SPEECH ACT THEORETIC SEMANTICS

by

DANIEL W. HARRIS

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2014

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Gary Ostertag

_____ _____

Date Chair of Examining Committee

Iakovos Vasiliou

_____ _____

Date Executive Officer

Ernie Lepore

Michael Devitt

David Rosenthal

Stephen Neale

_____

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

SPEECH ACT THEORETIC SEMANTICS

by

Daniel W. Harris

Adviser: Professor Stephen Neale

Abstract:

I defend the view that linguistic meaning is a relation borne by an expression to a type of speech act, and that this relation holds in virtue of our overlapping communicative dispositions, and not in virtue of linguistic conventions. I argue that this theory gives the right account of the semantics–pragmatics interface and the best-available semantics for non-declarative clauses, and show that it allows for the construction of a rigorous compositional semantic theory with greater explanatory power than both truth-conditional and dynamic semantics.

PREFACE & ACKNOWLEDGEMENTS

The ideas of Paul Grice have come up often during the recent boom of work on the semantics–pragmatics interface. In experiencing this boom, first as an undergraduate at Simon Fraser University in 2005, then in graduate seminars at the City University of New York, New York University, and Rutgers, as well as in summer schools at the Central European University in 2010 and the University of Riga in 2012, I have often thought that if we really want to understand the semantics–pragmatics interface, it is not enough to be Griceans when it comes to pragmatics; we must also work out the details of a properly Gricean semantics.

Part of Grice's ambition in introducing the program of intention-based semantics, initially in his 1957 paper, 'Meaning,' and later in his 1967 William James Lectures (1989: chs.1–7), was to give a theory of the timeless meanings of utterance-types, and this category includes (but is not exhausted by) the meanings of linguistic expressions. The most worked-out version of Grice's take on linguistic meaning can be found in his 1968 paper, 'Utterer's Meaning, Sentence Meaning, and Word Meaning' (1989: ch.6). This paper is a difficult read, and it languished, unanthologized until the publication of *Studies in the Way of Words* in 1989, in the pages of the short-lived and relatively obscure journal, *Foundations of Language* (1965–1975). I suspect that the essay's relative obscurity partly explains the fact that Grice has had little influence on contemporary theorizing about linguistic meaning.[1] Grice's 1968 article has influenced the present work immensely. Grice

[1] Of course, Grice's theory of conversational implicature has had a massive influence on semantics, but this influence has been *from outside*, via its influence on pragmatics, by placing limits on what is generally taken to be the scope of semantic explanation. On occasion, Gricean reasoning has had a more direct influence on semantics—for example, via Schiffer's (1995) objections to hidden-indexical theories of propositional-attitude reports. Another example is Potts' (2005) use of Grice's definition of conventional implicature in the course of bringing that notion into mainstream semantics. (Ironically, of course, although Potts attributes all four of the clauses of his definition to Grice, he argues that none of Grice's own examples are genuine cases of conventional implicature.)

does the following things in the paper: (a) he lays the foundations for a theory of the nature of sentence meaning and semantic underspecification that I develop and defend in Chapter Two, (b) he builds on his intentionalist theory of illocutionary force, and hints at how it can be used to give a semantics of grammatical mood that influences the theory I build in Chapter Three;[2] (c) he sketches a metasemantic theory that is the main inspiration for the theory I give in Chapter Four; and, finally, (d) he outlines a proto-speech-act-theoretic account of the composition of complex utterance-type meanings, which is one of the main influences on the account of the compositionality that I give in Chapter Five. To be sure, the sketches, outlines, and hints in Grice's 1968 article are incomplete and confused in various respects. I have had to extrapolate from, build on, and, in places, heavily revise Grice's ideas in order to arrive at the theory spelled out here. And, of course, my own views have also benefitted from a wide variety of other influences, some of which I'll mention here. But it is unlikely that I would have thought up most of the components of what follows if I had not read, and re-read, Grice's 1968 paper.

The crucial thing about the ideas in that paper, as well as the ideas in this dissertation, is that their appeal is best seen from a perspective on the nature of language, communication, and the mind–language interface that is thoroughly Gricean. This is a perspective that I have inherited not only from Grice, but also from a succession of Gricean mentors. I bought my copy of *Studies in the Way of Words* for Mike Harnish's undergraduate seminar at Simon Fraser University in 2005.[3] In my term paper for that course, which anticipates some of Chapter Two of this dissertation, I

---

[2] The initial formulation of Grice's theory of illocutionary force can be found in his 1969 essay, 'Utterer's Meaning and Intention' (1989: ch.5). It builds on some ideas in Strawson (1964).

[3] I first read Martinich's (1985/2001) anthologized versions of 'Meaning' and 'Logic and Conversation' a year earlier, in Martin Hahn's Philosophy of Language course at SFU, which was where I became interested in the philosophy of language in the first place. The following autumn, Martin agreed to supervise a directed reading on Frege with me, and, in return, got to read a convoluted term paper in which I attempted to solve every problem in the philosophy of language all at once.

defended the view that the reference of a demonstrative is fixed by the speaker's communicative intentions.

The reading list for Mike's seminar also included Stephen Neale's epic review of *Studies*, 'Paul Grice and the Philosophy of Language' (1992), which I have re-read several times since. It took a couple of years after I came to the Graduate Center to study with Stephen before he was able to convince me of the Gricean program in all of its nuance. By the middle of 2011, I had become convinced of the fruitfulness of the methodological framework that Stephen calls linguistic pragmatism, and I had decided to work out the implications of that framework for linguistic meaning. In the fall of 2011, I met with Stephen nearly every week, and we took turns discussing the early bits of this dissertation as well as some articles that I was helping him to prepare for publication in a volume of collected papers. In several of those articles (2004, 2005, 2007a), Stephen outlines what he calls the *blueprint theory of meaning*, and in 'Term Limits Revisited' (2008), he sketches an "act-syntactic" approach to the meanings of noun phrases—both ideas that he presented in greater detail in his seminar on linguistic pragmatism in the spring of 2011. I was convinced that these ideas pointed in the direction of the right sort of approach to semantics—one that also fit nicely with Grice's (1968) approach. I had also been confused about how Stephen's more conventional work on semantics—such as what he says in *Descriptions* (1990) and *Facing Facts* (2001)—could possibly be compatible with his radically Gricean views about the semantics–pragmatics interface. So, with Stephen's encouragement, I decided to work out a pragmatist approach to the foundations and methodology of natural-language semantics that could incorporate both the rigor of contemporary truth-conditional and dynamic semantics and the foundational depth of Griceanism. I am filled with gratitude for Stephen's guidance and enthusiasm, without which this dissertation would not have been possible.

Others—Gricean or otherwise—who, in their roles as professors and mentors, and through their seminars, questions, comments, and encouragement, have helped to shape the views pre-

sented here include Kent Bach, Liz Camp, Michael Devitt, Peter Hanks, Kirstie Laird, Sandra Lapointe, Ernie Lepore, Peter Ludlow, Jeff Pelletier, Craige Roberts, David Rosenthal, Stephen Schiffer, Dan Sperber, Jason Stanley, Zoltán Szabó, and Deirdre Wilson.

I would particularly like to single out Michael Devitt, Ernie Lepore, Gary Ostertag, and David Rosenthal, both for agreeing to be members of my dissertation committee and for being supportive and persistent forces for good in my intellectual life while I've been a graduate student.

I took several of Michael's seminars early in my career at the Graduate Center, and the subtle pattern of ways in which we agree and disagree has forced me to attend to many details of my views that might otherwise have gone unarticulated. Michael has also been an invaluable model for how to stay focused on what is important, both philosophically and professionally, and I am grateful for this.

Ernie's recent and ongoing work—much of it jointly thought up with Matthew Stone and Una Stojnic—conceives of the overall goals of semantics and pragmatics in communication-centric terms that I find largely congenial, but our projects diverge on the details at nearly every choice point. Our resulting debates have been as enjoyable as they have been productive, and I look forward to continuing them. I've also learned from Ernie that a crucial part of doing philosophy is working to create the social conditions for philosophy, that healthy and productive philosophical communities are fueled by a constant supply diverse viewpoints, contrasting personalities, and mutual fascination, and that the right kinds of attitude and effort can reveal that philosophy is not a zero-sum game. I am fortunate to have learned these lessons at the start of my career.

My debates with David have tended to center around the most theoretically basic commitments that underlie my project—the idea that intentional mental states ground the properties of speech acts, for example, and the idea that we must explain linguistic meaning compositionally at all. I like to think that I have come some way toward internalizing David's foundational concerns, so that I will continue to worry about and make progress on them in the years to come. I am

grateful for this, and I feel lucky to have had David as a mentor and a friend for the past several years.

Gary has been a patient, careful, and energetic interlocutor since I first met him in my second year at the Graduate Center. He read many early versions of my work, and has always been adept at pointing out large and small problems in my arguments while at the same time keeping my own explanatory goals charitably in focus. Gary's home base at the Kripke Center is a place where I can reliably go to upgrade both my philosophical acumen and my musical taste in the most enjoyable ways possible.

I am also greatful to all of the following people for their influence on the ideas contained in this dissertation, either through philosophical conversation, via their excellent questions at talks, or by commenting on drafts of chapters: Josh Armstrong, Endre Begby, Ray Buchanan, Alexis Burgess, Nate Charlow, Lars Dänzer, Imogen Dickie, Daniel Fogal, Peter Hanks, Megan Henricks Stotts, Thomas Hodgson, Zoe Jenkin, Marilynn Johnson, Jessica Keiser, Phil Kremer, Oliver Marshall, Rachel McKinney, Eliot Michaelson, Lisa Miracchi, Friderike Moltmann, Adam Morton, Matt Moss, Sarah Murray, Myrto Mylopolous, David Pereplyotchik, Ariadna Pop, Jesse Rappaport, Gurpreet Rattan, Indrek Reiland, Mark Schroeder, Will Starr, Matthew Stone, and Elmar Unnsteinsson.

I presented earlier versions of some of the work contained here at the Atlantic Canadian Philosophical Association's Annual Meeting at Kings College in Halifax in October of 2010, at the UCLA/USC Graduate Philosophy Conference at UCLA in February of 2011, at the conference on Meaning, Context, and Implicit Content in Cérisy, France in June of 2011, at the New York Philosophy of Language Workshop in September of 2011, March of 2013, and January of 2014, at APA Pacific Division Meetings in Seattle in April of 2012 and in San Francisco in March of 2013, at the CUNY Cognitive Science Speaker Series in March of 2012, at the Society for the Study of the History of Analytical Philosophy's annual meeting at McMaster University in May of 2012, at

the Summer School in Cognitive Science and Semantics in Pumpuri, Latvia, in July of 2012, at the Western Canadian Philosophical Association's annual meeting in Victoria in October of 2012, at the University of Western Ontario's PhilMilCog Graduate Philosophy Conference in May of 2013, and at talks at the University of Pennsylvania and Hunter College, both in January of 2014. My thanks to the audiences at each of these venues for their thought-provoking feedback.

Work on this dissertation has been financially supported by an Enhanced Chancellor's Fellowship, a Dissertation Fellowship, and two Doctoral Student Research Grants from the CUNY Graduate Center and by a Doctoral Fellowship from the Social Sciences and Humanities Research Council of Canada (#752-2009-0684).

My thanks to Daniel Fogal, Rachel McKinney, Eliot Michaelson, Matt Moss, David Pereplyotchik, Dan Shargel, and Elmar Unnsteinsson for being the cornerstones in a social and intellectual support structure that I have leaned on continually on my way through graduate school. Each of them is an excellent philosopher, a cooperative communicator, and a good friend.

Finally, I thank my parents, Fred and Mary, my brother, Rob, and my partner, Margot, for believing that I could be a philosopher and doing everything in their power to help. I am unable to articulate how important this has been.

TABLE OF CONTENTS

CHAPTER ONE: MEANING, SPEECH ACTS, & METHODOLOGY

What is linguistic meaning? In the context of contemporary philosophy of language and linguistics, to answer this question would be to articulate and explicate a concept of meaning that can play the various explanatory roles demanded of it by pragmatics, semantics (lexical and compositional), and metasemantics—a concept that can play nicely with our best theories of the workings of the human mind, the nature of representation, and the intricacies of communication and social interaction. The articulation and explication of such a concept is the task of this dissertation.

This chapter will serve as an introduction to the project, its place in the dialectical landscape of current thinking about meaning and communication, and its methodological background. In §1.1, I give a brief introduction to my own theory of linguistic meaning, which I call speech-act-theoretic semantics. In §1.2, I introduce the Gricean theory of speech acts that I favor. The main purpose of this dissertation is not to defend Griceanism, but to articulate and defend an approach to semantics that can interface with a Gricean approach to pragmatics. Nevertheless, because the Gricean approach to speech acts is an important component of speech-act-theoretic semantics as I conceive of it, it is important that I say something about why I favor the approach. In §1.3, I outline my overall methodological approach to the study of linguistic meaning, situating it within Gricean speech-act theory.

1.1 Speech-Act-Theoretic Semantics: The Basic Picture

Speech-act-theoretic semantics identifies the meaning of an expression with the type of speech act that the members of a community are disposed to perform with the expression and interpret others as performing with it. This approach encompasses approaches to several broad components of the study of linguistic meaning, each of which breaks away from contemporary orthodoxy.

Metasemantics addresses the question: *In virtue of what do expressions have meaning for a group of speakers?* The dominant answer to this question is that expressions have meaning in virtue of the linguistic conventions in which language-users participate. In Chapter Four, I will offer a series of objections to this view and defend the alternative view that an expression has meaning for a group of speakers in virtue of their overlapping dispositions to perform speech acts of a certain type with the expression, and to interpret utterances of the expression as performances of the same type of speech act. We can think of the type of speech act thus related to an expression as its *meaning* for the relevant community.

Compositional Semantics addresses the questions: *What are the meanings of complex expressions—such as sentences—and how are complex expressions' meanings determined by their structures and the meanings of their parts?* The two mainstream answers to this question are *truth–conditional semantics*, on which sentences' meanings are properties that fix their truth conditions relative to contexts, and *dynamic semantics*, on which sentence meanings are context-change potentials. I will raise objections to both of these genres of semantics in Chapters Two and Three, and argue that sentence meanings are types of illocutionary acts. On the most influential versions of both truth-conditional and dynamic semantics, simpler meanings combine via functional application to yield complex meanings. In Chapter Five, I'll defend the alternative view that semantic composition is a matter of the way in which our dispositions to perform word-sized speech acts add up to our dispositions to perform sentence-sized speech acts, and that facts about this adding-up are intimately connected to combinatory speech acts such as predication.

Lexical Semantics addresses the question: *What are the meanings of semantically primitive expressions?* Semanticists answer this question in a variety of ways; most influentially, they say that a word's meaning is its character—a function from possible contexts to the word's contents at those contexts. In Chapter Two, I argue that this answer can't account for the ways in which words can be used to do different things on different occasions, and defend the alternative answer that the

meanings of semantically primitive expressions are types of acts of referring to objects, properties, and relations.

Speech-act-theoretic semantics incorporates the advantages of use-theoretic approaches to meaning by making the meaning of an expression a matter of what people are disposed to do with it, rather than merely a matter of what the expression stands for. In performing speech acts with particular expressions, we don't just represent the world and what's in it; we also ask questions, issue directives, signal presuppositions, imply things beyond what we say, express non-cognitive mental states, and signal socially significant facts about ourselves. Since these are all aspects of the speech acts that we are disposed to perform with particular expressions, speech-act-theoretic semantics can find a place for each in its account of those expressions' linguistic meanings. This gives speech-act-theoretic semantics significantly more explanatory power than mainstream semantics when it comes to dealing with the aspects of natural languages that set them apart them from the artificial languages of logic and mathematics.

