in the Study of Language*, 277–322. Amsterdam: Mathematical Centre Tracts.
- Kaplan, David. 1989. Afterthoughts. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from Kaplan*, 564——. Oxford: Oxford University Press.
- Kratzer, Angelika. 2006. Decomposing attitude verbs Handout. Honoring Anita Mittwoch on her 80th birthday. The Hebrew University of Jerusalem.
- Maier, Emar. 2017. The pragmatics of attraction: Explaining unquotation in direct and free indirect discourse. In Paul Saka & Michael Johnson (eds.), *The Semantics and Pragmatics of Quotation*, Berlin: Springer.
- Maier, Emar & Sofia Bimpikou. 2019. Shifting perspectives in pictorial narratives. *Sinn und Bedeutung* 23.
- McCloud, Scott. 1993. *Understanding Comics*. Tundra Publishing.
- Recanati, François. 2001. Open quotation. *Mind* 110(439). 637–687. doi:10.1093/mind/110.439.637.
- Rooth, Mats & Dorit Abusch. 2017. Picture descriptions and centered content. *Proceedings of Sinn und Bedeutung* 21.
- Saraceni, Mario. 2003. *The Language of Comics*. Psychology Press.
- Schlenker, Philippe. 2018. Iconic pragmatics. *Natural Language & Linguistic Theory* 36(3). 877–936. doi:10.1007/s11049-017-9392-x.
- Wildfeuer, Janina. 2019. The Inferential Semantics of ComicsPanels and Their Meanings. *Poetics Today* 40(2). 215–234. doi:10.1215/03335372-7298522.
- Zimmermann, Thomas Ede. 2016. Painting and opacity. In Freitag et al. (ed.), *Von Rang und Namen: Philosophical Essays in Honour of Wolfgang Spohn*, 425–453. Mentis Verlag.

# Walk-denoting music: refining music semantics

Léo Migotti<sup>1</sup> and Léo Zaradzki<sup>2</sup>

<sup>1</sup> Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University, Paris, France

<sup>2</sup> Laboratoire de Linguistique Formelle, Université Paris Diderot, Paris, France

## Abstract

Music has recently been argued to have a referential semantics (Schlenker (a,b)), *i.e.* to trigger inferences about an extra-musical reality. In this view, because the set of possible denotations is often very large, the meaning of music is often very abstract. Here we consider a very particular kind of musical sequences, which we call walk-denoting as they strongly evoke walking-situations — namely situations in which at least one character is walking. We show that the current model for music semantics is doubly insufficient. First, it makes wrong predictions with respect to the considered musical snippets. Using the method of minimal pairs, we come up with an enhanced model that accounts for inferential differences that the previous one left aside. Second, it relies on the non-trivial assumption that all notes are interpreted as events, while alternative theories seem to be just as plausible. Because a rewriting of our prototypical musical snippet adding a quaver did not seem to affect the denotation, the possibility that some musical events denote nothing needs to be investigated. Finally, we sketch the overall theoretical landscape through two main theories, which either consider that all musical events are interpreted, or that some of them might not be.

## 1 Introduction

Recent investigations about the application of formal linguistics methods to non-linguistic objects such as music strongly suggest that music can convey information about the world through semantic rules that bridge the characteristics of music and the ones of what it can evoke or represent. While evidence has been provided regarding the interpretation of some musical features from a purely semantic perspective (Schlenker (a,b)), little attention has been given to the systematic link that exists between the internal structure of music, *i.e.* its syntax, and the information it conveys, *i.e.* its semantics. Yet, we know that both music and the situations evoked can be represented hierarchically (Jackendoff (2009); Schlenker (b)). It then seems to make much intuitive sense to posit that if music has a semantics, its syntax, in relation with that of the denoted situation, has to play a role as well. We first present Schlenker's theory of musical semantics. We discuss a case-study about musical snippets evoking walking-situations, and we then highlight some of the limitations of Schlenker's model. We argue that Schlenker's theory lacks conditions on the rules linking music and situations structures. We finally present two possible theoretical accounts of this relationship. We will keep the choice to be made for further theoretical and experimental research.<sup>1</sup>

---

<sup>1</sup>We would like to deeply thank Philippe Schlenker for providing us with reliable introspective judgments as well as crucial theoretical insights, and for reviewing this article. We also thank Emmanuel Chemla who helped us clarifying our theoretical hypotheses. Thank you also to all our informants for taking the time to describe the detailed spontaneous inferences they drew from our musical stimuli.

## 2 A formal model for music semantics

Because music can evoke or make us think about certain events, be they real or not, and describe some situations better than others, Schlenker (a,b) argued that it must have a semantics. Indeed, music does not only convey information about its form and its internal structure, it also conveys information about an extra-musical reality: some music evoke sad or happy situations, others might well describe a landscape or an animal. The set of all the hypothetical situations a music can appropriately describe is therefore taken to be its meaning. The following section draws from Schlenker (a,b) and presents the concepts, terminology and notation that are needed for our theoretical proposals in section 4.

The first core idea we rely on is that music is able to convey information about the world because certain musical parameters such as timber, pitch, loudness or harmonic stability (among many others) are semantically interpreted: each of them bears some of the meaning of music (Schlenker (a,b)). For instance, pitch might well provide information regarding the size of a character involved in a scene that is depicted by the music. From now on we will talk about *inferences* to refer to this information music provides and listeners get. Also, we will call what these inferences are about *virtual sources*<sup>2</sup>.

Because music shares many features with sound, but cannot be reduced to it, the musical parameters responsible for the triggering of inferences can be split into two main categories. The first category gathers the parameters music and sound have in common, which makes it possible to derive semantic rules from normal auditory cognition. Loudness is of this sort: as any sound in nature has a certain level of loudness, we know from our world experience that loudness can be linked to certain properties of the actual source of the sound, arguably either its distance to the listener, or its level of energy. Applying this to music, we get a semantic rule on loudness interpretation, according to which the loudness of any musical event is ambiguous and either interpreted in terms of distance or energy of the *virtual* source. The second category gathers parameters that have no trivial counterpart in the non-musical world, and are intrinsically linked to tonal and harmonic properties of music. For instance, the harmonic stability of a given chord in a given key follows from tonal rules. As we do not experience harmonic stability of non-musical events, we cannot derive a semantic rule from auditory cognition; rather, Schlenker argues that this parameter is interpreted as the *actual* stability of the source, or that of the emotional state in which the listener is put.

The second core idea is that our music semantics needs to state rules linking the musical parameters to their semantic interpretation. From now on, we will use ‘*musical event*’ to refer to any note or chord, and ‘*denoted situation*’ to refer to any complex situation pertaining to the set of situations a musical snippet can evoke. Formally, we define, just as Schlenker, a musical snippet as an  $n$ -tuple  $M = (m_1, \dots, m_n)$ , and a possible situation  $S$  as an  $n$ -tuple  $S = (e_1, \dots, e_n)$ , where  $(m_1, \dots, m_n)$  is the succession of notes, each of which represents the corresponding event carrying the same index in the situation.

In formal linguistics, the common view is that the meaning of a sentence is the set of all situations of which the sentence is true. Transposing this to music, the meaning of  $M$  is the set of all situations which  $M$  is true of. We thus need a notion of musical truth. We say that  $M$  *denotes*  $S$  ( $M \models S$ ) if  $S$  is one — among many others — possible denotations for  $M$ . The final step is therefore to find rules to compute the truth-value of a music, given a specific situation it might denote. Schlenker posits that those rules are order-preservation rules: for  $M$  to denote

---

<sup>2</sup>Their naming must not, however, confuse us about their nature: virtual sources are not actual sources of the sound; they are virtual objects that may or may not produce sounds — if they do, the music does not even need to match this sounds — involved in the denotation of music.

$S$ , musical parameters involved in each event in  $M$  need to be ordered in the same way as their interpretation in the denotation. For instance, loudness levels and corresponding levels of energy or distances from the listener must be ordered in the same way.

Although the above model makes clear intuitive sense, we argue that it makes wrong predictions regarding the possible denotations of some musical snippets. Specifically, we claim that:

1. It makes incorrect predictions regarding the possible denotations of some specific musical snippets we think trigger strong inferences about a virtual source walking, as shown in next section.
2. It relies on the assumption that each musical event is systematically interpreted, regardless of its structural role, while it seems reasonable to posit that some musical events are more important than others.

### 3 Walk-denoting music and walking-situations

As stated, the above theory fails to account for some strong inferences we believe to be triggered by the prototypical musical snippet about walking-events in Figure 1<sup>3</sup>.

A walking-situation is a situation in which at least one of the virtual sources is walking. A walk-denoting music is a musical snippet that can denote a walking-situation. For levels of stability to match, and a music to denote a walk, we thus need to have a musical event to walking-event matching as shown in Figure 1: bass notes represent footsteps, while second and fourth chords represent the ‘bounces’ occurring during the transition from one foot to the other.

Based on our own introspective judgments, as well as on that of informants, we argue that the music contained in Figure 1 triggers very strong inferences about walking-situations. One might argue that listeners get these inferences because they are constantly experiencing walking-situations. This argument does not explain, however, the existence of contrastive judgments between musical snippets which, based on Schlenker’s model, should all be able to denote a walking-situation, while the inferential judgments we got from our informants do not match the theoretical predictions from the model. Let us consider the score in Figure 2<sup>4</sup>.

In order to compute the possible denotations of this piece, we first need to understand what the meaningful parameters are, both in the music itself, and in the virtual walking-situation. Intuitively, we argue that the most prominent parameter which is involved in a walking-situation and varies throughout it is stability: each footstep appears to be a quite stable event, while the transition from one foot to the other are relatively less stable, be it only because a foot is lifted

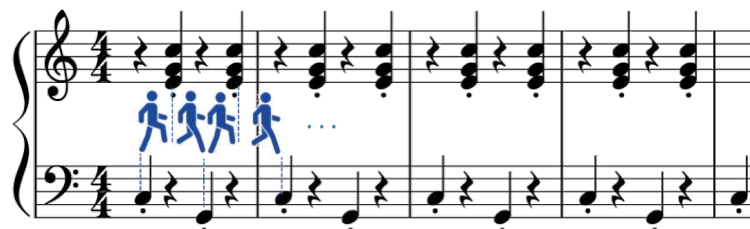


Figure 1: Prototypical walk-denoting music

<sup>3</sup>audio file: <https://www.youtube.com/watch?v=BWbqZ1BiRzI&feature=youtu.be>  
All scores are directly clickable to access the audio file.

<sup>4</sup>audio file: [https://www.youtube.com/watch?v=SvmXWX\\_xAeU&feature=youtu.be](https://www.youtube.com/watch?v=SvmXWX_xAeU&feature=youtu.be)





Figure 2: A music failing to denote a walking-situation



Figure 3: Violating alternation condition

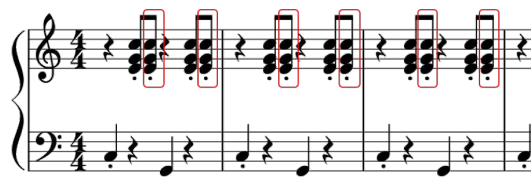


Figure 4: Prototype with extra quavers

in the air. Thus, the corresponding musical parameter *in music* must be harmonic stability. Based on rules of preservation of ordering, the theory predicts that music in Figure 2 must be able to denote a walking-situation, which it does not, according to our introspective judgments. Thus, we need to refine the formal theory that made this incorrect prediction possible.

In order to do so, we varied a whole set of musical parameters once at a time, and selected the most relevant ones according to a few informants. This led us to hypothesise that, for a musical snippet to denote a walking-situation, it has to involve the steady repetition of two different chords, that are both intrinsically stable, and sufficiently close to each other in the tonal space.

A way to test whether this set of conditions is accurate was then to build minimal pairs, *i.e.* couples of stimuli made of the above prototype and a composed musical snippet based on this very prototype but violating one of the five above conditions. Our prediction is thus that violating any of the condition would trigger an inferential preference for the prototypical snippet, that satisfies all conditions<sup>5</sup>.