Speech-act-theoretic semantics retains the chief explanatory advantage of truth conditional semantics and dynamic semantics by accounting systematically for the compositionality of linguistic meaning in terms of the compositionality of speech acts. Just as the meaning of a sentence depends systematically on the meanings of its parts, the properties of sentence-sized speech acts (like the act of asserting that Ludwig is a philosopher) depend systematically on the properties of the word-sized speech acts which are involved in their performance (like the act of referring to Ludwig and the act of predicating the property of being a philosopher of him).

The idea of explicating linguistic meaning in terms of speech acts is not new.[1] Other proposals for speech-act-theoretic approaches to semantics include those of Alston (2000), Barker (2006), Searle (1969). Given particular approaches to speech-act theory, it is also possible to understand

[1] Unless I indicate otherwise, I will reserve the phrase 'speech-act-theoretic semantics' exclusively for my own view.

some versions of expressivism and dynamic semantics as speech-act-theoretic approaches to semantics.

My own approach to semantics differs from each of these views in a variety of ways, but one of the most crucial points of disagreement revolves around their different accounts of the nature of speech acts. Searle and Alston defend variations on the Austinian view that illocutionary acts are constituted, at least in part, by linguistic convention. Expressivism revolves around the idea of identifying speech acts in terms of the kinds of mental states they express. Dynamic semantics is based on the Stalnakerian idea of identifying speech acts by their effects on discourse context. I'll briefly critique each of these conceptions of speech-act theory in §1.2.

Precursors of the kind of theory I'll defend here can be found in the work of Paul Grice (1968), Stephen Neale (2008, forthcoming), and Stephen Schiffer (2003). What ties their ideas together with mine is a Gricean approach to communication and speech-act theory, according to which performing a speech act is a matter of producing an utterance with addressee-directed communicative intentions. The idea that the speech-act types with which I identify expressions' meanings are individuated in terms of the kinds of communicative intentions underlying them will be a crucial component of the semantic accounts that I develop in Chapters Two and Three. But although this idea has been defended by Grice, Neale, Schiffer, and others, nobody until now has attempted to fully articulate speech-act-theoretic semantics, systematically show why it is better than the contemporary alternatives, defend it against objections, or work out the details of how it can help us to understand the meanings of a range of particular linguistic phenomena. These my the goals.

1.2 A Gricean Theory of Speech Acts

Speech-act-theoretic semantics draws on the explanatory resources of a Gricean approach to the mind–language interface, human communication, and—in particular—speech acts.[2] A Gricean theory of communication includes two interwoven components: *the theory of speaker meaning*, which spells out speakers' contributions to communication, and *the theory of interpretation*, which spells out addressees' contributions. Both of these components are psychologistic, in that they explicate the concepts involved in theorizing about meaning and communication in terms of the mental states of speakers and hearers, respectively. To mean something, Grice tells us, is to produce an utterance with some communicative intention. Roughly: a communicative intention is an intention to produce a certain response in an addressee, in part by means of the recognition of that intention.[3] An addressee successfully interprets the speaker if she recognizes their intention to produce that response. The job of a theory of interpretation is to explain the non-demonstrative inferential process by which audience members come to recognize communicative intentions.

The two halves of the Gricean theory of communication are psychologically intertwined and constitutively cooperative. Utterers can intend to affect their addressees only in ways that they believe are possible, and, because communicative intentions are reflexive, in the sense that they involve intending that one's addressee recognize that one intends to produce a certain response,

---

[2] The canonical texts of the Gricean tradition include Bach (1987), Bach and Harnish (1979), Bennett (1976), Grice (1957, 1968, 1969, 1975, 1989), Loar (1981), Neale (1992, 2004, 2005, *forthcoming a*), Schiffer (1972, 1981, 1982), Sperber and Wilson (1986/94), Strawson (1964, 1969), Wilson and Sperber (2012). The version of Gricean speech-act theory I will construct here differs from the accounts given by all of these authors in greater or lesser ways, but it comes closest to the ideas of Grice (1968, 1969), Strawson (1964), and Schiffer (1972: ch.4). I mostly agree with the version of Griceanism that can be found in Neale (1992) and Neale (*forthcoming a*), and much of the present chapter is influenced by those articles and by conversations with Neale.

[3] The precise formulation of communicative intentions is a matter of considerable debate among Griceans; see Bach and Harnish (1979), Bennett (1979), Grice (1957, 1969), Neale (1992), Schiffer (1972), Sperber and Wilson (1986/94), and Strawson (1964) for some of the most influential arguments and proposals.

an utterer can have a communicative intention only if she believes that it is possible for her addressee to recognize that she intends to produce the relevant response. These doxastic constraints on intentions will depend on an utterer's beliefs about how her addressees will interpret her; addressees' interpretations, which proceed via nondemonstrative inference, must be based on their beliefs about utterers' larger plans and other psychological states. One of Grice's great insights was to posit a shared stock of norms that govern both sides of this communicative equation, and indeed all cooperative behavior, by guiding the processes behind both intention-formation and intention-recognition. According to Grice, the overarching norm is *the Cooperative Principle*: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (1989: 26). Grice goes on to formulate a collection of maxims of conversation that govern some forms of communicative cooperation specifically.[4] Although Grice's best-known application of the maxims of conversation was to his theory of conversational implicature, the cooperative principle and the maxims should be understood as guiding all aspects of the psychological processes underlying communication of the sort that Grice was interested in.

A useful and equivalent way of describing communication is in terms of the performance and interpretation of *communicative illocutionary acts*. Following Austin (1962), I'll say that an illocutionary act is an act that one performs in producing an utterance, such as an act of asserting a proposition, asking someone a question, directing someone to do something, marrying someone, voting in an election, and so on. The central idea of Gricean speech act theory is that there is a theoretically significant category of *communicative* illocutionary acts whose performance is constituted solely by the production of an utterance with certain communicative intentions. Within

[4] The status and formulation of the maxims is another matter of controversy among Griceans and other theorists. For some influential revisionary proposals, see Bach and Harnish (1979) and Sperber and Wilson (1986/95). One clear limitation of Grice's formulation of the maxims that I will discuss in Chapter Three is that his maxims of quality and quantity don't apply to non-assertive speech acts.

this genre of speech act theory, there is much disagreement about details, but the theory I'll present here draws on, adjusts, and elaborates ideas that can be found in Grice (1968, 1969), Schiffer (1972), and Strawson (1964). To perform a communicative illocutionary act, on this view, requires meeting the following schematic condition:

(GI) Gricean Illocution

S performs an illocutionary act α in uttering *u* only if *S* utters *u* intending:

(1) to produce thereby a certain response Δ in a certain addressee A;

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly based on of her recognition of (1) and (in some cases) partly on the basis of some further reason ρ.

(GI) is not sufficient for the performance of an illocutionary act by anyone's lights. Strawson (1964) adds a fourth clause requiring A to also recognize S's intention (2). Schiffer, after arguing that this is still not sufficient, argues that the definition must be infused with the notion of mutual knowledge (1972: chs.1–3). Grice also adds further conditions to block counterexamples (1969). I propose that we take (GI) as a necessary condition on the performance of an illocutionary act, and stipulate that a theory of illocution counts as Gricean only if it builds on (GI) by the addition of further conditions or elaborations.

(GI) provides framework within which to give a constitutive account of the content and force of communicative illocutionary acts. Given that S exemplifies (GI) and thereby performs an illocutionary act α, the content *p* and force *F* of α will be constituted by features of the response Δ that S intends to produce in her addressee, as well as (in some cases) the further reasons ρ on the basis of which she intends to produce Δ in A. Specifically: *p* will be determined wholly by some feature of Δ, and F will be determined by some combination of features of Δ and ρ. This is all rather abstract; here are some examples that would fit into the framework:

(GI1) Asserting that *p*

S asserts that *p* in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby in A a belief that (A believes) *p* (Δ);[5]

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and A's belief that S knows that *p* (ρ).

(GI2) Suggesting that *p*

S suggests that *p* in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby in A an act of considering whether *p* (Δ);

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and A's belief that S has a non-negligible credence in *p* (ρ).

[5] The question of whether asserting *p* requires intending to produce in one's addressee (a) a belief that *p* or (b) a belief that one believes that *p* is a matter of contention among Griceans. Grice originally went with (a) (1957) but changed his mind to (b) in later work (1968) in order to deal with putative counterexamples wherein a speaker apparently asserts *p* despite the fact that her beliefs rule out the possibility of convincing her addressee to believe *p*. But there are other ways to handle such cases. For example, it is plausible that we sometimes perform speech acts involving intentions of type (a) and sometimes perform speech acts involving intentions of type (b). Moreover, it seems plausible that colloquial usage papers over this theoretically interesting distinction by allowing us to call both types of speech act 'assertions'. These considerations illustrate the broader point, seemingly not fully appreciated by Grice, Strawson, and Schiffer (whose work, after all, grew out of ordinary language philosophy) that we should not expect the theoretically interesting distinctions between types of illocutionary act to be reflected in the colloquial vocabulary that we use to talk about speech acts. Folk speech-act theory may be a confused mess, and a theoretically useful taxonomy of types of communicative illocutionary act and the intentions that underpin them should ultimately be based on the description and explanation of variation in the ways that we communicate rather than our everyday ways of talking about talking.

(GI3) Ordering that *p*

S performs an orders A to *p* in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby in A an intention that p[6] (Δ);

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and A's belief in S's authority to require S to *p* (ρ).

---

[6] The careful reader might balk at my way of formulating ascriptions of intentions and directive speech acts using that-clauses, and might prefer Schiffer's (1972) way of schematically representing intentions and directive acts: 'A intends to ψ' and 'S directs A to ψ', where ψ ranges over properties rather than propositions. This reader may be correct, and nothing about my broader theory would stop me from going along with the suggestion, but the issue is complicated, and requires a brief digression about the LFs of propositional attitude reports with infinitive complements and first-personal thought contents.

The LF of a standard intention ascription 'A intended to dance' can be represented as [A_i intended [PRO_i to dance]], where the matching subscripts on 'A' and 'PRO' represent that the two expressions bear a semantically special relationship to one another which could be thought of as a kind of *de jure* coreference or perhaps binding. The result of this special relationship is that intention reports seem to have an *essentially indexical* or *first-personal* quality, which can be seen from the fact that 'S intends to Dance' has different truth conditions than 'S intends that S dances' (since the latter could be true even if S doesn't realize that he himself is S). It seems to follow that we need a special treatment of intention reports (and, similarly, for desire reports, and so on). One response to this explanandum is to posit a special kind of first-personal proposition and say that intending sometimes relates a speaker to this kind of proposition (*cf.* Stanley 2011: ch.3). Another option is to say that essentially first-personal contents are self-locating properties (*cf.* Lewis 1979b). If we take the former view—that the contents of intentions are propositions, albeit sometimes propositions of a special kind—then I maintain that my use of that-clauses is fine, so long as it is understood as technical, and so long as we keep in mind that not just any proposition can be slotted in for 'that *p*'. If the contents of intentions are self-locating properties, on the other hand, then my locutions are quite misleading, and usage more like Schiffer's should be preferred. I will stick to the awkward-sounding locution in what follows, with the proviso that I could easily change things around if need be. (A slightly modified version of all of this goes for directive act reports as well.)

This issue will turn out to be intimately connected to the issue of whether the mood-free cores of imperative sentences are sentence radicals that can be used to express propositions or predicates that can be used to refer to properties. As it turns out, this is a matter of some controversy in semantics at the moment (see footnote 3 in Chapter Three). One reason that I provisionally adopt the assumption that the contents of intentions and directive acts are propositions, together with my awkward-sounding technical usage in reports of intentions and directive speech acts, then, is that these simplifying assumptions go well with my simplifying assumption in Chapter Three that imperatives contain mood-free sentence radicals. But again: little will hang on that assumption, which I will adopt mainly because it allows me to compare my theory to others more easily.

(GI4) Advising that p

S suggests that A *p* in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby in A an intention that p (Δ);

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and A's recognition of S's belief that *p* is in S's interest (ρ).

(GI5) Asking whether

S asks whether *p* in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby in A an intention to inform S whether *p* (Δ);

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and A's recognition of S's desire to know whether *p* (ρ).

(GI6) Asking *wh-*

S asks *wh-* is *F* in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby in A an intention to inform S that *x* is F (Δ);[7]

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and A's recognition of S's desire to know *wh-* is F (ρ).

These examples are loosely based on the ideas of Grice (1968), Schiffer (1972: ch.4), and Strawson (1964), but mix together and deviate from all of their accounts in some respects. There are many

[7] Even more than the (GI1–5), (GI6) must be treated as particular provisional, both because different kinds of wh-questions place different constraints on good answers and because of the difference between *mention some* and *mention all* readings of wh-questions (see, e.g., Groenendijk and Stokof 1982). (GI6) is at best a first shot at an explication for mention-some questions. It's likely that different explications will have to be given for mention-all and mention-some questions relative to each type of wh- question.

details over which we could (and probably should) continue to debate. (GI1–6) should be treated as a rough first shot at explicating a few types of illocutionary acts, articulated for the purposes of illustrating the overall Gricean approach.

(GI1–6) illustrate several significant points about the Gricean theory of force. First, (GI1–6) are all (partial specifications of) kinds of Gricean communicative intention. Performing an illocutionary act is, for a Gricean, a way that a speaker can mean something.

Second, illocutionary acts can be divided into the mutually exclusive categories of *assertive acts* (*assertives*) and *directive acts* (*directives*) on the basis of whether the response Δ that a speaker intends to produce in an has a mind-to-world (cognitive) direction of fit or a world-to-mind (connative or action-guiding) direction of fit, respectively (*cf.* Grice 1968; Schiffer 1972: 95). The term 'assertive act' is somewhat misleading: although assertion is one kind of assertive act, there are other assertive acts that aren't assertions—such as (if these are all genuinely distinct kinds of illocutionary act): acts of supposing, suggesting (that *p*), assuring, reporting, estimating, hypothesizing, and so on. What unites all of these assertive acts is that their characteristic communicatively intended response Δ is a *cognitive* mental state of some kind. Performing them, that is, requires communicatively intending to produce a cognitive response, such as a belief, an act of supposing or considering-whether, or a change in credence. In the above explications, for example, asserting involves intending to bring about a belief and suggesting requires intending to bring about a mental act of considering whether. Assertives stand in contrast with directives, whose performance requires intending to bring about a non-cognitive, connotative, or action-guiding

psychological state of some kind.[8] I have followed Grice (1968, 1969) in taking these action-guiding mental states to be intentions.[9] To command, request, advise, instruct, and so on, are all, in part, to communicatively intend to produce an intention in one's addressee; the content of this communicative intention is the content of one's directive speech act.

Third, we can draw more finely grained distinctions within each of the broad categories of assertoric and directive illocutionary acts based on features of both the type of response Δ that the speaker communicatively intends to produce and the reasons ρ on the basis of which she intends to produce Δ.[10] Within the broad category of assertive speech acts, for example, asserting *p* is distinguished from suggesting that *p* on the basis of both Δ ('belief that (...) *p*' versus 'consideration whether *p*') and ρ ('belief that the speaker knows *p*' versus 'belief that the speaker has non-negligible credence that *p*'). Within the broad category of directives, commanding A to *p* is distinguished from advising A to *p* solely on the basis of ρ ('recognition of the speaker's authority...' versus 'belief that the speaker believes that *p* is in A's best interest').

Fourth, questions are a sub-genre of directives: to ask a question is to direct an addressee to supply one with information. Whereas performing a non-question directive whose content is *p*

---

[8] This way of drawing the distinction between assertoric and directive illocutionary acts is not universal to all Gricean accounts. For example: Bach and Harnish distinguish the similar (but not quite coextensive) categories of *constative* and *directive* illocutionary acts on the basis of whether they are intended to express the speaker's belief or his "attitude toward some prospective action by the hearer" (1979: 47). Although I am doubtful of the utility of Bach and Harnish's notion of expressing a mental state, I have no special reason to think that their version of Gricean speech act theory is incompatible with the kind of speech-act-theoretic semantics that I will defend in later chapters. One reason for adopting the Grice–Strawson–Schiffer approach rather than the Bach–Harnish approach is that the former makes comparing Gricean and dynamic theories of mood and force particularly straightforward, and this comparison will be at the heart of Chapter Three.

[9] Schiffer adopts an alternative approach on which performing a directive speech act requires intending to bring about one's addressee's action itself (1972: ch.4). Other possibile intended responses that would fit within the same general framework are desires or obligations of some kind, although I don't regard either of these two options as particularly attractive.

[10] Here I follow Schiffer quite closely (1972: ch.4, §2). My 'Δ' corresponds to Schiffer's '*p*' (in the case of assertorics) and 'ψ' (in the case of directives), and my 'ρ' corresponds to Schiffer's '$\rho(t)$'.

requires communicatively intending to produce an intention to being about *p* in one's addressee, and particular types of illocutionary act within this genre are distinguished on the basis of the value of ρ, questioning requires communicatively intending to produce an intention to inform one whether *p* (or *wh-* is F) in one's addressee. Different types of questioning, including the distinctions between asking whether and asking *wh-* as well as the distinctions between various kinds of asking *wh-*, can be distinguished on the basis of the nuances of Δ. There may also be different types of questioning that are distinguished on the basis of ρ; for example, whereas *asking* involves intending that the addressee answer because of the speaker's desire for information (making asking akin to requesting information), it is plausible that the act of *interrogating* also involves intending that the addressee answer because of the speaker's authority, or in order to avoid unpleasant consequences of not answering (making interrogating akin to commanding that one's addressee provide information) (*cf.* Schiffer 1972: 103).

All of the illocutionary acts I have discussed so far are either assertive or directive acts, and I have claimed that these two kinds of acts are mutually exclusive. Do they also mutually exhaust the category of illocutionary acts? This is a matter of controversy. Clearly, there are speech acts that have traditionally been thought of as illocutionary acts but that don't fit into the framework I've outlined so far. One kind of example ivolves some purported illocutionary acts, such as the act of greeting someone by saying 'Hi!' or the act of insulting someone by saying 'Screw you!', that don't seem to have contents in the same way that the illocutionary acts I've discussed have content. (Of course, *wh*-questions don't have propositional contents, but they seem to have contents of *some* kind—perhaps *types* or *sets* of propositions). In at least some cases, these contentless speech acts also don't seem to involve intending to produce either a cognitive or action-guiding attitude in an addressee. It might make sense to follow Searle in holding that greeting involves intending to produce the knowledge in the addressee that she is being greeted (1969: 46)—thereby making greeting an assertive speech act by the criteria I've outlined. And we might be

tempted to extend this account to insults, taking them to be (in part) a matter of communicatively intending to produce in an addressee the recognition of being insulted. But greeting, insulting, and some other speech acts seem aimed more at producing responses that are neither cognitive nor action-guiding—such as, perhaps, the feeling of being welcome or offended. If this is right, then the range of values for Δ in the above schema may have to be expanded to include affective or phenomenological responses. This would mean that assertives and directives do not mutually exhaust the illocutionary acts, but it poses no threat to the framework as a whole.

Some illocutionary acts—including many of Austin's favorite examples, such as the act of marrying someone by uttering 'I do'—can be performed only if the right social or institutional background conditions obtain. Following Bach and Harnish (1979), I will call these *conventional illocutionary acts*, to be distinguished from communicative illocutionary acts like asserting, asking, and requesting. The existence and purported extent of conventional illocutionary acts has often featured in objections to Gricean speech act theory, in part because they seem to conflict with the Gricean idea that what we do with language is a matter of our intentions, and in part because many speech act theorists—most prominently Austin (1962) and Searle (1969)—have claimed that all, or nearly all illocutionary acts should be thought of as conventional because they are constituted by *linguistic* conventions. This is a claim that I wish to wholeheartedly deny.

Austin's focus on paradigm cases of conventional illocutionary acts, such as marrying someone and christening a ship, is seemingly what led him to think of illocutionary acts in general as deeply conventional phenomena. This is also what has led most Griceans to bracket off conventional acts as a special case that is not of particular interest to the study of human communication. Indeed, Schiffer finds it tempting to deny that convention-constituted speech acts are genuine illocutionary acts at all (1972: ch.4), presumably because he is not convinced that they form a unified explanatory category together with the communicative illocutionary acts that are his primary focus. I see the distinction between this view and Bach & Harnish's as a verbal one. On the

latter view, conventional illocutionary acts are seen as a distinct sub-genre of illocution, distinguished specifically by the fact that their performance requires that background conditions beyond the speaker's intentions obtain. Convetional illocutionary acts do not lack theoretical interest. Indeed, I believe that understanding the nature of conventional illocution will be key to solving a wide variety of philosophical problems about speech that is native to institutional contexts, including special problems about the nature of legislative speech and legal interpretation, confession, and legal testimony.[11] The reason for bracketing conventional illocutionary acts in the study of human communication, then, is not that they are theoretically uninteresting. Rather, conventional acts should be bracketed because they introduce theoretical elements that may be distracting and misleading when it comes to human communication in its paradigmatic form.

An illustration of how conflating conventional acts with communicative acts can make a mess of one's understanding of speech acts is one of Thomason's (1990) arguments that performing a speech act involves intending to produce an effect, not in an addressee, but instead in a *conversational record*—a common stock of mutual presuppositions shared by the participants in a conversation (cf. Stalnaker's (1978) *common ground*).[12] Thomason argues for this dynamic account of assertion rests on the basis of an analogy between normal assertion and formal testifimony in court. When testifying, a witness's goal usually isn't to affect the beliefs of whomever we're addressing. A witness who's been well-prepared by the lawyer who is asking the questions won't be

---

[11] These issues are illuminated in the dissertation of my colleague, Rachel McKinney, which is still in progress. In forthcoming work, McKinney and I draw out the implications of the communicative–conventional distinction for the understanding of police interrogation, false confession, and other forms of apparently unintended illocutionary acts.

[12] The idea that speech acts can be individuated in terms of their effects—intended or otherwise—on the shared context of conversation is one of the foundational ideas of the dynamic theory of speech acts, which has recently come to dominate pragmatic theory, particularly among linguists. I will discuss dynamic pragmatics in detail in Chapter Three (§3.4), where I'll argue that dynamic accounts collapse into Gricean accounts precisely because performing a speech act involves intending to change the mental states of an addressee rather than the mutually represented context of conversation.

saying anything that the lawyer doesn't expect, for example. Instead, the goal of a witness should be to place certain information onto the official court record. Thomason argues that something similar is going on in normal assertion—that our primary aim to update the public, mutally represented conversational record rather than any particular addressee's mental states. But this analogy between testimony and assertion is misleading precisely because testimony, unlike assertion, is a conventional illocutionary act—an act whose performance is possible only in an elaborate set of institutional circumstances, and whose purpose is defined in part by the goals of the institutions that give rise to those circumstances. We should therefore expect that to testify in court is to do something very different than to assert. This explains why the intentions behind one would be so different than the intentions behind the other.

Of course, there is plenty of room to disagree about the exact placement of the boundary between purely intention-constituted speech acts and those that are partly constituted by social or institutional conditions. One of the broadest and best-known disputes about the nature of both illocutionary acts and speaker meaning is best understood as a dispute over this boundary. Although some of the core ideas of the Gricean, intention-based theory of communicative illocutionary acts are at least suggested by Grice (1957), the first explicit articulation of the theory is Strawson's (1964). In a well-known response to Strawson, Searle protests that the Gricean picture vastly underestimates the role of convention in the performance of illocutionary acts, and also in the nature of speaker meaning itself:

> [Grice's theory of speaker meaning] fails to account for the extent to which meaning can be a matter of rules or conventions. This account of meaning does not show the connection between one's meaning something by what one says, and what that which one says actually means in the language. [...]
>
> Suppose that I am an American soldier in the Second World War and that I am captured by Italian troops. And suppose also that I wish to get these troops to believe that I am a German soldier in order to get them to release me. What I would like to do is to tell them in German or Italian that I am a German soldier. But let us suppose that I don't know enough German or Italian to do that. So I, as it were, attempt to put on a show of telling them that I am a German soldier by reciting those few bits of German I know, trusting that they don't know enough

> German to see my plan. Let us suppose that I know only one line of German which I remember from a poem I had to memorize in a high school German course. Therefore I, a captured American, address my Italian captors with the following sentence: *Kennst du das Land wo die Zitronen blühen?* [*fn*. If it seems implausible that one could intend to produce the desired effects with such an utterance in these circumstances, a few imaginative additions to the example should make the case more plausible, e.g., I know that my captors know that there are German soldiers in the area wearing Italian uniforms, I know that they have been instructed to be on the lookout for these Germans and to release them as soon as they identify themselves. I know that they have lied to their commander by telling him that they can speak German when in fact they cannot, etc.] Now, let us describe the situation in Gricean terms. I intend to produce this effect of believing that I am a German soldier, and I intend to produce this effect by means of their recognition of my intention. I intend that they should think that what I am trying to tell them is that I am a German soldier. But does it follow from this account that when I say, *Kennst du das Land…etc.*, what I mean is, "I am a German soldier"? Not only does it not follow, but in this case I find myself disinclined to say that when I utter the German sentence what I mean is "I am a German soldier", or even "Ich bin ein deutcher Soldat", because what the words mean and what I remember that they mean is "Knowest thou the land where the lemon trees bloom?" Of course, I want my captors to be deceived into thinking that what I mean is: "I am a German soldier", but part of what is involved in that is getting them to think that that is what the words I utter mean in German. In the *Philosophical Investigations*, Wittgenstein (discussing a different problem) writes "*Say* 'it's cold here' and *mean* 'it's warm here'" [§510]. The reason we are unable to do this without further stage setting is that what we can mean is at least sometimes a function of what we are saying. Meaning is more than a matter of intention, it is also at least sometimes a matter of convention. (1969: 44–5)

Searle admits that illocutionary acts aren't necessarily convention-based, since "some very simple sorts of illocutionary acts can indeed be performed apart from any use of any conventional device at all, simply by getting the audience to recognize certain of one's intentions in behaving a certain way" (1969: 38). But he takes such cases to be both rare and marginal, and so outlines a theory of illocutionary acts according to which they are typically constituted by linguistic rules, much as the act of playing chess is constituted by the rules governing the movement of chess pieces. The best summation the Gricean response to this argument that I know of comes, ironically, from Searle himself, who apparently intends it as the linchpin in his *reductio ad absurdum* of the Gricean view:

> One might say that on Grice's account it would seem that any sentence can be uttered with any meaning whatever, given that the circumstances make possible the appropriate intentions.

> But that has the consequence that the meaning of the sentence then becomes just another circumstance. (1969: 45)[13]

But for the Gricean, the meaning of one's utterance-type really *is* "just another circumstance" in the relevant sense. The role of utterance-type meaning, much like the roles played in communication by other circumstances external to the performance of a speech act itself, is to enable the addressee to infer the speaker's intentions, and thereby to enable the speaker to have those intentions (given her estimate of what the addresse is capable of working out). The meanings of the expressions we utter needn't be constitutive of what we say or otherwise do in uttering them in order to play this dual role. And, as I will argue in Chapter Two, there are excellent reasons to think that expressions' meanings aren't even partly constitutive of what we say or otherwise do with them.[14]

---

[13] As Neale (*forthcoming a*: *fn*.27) points out, it can be seen from the two quoted passages that Searle's objection to Griceanism is based, at least in part, on the conflation of several importantly distinct notions of meaning. For his objection to an intention-based account to hold any weight, it must be intended as an objection to Grice's notion of *utterer's occasion meaning* (a.k.a. 'speaker meaning'). But some of what Searle says suggests that what he is really objecting to is the idea that a speaker's intentions can override what Grice would call the timeless meanings of utterance types or the applied timeless meanings of utterance types relative to particular occasions of use (on the distinction between these three notions, see Grice 1969). One clue is that Searle attributes meaning using direct quotation. Another is the locution that Searle employs in summing up his point (emphasis added): "on Grice's account it would seem that *any sentence can be uttered with any meaning* whatever". This phrasing suggests that what Searle is really interested in is whether a speaker's intentions can determine what a sentence means, either in general or relative to a particular occasion. But the distinction between utterer's meaning and (applied) timeless meaning is absolutely central to the Gricean project, and, I think, Searle's argument becomes less intuitive if we attempt to rephrase his point to make it squarely about speaker meaning.

[14] For a nice historical discussion of Grice's aversion to both convention and the terminology of speech-act theory, see Chapman (2005: 76). I think of speech acts (and communicative acts more generally) as a topic in pragmatics and the philosophy of language that admits of competing approaches, including the convention-based tradition of Austin (1962) and Searle (1969), the intention-based, Gricean approach outlined here, the expressivist approach of (e.g.) Davis (2003) and Green (2007), the dynamic approach of (e.g.) Portner (2004), Roberts (1996/2012), and Stalnaker (1978), and the conceptual-role approach of (e.g.) Brandom (1998), Sellars (1969), and Vendler (1972), and others. Chapman suggests that Grice avoided the terminology of speech-act theory because he took it to be native to Austin's own, convention-laden approach to the topic.

I want now to draw out some interesting consequences of the the fully intentionalist, Gricean account of communicative illocutionary acts that I've sketched. One is that the theory, as stated so far, has nothing to do with either language or utterance-type meaning, and therefore nothing to do with linguistic meaning. At least in principle, on my view, it is possible to perform any communicative illocutionary act with an utterance type that is neither linguistic nor meaningful. All that's required is that one have the right communicative intentions, and all this requires is that the speaker takes her addressee's present psychological state to be such that she can produce the right sort of response. As I'll argue in §1.3, the roles of language and linguistic meaning lie in their capacity to explain how certain utterance types—the linguistically meaningful ones—make it routine for communicators to efficiently perform and interpret extremely information-rich speech acts in a way that is less wedded to their particular circumstances.

A second interesting consequence of intentionalism about speech acts is that the individuation conditions of utterances and illocutionary acts differ in such a way that it is possible to perform more than one illocutionary act with a single utterance. An *utterance*, on this view, can be thought of as any piece of behavior that is designed and intended for the interpretation of one or more addressees.[15] But illocutionary acts are individuated in terms of the more fine-grained communicative intentions with which an utterance is produced, and it is possible to produce a single utterance with more than one communicative intention, thereby performing more than one communicative act. These distinct communicative acts may both be addressed to the same per-

[15] An addressee may be a particular person or group of people about which the utterer has a de re communicative intention, or it may be one of many people who meet a condition specified in the utterer's de dicto communicative intention, as when I record my answering machine message intending to produce an intention to leave me a message in *whomever calls my phone and gets the voicemail*.

son—as when one both says something and implies something else—or they may be aimed at distinct addressees. Consider the following example:[16]

> Party Yawn
>
> Dan is at a party with Smith and Jones. Dan and Jones have discussed the fact that they aren't tired, but are getting bored, would like to go to another party, but don't want Smith to come. Smith is oblivious to all of this. In full sight of both of both Smith and Jones, Dan produce an utterance consisting of an exaggerated and obviously fake yawn and glance at his watch. In doing so, Dan have two distinct communicative intentions: (i) Dan communicatively intends to produce a belief in Jones that he believes it's time to go to the next party, and (ii) Dan communicatively intends to produce a belief in Smith that he (Dan) believes it's time to go home and sleep.