From a theoretical perspective, it appears that these five parameters can be classified in two groups that make cognitive sense, and that can be derived from theoretical considerations. As a walk itself is defined as the alternation of two footsteps, a first natural class of conditions follows from the fact that any music that denotes a walking-situation must also be composed of exactly two events. The second and third conditions on repetition and regularity can be derived from the same physical fact: a normal, stereotypical walk is necessarily the repetition of footsteps (which explains why the musical events must themselves be repeated), and that repetition needs to be approximately symmetric (which explains why the repetition of musical events shall never be broken and remain steady). A second class of conditions appears to be linked to the fact that the right footstep is necessarily different from the left one, but that both events are not so different and are also both rather stable, although one might be a little bit more stable than the other; thus, the corresponding two musical events must be minimally different as well.

These conditions being stated, we were however concerned with the music snippet in Figure 4<sup>6</sup>. Indeed, introspective judgments given by our informants as well as our own suggested that the denotation of this rewritten version of the prototype, in which a quaver was added on each offbeat, was not affected; or that it was affected in a very subtle way, that did not correspond

<sup>5</sup>In order to check our theoretical intuitions, we are currently running an experiment which aims at checking whether these conditions actually play a role in the triggering of inferences about walking-situations in listeners, by presenting participants the minimal pairs that are available at <https://www.youtube.com/watch?v=04Puddu3wXQ&feature=youtu.be>.

<sup>6</sup>audio file: <https://www.youtube.com/watch?v=gR-wHm4nFYk&feature=youtu.be>

to a radical change in the denotation or a new event. We thus wondered how this could be the case, from a theoretical perspective.

## 4 Theoretical refinements

In addition to showing that Schlenker’s interpretive rules are incorrect or at least underspecified, walk-denoting music examples raise one fundamental question. The example in Figure 4 suggests that some notes may not be interpreted as concrete events in the denoted situation<sup>7</sup>. If this is so, then we have to answer two new questions:

1. If a note is not interpreted as a concrete event, what is its semantic role?
2. How do we determine which notes must be interpreted and which may remain uninterpreted?

In the rest of the article, we discuss theoretical issues related to these questions and present competing theories of musical meaning.

### 4.1 Hierarchical structures

We know from Lerdahl and Jackendoff (1983) that there are ways to assign hierarchical structures — *e.g.* tree-like structures — to musical pieces. Many different views address the question of which structure is the best to account for musical hierarchy and dependencies, such as time-span reduction or prolongational reduction from Lerdahl and Jackendoff (1983). From a different perspective, Harasim, Rohrmeier and O’Donnell argue that the internal and harmonic structure of any given tonal piece can be accounted for with the three notions of prolongation, preparation and substitution (Harasim et al. (2018)). On the other hand, it has been proposed that events too feature a hierarchical structure that can be represented as a tree (Jackendoff (2009)).

In a famous series of experiments where they asked subjects to segment taped common situations in sub-units, Zacks, Tversky and Iyer showed that people conceive situations as partonomic hierarchies (Zacks et al. (2001)). In particular, people form mental groups of events that recursively embed into one another, often in a goal-directed fashion. On the musical side, Lerdahl and Jackendoff gave formal rules derived from Gestalt principles to determine a so-called *grouping structure* for musical pieces. This grouping structure is exactly of the same partonomic kind as those evidenced by Zacks and colleagues. For example, Figure 5 (taken from Lerdahl and Jackendoff (1983)) shows the grouping structure that the system derives, in accordance with



Figure 5: A possible grouping structure for the opening theme of Mozart’s 40<sup>th</sup> Symphony.

<sup>7</sup>While this is a constructed example, examples exist in music. To give only a famous one, in the opening of *Peter and the Wolf* by S. Prokofiev, after the first occurrence of Peter’s theme in first violins, it is played again by the seconds, while the firsts play high-pitched offbeats. We believe that it does not add any event to the denoted situation (*e.g.* Peter gamboling through the meadows) but only gives it a more carefree character. The example can be heard here: [https://www.youtube.com/watch?v=uwKgh8QH\\_mc&t=2m39s](https://www.youtube.com/watch?v=uwKgh8QH_mc&t=2m39s).

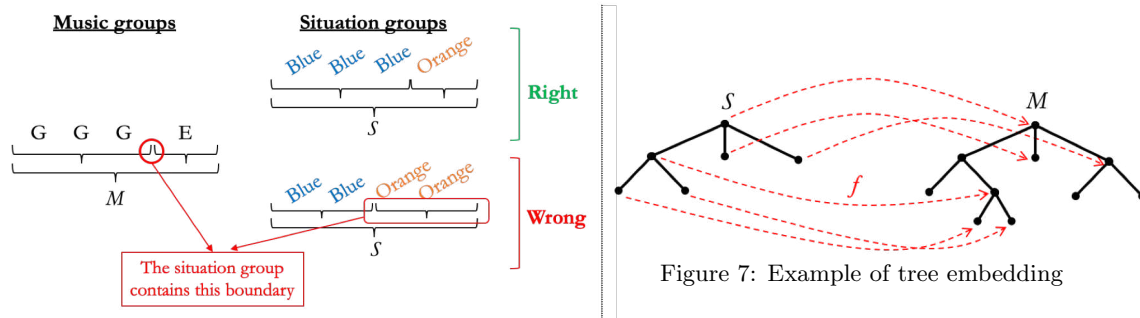


Figure 6: Example of grouping mismatch.

listeners' common intuitions, for the opening theme of Mozart's 40<sup>th</sup> Symphony. As shown by the curly brackets, the first three notes form a first group together, and the next three form another group. These two groups are then grouped together at a higher level to form a new group, and so on.

Elaborating on ideas already suggested in Schlenker's work, we posit that for a musical snippet  $M$  to denote a situation  $S$ , the grouping structures of both must match, in the following sense: a group on the situation side must not contain a group boundary on the musical side. More precisely, if  $e_1, \dots, e_k$  and  $a$  are events of  $S$  associated in  $M$  with musical events  $m_1, \dots, m_k$  and  $b$ , and if there is a group to which  $e_1, \dots, e_k$  belong but  $a$  does not belong, then there must be no musical group to which  $b$  together with some but not all the  $m_i$  belongs. As an illustration of this phenomenon, let us take the opening motif of Beethoven's 5<sup>th</sup> Symphony. This consists in four notes: G G G E. According to Lerdahl and Jackendoff's rules<sup>8</sup> they are grouped as  $[[G, G, G], E]$ . In the corresponding section of *Fantasia 2000*<sup>9</sup> each note is illustrated by a coloured lightning: the Gs are interpreted as blue lightnings, and the E as an orange one. It seems to us, though, that image and sound fit far less well if we colour, say, the third lightning in orange too. As shown in Figure 6, this is because the orange group would then contain the G/E boundary.

While we believe that preserving the hierarchical structures is a rather natural requirement for the interpretive rules, we will not argue further here, and leave this study for future research. We here assume the following strong version of preservation. First, we posit that musical snippets are associated with a hierarchical structure which is mathematically implemented as a directed rooted tree, possibly with vertices labeled as heads at each level<sup>10</sup>. Such a modeling is compatible with many approaches to musical syntax, either based on grouping structures or rather on harmony. Thus, we do not need to commit to any particular formal system here. Second, in line with Jackendoff (2009), we posit that situations too are associated with a hierarchical structure implemented as a directed rooted tree with heads. While this is a stronger assumption than what had been experimentally proved by Zacks *et al.*, we take it as a working hypothesis. Third, we claim that a necessary condition for a musical snippet  $M$  to denote a situation  $S$  is that the tree of  $S$  can be embedded into the tree of  $M$ . Formally, if  $M$  is a tree  $(V, E)$  and  $S$  is a tree  $(V', E')$ , where  $V$  and  $V'$  represent the sets of vertices and  $E$  and  $E'$  represent the sets of edges<sup>11</sup>, a necessary condition for  $M \models S$  is the existence of an injective

<sup>8</sup>and also to general Gestalt principles

<sup>9</sup>an animated film where image is intended to be a denotation of the music; see: <https://www.youtube.com/watch?v=nMnlxYkZKaU>

<sup>10</sup>We discuss this notion of heads in more details below. At first glance, let us say that heads are musical events that are more prominent or structurally more important than non-heads.

<sup>11</sup>Since the tree is directed, edges are ordered pairs of vertices.

root-preserving function  $f : V' \rightarrow V$  such that:  $\forall x, y \in V', (x, y) \in E' \Rightarrow f(x) \sim f(y)$ , where  $f(x) \sim f(y)$  means that there exists a path in  $M$  from  $f(x)$  to  $f(y)$ . What  $f$  does is that it takes any event of  $S$  and maps it onto a musical event meant to be its musical representation, in an injective way. Moreover, if an event of  $S$  is subordinated to another, then the same subordination relationship should hold between their musical representations. Figure 7 shows an example of such a function.

## 4.2 Do all notes have the same semantic status?

We now turn to the question of uninterpreted notes. The discussion from the end of section 3 suggested that all notes of a musical snippet are not necessarily interpreted as events in the denoted situation. While we do not have data to decide whether or not this can be, we will here present two competing positions about it.

In Figure 4 we saw that in the walk-denoting case, bass notes were interpreted by steps and other beats by bounces; as for the additional offbeats, it was not clear. What we can say is that most important musical events are bass notes, the other beats are musical ‘bounces’ of these bass notes, and offbeats are kind of squared bounces (bounces’ bounces). Thus, most important musical events are interpreted as important events in the situation too. This suggests that the musical salience of notes plays a role. This is why we added heads to our tree-implementations: heads are special vertices meant to represent the most salient events of a musical snippet or a situation. We will examine below alternative ways of implementing salience and how it can be computed. Now we sketch two informal opposite theories, to be refined below:

- **Theory A** (strong<sup>12</sup> theory): Every note is necessarily interpreted in the situation — though maybe by a very abstract event — that is, every note matches to an event in the situation (in this regard, this is the closest position to Schlenker’s). Moreover, more important musical events should correspond to more important events in the situation<sup>13</sup>.
- **Theory B** (weak<sup>12</sup> theory): Each note is not necessarily interpreted as an event in the situation. This does not mean that these uninterpreted musical events are semantically vacuous, but they do not refer to a concrete event, rather they modify the denoted situation as some adverb would in linguistics. Yet the musical heads always need to be interpreted — though not necessarily by head events. Moreover, more important interpreted musical events should correspond to more important events in the situation<sup>13</sup>.

As an illustration, Figure 8 shows possible trees associated with the music of Figure 4 and with the corresponding walking-situation, and an embedding of  $S$  into  $M$  as described by Theory B (for the sake of readability, we drew the arrows only for the first branch of  $S$ , but the same exist for the other branches). As one can see, not every note in  $M$  is matched to an event in  $S$ , but every head note is.

As mentioned above, both theories, though opposite, deal with the salience of notes or events. Salience can be implemented in our treeish framework by the mean of heads. These heads may be obtained by formal rules, exactly as the hierarchical structure is. However, we think that this binary notion of heads is too coarse. We may wish, for instance, to distinguish between three degrees of salience or more. We may then add, for instance, an *ad hoc* notion of secondary heads, and require them to be interpreted, but not by first-order heads, and so

<sup>12</sup>The names *weak* and *strong* theories reflect the fact that Theory A requires more constraints.

<sup>13</sup>This last idea is related to what Schlenker has suggested in Section 8.5 of Schlenker (b). There are stronger or weaker ways to express this condition; we will be more precise below. Note that it is also possible to completely drop it.