In this case, Dan performs two communicative acts by producing a single, nonlinguistic, utterance: he addresses to Jones a nonlinguistic assertion that it's time to go to the next party, and he addresses to Smith a nonlinguistic assertion that that it's time for him to go to bed.[17] That Dan is able to do this results from the fact that he takes Jones and Smith to have very different background information on the basis of which they'll interpret him.

The Party Yawn case illustrates two other significant points about communicative acts. First, communicative acts needn't have anything to do with either language or utterance-type meaning. The point about language is obvious enough. The point about utterance-type meaning is less ob-

---

[16] Stephen Neale has presented a similar example in talks and seminars in order to make the point that the saying–implicating distinction does not depend on language. Some aspects of this example were suggested to me by my colleague, Marilynn Johnson.

[17] It might be better to describe what's going on in Party Yawn by saying that Dan indirectly asserts to Smith that it's time to go home by directly asserting to Smith that he's tired, and that he indirectly asserts to Jones that he wants to go to the next party by making as if to indirectly assert to Jones that it's time to go home, which he in turn does by making as if to directly assert to Jones that he's tired. I ignore some of this complexity in the text.

vious, but it plausibly follows from the fact that both of Dan's speech acts might have led to successful communication in the Party Yawn case even if Dan, Jones, and Smith belonged to a community whose members had developed no special conventions or dispositions for communicating anything in particular by yawning or looking at their watch. Dan's intentions, in that case, could still be worked out merely by recognizing that people are tired when they yawn, and that they look at their watch when they're worried about how late it's getting.[18] The other point that Party Yawn illustrates is that communicative acts can be either direct or indirect: Dan informs Jones that it's time to go to the next party by *making as if to* address to her a nonlinguistic assertion that it's time to go home and sleep. He indirectly performs one communicative act by making as if to perform another one. But Dan's act of nonlinguistically asserting to Smith that it's time for him to go home and sleep is not indirect in this way.

Taken together, the two points in the last paragraph illustrate two other interesting aspects of Gricean speech-act theory. First, Party Yawn illustrates the fact that communicative illocutionary acts have their force and content only relative to some addressee(s). In the vocabulary of Wilfrid Sellars (1969)—but contrary to his theory of speech acts—every communicative illocutionary act is a *talking-to*, and not a mere *expression of thought*. If this weren't the case, then we would be forced to say that Dan both asserts and makes as if to assert that it's time to go home and go to bed in the Party Yawn case. Griceanism easily handles the addressee-relativity of communicative acts by relativizing the communicative intentions that constitute communicative acts to addressees. In addition to Sellars, many alternative approaches to speech-act theory lack the resources to explain this kind of addressee-relativity.[19]

---

[18] Roughly the same point—that the notion of speaker meaning is prior to both language and utterance-type meaning—is made by Grice (1968) and Schiffer (1972: ch.3; 1987: §9.3).

[19] For example, Bar-On (2004), Davis (2003), Green (2007), and Pagin (2010) all defend versions of speech-act theory in which the content and force of a speech act are constituted by features of the psychological states that are thereby *expressed* or *indicated*, where these are not taken to be addressee-relative notions.

Party Yawn also entails that the distinction between direct and indirect communicative acts is explanatorily prior to linguistic and semantic notions. It is typical, even among Griceans, to explicate the distinction between direct and indirect speech acts, and, in particular, the distinction between what a speaker says and what she implicates (which I view as a special case of the direct/indirect distinction), partly in terms of either linguistic or semantic notions. For example, Schiffer holds that what a speaker says in uttering a sentence is the part of what she means that (unlike what she implicates) "fits" the sentence's meaning (1972). In a similar vein, Sperber and Wilson argue that the explicit portion of what a speaker means in uttering a sentence should be distinguished from the implicit portion by virtue of the fact that the former, but not the latter, is a "development" of the sentence's logical form (1986/94).[20] But accounts like this suffer from serious problems, not the least of which is the difficulty of giving an account of what it means for a proposition to 'fit' a sentence's meaning or 'be a development of' its logical form. Further tension for these views arises from malapropisms and slips of the tongue, as when a speaker utters, 'I hissed my mystery lecture', with the intention of asserting that she missed her history lecture. As Neale has argued, we have good reason for thinking that what the speaker says in this case is just what she intended to say, rather than some proposition that fits or develops the meaning of the sentence she utters (2004, 2005, *forthcoming a*). This full-blown intentionalism follows from Grice's view that one says something only if one means it (1989: *ch.*5). Moreover, as Neale points out, there are good reasons to go along with Grice on this issue, since the most theoretically useful notion of what a speaker says will make it the object that a hearer must identify in order for communication to succeed. If I utter, 'I hissed my mystery lecture', intending to say that I missed my history lecture, I will succeed in communicating only if my addressee charitably treats me as saying what I intended to say, rather than what I may superficially appear to have said. Moreover,

---

[20] See also Bach (1987, 2005) and Devitt (*forthcoming*) for the view that what a speaker says in uttering a sentence is partly constituted by the sentence's meaning. I will discuss their views in greater detail in Chapter Two.

this case shouldn't be treated as one in which I made as if to say that I hissed my mystery lecture in order to implicate that I missed my history lecture. After all: I might have intended merely to make as if to say that I missed my history lecture, in order to implicate something further—for example, that I don't remember the dates of the French Revolution. But I didn't make as if to say that I hissed my mystery lecture in order to make as if to say (or make as if to implicate) that I missed my history lecture in order, in turn, to implicate that I don't remember the dates of the French Revolution. And, if I really did *say* that I missed my history lecture, then views like Schiffer's and Sperber and Wilson's have a serious problem. Unless one adopts the view that the meaning or logical form of the sentence a speaker utters can itself be sensitive to her intentions (thus, seemingly, commiting oneself to an unattractive form of grammatical Humpty Dumptyism), it is hard to see how what I said can be a development of the logical form of the sentence I actually uttered.[21] Finally, if agents can perform both direct and indirect communicative acts without using language or meaningful utterance-types at all, then there is reason to think that the saying/implicating distinction, as well as the distinction between direct and indirect speech acts, are just special cases of the broader distinction between direct and indiect communicative acts, and that, therefore, all of these distinctions can be made without appealing to linguistic or semantic notions.

If all of this is right, then how should we draw these distinctions? My answer is that performing an indirect communicative act requires having different kinds of intentions than performing a direct communicative act. Roughly, to perform an indirect communicative act with the content *p* requires, in addition to the usual communicative intentions, that one's addressee recognize one's intention to bring about a mental state whose content is *p*, that one intend them to recognize this

[21] Intentionalism about saying is often attacked on the grounds that it is just too darn unintuitive that I could say something that doesn't fit with the meaning of the sentence I utter. Where does the strength of this intuition come from?

partly on the basis of their recognizing a second communication intention with which one performs the speech act, and in virtue of which one performs a direct speech act.[22]

(MAI) Making As If

S performs an illocutionary act α by making as if to perform an illocutionary act β in uttering *u* iff:

(i) β is an illocutionary act of the type that can be performed by communicatively intending to produce a response γ in an addressee, in part on the basis of the recognition of this intention and in part on the basis of some further reason σ; and

S utters *u* intending:

(a) to produce thereby in an addressee A an act of entertaining, and rejecting, the possibility that S intends to produce γ in A partly on the basis of σ.

(1) to produce a certain response Δ in A;

(2) A to recognize S's intentions (a) and (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (a) and (1) and (in some cases) partly on the basis of some further reason ρ.

An indirect speech act is a speech act that one performs by either performing or making as if to perform some other speech act. (We can say that S performs an illocutionary act α by performing another illocutionary act β just in case (1) S performs both α and β, and (2) S meets a modified version of (MAI) such that condition (a) is replaced by whatever S would have to intend in order to perform β.

Grice's notions of saying and implicating are special cases of direct and indirect illocutionary acts, respectively. There are several ways of specifying how these special cases are delimited that would fit with Grice's usage, but here is one: to *say* is to perform a direct act of asserting by means

[22] The formulation of (MAI) is influenced by unpublished work by Jessica Keiser.

an utterance of some meaningful, linguistic type, and to implicate is to perform an indirect act of asserting. Ultimately, I don't think that it matters very much how saying and implicating are explicated, because the notions of direct and indirect communicative illocutionary acts that I have explicated are more general and more theoretically interesting.[23] What makes an utterance type linguistic, on my view, is that it bears a special relation to either the speaker's language faculty (Chomsky 1986) or to a productive system of other utterance-types (Devitt 2007; Lewis 1975a). (Although I am tempted toward the Chomskian half of this disjunction, I can officially remain neutral between Chomsky's psychologism, Devitt's nominalism account, and Lewis's platonism about the metaphysics of language.) All of the concepts involved in explaining communication that I have outlined so far, including the distinction between direct and indirect speech acts, are explanatorily prior to the notion of utterance-type meaning. Semantic concepts and linguistic concepts, are, moreover, explanatorily independent: neither is needed to explain the other, and each is needed to explain concepts for which the other is not needed. The two collections of concepts come together in the notion of linguistic meaning, which is explanatorily posterior to both.

What do I mean by saying that some concept or collection of concepts is *explanatorily prior to* or *explanatorily independent of* another one? I have in mind the following definitions:

Weak Explanatory Priority ($\leq$)

c_1 is weakly prior to c_2 ($c_1 \leq c_2$) iff c_2 is to be explicated, in part, in terms of c_1.

(Strong) Explanatory Priority ($<$)

c_1 is strongly prior to c_2 ($c_1 < c_2$) iff $c_1 \leq c_2$ but $c_2 \not\leq c_1$.

[23] It is plausible that the notion of saying still has an important role to play in the methodology of semantic theory, since speakers seem to have robust intuitions about what a speaker said, as opposed to what she merely implied. For doubts about the usefulness of these intuitions, see Cappelen and Lepore (2005).

Explanatory Connectedness ($\approx$)

c_1 is explanatorily connected to c_2 ($c_1 \approx c_2$) iff $c_1 \leq c_2$ and $c_2 \leq c_1$.

Explanatory Independence ($\not\approx$)

c_1 is explanatorily independent of c_2 ($c_1 \not\approx c_2$) iff $c_1 \not\leq c_2$ and $c_1 \not\leq c_2$.

In saying that psychological concepts are explanatorily prior to semantic ones, for example, I mean that they are *stongly* explanatorily prior, in the sense outlined here. Griceanism thus posits an explanatory hierarchy of the key concepts in the philosophy of language and mind, starting from a collection of psychological concepts that are more basic than the rest, and explicating new concepts from there on up as new explanatory roles arise for them.[24] Although I won't attempt to argue for each priority claim here, I see this Gricean hierarchy mapped out as follows:

$$C_1 < C_2 < C_3 < C_4 < \left.\begin{matrix} C_5 \\ \not\wedge \\ C_A \end{matrix}\right\} < C_{6B}$$

(C_1) Intentionality, mental content, mental attitude (belief, intention, etc.).

(C_2) Higher-order intentionality.

(C_3) Communicative intentions, inferential communication, speaker meaning, interpretation, communicative illocutionary acts, illocutionary act content, illocutionary force, referring, predicating.

(C_4) Direct communicative acts, indirect communicative acts, saying*, implicating

[24] Of course, Griceans are free to think that psychological concepts should themselves be explained in terms of still other concepts. All that Griceans are committed to vis-a-vis these concepts is that the best accounts of them won't appeal to linguistic or semantic notions.

(C_5) Communicative dispositions, utterance-type meaning, speech communities.

(C_A) Language, syntactic structure, sentence, word, linguistic expression, etc.

(C_{6B}) Linguistic meaning, (word meaning, sentence meaning, phrase meaning)

So far, I have said nothing about how utterance-type meaning and linguistic meaning fit into the Gricean hierarchy, except to say that they are explanatorily posterior to the concepts involved in explaining folk psychology and communication, and that they play no role in explicating those concepts or the distinctions between them. I will spell out these connections in §1.3.

1.3 The Explanatory Role of Meaning in Communication

In order to elucidate the nature of utterance-type meaning, and linguistic meaning in particular, we must first identify the explanatory roles of these concepts as they occur in some broader theoretical enterprise. This methodological outlook is nicely summed up in what I'll call *Lewis's dictum*:

> In order to say what a meaning is, we may first ask what a meaning does, and then find something that does that. (1970: 195)

Lewis's dictum recommends a broad methodological policy: we should investigate the nature of meaning by first investigating its explanatory role. This policy is a sound one because meaning is a theoretical posit. Like other posits, our best and only way of finding out about meaning is by giving it a role to play in a theory that explains something else. Lewis's dictum, and the methodological policy it sums up, is a methodological cornerstone in much of what follows.

What is the explanatory role of meaning? In Lewis's terms: what does meaning do? Lewis attempts to answer this question by saying that meaning "is something that determines the conditions under which a sentence is true or false" (1970: 195). One serious problem with this answer is that it apparently does not apply to nondeclarative sentences, which don't seem to have truth

conditions. Lewis attempted to address this worry by reducing the semantic properties of nondeclaratives to those of declaratives, but, as I'll argue in Chapter Three, Lewis's solution does not work. Even if we restrict our attention to declarative sentences, however, Lewis's answer is still problematic. Truth deflationists, for example, might accept that (some) sentences' meanings determine their truth conditions, but would deny that this fact has any explanatory significance. Moreover, as I'll argue in Chapter Two, and as others have previously argued,[25] the meanings of declarative sentences typically underspecify the propositions that they're used to express, and this underspecification cannot be papered over by relativizing a sentence–content relation to contexts or indices, as Lewis and other truth-conditional semanticists have attempted to do.

I will therefore posit a different explanatory role for meaning—one that even Lewis is committed to. This is the role of facilitating communication. Whatever other reasons we have for positing meaning, we certainly need it in order to explain how speakers are able to communicate so much more efficiently with some types of utterances (i.e., the meaningful ones) than with others.

For example, I can walk into any room full of English-speaking strangers and, apropos of nothing, utter the sentence, 'Have any of you ever fantasized about high-fiving a Scientologist?', and there is a good chance that many of the strangers would correctly understand me to be asking whether any of them has ever fantasized about high-fiving a Scientologist. If I were to find myself in a room full of monolingual Japanese speakers, on the other hand, my attempt to ask the same question would go much less smoothly. Even if I were to somehow succeed (which seems unlikely), asking my question in the second room would take enormously more time and effort than it would have in the first. When I refer to the *efficiency* of linguistic communication, this is the kind of datum that I have in mind: any instance of smooth and successful communication when these qualities clearly depend on the fact that the sentence (or other utterance-type) involved is

[25] E.g., Bach (1987: *chs*.3–4), Carston (2002), Neale (2004: 81–97), Schiffer (2003: ch.3–4), Sperber and Wilson (1986/94), Wilson and Sperber (2012: ch.1).

meaningful for a group that includes the speaker and the addressee. At least part of why we need a concept of meaning is to explain the efficiency of linguistic communication in this sense. [26]

One consequence of this methodological outlook is that meaning is a *public* property of utterance *types*, including linguistic utterance types, such as words, phrases, and sentences.[27] Meaning is public in the sense that utterance types have meaning for the speech community in which utterances of that type are likely to lead to efficient communication.[28] I won't posit any notion of meaning that is possessed by particular utterances, expression tokens, expression tokenings, expressions-in-contexts, or expression–context pairs. In these respects, I differ from most contemorary semanticists.