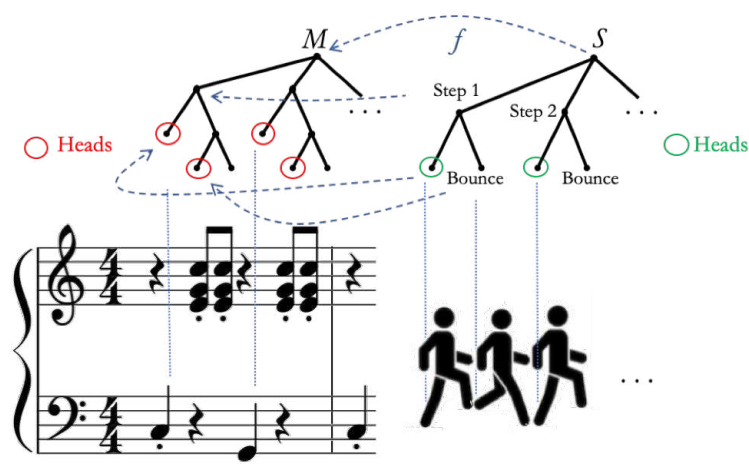


Figure 8: An illustration of Theory B on the example from Figure 4.

on. Since the number seems to be unbounded, we rather replace the notion of heads by that of *weight*<sup>14</sup>. We then suppose that  $M$  and  $S$  are rooted directed trees coming together with weight functions. A weight function on  $M$  (resp.  $S$ ) is just a function  $p : V \rightarrow \mathbb{R}_+^*$  (resp.  $p' : V' \rightarrow \mathbb{R}_+^*$ ). How these are obtained may rest on formal rules akin to those developed in [Lerdahl and Jackendoff \(1983\)](#), but we did not investigate it yet. We can require that  $f$  preserves weights in a sense we now make precise.

Let us begin with Theory A, which is simpler. Since every note is interpreted, we can associate every vertex  $x \in V$  with a vertex  $g(x) \in V'$  such that  $f(g(x)) = x$  ( $g$  is the reverse function of  $f$ ). We then require, for  $M$  to be true of  $S$ , that  $\forall x, y \in V$ ,  $p(x) \leq p(y) \Rightarrow p'(g(x)) \leq p'(g(y))$ . This ensures that more important musical events will be interpreted as more important events in the situation.

Things are bit more complex with Theory B, where all notes need not to be interpreted (whence the function  $g$  doesn't necessarily exist). Since heads are now implemented as weighted events, we will require that weighted musical events are interpreted. This can be done by requiring that if a vertex  $y \in V$  has greater weight than another vertex  $x \in V$  and if  $x$  is interpreted, then  $y$  is interpreted too (and by a more weighted vertex in  $V'$ ). Nevertheless, we think that this global condition is too strong and should be local: it may happen that some part of the musical snippet describes the situation in a very fine-grained fashion (*i.e.* almost every note is interpreted) while another does it in a more coarse-grained fashion, and yet the former's notes do not have bigger weights than the latter's. Local conditions in trees have been formalised in linguistics through the notion of *c-command*<sup>15</sup>. Formally, our localised condition becomes:  $\forall y \in V$ ,  $(\exists x \in \text{Im}f \text{ s.t. } y \text{ c-commands } x \text{ and } p(y) \geq p(x)) \Rightarrow y \in \text{Im}f$ . We also add the weight-preservation condition as in the other theory.

An alternative notion to weight would be that of reduction<sup>16</sup>. Intuitively, a musical passage can be reduced to another if it is heard as an elaboration of it — *e.g.* an ornamented version. The point is that if the tree-structure of a musical snippet encodes its reduction steps — as it

<sup>14</sup>though the same intuition is behind

<sup>15</sup>Abbreviation for constituent-command. Here we say that a node  $x$  c-commands a node  $y$  if it dominates it or is a sister of one of  $y$ 's ancestors.

<sup>16</sup>At least two notions of reduction are discussed in [Lerdahl and Jackendoff \(1983\)](#): time-span reduction and prolongational reduction. For more details, see Chapter 5 and following.

is the case of the structures that [Lerdahl and Jackendoff \(1983\)](#) deal with<sup>17</sup> — weight functions may be redundant because it seems that low-weighted musical events are those which disappear after a few reduction steps. One advantage of replacing weight by reduction steps is that it now comes along with the structure and does not need to be computed separately. One drawback is that we lose the local character of weight functions, because every local branch will now reduce at the same speed<sup>18</sup>. We leave a closer investigation of these theoretical possibilities of implementing heads for future research.

### 4.3 The meaning of uninterpreted notes

Assuming now that some notes remain uninterpreted, as in Theory B, what would be their semantic role? Let us give a few clues.

We saw in the case of [Figure 4](#) that the added notes change the *character* of the walking-situation, but do not seem to add extra events. There could be several variants of such a phenomenon, regarding how these extra notes affect the semantics of the whole.

According to one variant<sup>19</sup>, each extra note modifies the meaning of its local branch. For instance it could be that the relevant reduction level in the case of [Figure 4](#) is the beat level, so that each beat is viewed as a semantic atom, packaging all the information of its musical sub-events. The second and fourth beats of each bar thus have a semantics computed from the two quavers it is made of, and this may indicate that it is, for example, a particularly supple bounce. That is, each bounce is further characterised by this extra note. According to an alternative variant, extra notes are first ignored, leading to the same denotation set as with the simple snippet from [Figure 1](#). Only in a second stage will the extra notes add the inference that the walking-situation is more bounced or more energetic; or it will give the listener clues about the mental state of the walking character (according to our informants, he is cheerful and happy, or even wanting to dance). Using a (possibly dubious) analogy from language, the former variant predicts that low-weighted notes behave more like adjectives modifying a noun phrase (here: a musical event), while the latter predicts they behave more like adverbs, modifying the whole sentence. As we have no clue for favoring one over the other — leaving aside the fact that both could partly hold together — we will not say more than just the discussion of this example, and leave these issues for future research.

### 4.4 Is every event musically represented?

As a reverse question, we may ask if every event in the situation should be represented by a note in  $M$ . This seems to be hardly the case, since a situation can never be exhaustively described even with language.<sup>20</sup> This seems to invalidate our theory, since we required that the situation tree is ‘contained’ in the musical tree, which is much smaller. However, we can get out of this problem by considering *reductions* of the situation tree<sup>21</sup> or by saying that we will assimilate the situation with what is relevant in its perception by a given agent. We will then ask that

---

<sup>17</sup>The trees here are not directed rooted trees as generally understood in mathematics, but rather something equivalent to directed rooted trees with heads, though a bit more complex.

<sup>18</sup>Typically, with time-span reduction, each group of two quavers will be at some point replaced by a single crotchet. But it could well be that in some branch each quaver is interpreted, while in some other only one of them is.

<sup>19</sup>Thanks to Philippe Schlenker for suggesting this variant.

<sup>20</sup>Think, for example, of the whole subtlety in the gestures of a character, or of all the particle interactions that take place everywhere.

<sup>21</sup>Just like there are notions of reduction for music, we can posit that the same things exist with situations, as suggested in [Jackendoff \(2009\)](#).



every head event in the situation is represented by a head event in the musical snippet.

All this being said, let us state as a summary a final formal formulation of one variant of Theory B. Theory A could be straightforwardly formulated in a very similar way.

**Theory B.** *Let  $M = (V, E, p)$  be a musical snippet and  $S$  a situation.  $M \models S$  if, and only if,  $M \models S$  in Schlenker's sense, and moreover there exists a reduction  $S' = (V', E', p')$  of  $S$ , and an injective root-preserving function  $f : V' \rightarrow V$  such that:*

1.  $\forall x, y \in V', (x, y) \in E' \Rightarrow f(x) \sim f(y)$
2.  $\forall x, y \in V', p'(x) \leq p'(y) \Rightarrow p(f(x)) \leq p(f(y))$
3.  $\forall y \in V, (\exists x \in \text{Im}f \text{ s.t. } y \text{ c-commands } x \text{ and } p(y) \geq p(x)) \Rightarrow y \in \text{Im}f$

## 5 Conclusion

The investigation of walking-situations and walk-denoting music enlightened the necessity to come up with a theory of musical events. This theory needs to provide rules involving the structural role of each musical event determined through the rigorous analysis of musical structures, and explain how this impacts the very possibility for each event to be interpreted, *i.e.* to have a counterpart in the denotation. Besides, our goal was to account for how this interpretive semantic mechanism works (*i.e.* what happens to the musical structure when interpreted, and how the formal tree-like structure of the music is related to the formal tree-like structure of the events it denotes). Further research will investigate the experimental extensions of this theoretical work, in order to check whether it has some cognitive reality in listeners.

## References

- Daniel Harasim, Martin Rohrmeier, and Timothy J. O'Donnell. A generalized parsing framework for generative models of harmonic syntax. In *Proceedings of the 19th International Society for Music Information Retrieval Conference, ISMIR 2018, Paris, France, September 23-27, 2018*, pages 152–159, 2018.
- Ray Jackendoff. Parallels and nonparallels between language and music. 26(3):195–204, 2009. ISSN 0730-7829. doi: 10.1525/mp.2009.26.3.195.
- Fred Lerdahl and Ray Jackendoff. *A generative theory of tonal music*. The MIT Press, Cambridge, MA, 1983. ISBN 0262120941.
- Philippe Schlenker. Outline of music semantics. 35(1):3–37, a. ISSN 0730-7829, 1533-8312. doi: 10.1525/mp.2017.35.1.3. URL <http://mp.ucpress.edu/lookup/doi/10.1525/mp.2017.35.1.3>.
- Philippe Schlenker. Prolegomena to music semantics. 10(1):35–111, b. ISSN 1878-5158, 1878-5166. doi: 10.1007/s13164-018-0384-5. URL <http://link.springer.com/10.1007/s13164-018-0384-5>.
- Jeffrey Zacks, Barbara Tversky, and Gowri Iyer. Perceiving, remembering, and communicating structure in events. *Journal of experimental psychology. General*, 130:29–58, 04 2001. doi: 10.1037//0096-3445.130.1.29.

# Joker face. Recognizing irony in the visual mode in spoken and signed language

Francesca Panzeri<sup>1</sup>, Beatrice Giustolisi<sup>1</sup>, Lara Mantovan<sup>2</sup> and Chiara Calderone<sup>2</sup>

<sup>1</sup>University of Milan-Bicocca, Italy

<sup>2</sup>Ca' Foscari University of Venice, Italy

francesca.panzeri@unimib.it, beatrice.giustolisi@unimib.it,  
lara.mantovan@unive.it, chiara.calderone@unive.it

## Abstract

A remark such as “The party was really funny!” can be interpreted literally, as a praise, or ironically, as a criticism. To avoid misunderstandings, the ironist can pronounce the comment with a particular intonational contour, the so called ironic tone of voice, and accompany it with specific facial expressions. Also in Italian Sign Language (LIS) non-manual markers can signal the ironic intent of the signer, and thus facilitate the detection of the intended meaning. We report the results of two experimental studies that aim at identifying the role played by purely visual cues in the detection of irony in spoken Italian, and at verifying whether ironic remarks in LIS are recognizable also by those who are not competent in that language. We discuss the findings in the light of the debate concerning the affective or grammatical status of facial expressions, in spoken and signed languages.

## 1 Introduction

When people communicate, they rely on verbal and nonverbal cues, that is, acoustical and visual markers, to transmit and interpret the message. Prosodic features may encode linguistic or affective information (but see Seddoh, 2002 for a critique of this dichotomy). As for linguistic prosody, different speech acts, such as statements, questions, and orders, can be disambiguated through specific intonation, and the interpretation of focused structures relies on the placement of stress on particular speech units (see Prieto, 2015 and references therein). In these cases, the contribution of prosody is viewed as linguistic because it has an effect on the interpretation of statements. Speakers may also use specific acoustic profiles to signal their own emotional attitude (e.g., anger, joy, contempt) towards a proposition, with prosody transmitting purely affective information (Hammerschmidt & Jürgens, 2007). Visual markers, that is facial expressions and bodily movements, are less studied. Humans have evolved an ability to infer from the speaker’s facial expressions her emotions and other socially relevant states (McCullough et al., 2005 and references therein); still, it is not clear whether, in spoken languages, facial expressions transmit only the speaker’s attitude, or whether they can also serve linguistic functions.

Specific facial expressions typically accompany the intonational contour associated to questions, orders, and to the marking of prominent or new information; it is still debated whether visual cues are simply parasitic on acoustical data, offering weak and redundant information with respect to the stronger auditory signals (as suggested by Scarborough et al., 2009; Krahmer et al., 2002; Srinivasan & Massaro, 2003), or whether they can provide independent evidence for the interpretation of the sentence. Recent studies found that hearers do rely on visual information to disambiguate between contrastive focus statements and echo questions (Borràs-Comes & Prieto, 2011), between double or simple negation interpretations (Prieto et al., 2013), and to identify the illocutionary force of a speech act (Domaneschi et al., 2017).