A further consequence of my line of thought so far is that sentence meaning is not a kind of intentional or propositional content. There is no such thing, that is, as semantic content, if we take that to be any property of sentences, sentence tokens, or utterances of sentences that determines

---

[26] This way of framing the explanatory role of meaning is well-articulated and defended by Neale (forthcoming a: §3), where he argues that all theorizing about language and linguistic communication must be done in the service of a Master Question:

> We produce various noises, marks, and gestures, and in so-doing we manage to express or sharpen our thoughts, to communicate information about the external world or about our own beliefs, desires, plans, commitments, hopes, fears, feelings etc. efficiently (i.e. quickly, systematically and consistently). What facts about the noises, marks, and gestures we produce, what facts about our cognitive apparatus, and what facts about our physical circumstances and social relations explain this extraordinary capacity?

[27] For the view that the bearers of meaning are utterance types, see also Grice (1968), Kaplan (1989a,b), Lewis (1975), Neale (2004, 2005), Schiffer (1972, 1993, 2003, 2006), Stalnaker (1978). I won't attempt to give identity conditions for linguistic expressions or other utterance types. I take this to be a matter to be settled by phonology and psycholinguistics. My only significant commitment in this regard, which I share with the philosophers just mentioned, is that linguistic expressions and other utterance types can be individuated independently of their semantic properties.

[28] Does this mean that there is no sense to be made of the idea that an expression could have meaning for a single speaker? Not necessarily, since we might this notion of meaning to make sense of the idea that a speaker can communicate with a later time slice of herself, for example, by writing herself a note. But this isn't what philosophers usually have in mind when they speak of the meaning of an expression in a speaker's idiolect.

their truth conditions. This follows from the fact that the bearers of semantic properties are linguistic expressions, together with the fact—well documented in recent literature on the semantics–pragmatics interface—that the semantic properties of linguistic expressions massively underdetermine the contents they are used to directly and literally express on particular occasions. I will return to this point in greater detail in Chapter Two.

The confused idea that linguistic expressions possess a kind of intentional content has pervaded the study of language for the last one hundred and twenty-five years. One source of this idea is an equally confused and pervasive account of the role of meaning in linguistic communication. This is what Sperber and Wilson (1986/94) call the *code model*, according to which the role of meaning is to encode propositional contents in linguistic vehicles, which are then passed from speakers to hearers, who decode them. Variants of this model show up in most of the pragmatic theories that truth-conditional semanticists like best. Lewis holds that hearers interpret speakers' utterances by relying on the assumption that members of their speech community utter a sentence only if the proposition that the sentence expresses in the language that they share is true (1969, 1975a). Dynamic theorists take communication to succeed when a speaker adds some new content to the shared context of conversation, and dynamic semanticists hold that sentence meanings accomplish this because they are functions that map one state of the context into another. I will criticize these implementations of the code model in Chapters Four and Three, respectively. For devastating criticisms of the code model in general, see Sperber and Wilson (1986/94: ch.1).

The correct conception of the role of meaning in communication must be built on top of the right theory of communication. On the Gricean model I put forward in §1.2, communication succeeds when a speaker performs a communicative illocutionary act and her addressee recognizes the first component of her communicative intention, thereby fulfilling the second component. If I address an assertion that the dog ate my homework to you, for example, we will thereby

successfully communicate if you recognize that I intend to produce in you the belief that (I believe that) the dog ate my homework. How are we able to recognize one another's intentions in this way? The details are fuzzy, but we can be sure that the psychological process involves *inference to the best explanation* (a.k.a. *abductive* or *nondemonstrative* inference). In essence: to interpret a speaker is to rationalize their utterance by inferring the best explanation of why they produced it. We do this by positing an intention with which the utterance was produced. This sort of mindreading is one of the most discussed topics in cognitive science,[29] and the nondemonstrative inference that fuels it has proven to be among the mind's most opaque black boxes.[30] Nevertheless, I feel confident in the following principle of speculative psychology: whatever psychological processes underlie our ability to infer the best explanation of some observed phenomena, those processes are more likely to succeed, and to do so efficiently, when the search-space of possible explanations is narrower than when it is wider. It follows that anything that cuts down the size of the search space will make nondemonstrative inference more efficient. I maintain that this principle provides the key to understanding the role of meaning in communication.

Returning to an example from earlier, suppose that I want to ask a room full of speakers whether they've ever fantasized about high-fiving a scientologist. What makes the process of realizing this desire so much more efficient if the room is full of English speakers than if the room is full of monolingual Japanese speakers (given that I don't speak Japanese)? The answer is that the sentence, 'Have you every fantasized about high-fiving a scientologist?' is meaningful for me and the English speakers, and the meaningfulness of this sentence (call it X) is the sort of fact that

---

[29] For an excellent review of the literature on psychological and philosophical approaches to mindreading, see Goldman (2012).

[30] For pessimism about the very possibility of a cognitive-scientific approach to the nondemonstrative inference underlying interpretation, see Chomsky (2000), Davidson (1986), and Fodor (1983, 1987, 2001). For optimism about such a cognitive-scientific approach, and for a detailed defense of the idea that the role of linguistic meaning is to narrow the search space of the nondemonstrative inferences involved in interpretation, see Sperber and Wilson (1986/94).

makes the following two counterfactuals true: (a) if an English speaker were to have the communicative intentions involved in asking the question I want to ask, they would be likely utter X, and (b) if an English speaker were to observe an utterance of X, they would be likely to take the speaker to have the kind of intentions involved in asking the question I want to ask. The truth of these counterfactuals accomplishes two things. First, my English-speaking addressees will have a much easier time recognizing the intentions behind my utterance because they have a much narrower search space of possible explanations to look through. Barring very peculiar circumstances, they can ignore the possibility that I am trying to tell them that Seinfeld is on TV, for example. The absence of an utterance type that is similarly meaningful for me and my Japanese addressees means that I have no comparable method for narrowing their interpretive search spaces, and so they have a much wider and more varied space of possible explanations for my utterance. The kinds of evidence of my intentions with which I can provide them is radically impoverished vis-a-vis my English-speaking addressees, and so we would have to work much harder to overcome that lack. The truth of the two counterfactuals also make it the case that I can reasonably believe that my attempt to communicate will succeed, and, given that intentions are doxastically constrained, this might unlock the possibility of having those intentions in the first place.

So the meaning of an utterance type plays a dual role, corresponding to the two halves of a Gricean theory of interpretation. It allows tokens of the utterance type to serve as a rich sort of evidence of the speaker's intentions, thereby narrowing the space of interpretive options for an addressee and making interpretation more efficient. In playing this role in interpretation, it also unlocks makes it rational for speakers to have a range of communicative intentions that it would otherwise have been irrational or impossible for them to possess. This is how utterance-type meaning makes communication more efficient, on my view. My goal in the following chapters is to give a theory of meaning that allows it to play this role.

CHAPTER TWO: SEMANTICS WITHOUT SEMANTIC CONTENT

Introduction

Contemporary semantics is built around the idea that the meaning of a sentence is a property that determines its truth condition in a context, and that a word's meaning is a property whose role is to partly determine the truth conditions of the sentences in which the word appears. By focusing on the meanings of expressions that can be used to say different things on different occasions, I will argue in this chapter that words and sentences don't have truth-condition-determining properties of these kinds.

In §2.1, I will introduce the basic ideas and motivations behind truth-conditional semantics, as well as some of the foundational problems raised by the fact that sentences' meanings underspecify truth conditions. In §2.2, I'll focus in on the case of demonstratives in order to argue that sentence meaning must play the role of constraining speakers' communicative intentions rather than determining truth conditions. In §2.3, I will introduce the basic moving parts of speech-act-theoretic semantics and show how to account for semantic underspecification within this kind of semantic theory.

2.1 Truth-Conditional Semantics and Underspecification

The dominant research paradigm in semantics has, for about half a century, centered around the idea that a sentence's meaning is a property that fixes the sentence's truth condition. Different genres of semantic theory spell out the nature of these truth-condition-determining properties in different ways. According to the approach of Donald Davidson (1965, 1967, 1970), a sentence has meaning just in case there is an extensional truth theory whose theorems include a T-sentence for

the sentence and that truth theory meets further conditions that make it *interpretive*.[1] Within a school of approaches favored by most working semanticists now, the job of a semantic theory is to assign to sentences model-theoretic objects, such as functions from worlds to truth values (intensions) (Lewis 1970; Montague 1974; Stalnaker 1978; von Fintel and Heim 2011), sets of centered worlds (Lewis 1979b; Stalnaker 2008), pairs of intensions (Chalmers 1996), or situations or sets of situations (Barwise and Perry 1983). These model-theoretic objects in turn determine the sentences' truth conditions. An alternative approach originally due to Gottlob Frege (1892) and Bertrand Russell (1903), but more recently defended and developed by David Kaplan (1989a) and Scott Soames (1987), takes the job of a semantic theory to be that of assigning to sentences structured propositions—abstract entities composed of either objects, properties, and relations or modes of presentation of them; these structured entities in turn determine sentences' truth conditions.

I will group all of these approaches together under the covering term, *truth-conditional semantics*. Although they go about the business of semantics in different ways, and explicate an expression's meaningfulness in terms of different kinds of properties that it might have, the approaches I've alluded to all share an overall conception of the *explanatory role* that linguistic meaning is supposed to play. They confront the question that is raised by Lewis's dictum—*what does meaning do?*—with Lewis's own answer: "A meaning for a sentence is something that determines the conditions under which the sentence is true or false" (1970: 193–94). This overarching genre of semantic theory has dominated thought about linguistic meaning among linguists and analytic philosophers. Individual strands within the genre are distinguished by disagreements over how best to understand the properties in virtue of which sentences have their truth conditions.

---

[1] The question of what makes a truth theory interpretive turns out to be a difficult one for Davidsonians to answer. See Davidson (1973, 1975), Lepore & Ludwig (2005: chs.6–13), and Neale (2001: §2.2) for discussion.

In order to achieve some generality in the following discussion, I define a notion of *semantic content* as follows:

(SC1) Semantic contents are properties of *linguistic entities*, such as expressions, expressions in contexts, utterances of expressions, or expression tokens. More specifically: a linguistic entity's semantic content is its *semantic value*—i.e., the property that a compositional semantic theory aims to ascribe to it.[2]

(SC2) The semantic content of a sentence determines its truth condition, and the semantic content of a sub-sentential expression determines its contribution to the truth conditions of sentences in which it appears.

By explicating semantic content via its role in (SC1–2), I am construing the notion more broadly than is usual. Typically, an expression's semantic content is taken to be its *propositional* content (in the case of a sentence) or, in some sense that would have to be spelled out further, its contribution to sentences' propositional contents (in the case of sub-sentential expressions). But not every notion that fits (SC1–2) is best thought of as picking out a kind of propositional content. Some semanticists posit properties that satisfy (SC1–2), but deny that these properties ought to be cashed out as relations to propositions or their components. The best-known example is Davidson, who argued that *meanings*, construed as entities, "do not seem to…oil the wheels of a theory of meaning" (1967: 20). Nonetheless, Davidson did think that expressions have properties that satisfy

[2] An expression's semantic value, as I use the term, can be thought of as the meaning of an expression in the technical sense that is relevant to semantic theory. Schiffer (2003: ch.3) points out that most linguistic expressions are such that there is no entity that we can felicitously say, of the entity, that it is the expression's meaning, and concludes that talk of the meanings of words and sentences is confused. This may be a correct characterization of ordinary usage or folk-semantic theory, but at least many technical semantic theories deviate from ordinary usage in this way, and I see no problem with saying that some entity is an expression's meanings in at least a technical sense. Still, I will sometimes rely on Lewis's (1980) notion of a *semantic value*, which he introduced specifically so as to remain neutral about what kind of entity or property semantic theories should ascribe to expressions.

(SC1–2). According to a Davidsonian, the sentence 'snow is white' has the property of being the subject of a T-theorem of an interpretive truth theory for a language, for example. My extended usage of 'semantic content' to cover properties like these is unorthodox—and would probably have irked Davidson—but the usage is justified because it allows me to discuss aspects that Davidsonian semantic theories have in common with other forms of truth-conditional semantics that do trade in semantic entities.[3]

My main goals in this chapter can be summed up as follows: I will show that there is no such thing as semantic content (§2.2), and that semantics can (and must) get along just fine without it (§2.3). Because my argument contra semantic content will revolve around semantic underspecification, I will spend the remainder of this section exploring that notion and some of the debates about the relationships between descriptive semantics, foundational semantics, and pragmatics that have arisen from it.

With my unusually broad concept of semantic content in hand, we can sum up the central tenet of truth-conditional semantics by saying that the job of a semantic theory is to systematically and finitely ascribe to each meaningful expression its semantic content. Once we pay a bit of attention to actual natural language expressions, however, a pair of obvious problems with this view quickly emerge. One is that sentences in non-declarative moods—including interrogatives ('Who are you?') and imperatives ('Buy me a drink!')—apparently aren't truth-apt, and so don't have truth conditions or semantic contents. I will consider this worry, and the semantics of non-declaratives more generally, in Chapter Three. The second problem is that most declarative sentences, conceived of as expression types, can't be said to possess semantic contents either, because

[3] Nonetheless, Davidson is not my main target in what follows, and some of my arguments might have to be reformulated in order to apply cleanly to his version of semantics.

they exhibit *semantic underspecification*.[4] Underspecified sentences' meanings don't fully specify truth conditions, so that knowing everything there is to know about an underspecified sentence's meaning would not, by itself, allow one to identify any truth condition associated with any utterance produced with it or speech act performed with it. It follows that at least many linguistic expressions don't—at least by themselves—possess semantic contents.

In order to illustrate semantic underspecification and the strategies that truth-conditional semanticists have used to address it, I will consider a range of examples.

Indexicals

(1) a. I am Canadian.

b. Donald Trump lives here.

c. It is now the twenty-first century.

Each of these three sentences expresses determinate truth conditions in particular contexts (or can be used to literally and directly express propositions with determinate truth conditions on particular occasions), but none of them has truth conditions on its own.[5] The semantic underspecification of each of (1a–c) can be traced (at least in part) to the underspecification of the in-

---

[4] Two more common terms for this phenomenon are *context sensitivity* and *semantic underdetermination*, but, as Neale (*forthcoming a*: §3.2) has pointed out, both of these expressions are misleading. The latter term presupposes that the role of linguistic meaning is to determine the truth conditions of a sentence (perhaps relative to context), and the former presupposes that underspecified expressions' semantic contents are determined by something that can be called 'context'. Both of these points are precisely what is at issue in this chapter.

For helpful discussions of semantic underspecification, see Carston (2002) and Neale (2004: pp.87–90).

[5] The parenthetical qualification is necessary because I'll argue in §2.2 that only speakers can express propositions, and that linguistic objects, such as expressions or utterances of them, cannot. In order to avoid sounding as though I am begging the question agains the idea of semantic content, I will resort either to parenthetical disjunctive statements (as here), or to the locution, 'the truth conditions (content) associated with an utterance', which should be understood as neutral between 'the truth conditions (content) of an utterance' and 'the truth conditions (content) of the proposition expressed by a speaker with an utterance'.

dexical constituent that I have highlighted in bold. Expressions like 'I', 'here', and 'now' are *indexical* in the sense that what they refer to (or what speakers refer to with them) vary from one context of utterance to another. Clearly, this is incompatible with the most naïve form of truth-conditional semantics, on which the job of a compositional semantic theory is to assign semantic contents directly to sentence types.