In signed languages, facial expressions can encode both linguistic and affective information, and they have thus been considered the analogue of intonation (Sandler & Lillo-Martin, 2006). The so-called non-manual markers (NMMs, facial expressions, head and body movements) may signal different lexical and syntactic structures, such as relative clauses, questions, conditionals, adverbials, and topics (Pfau & Quer, 2010). It is noteworthy that the very same facial movements can be used to express both linguistic/grammatical meanings and paralinguistic/affective information: furrowed eyebrows, for instance, signal content questions in different sign languages, but they are also associated to the expression of anger in general human communication (de Vos et al., 2009). Once that specific facial expressions have been related to linguistic constructions in a signed language (e.g., eyes wide in commands, head nods and tilts in advices for American Sign Language, Brentari et al., 2018), it can be hard to recognize whether they solve an affective or linguistic function.

Different diagnostics have been proposed to distinguish the grammatical from the affective status of NMMs: linguistic NMMs are more conventionalized, and exhibit less individual variation (Brentari et al., 2018); they are less gradient, because they give rise to a categorical interpretation (a sentence is either a question or not, whereas a person can be angry at various degrees, de Voos et al., 2009); they exhibit an abrupt onset and offset (Baker-Shenk, 1983); and they are aligned with manual signs and thus have a systematic distribution (Dachkovsky, 2005, de Vos et al., 2009). Besides these articulatory features, linguistic and affective NMMs differ in that the former, by definition, concern the interpretation of sentences, and thus have an influence on their acceptability, and are required by grammar (Reilly & Bellugi, 1996). Focusing on their accessibility, affective NMMs are potentially recognizable also by non-signers, since at least some basic facial expressions are claimed to be universal or at least cross-culturally identifiable (Darwin, 1872; Ekman, 1973/2006), whereas grammatical NMMs are claimed to be more arbitrary and thus inaccessible to non-signers (Hermann & Pendzich, 2014), with some exceptions (for instance, head nods for assertion and headshake for negation).

In the present work, we aimed at further exploring the questions of the role played by visual cues in disambiguating ironic from sincere statements, in spoken and in signed languages. A sentence like “What a great job” can be uttered sincerely, and thus be interpreted literally, or ironically, with the speaker signaling that she does not believe that the job was well done. To avoid misunderstandings, speakers can make their ironic intent manifest using irony markers (Attardo et al., 2003). Besides choosing particular linguistic expressions, such as extreme adjectives, or rhetorical questions, the ironist can pronounce the comment with the so-called ironic tone of voice (Cheang & Pell, 2008), and accompany it with particular facial expressions (e.g., winks, nudges, Muecke, 1978). Acoustical and visual cues can then help the hearers recognizing the intended meaning, disambiguating thus between a literal and an ironic interpretation. It has been found that hearers can correctly recognize ironic remarks, even in the absence of a preceding context, relying only on acoustical cues (Bryant & Fox Tree, 2002). To our knowledge, the contribution provided by only corporal movements has not been tested yet.

In a preceding study (Mantovan et al., 2019), our research group investigated the expression of irony in Italian Sign Language (LIS): having compared minimal pairs of literal compliments / ironic criticisms (such as “That’s very beautiful!”) and of literal compliments / ironic criticisms (such as “What an awful house!”), we noticed that specific NMMs signaled the signer’s meaning (with more raised eyebrows, head nods and lateral tilts in ironic remarks than in their literal counterparts) and also the signer’s attitude (with mouth corners up accompanying both literal and ironic compliments, and corners down co-occurring with sincere and sarcastic criticisms). Nevertheless, the linguistic or affective status

of these markers was not questioned, nor it was tested whether the presence of these markers could effectively facilitate the recognition of irony.

## 2 The studies

We discussed how in spoken languages ironic remarks are typically accompanied by a specific intonation and by particular facial expressions. Interlocutors are able to recognize ironic remarks relying only on the ironic tone of voice, whereas the contribution of purely visual information in the disambiguation of sincere/ironic comments have not been tested yet. Our first study aims at filling this gap, by comparing the rate of recognition of ironic statements in three different conditions: when interlocutors can rely on both visual and acoustical information; when they can only hear the remark; and when they can only see the facial expression and bodily movements of the speaker. Our goal is to determine whether irony can be detected with purely visual information.

As for signed languages, in LIS ironic remarks tend to be marked with particular NMMs. As we saw in the previous section, NMMs can serve a linguistic or an affective function. One of the diagnostics criteria to distinguish these roles questioned their accessibility: grammatical NMMs can be considered as more arbitrary, and thus recognizable at a higher rate by signers; affective NMMs on the other hand express the signer's emotional state, and thus they should be recognizable independently from linguistic knowledge, at least within the same cultural group. The second study aims at gathering some evidence in favor, or against, the grammatical status of ironic NMMs in LIS: a group of hearing participants who know LIS (yes-LIS), and another group who does not know LIS (no-LIS), were asked to detect ironic comments, in LIS and in Italian; we compared their rate of recognition of irony in the two modalities. We hypothesized that if ironic NMMs in LIS had a linguistic status, they would facilitate the recognition of irony in the yes-LIS group more than in the no-LIS group.

### 2.1 Methods

We used a Discourse Completion Task (Félix-Brasdefer, 2010) to obtain a semi-spontaneous elicitation of five minimal pairs of remarks by two groups of Italian adults: four deaf native signers and four Italian speakers who were videotaped while uttering the remarks. The same remark (e.g., “The party was really funny!”) was elicited in one session after a context that triggered its literal interpretation and, in a different session, that took place at least two weeks after the first session, after a context that supported its ironic interpretation. All the remarks had a literal positive interpretation, and thus they corresponded to compliments in their sincere reading, and to ironic criticisms in their ironic interpretation (see Mantovan et al., 2019 for further details on the elicitation procedure for native signers: the same procedure was used also for Italian speakers). The 20 remarks produced by Italian speakers were then edited to obtain three versions: the original version with video and audio (VA), one with only the video (V), and one with only the audio (A). With this material, we prepared two web-based experiments.

Study 1 focused on spoken Italian only. Participants were shown the two utterances of the same remark, and were asked to indicate when the speaker was being ironical. The 20 remarks (the 5 minimal pairs of literal compliments / ironic criticisms uttered by four Italian native speakers) were presented in three different conditions: the first block comprised only pairs of videos, without audio (with the remark written on the screen); the second block only pairs of audio tracks; the final block displayed pairs of videos with audio. Every participant, then, was presented with a total of 60 items (20 pairs for each condition).

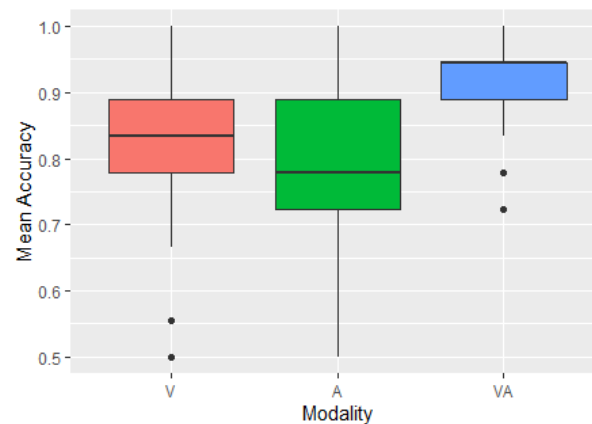
Study 2 included instances of ironic remarks in the two modalities, spoken Italian, and LIS. For this study, we only presented pairs of videos (without audio, but again with the uttered remark written on the screen) of the 8 actors (four native signers, and four Italian speakers) uttering the same remark in a sincere and in an ironic way, and participants had to recognize the ironical remark. The total number of items for this study was 40 (20 pairs in Italian, and 20 pairs in LIS).

The two studies were implemented on Qualtrics and distributed on-line.

The participants of Study 1 were 81 Italian adults (61 female), with a mean age of 23 years. The participants of Study 2 were 141 Italian hearing adults, 73 with no exposure to LIS (58 female, mean age: 26 years), and 68 who know LIS (63 female, mean age: 30 years).

## 2.2 Results

Since in Study 1 our goal was to compare the relative strength of visual and auditory information in the detection of irony, we first verified whether ironic remarks were indeed recognizable when both these cues were present. We thus looked at accuracy rates in the third condition, VA, and found that for two items accuracy was at chance level, and we thus eliminated these items. Moreover, since the experiment was web-based, we also wanted to exclude the possibility that participants were overall responding at chance level. We thus controlled that each participant responded correctly, always in the last VA block, to at least 12 items out of the remaining 18 (66% accuracy). Seven participants out of 81 did not meet this requirement, and were eliminated. The following analyses were then ran on 74 participants who responded to 18 items in the three different modalities. The rate of recognition of irony in the only video condition (V) was 83.71% (SD = 36.94), in the only audio condition (A) was 79.43% (SD = 40.44), and in the last condition, where both visual and auditory cues were present (VA), was 90.72% (SD = 29.02). How mean accuracy varied across modalities is depicted in Figure 1.

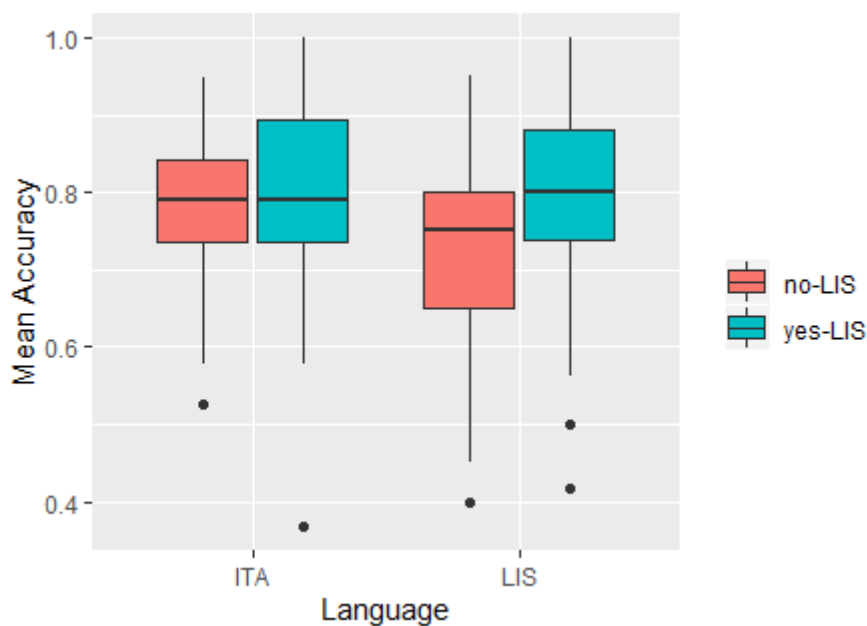


**Figure 1:** Boxplots depicting accuracy in the recognition of ironic remarks across modalities in Study 1.

Accuracy was analyzed using generalized mixed-effects models. Participant and items were entered as random intercepts in the null model. Adding modality as fixed factor significantly increased the goodness of fit of the model ( $\chi^2 = 72.41$ ,  $p > 0.001$ ). Post-hoc comparisons indicated that accuracy in the VA condition was higher than accuracy in both the V condition and the A condition (both  $p$ s  $< 0.0001$ ); and also that accuracy in the V condition was higher than accuracy in the A condition ( $p = 0.009$ ).

The aim of Study 2 was to verify whether the rate of recognition of irony markers in LIS was related to the knowledge of the language. We first checked whether the 40 critical items (5 remarks uttered by 4 Italian speakers and by 4 native LIS signers) were recognizable above chance level for at least one of the group of participants. One item had an accuracy around 35% for both groups and was thus eliminated from further analyses.

We then calculated the accuracy of irony recognition in the two languages (LIS and Italian), by the two groups of hearing participants, those who know LIS (yes-LIS), and those who do not know it (no-LIS). We found that the yes-LIS participants had an accuracy of 79.68% (SD = 40.25) in Italian, and of 79.23% (SD = 40.58) in LIS; the no-LIS group had an accuracy of 77.94% (SD = 41.48) in Italian and of 73.01% (SD = 44.40) in LIS. The results are illustrated in Figure 2.



**Figure 2:** Boxplots representing accuracy in the recognition of ironic remarks in Italian (ITA) and in LIS by hearing participants who know LIS (yes-LIS) and who do not know LIS (no-LIS) in Study 2.

Accuracy was analyzed using generalized mixed-effects models. Participant and items were entered as random intercepts, whereas language (Italian vs. LIS) and knowledge of LIS (yes-LIS vs. no-LIS) as fixed factors. The two main effects were not significant (language:  $p = 0.33$ ; knowledge of LIS:  $p = 0.43$ ), whereas the interaction was significant ( $\beta = 0.31$ ,  $SE = 0.14$ ,  $z = 2.18$ ,  $p = 0.03$ ): ironic remarks in LIS were recognized better by the hearing participants who know LIS compared to those who do not know LIS. The group (no-LIS and yes-LIS) per language (LIS and Italian) effect is plotted in Figure 3.

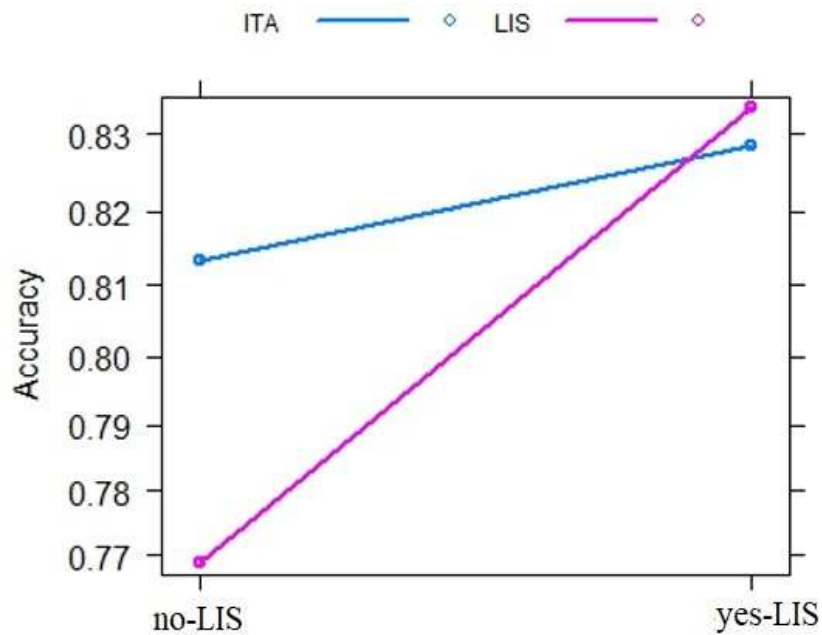


Figure 3: Participants (no-LIS and yes-LIS) by language (Italian and LIS) effect plot.

### 3 Discussion

In Study 1, we found that purely visual cues did enable interlocutors to correctly recognize ironic remarks, with an accuracy rate of 84%, well beyond chance level. Moreover, when participants had to distinguish ironic and sincere utterances of the same remark, they performed better when they were relying on purely visual cues than when their judgment was based on acoustical information. These results highlight the contribution of facial expressions and bodily movements in facilitating the correct interpretation of the intended meaning, and question the superiority of prosodic features in helping interlocutors in the interpretive process.

As for the debate concerning the affective or linguistic status of acoustical and visual cues in the detection of irony in spoken languages, our first study cannot settle the issue. Indeed, the recognition of the speaker's sincere or ironic intent has an effect on the interpretation of the sentence: the comment "The party was really funny!" constitutes a literally positive compliment under one reading, and ends up being a negative criticism on the other reading, with the speaker signaling that she does not believe that the party was funny. Traditionalist approaches to irony claim that with this type of figurative language the comment gets a reversed interpretation, and within this perspective, then, prosody and facial expressions would play a linguistic role, since they disambiguate the intended meaning. Nevertheless, both the Echoic account (Wilson & Sperber, 2012) and the Pretence account (Clark & Gerrig, 1984) to irony claim that the ironic remark is simply mentioned, preserving its literal meaning, and the ironist manifests to her interlocutors her scornful and negative attitude towards the proposition expressed, or towards the person who would hold such a view. The intonational contour would then have an affective nature, because the speaker would be

expressing acoustically her contempt towards the proposition she is echoing, or she would be exaggeratedly imitating the person she is pretending to be. Also facial expressions can be used to express the speaker's negative attitude, and thus they would only indirectly have an effect on the interpretation of the sentence.

Our second study, on the other hand, aimed at addressing the question of the affective or linguistic status of ironic NMMs in LIS. Following Brentari et al. (2018), we hypothesized that if the expression of irony were accompanied by grammaticalized NMMs, the detection of irony would be easier for those participants who know the language, and harder for those who are not familiar with it. We found that the participants who know LIS could detect ironic remark at a higher rate than those who do not know it. Still, this latter group of hearing individuals who are not acquainted with LIS could recognize irony with a 73% accuracy, which was lower than their accuracy for spoken Italian (78%), but still above chance level.

Taken together, the results of our two studies might suggest a more nuanced picture of the linguistic and affective status of prosody and facial expressions, both in spoken and in signed languages. Crasborn (2006) explicitly claimed, for signed languages, that "It is a commonly accepted intuition that linguistic non-manual signals have grammaticalized from emotional facial expressions, but few studies have explicitly addressed this hypothesis". Along the same lines, Benitez-Quiroz and colleagues (2016) argued that the facial expressions of negative moral judgment (anger, disgust and contempt) have evolved into a facial expression of ("linguistic") negation, which has then been grammaticalized as a non-manual marker in American Sign Language. This line of reasoning hypothesizes that some components of human languages, at least those related to prosody in spoken and signed language (i.e., NMMs), have evolved from facial expressions that encode the speaker/signer's emotions. If this is correct, then it should not be surprising that even facial expressions, and not only prosody, permit the detection of the speaker's ironic intent in spoken languages, and that even non-signers can recognize the expression of irony in LIS above chance level.

### 3.1 Limits and future directions

The studies we presented have important limitations that need to be acknowledged. First of all, participants were presented with two different utterances of the same remark, one sincere and one ironic, and were asked to recognize the sarcastic one. This is not an ecological setting, and the high accuracy in this task could come from a direct comparison of the two utterances. We are currently running a follow up experiment, in which the ironic and sincere remarks are presented one at a time, again in the three modalities (only video, only audio, and with both visual and acoustical cues). With this new design, we expect the overall accuracy to be lower, and that the contribution of purely acoustical information could override the role of visual cues, since the ironic intonation could be more conventionalized than facial expressions in the manifestation of irony.

Moreover, in our studies, we presented only remarks that were literally positive (such as "The party was really funny!"), and thus constituted compliments when the speaker was sincere, and criticisms when she was being ironical. As a consequence of this, we cannot be sure whether our participants were recognizing the speaker's ironical meaning or her scornful attitude, as manifested by the expression of her negative emotional state. It would be interesting to test also literally negative remarks, such as "This house is really awful!", that amount to criticisms when the speaker is sincere, and to compliments when she is being ironical. In our previous study on the expression of irony in LIS, we found that the position of the corners of the mouth was signaling the signer's attitude: when the signer had corners-up (a sort of smile), she was making a compliment, when the mouth corners were down she

was conveying a criticism. We hypothesized that these NMMs could be indirect signals of irony: when corners-up were co-occurring with a sign with a negative meaning (e.g., with AWFUL), or when corners-down accompanied a positive sign (e.g., FUNNY), the interlocutor could understand that the intended meaning should be reversed, and thus that the signer was being ironical. Testing the recognition of ironic compliments / literal criticisms in the three modalities could help us gaining a better understanding of the role played by visual and acoustical cues.

As for the second study we ran, on the recognition of irony in LIS by hearing participants who know and who do not know LIS, some weaknesses should be discussed. We found that the yes-LIS groups had an accuracy in the recognition of irony in LIS (but not in Italian) that was significantly higher than the no-LIS group, and we interpreted this result as possible evidence of the (partially) grammatical status of irony NMMs in LIS. Nevertheless, as pointed out by one reviewer, our participants were hearing persons studying LIS as second language (L2), and the mastery of linguistic NMMs is a late acquisition for L2 learners, at least for their correct production. We are currently running more fine-grained analyses that take into account the level of competence in LIS for each yes-LIS participant, to test whether knowledge of LIS is related to the accuracy of irony recognition.

Another analysis we are currently doing aims at disentangling the contribution of specific NMMs and other prosodic characteristics (such as the prolonged articulation of the evaluative sign) that are possibly related to the expression of irony in LIS. We plan to annotate their presence in each critical remark, and check whether there are cues that facilitate the detection of irony in LIS, and further verify whether these cues play a different role for the yes-LIS and the no-LIS participants.

## References

- Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor*, 16(2), 243-260.
- Baker-Shenk, C. (1983). *A Microanalysis of the Nonmanual Components of Questions in American Sign Language*. Dissertation, University of California, Berkeley, CA.
- Benitez-Quiroz, C. F., Wilbur, R. B., & Martinez, A. M. (2016). The not face: A grammaticalization of facial expressions of emotion. *Cognition*, 150, 77-84.
- Borràs-Comes, J., & Prieto, P. (2011). 'Seeing tunes.' The role of visual gestures in tune interpretation. *Laboratory Phonology*, 2(2), 355-380.
- Brentari, D., Falk, J., Giannakidou, A., Herrmann, A., Volk, E., & Steinbach, M. (2018). Production and comprehension of prosodic markers in sign language imperatives. *Frontiers in psychology*, 9, 770.
- Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and symbol*, 17(2), 99-119.
- Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech communication*, 50(5), 366-381.
- Clark, H.H. & Gerrig, R.J. (1984). On the Pretense Theory of Irony. *Journal of Experimental Psychology*, 1, 121-126.
- Crasborn, O.A. (2006). Nonmanual structures in sign language. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics (2<sup>nd</sup> edition)*, Oxford: Elsevier, 668-672.
- Dachkovsky, S. (2005) *Facial expression as intonation in ISL: The case of conditionals*. Dissertation, University of Haifa.
- Darwin, C. (1872). *The expression of emotions in animals and man*. London: Murray.
- De Vos, C., Van der Kooij, E., & Crasborn, O. (2009). Mixed signals: Combining linguistic and affective functions of eyebrows in questions in Sign Language of the Netherlands. *Language and speech*, 52(2-3), 315-339.

- Domaneschi, F., Passarelli, M., & Chiorri, C. (2017). Facial expressions and speech acts: experimental evidences on the role of the upper face as an illocutionary force indicating device in language comprehension. *Cognitive processing*, 18(3), 285-306.
- Ekman, P. (1973/2006). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review*, Los Altos (CA): Malor Books.
- Félix-Brasdefer, J.C. (2010). Data collection methods in speech act performance: DCTs, role plays, and verbal reports. In A. Martínez-Flor, E. Uso-Juan (Eds.), *Speech Act Performance: Theoretical, Empirical, and Methodological Issues*, Amsterdam/Philadelphia: John Benjamins, 41-56.
- Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice*, 21(5), 531-540.
- Herrmann, A. & Pendzich, N.K. (2014). Nonmanual Gestures in Sign Languages. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill & J. Bressemer (Eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 2, Berlin: de Gruyter Mouton, 2147–2160.
- Krahmer, E., Ruttkay, Z., Swerts, M., & Wesselink, W. (2002). Pitch, eyebrows and the perception of focus, in *Speech Prosody 2002 Conference*.
- Mantovan, L., Giustolisi, B., & Panzeri, F. (2019). Signing something while meaning its opposite: The expression of irony in Italian Sign Language (LIS). *Journal of Pragmatics*, 142, 47-61.
- McCullough, S., Emmorey, K., & Sereno, M. (2005). Neural organization for recognition of grammatical and emotional facial expressions in deaf ASL signers and hearing nonsigners. *Cognitive Brain Research*, 22(2), 193-203.
- Muecke, D. C. (1978). Irony markers. *Poetics*, 7(4), 363-375.
- Pfau, R., & Quer, J. (2010). Nonmanuals: Their grammatical and prosodic roles. In D. Brentari (Ed.), *Sign languages*, Cambridge: Cambridge UP.
- Prieto, P. (2015). Intonational meaning. *Cognitive Science*, 6(4), 371-381.
- Prieto, P., Borràs-Comes, J., Tubau, S., & Espinal, M. T. (2013). Prosody and gesture constrain the interpretation of double negation. *Lingua*, 131, 136-150.
- Reilly, J. S., & Bellugi, U. (1996). Competition on the face: Affect and language in ASL motherese. *Journal of child language*, 23(1), 219-239.
- Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*, Cambridge: Cambridge UP.
- Seddoh, S. A. (2002). How discrete or independent are “affective prosody” and “linguistic prosody”? *Aphasiology*, 16(7), 683-692.
- Scarborough, R., Keating, P., Mattys, S. L., Cho, T., & Alwan, A. (2009). Optical phonetics and visual perception of lexical and phrasal stress in English. *Language and Speech*, 52(2-3), 135-175.
- Srinivasan, R. J., & Massaro, D. W. (2003). Perceiving prosody from the face and voice: Distinguishing statements from echoic questions in English. *Language and Speech*, 46(1), 1-22.
- Wilson, D., & Sperber, D. (2012). *Meaning and relevance*, Cambridge: Cambridge UP.