The most influential strategy for dealing with semantic underspecification within the framework of truth-conditional semantics was developed in order to account for indexicals like these. According to this strategy, which was developed by Kaplan (1989a), the process by which a semantic theory assigns truth conditions to sentences takes two steps: each of the sentence's constituents is assigned a semantic content relative to a context of utterance, and then the semantic contents of the sentence's constituents are composed to determine the sentence's semantic content in that context.[6] On this view, it is expressions-in-contexts (or utterances of expressions), rather than expression types, that possess semantic contents. Kaplan (1989a) models the contexts relative to which expressions are assigned semantic contents as ordered quadruples (e.g., $c = \langle c_A, c_T, c_P, c_W \rangle$), whose coordinates represent the *speaker*, *time*, *position*, and *world* of a context, respectively. Kaplan's compositional semantics associates each indexical expression with a *character*—a function from contexts to semantic values. The character of 'I', for example, maps each context c to the agent c_A of the context. This captures the intuitive idea that the referent of an utterance of 'I' is its speaker.

As Neale (2004, 2005, *forthcoming a*) has repeatedly pointed out, it is important to distinguish the *metaphysical* question of what *fixes* or *constitutes* the properties of a given expression, utter-

---

[6] Kaplan is inconsistent about what sentences' semantic contents are. Informally, he represents them as Russellian structured propositions, so that, for example, 'Kaplan is funny' would be represented by Russellian proposition, ⟨Kaplan, FUNNY⟩, in which the first coordinate is Kaplan himself and the second coordinate is the property of being funny. In his model theory, however, Kaplan identifies sentences' contents with intensions, which he models as sets of world/time pairs (1989a: §XVIII).

ance of an expression, or speech act performed with an expression, from the *epistemic* question of how a hearer *identifies* those properties. Although we could conceivably construe Kaplan's theory of indexicality, or something similar, as part of an answer to a question of the latter, epistemic sort, this is not how Kaplan or his followers have understood it. Kaplan (1989a) intends his theory of indexicality to be an answer to the metaphysical question: what fixes the referent of an indexical in a given context of utterance? In the same vein, Lewis holds that sentences' meanings "determine" their truth conditions "in various possible states of affairs, at various times, at various places, for various speakers, and so on" (1970: 193–4). The Kaplanian–Lewisian strategy for dealing with semantic underspecification is to build a notion of the objective context of utterance into the models underlying one's compositional semantic theory, such that some fact about that context, together with the characters of each sentence's constituents and the principles governing semantic composition, fixes the semantic value (and so the truth conditions) of every meaningful sentence. Although Kaplan's contexts are simple and well-behaved ordered tuples that represent apparently objective facts about the situation in which someone produces an utterance, the same basic idea lies behind a wide variety of other proposals about the nature, structure, and basis of utterance contexts. Philosophers and linguists have argued that a wide variety of further elements of context must be posited, including the mutual presuppositions of a conversation's participants (Stalnaker 1978), a collection of assignment functions or discourse referents to keep track of entities about which the participants in a conversation are talking (Heim 1982, 1983; Kamp 1981; Karttunen 1976), a salience ranking on the domain of disourse of on the discourse referents (Lewis 1979a), representations of the questions under discussion at a given time in a conversation (Roberts 1996/2012), and representations of the norms (Gibbard 1990; Lewis 1975b, 1979a), plans (Gibbard 2003; Charlow 2013), preferences (Starr *ms*), To-Do lists (Portner 2004), or other action-guiding states that the participants in a conversation have agreed upon for the purposes of a conversation. Although the contexts thus posited differ in composition and complexity, the ba-

sic idea is always that the meanings of expressions and the context of utterance come together to jointly fix the semantic values of expressions in contexts. This strategy is most plausible if we restrict our attention to the word 'I', which is sometimes called a *pure* or *automatic* indexical, on the grounds that its character and the context in which it is uttered seem to automatically fix its referent in that context (see, e.g., Braun 2007: §1.3). It is often claimed that 'yesterday', 'today', 'tomorrow', 'here', and 'now' are pure indexicals in this sense.

I doubt that there are really any pure indexicals in this sense, but it can easily be seen that 'here' and 'now' are impure. Even if we assume that 'here' can only refer to a place that the speaker is in (or that a speaker must use it to refer to a place she is in), this does not fully settle which place a given utterance of 'here' refers to (or what a speaker refers to with it). This is because a speaker's location is not any single thing. I am in my apartment, in my building, in Fort Greene, in Brooklyn, in New York, and so on, and any one of these locations could be the referent of 'here' that's relevant to the truth conditions of (1b) in a given context (or of what I might say with (1b) on a given occasion). Moreover, nothing about the meaning of 'here' together with the speaker's physical location could determine which of these locations is the one that features in the truth conditions of (1b) (or what a speaker says with it).

What *does* fix the location that figures in the truth conditions associated with an utterance of (1b)? On the view that I will defend using other examples in §2.2, the relevant location is determined by the communicative intention with which a speaker utters (1b) (cf. Neale 2004; Schiffer 1981; Sperber and Wilson 1986/94). One consequence of this view, I'll argue, is that we can make no sense of the idea that the expression 'here', or an utterance of it, refers. Instead, only people refer, and they may do so by uttering expressions like 'here' (cf. Bach 1987; Neale 2004, 2005, *forthcoming a*; Schiffer 1981, 2003; Strawson 1950). And a consequence of this is that 'here' has no property that deserves to be called 'semantic content'. Before turning to that argument, I'll consider a a few other examples of semantic underspeficiation.

DP Restriction

(2) a. The book is over one-hundred pages long.

b. Every beer is in the bucket.

The bold-faced constituents of (2a–b) are DPs ('Determiner Phrases'). Each consists of a determiner ('the' and 'every') and an NP nominal ('book' and 'beer'). A much-discussed fact about DPs is that the truth conditions associated with sentences containing them must be understood as if the things to which the DPs' nominals applied were somehow *restricted*. Normally, the truth condition associated with a direct and literal utterance of (2a) will not be about the one and only book, for example, but will be about the one-and-only *book-with-some-further-property*. A speaker might use (2a) to express the proposition that *the book they are writing* is over one hundred pages long, for example, or that *the book their nephew is reading* is over one-hundred pages long, and so on. Similarly, the truth conditions associated with a direct and literal utterance of (2b) would normally not concern every book in the world, but instead every book that meets some further condition. The problem of saying what fixes these implicit restrictions is the problem of *DP restriction*.[7] As it applies to quantifier phrases like 'every beer' in (2b), this problem is usually called the problem of *quantifier domain restriction* (e.g., Stanley and Szabó 2000); as it applies to definite descriptions like 'the book' in (2a), it is usually called the problem of *incomplete descriptions* (e.g., Neale 1990: §3.7).

Two traditional solutions to the problem are what Neale calls the *implicit* and *explicit* approaches (1990: §3.7; 2004: §§5–8), which he characterizes as follows:

---

[7] Stanley (2007) uses the term 'nominal restriction' because he (following Stanley and Szabó 200) locates the restriction syntactically inside the DP's nominal. Westerståhl (1985) instead argues that the restrictor is contributed by the DP's determiner. In calling the phenomenon 'DP restriction', I wish to remain neutral between these two options, as well as the possibility that the restriction is contributed by a property of the DP as a whole.

> How are we to explain the incontrovertible fact that *A* can use a description 'the φ' in an utterance of the simple form 'the φ is ψ' and thereby perform a perfectly felicitous speech act, indeed *say something true*, even though *A* and *B* (the hearer) both know that φ(*x*) is true of more than one thing? The question generalizes: How are we to explain the fact that (roughly) for a range of determiners, *D*, *A* can use '*D* φ' in an utterance of the simple form '*D* φ is ψ' and thereby perform a perfectly felicitous speech act, indeed *say something true*, even though *A* and *B* both know that φ(*x*) is true of some things that are not relevant to the truth or falsity of what *A* said?
>
> Many philosophers appear to think the answer to the question…is obvious. 'There's always an *implicit background restriction* on the domain over which a quantifier expression ranges,' is one old reply. Another is, 'An utterance of 'the φ is ψ' is sometimes *elliptical* for an utterance of 'the φ that ζ is ψ', where ζ is something the speaker could have made explicit but didn't.' Call these the *implicit* reply and *explicit* reply, respectively, based only on the appearance of the words 'implicit' and 'explicit' in the quoted remarks. (Neale 2004: 105–6)

For my purposes, the details of who endorses each of these approaches, why, and who is right, are somewhat irrelevant. On either approach, some precisification of this question arises: in virtue of what is a DP's nominal restricted in the particular way that it is? [8]

Assuming the implicit approach, this question can be rephrased as the question of what fixes the particular implicit background restriction over which a DP ranges on a particular occasion. One possibility is that a continually evolving domain of discourse is somehow built into the context against the background of which expressions are assigned their semantic contents (e.g., Barwise and Perry 1983). But then the question arises: in virtue of what does any particular object get into the contextually relevant domain of objects? Even this question turns out to be ill-posed, however, because no single contextually relevant domain will be enough. As several authors have pointed out, it is possible to utter 'The Russian voted for the Russian' in order to say that one specific Russian voted for a different specific Russian, and, so, if the implicit approach is to work, then the two instances of 'the Russian' in this sentence must get their semantic contents from distinct contextually relevant domains, each of which contains a different unique Russian (Neale

[8] Of course, this is not to say that there won't be any indeterminacy about how a nominal gets restricted in a particular case. See, e.g., Buchanan (2010) and Buchanan and Ostertag (2005) for helpful discussions of this sort of indeterminacy. Still: insofar as there is a fact about how a nominal is indeterminately restricted, we can ask what constitutes this fact.

2004: 124; Soames 1986b; Westerstahl 1985). Simply adding more background domains to the context isn't enough to solve this problem, moreover, since we still have to answer the question: in virtue of what is each distinct instance of 'the Russian' matched up with the contextually restricted domain relevant to *its* semantic value, as opposed to whatever other domains might be hanging around in the context? It is difficult to imagine how one could answer this question with respect to the truth conditions associated with a given utterance without appealing to the communicative intentions of the speaker.

Assuming the explicit approach, on the other hand, the question of what constitutes the relevant restriction can be rephrased as the question of what makes it the case that some more fully articulated DP counts as the one for which an actual utterance of some DP was elliptical. Neale argues that this fact could be determined only by the speaker's communicative intentions. Roughly: taking Dφζ to be a DP whose nominal φζ is a restriction of the nominal φ of Dφ, what makes φζ *the* restriction of φ that is relevant to an utterance of 'Dφ is ψ' on a particular occasion is that the speaker could have uttered 'Dφζ is ψ' in order to give more precise evidence of the communicative intentions that she actually had in uttering 'Dφ is ψ' on that occasion. Other than the speaker's intentions, it is hard to see what could do this work.

An added complication arises from Stanley and Szabó's (2000) and Stanley's (2002) arguments that DPs' nominals contain aphonic, but syntactially-realized context-sensitive variables that function to restrict DPs' domains. On this view, the LF of (2a–b) can be roughly approximated by (2*a–b):

(2*) a. [$_{DP}$ [$_{Det}$ the] [$_{NP}$ book $f(i)$]] [$_{VP}$ is over one-hundred pages long]

b. [$_{DP}$ [$_{Det}$ every] [$_{NP}$ beer $f(i)$]] [$_{VP}$ is in the bucket]

Stanley and Szabó explain the properties of the aphonic domain restrictor as follows (with adjustments to accommodate the present example and numbering):

> The value of "*i*" is an object provided by the context, and the value of "*f*" is a function provided by context that maps objects onto quantifier domains. The restriction on the quantified expression ["every beer"] in [(2*b)] would then be provided by the result of applying the function that the context supplies to "*f*" to the object that the context supplies to "*i*". (Stanley and Szabó 2000: 101).

Stanley and Szabó posit the variable *i* in their LFs in order to explain the fact that some type-*e* variable hidden in DPs' nominals can be bound by DPs higher up in an LF. They posit the aphonic indexical *f* in order to explain how the semantic content of a type-*e* variable like *i* can be transformed into something of a semantic type that can restrict a DP's domain by combining with its nominal.

If Stanley and Szabó are right that nominals contain aphonic domain-restricting variables, we still need to ask the question: in virtue of what do those variables have the contents that they do on particular occasions (*cf.* Neale 2004: §3; 2007a)? For example: what is it about two different occasions of use that makes it the case that the values of *f* and *i* are saturated in two different ways on those occasions, so that (for example) 'the book' picks out *the book on the table at Jason's house* on the first occasion and *the book on the shelf at Zoltán's house* on the second occasion? Stanley and Szabó's answer is that the values of these variables are "provided by context", but this just raises the further question: what fact about the context makes it the case that one value is "provided" rather than another? For roughly the same reasons mentioned in connection to the explicit and implicit approaches, it is difficult to see what could do this work other than some fact about the speaker's communicative intentions.

Some Other Examples of Underspecification

The bold-faced expressions in (3–9) exhibit some other heavily discussed kinds of semantic underspecification. We can see this by noting that each sentence (*n*) could express a range of contents in different contexts (or could be used to literally and directly express a range of different

contents on different occasions), two of which could be more explicitly paraphrased by (*n*a) and (*n*b).

(3) It's raining.

a. It's raining in New York.

b. It's raining in San Francisco.

(4) Elmar's shark is more than six feet long.

a. The shark that Elmar is eating is more than six feet long.

b. The shark that Elmar rides is more than six feet long.

(5) The coffee in Rome is expensive.

a. The coffee in Rome is expensive by the standards of food in Rome.

b. The coffee in Rome is expensive by the standards of coffee elsewhere.

(6) Rachel must be here.

a. Rachel must, for all we know, be here.

b. Rachel must, legally, be here.

Debates about these and other semantically underspecified expressions parallel the debate about DP restriction.

In line with Stanley and Szabó's syntactic approach to DP restriction, it is commonly argued that the expressions in (3–6) contain one or more aphonic variables, and that the context-sensitivity of these variables is responsible for the underpecification of the (3–6). The following examples are representative. Stanley argues that the LF of (3) contains a hidden variable that is saturated by "the contextually salient location" in the context of utterance (2000: 52). Peters and Westerstahl argue that possessives like 'Elmar's shark' in (4) contain implicit quantifiers which are themselves implicitly restricted (2013: 713). Kennedy argues that gradable adjectives like 'expensive' in (5) inherit their underspecification from implicit variables saturated by "the contextually

appropriate standard of comparison, whatever that is" (2007: 7). Kratzer argues that modals like 'should' in (6) have semantic contents only relative to a *conversational background* (a.k.a. modal base) and *ordering source*, that the values of these variables determine whether a modal should be read epistemically (as in (6a)), deontically (as in (6b)) in some other way, and that the conversational background and ordering source are elements of context (2012).[9] Similar claims have been made about a range of other semantically-underspecified expressions.[10]

As in the cases of indexicals and DP restriction, however, positing hidden variables and saying that they are saturated by context does not answer the question of what fixes those expressions' semantic contents. What, specifically, about the context makes it the case that the truth conditions associated with a given utterance of each sentence (*n*) above are such that they could be better captured by (*n*a) than by (*n*b), for example? Again, it is difficult to imagine what could do this work other than some fact about the communicative intentions of the speaker (Neale 2004: §3, 2007a,b, *forthcoming a*; Dowell 2011, 2013).

The question of what fixes the content of a given expression (or what fixes the content that a speaker expresses with it) is typically thought of as a question about *founational semantics* or *metasemantics*, as opposed to a question in *descriptive* or *compositional* semantics, which is the study of how sentences' meanings are determined by their structures and the meanings of their parts. One reason that semanticists are typically vague or noncommital on the question of what fixes the contents of semantically underspecified expressions is that they take descriptive semantics and foundational semantics to be distinct projects that can be successfully pursued in relative isolation

---

[9] Kratzer remains neutral about whether the conversational background and ordering source are the values of aphonic variables, or whether they get into the truth conditions of a modal claim in some other way. Hacquard (2006) develops Kratzer's account of modals into an elaborate aphonic-variable theory.