# Indexing Across Media

Mats Rooth and Dorit Abusch

Cornell University  
Ithaca, New York, USA  
mr249@cornell.edu da45@cornell.edu

## 1 Introduction

Indexing across pictures and language is illustrated in (1). In the first example, the pronoun in the sentence somehow picks up a discourse referent set up by the picture, and the picture and the sentence jointly put constraints on the same individual in a described situation. In (1b) a nominal phrase functioning as a title or caption gives information about an individual depicted in the picture. This paper analyzes indexing in such examples, starting from a dynamic semantics for indexing in pictorial narratives. The current section reviews the semantic framework. The basic analysis of indexing across media is laid out in Section 2, using a setup involving a formal language and interpretation for it. Section 3 looks at data where in a combination of a picture and some language, the linguistic part is a nominal (such as a title) rather than a sentence. Here a constraint on interpretation is observed, which in the theory is enforced in the syntax of discourse representations. Section 4 looks at data involving definite reference and quantification. Section 5 points out purely linguistic data that are analogous to the data from Section 3. Section 6 wraps up.<sup>1</sup>

- (1) a.  b. 
- He's a sailor.                      A castle owned by a duke
- Navy sailor  
drinking coffee.  
Openclipart.
- Castle on a hill.  
Public Domain  
Vectors.

We assume the semantics for pictures and indexing in pictorial narratives employed in previous work (Abusch, 2012, 2014, to appear; Abusch & Rooth, 2017; Rooth & Abusch, 2018; Maier & Bimpikou, 2019). The framework is reviewed briefly here.<sup>2</sup> A propositional semantics for pictures is based on geometric projection. The basis is a projection function  $\pi$  that maps a world and a viewpoint to a two-dimensional picture, using a mathematical, computational, and/or physically realized procedure such as perspectival projection or orthographic projection. A viewpoint is analogous to a camera position, or the station point in the classical theory of perspective. Where  $w$  is a world and  $v$  is a viewpoint, the function value  $\pi(w, v)$  is the picture that is projected from world  $w$  as observed from viewpoint  $v$ . Propositional semantic values are then obtained by inverting projection. The propositional semantic value of a given picture  $p$  is the set of worlds that project to  $p$  via  $\pi$ . There are a handful of independent arguments for employing viewpoint-centered semantics values, which are sets of pairs of a world and a

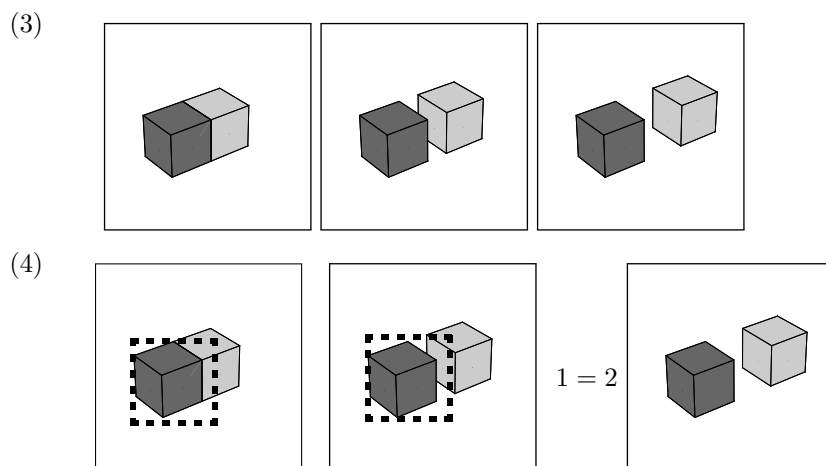
<sup>1</sup>The images in the paper that are quoted from books and other sources are used for educational and critical purposes, and are property of their respective owners.

<sup>2</sup>Abusch (to appear) is a thorough review.

viewpoint.<sup>3</sup> In this option, which is assumed here, the semantic value  $\llbracket p \rrbracket$  of picture  $p$  is the set of pairs  $\langle w, v \rangle$  such that  $w$  projects to  $p$  from viewpoint  $v$ ,  $\pi(w, v) = p$ . This is recorded in (2a).<sup>4</sup> (2b) is a variant where times are included in the model, the projection function has a time argument together with a world and a viewpoint, and the semantic value is a set of triples of a world, a time, and a viewpoint.

$$(2) \quad \text{a. } \llbracket p \rrbracket = \{ \langle w, v \rangle \mid \pi(w, v) = p \} \quad \text{b. } \llbracket p \rrbracket = \{ \langle w, t, v \rangle \mid \pi(w, t, v) = p \}$$

In the analysis of Abusch (2012, to appear) a picture or picture sequence is incremented syntactically with geometric areas, which introduce discourse referents. As an example, (3a) is a three-panel comic of two cubes moving apart. A basic semantics combines the semantics of the individual pictures with homomorphic temporal progression.<sup>5</sup> This basic semantics (in a possible-worlds model with worlds and times) does not entail that in the described situation, the cube corresponding to the gray area in the first frame of the comic is the same as the cube corresponding to the gray area in the second frame (and so on). In order to express these understood identities, Abusch (2012) suggested incrementing pictures with *areas in the picture* that introduce discourse referents. In (4), there is a bounding box around the dark area in the first picture, and similarly in the second picture. These serve to introduce discourse referents for depicted individuals, in this case a discourse referent for the cube depicted in the first picture, and another discourse referent for the cube depicted in the second picture. The identity  $1 = 2$  has the semantics of equality in the model, indicating that the individuals in the model that correspond to the two discourse referents are identical. Bounding boxes serving as proxies for individuals are used in machine learning databases and algorithms. For instance, (5) is an image from the Pascal VOC dataset with a picture of a bus and a bounding box for the bus (Everingham et al., 2012).



<sup>3</sup>These arguments include ones based on the semantics of discourse referents (as here), accounting for Necker ambiguities, and the use of perspectival phrases such as *in front of* in sentences describing pictures. See Abusch (to appear) and Rooth & Abusch (2018).

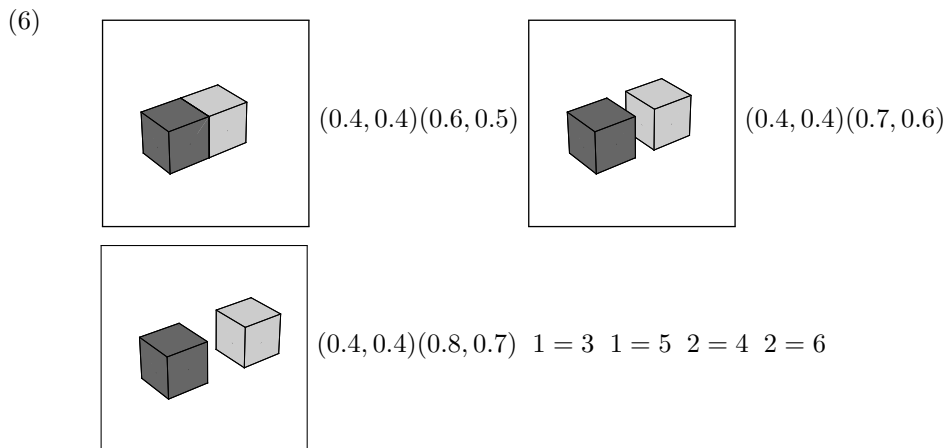
<sup>4</sup>A model  $M$  and parameters  $A$  of the projection function can be added outside the brackets,  $\llbracket p \rrbracket^{M,A} = \{ \langle w, v \rangle \mid w \in M^W \wedge \pi^A(w, v) = p \}$ .

<sup>5</sup>Abusch (2014) discusses temporal progression in visual narratives.

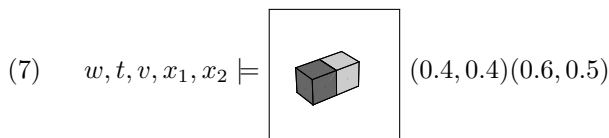


Pascal VOC dataset.

For simplicity, in this paper we use points in the area of the picture as geometric discourse referents. With the assumption that the first picture in (3) is a unit square, the pair  $(0.4, 0.4)$  measures 0.4 along the horizontal axis and 0.4 along the vertical axis to a point within the dark gray area. The pair  $(0.6, 0.5)$  measures 0.6 along the horizontal axis and 0.5 along the vertical axis to a point within the light gray area. (6) is a version of (3) that includes geometric discourse referents for all the depicted cubes. At the end there are equalities that use a recency convention. The equality  $1 = 3$  equates the most recently introduced discourse referent with the ante-penultimate one. This has the effect of equating the light cube in the final picture with the light cube in the middle picture.



Something like (6) is a formula of a formal language with a defined syntax, and a semantics that is stated in type theory and possible worlds semantics. Abusch (to appear) formalized the semantics inductively, using a format similar to (7), where a world (variable  $w$ ), a time (variable  $t$ ), a viewpoint (variable  $v$ ) and a string of individuals (here  $x_1x_2$ ) satisfy a formula.  $v$  is the viewpoint for the last picture in the formula.



The point of memorizing the viewpoint for the last picture is that this viewpoint is used in the semantics of discourse referents. Given a viewpoint  $v$  (understood as the viewpoint for the last picture), and a point  $d$ , understood as a point in the two-dimensional area of the last picture,  $v$  and  $d$  are used to pick out an object by tracing a directed line from  $v$  through the point  $d$  in the picture plane to the point where it intersects an object. An object that

witnesses the discourse referent is one that the directed line from  $v$  through  $d$  intersects before it intersects any other object. We write this condition as  $\bar{\pi}(w, t, v, d, x)$ , or when time is not being considered, as  $\bar{\pi}(w, v, d, x)$ .<sup>6</sup>

Let  $P$  be a visual narrative like (6), consisting of a sequence of pictures, with interleaved discourse referents and equalities between discourse referents. By collecting up the tuples that satisfy  $P$ , we obtain a semantic value for  $P$ , which is a set where each element is a tuple of a world, a time, a viewpoint, and witnesses for discourse referents. This is recorded in (8).

$$(8) \quad \llbracket P \rrbracket = \{ \langle w, t, v, x_1, \dots, x_n \rangle \mid w, t, v, x_1, \dots, x_n \models P \}$$

This is a set of cases in the sense of Lewis (1975). Lewis introduced case semantics to theorize about indefinite descriptions and anaphora in sentences with adverbs of quantification, such as the Murphy’s law examples (9). He showed that by assuming a case semantics for the two clauses in such sentences (i.e. the if-clause and the main clause) it is possible to arrive at a semantics for the whole compositionally.

- (9) a. If you drop an unbreakable object, it always lands on something more valuable.  
 b. If two cars are driving in opposite directions on a long road with a one-way bridge, they always meet at the bridge.

To deal with sentences that have free indices (such as the main clauses in (9)), it is necessary to say that a syntactic unit denotes a set of cases *relative to a case*. Where  $X$  is the syntactic unit and  $c$  is a case, for this we use the notation  $c\llbracket X \rrbracket$ .<sup>7</sup> We will always refer to cases that are of the form  $wv\mathcal{O}$  or (when time is being ignored)  $wv\mathcal{O}$ . Here  $\mathcal{O}$  is a string of objects that witness discourse referents,  $w$  is a world,  $t$  is a time, and  $v$  is a viewpoint. (10) gives some semantic values in this notation.