[10] Since tense gives rise to a form of semantic underspecification in much the same way as pronouns (see, e.g., Partee 1973) and every natural language sentence involves tense, it follows that every natural language sentence exhibits at least some underspecification.

from one another, so that how one answers questions in foundational semantics will have little impact on how one answers questions in descriptive semantics. The task of descriptive semantics, on this view, is to lay bare the semantically relevant structure of each meaninful sentence, assume that each underspecified expression is saturated by a semantic content of the right type, and compute the semantic content of the sentences from their structures and the semantic contents of their parts. This approach is typified by Heim and Kratzer's brief discussion of foundational semantics and context sensitivity (1998: 243):

> If you utter a sentence like
>
> [A] She is taller than she
>
> then your utterance is felicitous only if the utterance situation provides values for the two occurrences of the pronoun "she". Given that referring pronouns bear indices at LF, [A] has some representation such as [B],
>
> [B] She_1 is taller than she_2
>
> and we can think of an utterance situation as fixing a certain partial function from indices to individuals. An appropriate utterance situation for LF [B] is one that fixes values for the indices 1 and 2. That is, it is appropriate for [B] only if the variable assignment it determines includes 1 and 2 in its domain.
>
> Let "c" stand for an utterance situation or "(utterance) context" (we use these terms interchangably), and let "g_c" stand for the variable assignment determined by c (if any). We can thus formulate the following appropriateness and truth-conditions for LFs with free pronouns.
>
> [C] *Appropriateness Condition*
> A context c is appropriate for an LF φ only if c determines a variable assignment gc whose domain includes every index which has a fee occurrence in φ.
>
> [D] *Truth and Falsity Conditions for Utterances*
> If φ is uttered in c and c is appropriate for φ, then the utterance of φ in c is true if $[\![\varphi]\!]^{gc} = 1$ and false if $[\![\varphi]\!]^{gc} = 0$.

Because they are writing a book about descriptive semantics rather than foundational semantics, Heim and Kratzer make no attempt, either in this passage or elsewhere, to say what a context is or how it determines a variable assignment. The implication is that this isn't relevant to the task at hand. Stanley and Szabó (2000) are even more explicit in their defense of the independence of

descriptive and foundatinal semantics. They begin by considering two claims that someone might make about the semantics of the demonstrative, 'that' (2000: 72):

(7) Relative to *c*, "that" denotes *o*.

(8) *o* is what is demonstrated in context *c*.

Stanley and Szabó say that (7) is an answer to the "descriptive problem of context dependence"—the problem of "deriving", for an expression *e* relative to a context *c*, "the interpretation of *e* relative to *c*, given a prior characterization of what features of the context *c* have a bearing on the interpretation. (8), on the other hand, addresses the "foundational problem of context dependence for an expression e relative to a context *c*", which is the problem of "specifying what it is about the context in virtue of which certain entities (be they objects, properties, or propositions) play the role they do in the interpretation of an occurrence of *e*" (2000: 73). Stanley and Szabó go on to argue that it is crucial that we distinguish problems in descriptive semantics from problems in foundational semantics. They suggest, citing Braun (1994) and Reimer (1998) in agreement, that it is legitimate for those pursuing descriptive semantics to take "as given" the semantic contents of context-sensitive expressions (2000: 73), and argue that mixing up discussions of descriptive and foundational semantics should be avoided:

> In general, discussions of descriptive and foundational problems should be clearly distinguished from one another. To use an example discussed at length in Stalnaker (1997), in the case of proper names, an example of a foundational semantic debate concerns the viability of the causal theory of names, according to which the denotation of a name is due to the existence of a causal relation of the appropriate sort between it and its bearer. However, the causal chains are no part of the descriptive semantics of names. The descriptive semantics of names only involve linking them up with their bearers. Causal chains are rather part of a foundational semantic account of why names have the descriptive semantical interpretation they do. An example of a descriptive semantical debate is whether names are rigid designators, or are rather shorthand for non-rigid definite descriptions. Only confusion results from running such distinct questions together. (2000: 74)

Although I agree with Stanley and Szabó that descriptive and foundational semantics can be thought of as distinct projects (or at least as distinct parts of a single project), I disagree with several of their other assumptions about the nature of the boundary between descriptive and foundational semantics. One of these assumptions is that the semantic value of an expression—its meaning from the point of view of descriptive semantics—should be thought of as a kind of semantic content. On the view I will defend later in this chapter, the semantic value of an expression is the type of speech act that a speaker can perform with it, and the type of act of referring that is the semantic value of 'that' is individuated broadly, so that it can be tokened by referring to different objects on different occasions.

If I am right about this, then the proper foundational question about 'that' is: *in virtue of what is a certain broadly-individuated type of referential speech act the semantic value of 'that'?* Stanley and Szabó's version of the foundational question about 'that'—*in virtue of what does 'that' refer to an object o in a context c?*—is, by contrast, ill-posed, since demonstratives don't refer (only speakers do). In §2.2, I will argue that Stanley and Szabó's version of the foundational question about 'that' is illegitimate in this way by demonstrating that our best attempt to answer it collapses into an answer to an analogous question about what *speakers* refer to in uttering 'that'. This strategy reveals a second kind of disagreement with Stanley and Szabó: it shows that how we pose and answer questions in foundational semantics can have major consequences for the nature and methodology of descriptive semantics. Foundational and descriptive semantics must, then, be pursued in close connection with one another.

2.2 The Case Against Semantic Content

In order to avoid unnecessary complexity in what follows, I will assume that if sentences and their parts have semantic contents, these contents are Russellian propositions and their parts, respectively. On this view, for example, the semantic content of (9a) is the Russellian structured proposi-

tion (9b), whose parts—Bertrand Russell and the property of being British—are the semantic values of the sentence's parts, 'Bertrand' and 'is British', respectively.

(9) a. Bertrand is British.

b. ⟨Bertrand Russell, BRITISH⟩

Nothing of dialectical importance hangs on this Russelian assumption. It will simplify my discussion by allowing me to phrase Stanley and Szabó's foundational question in the way they do—as a question about a semantic relation borne by expressions to familiar objects and properties, rather than as a question about a relation borne by expressions to modes of presentation or various sorts of model-theoretic objects. But the same problems for understanding those relations would arise as arise for the Russellian semantic relation I'll discuss.

I will focus my attention in this section on the demonstrative pronouns 'this' and 'that' as they are used in (10) and (11):

(10) This stinks.

(11) That stinks.

Although any of the semantically underspecified expressions that I discussed in §2.1 would serve my purposes in this section, demonstrative pronouns are particularly useful for several reasons. One is that there is already a large literature discussing the question of what grounds demonstrative reference, and the authors of that literature have already done much of my work for me. A second advantage of focusing on demonstrative pronouns depends on a further simplifying assumption, which is already implicit in the fact that I refer to them as 'demonstrative pronouns' rather than as 'demonstrative determiners'. Following nearly everyone who has contributed to the literature to which I've just alluded (and, most influentially, Kaplan 1989a,b), I will take demonstratives like 'this' and 'that' as they appear in (10) and (11) to be syntactically unstructured, directly referential singular terms. This assumption is controversial, as various authors have enter-

tained the view that seemingly syntactically unstructured demonstratives are actually complex DPs with aphonic or deleted nominals (Neale 2004: §21; Elbourne 2005: ch.3; Hawthorne and Manley 2012: 218, *fn.*34). I am confident that bracketing these views will do no dialectical harm to this chapter because all such views reduce demonstratives to other kinds of semantically underspecified expressions of the kinds I discussed in §2.1, and I am confident that I could give parallel arguments for each of those kinds of expressions (indeed, I sketched the most important steps of these arguments in §2.1).

The upside of the simplifying assumption that simple demonstratives are syntactically unstructured and directly referential is that it allows me to move back and forth between talk of the *semantic contents* and the *semantic referents* of demonstratives. This is because a term is directly referential just in case its referent in a context is identical to its semantic content in that context (Kaplan 1989a). (Semantic reference can be defined by restricting the scope of (SC1) to directly referring expressions and replacing each instance of 'semantic content' with an instance of 'semantic reference'.)

My aim in this section, then, is to argue that there there do not exist entities that satisfy both of (SC1–2), and to do so by way of arguing that there is no coherent and explanatory notion of semantic reference. I am not the first to make these claims. Perhaps the best-known semantic-reference denier is P. F. Strawson, who famously summed up the position by saying that referring "is not something an expression does; it is something that someone can use an expression to do" (1950: 326). Various philosophers of language have since adopted weaker approximations of this view. Kent Bach maintains, for example, that "referring is ultimately not something that words do but something that speakers use words to do" (1987: 49, emphasis added). Although Bach does find room in his theory for a notion of semantic reference, he maintains that "speaker reference is a more fundamental notion than semantic reference" (ibid). This weakened version of the Strawson's slogan is as close as most philosophers of language are willing to get to the kind of full-

blown denial of semantic content that I am interested in here, but I will try to show that Bach's watered-down view does not go far enough.

Neale has at least flirted with Strawsonian semantic-content denial, in that he has periodically voiced skepticism that any notions of semantic reference and semantic content are needed for semantic theorizing.[11] In support of this skepticism, Neale offers the following argument for the dispensability of any non-speaker-based notion of reference.[12]

> ...we should separate (i) who or what A intended to refer to by an expression X on a given occasion, and (ii) who or what a rational, reasonably well-informed interpreter in B's shoes thinks A intended to refer to by X on that occasion. In cases where (i) = (ii), we can talk freely about what the speaker referred to. In cases where (i) ≠ (ii), we could argue about which of (i) or (ii) or some third thing has the 'right' to be called the person or thing referred to, but what would be the point? First, what third thing distinct from (i) and (ii) could be of any significance to a theory of interpretation? There is simply no rôle for a transcendent notion of what was referred to upon which (i) and (ii) converge when all goes well. [...] Second, why is a choice between (i) and (ii) needed in cases where (i) ≠ (ii)? Conceptually they are distinct, and they are both needed in a theory of interpretation. When all goes well, they coincide, and it's just too bad they don't always do so. There is no philosophical payoff in bestowing the honorific 'what was referred to' on one rather than the other, or on some third thing, when they diverge. (Neale 2005: 184–5).

In effect, Neale's argument takes the form of a challenge. In order to understand the nature of communication, we already need an intention-based notion of speaker reference, as well as the notion of what an (idealized) interpreter would understand a speaker as referring to. Neale challenges anyone who wishes to theorize about any further notion of reference (such as the notion of semantic reference) to articulate some explanatory role for that notion to play. Without such a well-articulated explanatory role, the notion of semantic reference is mere theoretical clutter with

---

[11] For Neale's skepticism regarding the notion of semantic reference, see (2008: 381; *forthcoming* a: 11, *fn*.20). Neale's skepticism about the notion of semantic content is more explicit, and comes out in his denial that there is any useful or coherent notion of *what a sentence says*, as opposed to the Gricean, speaker-based notion of *what a speaker says* (see Neale 2004, 2005, *forthcoming* b).

[12] For an almost identical argument against semantic content (in Neale's terminology: any notion of what is said that isn't an intention-based notion of what a speaker says), see Neale (2004: 78; 2005: 184.).

no explanatory work to do. Neale gives almost identical arguments against any notion of what is said that isn't an intention-based notion of what the speaker said (2005: 184; see also 2004: 78). I will refer to these objections as the *dispensability argument*.

The most interesting kind of response to the dispensability argument would address Neale's challenge directly by arguing that we *need* some notion of semantic content in order to construct an explanatory semantic theory. This response takes seriously the idea that we should countenance semantic content only if we have a role for it to play, but contends that we need semantic conent to play the role of the semantic values that a semantic theory assigns to expressions. According to half a century of accumulated wisdom, compositional semantics works by showing how expressions' semantic contents (relative to contexts) add up to sentences' semantic contents. So one conceivable response to Neale's dispensability argument would be that semantics can't be done without semantic content.

One of the broad purposes of this dissertation is to counter this objection to the dispensability argument—to produce a proof-of-concept for a kind of semantic theory that doesn't traffic in semantic content. On the Gricean view that I sketched in Chapter One, the role of linguistic meaning is not to determine truth conditions, but to enable efficient communication. The explanatory role I assign to linguistic meaning is twofold: first, linguistic meaning must provide evidence to interpreters of speakers' communicative intentions, and, second, linguistic meaning must thereby constrain the communicative intentions that a rational speaker can have in uttering expressions. Based, in part, on a similar line of thought, Neale has proposed that compositional semantics should be in the business of showing how the constraints on what we can communicatively intend in uttering a sentence are built up from the constraints on what we can communicatively intend in uttering words (and, in particular, from the referential intentions we can have in uttering referring expressions). Neale uses the term 'blueprints' to describe meanings, construed as constraints

on intentions (2004, 2005).[13] Neale is perhaps clearest about this point as it applies to the semantic properties of quantifiers and referring expressions:

> I prefer to watch those [varietals of noun phrase] that are useful emerge from *act-syntactic* reflections, by which I mean reflections on the most fundamental features of speech acts as they are regulated by (or projected through the lens of) syntax. … That may sound mysterious, but the basic idea is quite simple. We might put forward a preliminary act-syntactic, specification of the class of referring expressions that declares an expression E a referring expression if and only if (a) E is used regularly or systematically to perform such-and-such types of intentional acts—or at least could be so-used in a sense of 'could' that would require some explication—and (b) E has such-and-such syntactic distribution. (Neale 2008: 379)

Although others have made similar programmatic remarks about the idea that semantics should be a theory of constraints on intentions, nobody has shown how this idea can be fleshed out into a detailed and systematic compositional semantics.[14] It is therefore open to the defender of semantic content to protest that the proof is in the pudding—that until the details of this alternative kind of semantics are worked out, we should assume that the only workable conception of semantics is one that employs a notion of semantic content. Part of my goal in the remainder of this dissertation will be to show how to construct a semantic theory that builds on Neale's programmatic remarks, and thereby refutes the idea that semantics without semantic content is impossible.

Before I reach that positive proposal, I want first to offer a more direct objection to the notions of semantic reference and semantic content. Whereas Neale's argument from dispensability makes the claim that we can do without semantic content, my argument aims to show that we can't have semantic content even if we want it. I will demonstrate this by showing that any attempt to undestand the notion of semantic content for a variety of expressions will inevitably lead us to

---

[13] "A *semantic* theory specifies the constraints that word meanings and syntax place on what A can *say* by uttering X, a blueprint for X" (Neale 2005: 190).

[14] Schiffer has likewise proposed in various places that the meaning of a sentence is a constraint on either the communicative intentions a speaker can have in uttering the sentence or on the type of speech act that a speaker can perform with the sentence (e.g., 1987, 1993, 2003). But, unlike Neale, Schiffer does not attempt to generalize this point to the meanings of sub-sentential expressions.

some speaker-based notion instead, and that we should extend this result to all other expressions for methodological reasons. Call this the *argument from collapse*.

The argument begins with the following two questions:

(Q1) In virtue of what does an utterance of a simple demonstrative *d* refer to a given individual *x* in a context *c*?

(Q2) In virtue of what does a speaker refer to a given individual *y* by uttering a simple demonstrative *d* on a particular occasion?

If semantic reference and speaker reference (as applied to demonstratives) are really distinct notions with different explanatory roles, then the answers to these two questions must differ. I contend that our best answer to (Q1) will collapse into our best answer to (Q2).