- (10) a.  $wv\mathcal{O}\llbracket [\text{he}_1 \text{ has a dog}] \rrbracket = \{ c \mid \exists x.c = wvx\mathcal{O} \wedge \mathbf{dog}(w, x) \wedge \mathbf{have}(w, \mathcal{O}[1], x) \}$   
 b.  $wv\mathcal{O}\llbracket [1 = 2] \rrbracket = \{ c \mid c = wv\mathcal{O} \wedge \mathcal{O}[1] = \mathcal{O}[2] \}$

c.  $wv\mathcal{O} \left[ \begin{array}{|c|} \hline \img alt="A 3D wireframe box." data-bbox="255 552 315 594"/> \hline \end{array} \right] = \left\{ c \mid \exists v'.c = wv'\mathcal{O} \wedge \bar{\pi}(w, v') = \begin{array}{|c|} \hline \img alt="A 3D wireframe box." data-bbox="558 552 618 594"/> \hline \end{array} \right\}$

## 2 A Basic Analysis

We have seen that the semantics value of an enriched pictorial narrative (as formulated in Section 1) is the same kind of formal object as the semantics of a sentence of English containing indefinite descriptions and pronouns. In Abusch (2012, to appear); Abusch & Rooth (2017); Rooth & Abusch (2018), this is used to give an analysis of indexing in pictorial narratives, and analyses of additional phenomena, using the toolkit of dynamic natural language semantics. Here we observe that, once we move to the semantics, there is no difference between indexing within a medium and indexing across media. An index that is set up within a pictorial narrative

<sup>6</sup>There are questions about the optimal formulation geometric discourse referents (e.g. points vs. bounding boxes) and the optimal definition of what objects correspond to them. For instance, if the individuals in the model have part-whole structure and we use points, there may be unwelcome multiplicity in the value of  $x$ . Consider a point  $d$  in the head-area of a depicted character. Let  $x$  be a person in the model, and let  $x_h$  be that person’s head. If  $\bar{\pi}(w, t, v, d, x_h)$  holds, then also  $\bar{\pi}(w, t, v, d, x)$  holds. See for discussion Abusch (2014). Bounding boxes can be used to partially alleviate this problem. Ultimately though we are inclined to maintain that predications about the type of depicted objects are accommodated, e.g.  $\mathbf{person}(w, x)$ .

<sup>7</sup>Rooth (to appear) presents Lewis’s semantics for indefinites and adverbs of quantification along these lines.

can be picked up later in the pictorial narrative. But equally, it can be picked up with a pronoun in a sentence of natural language.

Consider the left column in (11), which we think of as a scenario where a parent reading a picture book to a child points out a character in a picture, gives some information verbally, continues by pointing at (or touching) the dog in the next picture, and then adds some more verbal information.

(11) a.



His name is Dick.

He has a dog.



His name is Spot.

b.



$\langle \cdot, \cdot \rangle$

$[\text{his}_1 \text{ name is Dick}]$

$[\text{he}_1 \text{ has a dog}]$



$\langle \cdot, \cdot \rangle$

$1 = 2$

$[\text{his}_1 \text{ name is Spot}]$

Images from William Gray, *Fun with Dick and Jane*, 1946.

The right column gives a counterpart in our formal language, where the finger-touching gestures are replaced by geometric discourse referents, and equalities between discourse referents are added. This formula has a linear structure with eight parts, which we name  $p_1$  (a picture), a point  $d_2$  introducing a discourse referent, a sentence  $s_3$  containing a pronoun, a sentence  $s_4$  containing a pronoun and an indefinite description, a picture  $p_5$ , a point  $d_6$  introducing a

discourse referent, an equality between discourse referents  $e_7$ , and finally a sentence  $s_8$ .<sup>8</sup>

The cross-medium narrative (11b) is to be interpreted in the uniform dynamic framework that was reviewed in Section 1. To simplify, in the current discussion we do not include times. (12) gives the semantics of the eight parts of the narrative. A picture  $p_i$  interpreted relative to  $wv\mathcal{O}$  introduces a new viewpoint  $v'$ , and checks that the world from the viewpoint projects to the picture.  $\mathcal{O}$  is not incremented. Thus  $wv\mathcal{O}[[p_i]] = \{z|\exists v'.z = wv'\mathcal{O} \wedge \pi(w, v') = p_i\}$ , where the new viewpoint is recorded in an output case  $wv'\mathcal{O}$ . A geometric discourse referent  $d_i$  non-deterministically chooses an object  $x$ , and checks the geometric constraint  $\bar{\pi}(w, v, d_i, x)$  that relates the viewpoint  $v$  for the last picture, the point  $d_i$ , and the value  $x$  for the discourse referent.  $\mathcal{O}$  is incremented with  $x$  to form  $x\mathcal{O}$ . Thus  $wv\mathcal{O}[[d_i]] = \{z|\exists x.z = wvx\mathcal{O} \wedge \bar{\pi}(w, v, d_i, x)\}$ . An equality  $m = n$  is semantically a test that checks equality of  $\mathcal{O}[m]$  and  $\mathcal{O}[n]$ , see (12g). The three sentences are given standard interpretations in dynamic semantics. Indexed pronouns look up their referents in  $\mathcal{O}$ , with indexing into  $\mathcal{O}$  following a recency convention. Thus the index 1 in sentence  $s_3$  (“his<sub>1</sub> name is Dick”) gets the value  $\mathcal{O}[1]$ , and  $[[s_3]]$  is a test which checks the name of  $\mathcal{O}[1]$ . The indefinite description in sentence  $s_4$  introduces a new value  $x$  that is entered as  $x\mathcal{O}$ , which is constrained to be a dog in  $w$ , and to be possessed by  $\mathcal{O}[1]$  in  $w$ .

- (12) a.  $wv\mathcal{O}[[p_1]] = \{z|\exists v'.z = wv'\mathcal{O} \wedge \pi(w, v') = p_1\}$   
 b.  $wv\mathcal{O}[[d_2]] = \{z|\exists x.z = wvx\mathcal{O} \wedge \bar{\pi}(w, v, d_2, x)\}$   
 c.  $wv\mathcal{O}[[s_3]] = \{z|z = wv\mathcal{O} \wedge \mathbf{name}(w, \mathcal{O}[1], \text{“Dick”})\}$   
 d.  $wv\mathcal{O}[[s_4]] = \{z|\exists x.z = wvx\mathcal{O} \wedge \mathbf{dog}(w, x) \wedge \mathbf{have}(w, \mathcal{O}[1], x)\}$   
 e.  $wv\mathcal{O}[[p_5]] = \{z|\exists v'.z = wv'\mathcal{O} \wedge \pi(w, v') = p_5\}$   
 f.  $wv\mathcal{O}[[d_6]] = \{z|\exists x.z = wvx\mathcal{O} \wedge \bar{\pi}(w, v, d_6, x)\}$   
 g.  $wv\mathcal{O}[[e_7]] = wv\mathcal{O}[[1 = 2]] = \{z|z = wv\mathcal{O} \wedge \mathcal{O}[1] = \mathcal{O}[2]\}$   
 h.  $wv\mathcal{O}[[s_8]] = \{z|z = wv\mathcal{O} \wedge \mathbf{name}(w, \mathcal{O}[1], \text{“Spot”})\}$

Thus the eight parts of the cross-medium narrative get interpretations in a uniform dynamic semantic framework. This immediately answers the question of how information from different media is combined: such information is combined in the way information in a single medium is combined in a dynamic framework, namely by dynamic conjunction. (13) is a formulation of dynamic conjunction in the current notation. Here  $x$ ,  $y$ , and  $z$  are cases of the form  $wv\mathcal{O}$ , and the definition essentially expresses relation composition.

$$(13) \quad x[[AB]] = \{z|\exists y[y\epsilon x[[A]] \wedge z\epsilon y[[B]]\}$$

Conjoining the parts in (12) using dynamic conjunction results in (14) as the semantics of the cross-medium narrative (11b), relative to a null context  $wv$  consisting of a world and an (irrelevant) viewpoint.

<sup>8</sup>The phenomenon of touching pictures to set discourse referents connects with the analysis of pointing in O'Madagain et al. (2019), where it is argued that pointing is usually sight-line pointing, and that such pointing is continuous with touching.



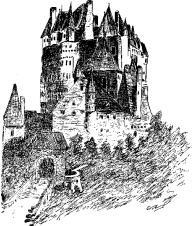
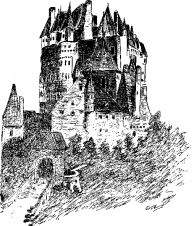
$$(14) \quad wv[p_1d_2s_3s_4p_5d_6e_7s_8] = \left\{ c[\exists x_3\exists v_2\exists x_2\exists x_1\exists v_1 \left[ \begin{array}{l} c = wv_2x_3x_2x_1 \quad \wedge \\ \pi(w, v_1) = p_1 \quad \wedge \\ \bar{\pi}(w, v_1, d_2, x_1) \quad \wedge \\ \mathbf{name}(w, x_1, \text{“Dick”}) \quad \wedge \\ \mathbf{dog}(w, x_2) \quad \wedge \\ \mathbf{have}(w, x_1, x_2) \quad \wedge \\ \pi(w, v_2) = p_5 \quad \wedge \\ \bar{\pi}(w, v_2, d_6, x_3) \quad \wedge \\ x_3 = x_2 \quad \wedge \\ \mathbf{name}(w, x_3, \text{“Spot”}) \quad \wedge \end{array} \right] \right\}$$

Some comments about the mechanics are in order. In a tuple  $c$  of the form  $wv_2x_3x_2x_1$ ,  $x_3$  is a witness for the discourse referent that was introduced last. That discourse referent is introduced by  $d_2$ , and corresponds to the dog in the second picture.  $x_2$  is a witness for the penultimately introduced discourse referent, which is introduced by the phrase [a dog] in  $s_4$ . These discourse referents are distinct, but they are identified by the equation  $1 = 2$  in (11b), which equates the ultimately and penultimately introduced discourse referents. This results in  $x_3 = x_2$  in the body of (14).  $v_2$  is the viewpoint used for  $p_5$ , and it is also used in selecting values for  $d_6$ , as expressed in the condition  $\bar{\pi}(w, v_2, d_6, x_3)$ . All the conditions in the body of (14) refer to the same world variable  $w$ , when they refer to a world at all. This indicates that the eight parts in (11b) are combining extensionally.

### 3 The Nominal Depiction Constraint

Look at the matrix of data in (15), where each cell has a picture combined with a nominal English phrase, rather than a sentence. The off-diagonal elements are somehow inconsistent or implausible. For instance, the top right combination with the caption “a castle owned by a duke” is intuitively inconsistent because what is depicted looks like a person, and not at all like a castle. Yet both a duke and a castle are mentioned in the phrase.

(15)

		Drawing of Louis Victor de Rochechouart, Duke of Mortemart, Duke of Vivonne by Antoine Maurin. Wikimedia Commons.
A duke who owns a castle	A castle owned by a duke	
		Castle on a hill. Public Domain Vectors.
A duke who owns a castle	A castle owned by a duke	

The data in (16) are similar. Even though the righthand combination is to some degree pragmatically coherent—the cat basket is empty because the cat that ordinarily occupies it is lost—this combination of a picture and a nominal caption conveys inconsistent information. In contrast, the combination (17), where the caption conveys similar information but is a sentence rather than a nominal, is slightly disjointed but consistent.

(16) a.



A lost cat

b.



A lost cat

Image from page 143 of *The Animals of the World. Brehm's Life of Animals*. 1895. Wikimedia Commons.

Image from page 659 of *Florists' Review*. 1912. Internet Archive Book Images.

(17)



A cat got lost and didn't come home.

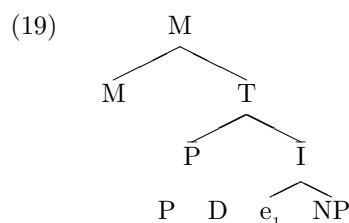
These data motivate the *nominal depiction constraint*: roughly, when a picture is accompanied by a nominal, the top-level index in the nominal is co-indexed with a discourse referent pointing into the picture. Or to put it differently, a witness for the top-level index of the nominal is depicted in the picture. For instance, assuming that the semantics of the phrase *a duke who owns a castle* distinguishes a discourse referent for a duke as the top-level index of the nominal, the LF of the top-left combination in (15) should involve a discourse referent pointing into the picture, and this discourse referent should be equated with the top-level index of the caption. This constraint will be imposed syntactically. The syntax of the mixed-medium narratives seen so far can be captured by the context free rules in (18). This creates left-bracing trees, consisting of pictures (syntactic category P), sentences (syntactic category S), geometric discourse referents (syntactic category D) and equalities (syntactic category E). M is the syntactic category of cross-medium narratives. A phrase of category S is assumed to be a sentence, as characterized syntactically and semantically by an interpreted grammar of English. So far, this does not introduce any nominal phrases into mixed-medium narratives.

- (18)
- M → P
  - M → S
  - M → M S
  - M → M P
  - M → M D
  - M → M E

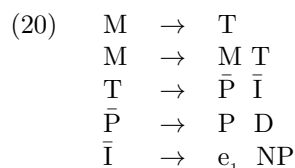
We treat a picture accompanied by a nominal phrase as a special construction that enforces co-indexing. To express this, we hypothesize that the nominal phrases have a predicative



syntactic and semantic type, here assumed to be NP. The tree shape in (19) enforces the required indexing. T is the syntactic category for the construction as a whole. It has two parts. The first part  $[\bar{p} P D]$  is a combination of a picture and a geometric discourse referent. It introduces a picture with a discourse referent pointing into it. Given the recency convention in the dynamic semantics, that discourse referent is accessed with the index 1. The second part  $[\bar{p} e_i NP]$  is a combination of an empty category with index 1 and a nominal predicate with syntactic category NP. It has the effect of applying the nominal predicate to the geometric discourse referent that is introduced in  $[\bar{p} P D]$ .



The phrase structure rules covering the construction are in (20). The important point in the analysis of the nominnal depiction constraint is that nominal phrases are not introduced freely. Rather they are introduced in a construction that stipulates indexing into a picture.

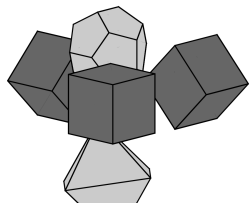


## 4 Quantification and Definite Reference

In (21a,b), the sentences can be conceived of as observations about the information conveyed by the accompanying picture. (22a,b) are combinations of the same form, but where the sentences give independent information of a kind that can not be conveyed by pictures. In (21a) and (22a), the DPs of the form [every cube] can conceivably be read as quantifying all the cubes in the world. But these DPs are more naturally read to quantify the cubes *that are depicted*. Half of the analysis of this reading is familiar. According to the analysis of Westerståhl (1989) quantificational determiners come with a context variable for a contextually determined domain of quantification. We write this here with a superscripted numerical index. The representation for the sentence in (21a) is then (23a), where the index for the domain of quantification is 1. The value of this index in context should be set in a way that (23a) gets the reading (23b).<sup>9</sup>

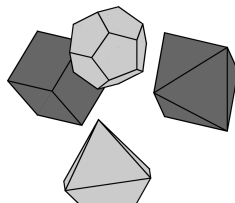
<sup>9</sup>The paraphrase needs to be analyzed too. See the next section.

(21) a.



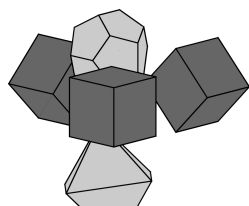
Every cube is dark.

b.



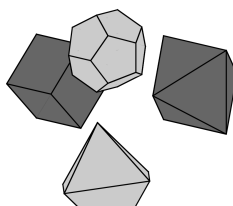
The cube is dark.

(22) a.



Every cube belongs to Jack.

b.



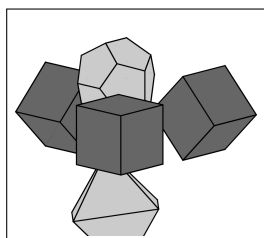
The cube belongs to Jack.

(23) a. Every<sup>1</sup> cube is dark.

b. Every cube that is depicted is dark.

These data lead to the hypothesis that pictures make available or can make available a group discourse referent for the depicted objects.<sup>10</sup> Following the strategy of expressing particular readings syntactically in the discourse representation, we propose an operator  $G$  that introduces a discourse referent for the group of objects that are depicted in the previous picture.  $G$  does not involve a point or a bounding box, because it is supposed to introduce a discourse referent for all the depicted objects. It is simply a syntactic constant. (24) is then the discourse representation for the depiction-restricted reading of (21a). The formula is structured linearly, beginning with a picture. Following that the operator  $G$  introduces a discourse referent for the set of objects depicted in the picture. This discourse referent was introduced last, and so is referenced with the index 1. In the sentence that completes the formula, the index 1 by virtue of its syntactic position contributes the domain of quantification for *every*.

(24)

 $G$  [every<sup>1</sup> cube is dark]

A semantics for  $G$  is defined as a quantified version of the semantics of geometric discourse referents. Suppose we are given a picture  $p$  with unit dimensions, a viewpoint  $v$ , and a world  $w$  such that  $\pi(w, v) = p$ . An object  $x$  in  $w$  is depicted in  $p$  if and only if there is a point  $d$  in

<sup>10</sup>Abusch (2012) also used group discourse referents in analyzing indexing in pictorial narratives.

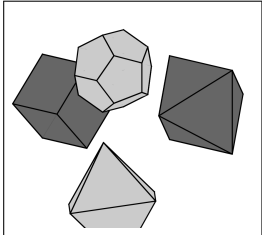
$[0, 1]^2$  such that  $\bar{\pi}(w, v, d, x)$ . Therefore we can say that  $x$  is a member of the group discourse referent created by  $G$  (given  $w$  and  $v$ ) if and only if there is some discourse referent  $d$  such that  $x$  is a witness for  $d$  relative to  $w$  and  $v$ . This leads to the definition of the semantics of  $G$  in (25). Where  $p_{24}$  is the picture in (24), (26) is the resulting semantics for (24), where the universal quantification is restricted to depicted objects via the conjunct  $y \in X$ .

$$(25) \quad wv\mathcal{O}[[G]] = \{c | \exists X.c = wvX\mathcal{O} \wedge X = \{x | \exists d.\bar{\pi}(w, v, d, x)\}\}$$

$$(26) \quad wv[[p_{24}G[\text{every}^1 \text{ cube is dark}]]] = \left\{ c | \exists X \exists v_1 \left[ \begin{array}{l} c = wv_1X \\ \pi(w, v_1) = p_{24} \\ X = \{x | \exists d.\bar{\pi}(w, v_1, d, x)\} \\ \forall y[\mathbf{cube}(w, y) \wedge y \in X \rightarrow \mathbf{dark}(w, y)] \end{array} \wedge \right] \right\}$$

Examples (21b) and (22b) have sentences with definite descriptions rather universal quantifiers. Here the observation is that the uniqueness presupposition of the definite description is satisfied among objects that are depicted in the picture. For instance in (21b), there is a definite description  $[_{DP} \text{the cube}]$ , and in worlds compatible with the picture, there is a unique cube that is depicted in the picture. These examples are analyzed in a parallel way, see (27).

(27)



$G$  [the<sup>1</sup> cube is dark]

## 5 Depiction Sentences

For some of the cross-medium data from Section 3 and Section 4, there are parallel data involving sentences that describe pictures. Recall  $p_{16a}$ , the picture of a cat lost in the woods, and  $p_{16b}$ , the picture of an empty cat basket. Referring to these pictures, sentence (28a) is true, and sentence (28b) is false, intuitively because it depicts a cat basket rather than a cat.

- (28) a. Picture  $p_{16a}$  depicts a cat.  
 b. Picture  $p_{16b}$  depicts a cat.

Sentence (29) is a version of what in Section 4 was cited as a paraphrase of the depiction-restricted reading of a mixed-medium sequence.

- (29) Every cube that is depicted in picture  $p_{21a}$  is dark.

These sentences can be used in a discussion among agents who can see the pictures. They can also be used to convey information to an agent who can not see the picture. This makes it implausible that the logical forms of these sentences include particular geometric discourse referents. The reason is that, without access to the picture, a listener can not be expected to accommodate a particular geometric discourse referent. But following the strategy used in the semantics of  $G$ , the discourse referent can be quantified in the semantics. This suggests a

semantic paraphrase along the lines of (30) for (28a). It says that there is a discourse referent such that, in every world and viewpoint compatible with the picture, the individual picked out by the discourse referent with respect to the world and viewpoint is a cat.

$$(30) \quad \exists d \forall w \forall v \forall x [\pi(w, v) = p_{16a} \wedge \bar{\pi}(w, v, d, x) \rightarrow \mathbf{cat}(w, x)]$$

This is a formalization of a de dicto reading of the sentence. Although we think this analysis works for pictures of cubes and dodecahedra in a modal space where worlds are occupied only by regular polytopes, cat pictures of the familiar kind do not have information strong enough to entail (30). After all, our own world contains realistic sculptures of cats that are not real cats. Also, depiction sentences have ambiguities along de dicto/de re lines, similar to the ambiguities studied for the verb *paint* in examples like (31) that are studied in Zimmermann (2006). There is much more to say about (28) and (29). Nevertheless, the connection between these examples and the nominal depiction constraint from Section 3 is intriguing, and that connection does fall out of the formalization (30).

$$(31) \quad \text{Edlon painted a bridge.}$$

## 6 Conclusion

The idea proposed here is to theorize about indexing across media by using a uniform dynamic semantic framework for the media. Indexing is analyzed at the semantic level, where the media are not distinguished. We defined a formal language and a semantic interpretation for it. Particular constructions and constraints were treated in the syntax of the formal language. While it would be possible to do the syntactic part without referring to possible worlds semantics and dynamic semantics, in the research strategy pursued here, the two go hand in hand.

## References

- Abusch, Dorit. 2012. Applying discourse semantics and pragmatics to co-reference in picture sequences. In *Proceedings of Sinn und Bedeutung*, vol. 17, .
- Abusch, Dorit. 2014. Temporal succession and aspectual type in visual narrative. *The Art and Craft of Semantics: A Festschrift for Irene Heim* 1. 9–29.
- Abusch, Dorit. to appear. Possible worlds semantics for pictures. In Daniel Gutzman, Lisa Matthewson, Cecile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.), *Companion to Semantics*, Wiley. Version of 2015 at [hdl.handle.net/1813/44654](http://hdl.handle.net/1813/44654).
- Abusch, Dorit & Mats Rooth. 2017. The formal semantics of free perception in pictorial narratives. In *Proceedings of the 21st amsterdam colloquium*, .
- Everingham, M., L. Van Gool, C. K. I. Williams, J. Winn & A. Zisserman. 2012. The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. [pascal-network.org/challenges/VOC/voc2012/workshop](http://pascal-network.org/challenges/VOC/voc2012/workshop).
- Lewis, David. 1975. Adverbs of quantification. In *Formal Semantics of Natural Language*, 178–188. Cambridge University Press.
- Maier, Emar & Sofia Bimpikou. 2019. Shifting perspectives in pictorial narratives. In *Proceedings of Sinn und Bedeutung*, vol. 23 2, 91–106.

- O'Madagain, Cathal, Gregor Kachel & Brent Strickland. 2019. The origin of pointing: Evidence for the touch hypothesis. *Science Advances* 5(7).
- Rooth, Mats. To appear. Adverbs of quantification. In Louise McNally & Zoltán Szabó (eds.), *A Reader's Guide to Classic Papers in Formal Semantics*, Springer.
- Rooth, Mats & Dorit Abusch. 2018. Picture descriptions and centered content. In Rob Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung*, vol. 21 2, 1051–1064.
- Westerståhl, Dag. 1989. Quantifiers in formal and natural languages. In *Handbook of Philosophical Logic*, 1–131. Springer.
- Zimmermann, Thomas Ede. 2006. Quaint paint. In Hans-Martin Gärtner, Sigrid Beck, Regine Eckardt, Renate Musan & Barbara Stiebels (eds.), *Between 40 and 60 puzzles for Krifka*, Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung (ZAS).