The main reason I've chosen to focus on simple demonstratives is that the existing literature on demonstrative reference has already done much of the work of establishing my conclusion. At one point, it was commonly held that the reference of a demonstrative was fixed by an accompanying demonstration (e.g., Kaplan 1989a). The paradigmatic example of such a demonstration is the act of pointing at one's referent, and some philosophers have even argued that it is some physical fact about this accompanying act of pointing that fixes the referent of a demonstrative (Quine 1968). As many philosophers have pointed out, however, demonstratives needn't be accompanied by an act of pointing, or by any other overt demonstration. Moreover, demonstration-theoretic accounts of demonstrative reference fail to explain the fact that even utterances of demonstratives that are accompanied by demonstrations sometimes refer to objects other than the object(s) at which the speaker points. One relevant kind of case involves deferred reference, as, for example, when a speaker points at a token object in order to refer to an associated type, or when a speaker points to a member of a group in order to refer to the group. It seems clear that any account of demonstrating capable of doing justice to these cases would have to ground the notion in

intention-recognition. Other philosophers and many linguists have attempted to answer (Q1) by arguing that the reference of a demonstrative in a context is whichever object is most salient in the context (Heim and Kratzer 1998; Mount 2008; Larson and Segal 1995; Wettstein 1984). If the notion of salience is taken intuitively, however, then it seems that salience is neither necessary nor sufficient for demonstrative reference, since it is possible for the intuitive referent of a demonstrative to be less salient than some other object (Michaelson 2013; Sperber and Wilson 1986/94). Moreover, salience-theoretic accounts of demonstrative reference have roughly the same problems with deferred reference as demonstration-theoretic accounts, since an utterance of a demonstrative may refer to a type whose token is salient, or to a group one of whose members is salient. Just like the notion of demonstrating, then, it seems clear that the notion of salience is a placeholder for something about the speaker's intentions.

The literature has now mostly converged on the view that demonstrative reference is fixed by the speaker's *directing* or *referential* intentions.[15] The fact that the best accounts of demonstrative reference ground reference in intentions is a key step in my argument, since, as I'll argue below, the best accounts of speaker reference likewise ground referring in intending. But this still leaves open the possibility that demonstrative semantic reference and demonstrative speaker reference are grounded in two different kinds of intentions. In the seminal paper in which he first distinguished the notions of semantic reference and speaker reference, Kripke (1977) defends this sort of view, though not with respect to demonstratives.

> In a given idiolect, the semantic referent of a designator (without indexicals) is given by a *general* intention of the speaker to refer to a certain object whenever the designator is used. The speaker's referent is given by a *specific* intention, on a given occasion, to refer to a certain object. If the speaker believes that the object he wants to talk about, on a given occasion, fulfills the conditions for being the semantic referent, then he believes that there is no clash between his general intentions and his specific intentions. (Kripke 1977: 112)

---

[15] E.g., Bach (1989, 1992); Kaplan (1989b); King (2001, 2012); King and Stanley (2005); Kripke (1977); Michaelson (2013); Siegel (2002); Soames (2010b).

Kripke applies his distinction between general and specific intentions to puzzles about both definite descriptions and proper names. For example, he argues that Donnellan's (1966) referential uses of definite descriptions are best understood as cases in which a speaker has an object-dependent specific intention about a referent that is not the same object that is picked out by the speaker's general intention to refer to whichever object satisfies the description. But Kripke's way of explicating the intentions underlying speaker reference and semantic reference can't help us to draw a distinction between speaker reference and semantic reference when it comes to demonstratives. This is because his account of the facts in virtue of which an expression *e* semantically refers to an object *x* take for granted that there is some fact about the speaker's language—some linguistic convention, say—that connects *e* to *o* in a way that is independent of the speaker's specific intentions, such that the speaker can intend to refer to *o* with *e* with the help of this fact. But it is precisely this sort of fact about demonstratives that I've been searching for in order to answer (Q1), and coming up empty.[16]

King (2012) and Michaelson (2013) both argue that demonstrative semantic reference is distinguished from demonstrative speaker reference by virtue of the fact that the intentions underlying semantic reference are somehow constrained in a way that the intentions underlying speaker reference are not. Here is King, for example:

> I think of the fact that a speaker intends an object to be the value of an occurrence of a demonstrative and the fact that a competent, attentive, reasonable hearer who knows the common ground of the conversation would take the speaker to intend that a certain object be its value to be objective features of a context of utterance. Call the former the *speaker fact* and the latter the *hearer fact*. A context is *appropriate* for a sentence containing demonstratives if each occurrence of a demonstrative in it is associated with a speaker fact and a hearer fact and these facts "involve" the same object (i.e. the intended object in the speaker fact is the object that would be taken to be intended in the hearer fact). For a given occurrence of a demonstra-

[16] To be clear: Kripke makes no attempt to apply his distinction between speaker reference and semantic reference to demonstratives, and there is no reason to think that he took the distinction to so apply. It follows from what I'll say later that Kripke's notion of semantic reference is without theoretical utility in the case of definite descriptions, proper names, or any other expression, as well. But I needn't offer specific arguments in favor of those points here.

> tive in an appropriate context, call this latter object the *coordinated object*. Then I view the meaning of a demonstrative as a function that maps an appropriate context to the coordinated object. I take the latter to be the semantic value of the occurrence of the demonstrative in the context. Finally, I take the lexical meanings of demonstratives to require that a use of a demonstrative be supplemented by a speaker's intention that is recognizble by an ideal hearer in just the way that the demonstration account held that the lexical meanings of demonsatratives require that a use be supplemented by a demonstration. (King 2012: 20)

In effect, then, King thinks that the semantic referent of a demonstrative *d* in a context *c* is whichever object *x* the speaker intends to be the value of *d* in *c*, provided that his his intentions meet certain recognizability constraints. I see two serious problems with this view. First, King's referential intentions are meant to ground the notion of semantic reference, but the content of each of his referential intentions is that some object *x* be the value of some expression *e*. But what is it for *x* to be the value of *e* if not for *e* to refer to *x*. This kind of circularity—grounding the notion of reference in an intention to refer—is a general problem with most intention-based accounts of reference, and I will return to it below. The second problem with King's account is that it cannot distinguish semantic reference from speaker reference. In personal correspondence, King has suggested that semantic reference differs from speaker reference because the intentions underlying the former, but not the intentions underlying the latter, are constrained by what a competent hearer could understand. But, as Griceans have long argued, communicative intentions (and, so, the referential intentions underlying speaker reference) are constrained in this way as well, and so it's not clear to me how King's proposal could help him to draw the distinction he needs.[17]

---

[17] My two worries about King's proposal are related. Below, I'll follow Bach (1987), Neale (*forthcoming a*), Schiffer (1981), and Stine (1978) in arguing that the referential intentions underlying speaker reference are best understood a kind of communicative intention, and that all other accounts that have been offered wind up construing referential intentions as intentions to refer (or something similar). But communicative intentions are reflexive; they are intentions to produce some response in an addressee in part via their recognition of the intention to produce this response. And so, one cannot rationally have a communicative intention to produce a response *r* unless one's beliefs leave open the possibility that one's addressee, if rational, will recognize this intention. So, the only non-circular way of spelling out the contents of referential intentions (that I know of) gives rise to the recognizability constraints on speaker reference that collapse it into King's notion of semantic reference.

The first of my two worries about King's view is quite general: most accounts of referential intentions are articulated by saying that a given utterance *u* of a demonstrative refers to an object *x* in virtue of the fact that the speaker intended that *u* refer to *x*.[18] This sort of formulation, which attempts to ground the notion of demonstrative semantic reference in the notion of a *speaker's intention to refer*, is inadequate as stated. If the concept of *reference* involved in the speaker's intention to refer is the concept of *speaker reference*, then the collapse of demonstrative semantic reference into demonstrative speaker reference is trivial, since for a demonstrative to semantically refer to *x* is just for its speaker to intend to refer to *x* with it. If the appealed-to concept of reference is the notion of semantic reference, on the other hand, then explications of this form are circular. Any account of referential intentions must articulate their contents in a way that does not include any notion of reference or referring.

To my knowledge, the only plausible attempts to non-circularly spell out the contents of referential intentions in this way are those of Bach (1987), Neale (*forthcoming* a), Schiffer (1981), and Stine (1978) all of whom understand referential intentions as kinds of, parts of, or aspects of Gricean communicative intentions. Here, for example, is Bach's explication:

> To refer to something is to use a singular term with the intention (part of one's communicative intention) of indicating to one's audience the object of the attitude one is expressing. (Bach 1987: 52)

Schiffer (1981) offers a more detailed and elaborate account of referring in the same, intention-based spirit. He also distinguishes between a generic notion of *referring to an object in producing*

---

[18] This is how Michaelson (2013) describes what he calls the 'standard account'. As he points out, the view is often attributed to Kaplan (1989b), as in the following passage from Siegel: "[A]ccording to [Kaplan's (1989b) view] what he calls a "directing intention" fixes reference (p. 582). This view differs from the one Kaplan held in [1989a], according to which intentions do not fix reference. A directing intention to refer to something is an intention to refer to a thing on which one has focused one's attention" (Siegel 2002: 6, *fn*.7). As Michaelson points out, however, Kaplan explicitly restricts his discussion to cases involving "perceptual demonstratives", and so he may not have held the more general view (1989b: 582).

an utterance and a notion of *referring to an object with a particular token expression that one utters*. Neale (*forthcoming a*: §4) reconstructs Schiffer's (1981) account as follows:

> Referring-In
> In (the course of) uttering *x*, *S* refers to *A* if, and only if, in uttering *x*, *S* means an *A*-dependent proposition.

> Referring-With
> In uttering *x*, *S* refers to *A* with *e* if, and only if, (1) *e* is properly contained in *x*, and (2) (∃*H*) (∃*R*) s.t. in uttering *x*, *S* intends *H* to recognize that *S* was referring to *A* in uttering *x*, at least partly on the basis of their mutual knowledge that *R*(*e*, *A*).

I think that some theory of referring along these lines—one that grounds speaker reference in communicative intentions—has to be right. I will provisionally adopt the following proposal as a necessary condition on the posession of a referential intention, leaving the possibility of some further conditions open.[19]

> (RI) Referential Intention
>
> S has a referential intention to refer to an object, property, or relation *x* with an utterance type *e* in producing an utterance *u* only if *e* is a proper part of an expression of which *u* is a token and there exists an *x*-dependent proposition *p*, addressee A, and binary relation *R* such that *S* produces *u,* thereby intending:
>
> (1) to produce a state of entertaining *p* in A;
>
> (2) A to recognize S's intention (1);
>
> (3) A's act of entertaing *p* to be at least partly on the basis of her recognition of (1);

[19] This account is watered down from Schiffer's in that it does not make use of the notion of mutual knowledge. I will argue in Chapter Four that linguistic meaning is not grounded in mutual knowledge, and I think that one could adjust this argument to support the conclusion that mutual knowledge is not required for referring to an object with a token expression. Much like (GI) from §1.2—my specification of the intentions required for performing a comunicative illocutionary act—(RI) might have to be augmented with further conditions before it would count as sufficient.

(4) A's act of entertaing p to be at least partly on the basis of A's sensitivity to the fact that $R(e,x)$.

Bach's, Neale's, Schiffer's, Stine's, and my own explications of referring are interesting because of what they have in common. First, they are the only non-circular theories of referential intentions in the literature, and the only theories to satisfyingly spell out the contents of referential intentions.[20] Second, by making referential intentions a kind of communicative intention, they allow for a satisfying account of why the notion of referring should be posited in order to explain communication. Although these account make speakers' communicative intentions explanatorily prior than her referential intentions (since the latter are explicated in terms of the former), referential intentions might be epistemically prior to full-blown communicative intentions from the point of view of an interpreter. You might come to recognize the communicative intentions behind my utterance of 'Kent is right' in part by first coming to recognize the referential intentions that I have in uttering 'Kent', for example. By uttering 'Kent' as part of my utterance of 'Kent is right', then, I can give you evidence that the content of my overall communicative intention is a Kent-dependent proposition, which takes you some of the way toward identifying the whole proposition. This is a notion of referring that is tailor-made to play a role in explaining the efficiency of linguistic communication in the way that I advocated in §1.3, and that relies on the same Gricean resources that I drew on there.

Third, communicative-intention-based accounts of referring are all clearly attempts to explicate *speaker reference* rather than anything in the neighborhood of semantic reference. Indeed, they are the best and most plausible explications of speaker reference available. In other words, these explications are the best-available answers to (Q2). But, as I have attempted to show, attempting to answer (Q1) inevitably leads to these same accounts. This narrative, which is borne

[20] Schiffer's/Neale's explication of *referring-with* is stated in terms of *refering-in*, but this is perfectly fine, since referring-in is not explicated in terms of any other notion of referring.

out by the literature on demonstrative reference, is the basis for my claim that the notion of demonstrative semantic reference collapses into the notion of demonstrative speaker reference.[21]

Suppose that the preceding argument is sound, and that—to echo Strawson—referring is not something that demonstratives do; it is something that speakers can do with demonstratives. In that case, how should we conceive of the linguistic meanings of simple demonstratives? In particular, how can we account for the fact that the simple demonstratives 'this' and 'that', and therefore the sentences (10) and (11), have different meanings? In keeping with the constraint-theoretic proposals of Neale and Schiffer, we can explicate the meaning of a demonstrative in terms of the constraints it puts on the referential intentions of a speaker who utters it. As a very rough first pass, we might say that the communicative intentions of a speaker who utters 'this' may do so only with communicative intentions about relatively proximal non-persons (in some perhaps metaphorical sense of 'proximal'), whereas a speaker who utters 'that' may do so only with communicative intentions about relatively distal non-persons (in some perhaps metaphorical sense of 'distal').[22] I will spell this out in my preferred, speech act-theoretic way in §2.3. (I will also replace the modal-sounding notion of *constraint* with a *disposition*-theoretic account of the connection between words and their meanings in Chapter Four.)

Although the argument from collapse is easy to run on demonstratives because the existing literature has already pushed toward the conclusion that demonstrative reference is grounded in referential intentions, I am confident that similar arguments could be run on other semantically underspecified expressions. I sketched how some of those arguments would go in §2.1. These ar-

---

[21] See also Bach (1987: chs.4+9, 2005, 2006, *ms*), Burton-Roberts (2006: 33–34), Fodor and Lepore (2005), and Schiffer (2005: 1141), who argue that speakers' intentions can't be considered part of the context of utterance in any sense of context that could be relevant to semantics. I am sympathetic to these arguments, but it is important to distinguish them from the argument from collapse, which is instead based on the conclusion that there is no explanatory role for a notion of semantic reference that is distinct from the notion of speaker reference.

[22] Cf. Evans (1980) and Neale (2007b).

guments would have to show, for some underspecified expression *e*, that there is no way to distinguish the facts that fix *e*'s content in a context from the facts that determine what a speaker directly and literally means in uttering *e* in the same context. The key step would be to show that if there is a fact about *e*'s content in a given context, this fact must be grounded in the speaker's communicative intentions. I sketched arguments of this sort for a few semantically underspecified expressions in §2.1.

If I am right about everything I've said so far, one possibility that still remains is that we should adopt a mixed content/constraint semantics. This is roughly the position of Kent Bach (2005) and Michael Devitt (forthcoming), both of whom hold that the semantic content of a sentence containing a demonstrative is a *gappy proposition* or *propositional radical*. The sentence (11), on this view, would express the propositional radical (12):

(11) That stinks.

(12) < $\_\_\_\_\__{distal}$, STINK>

The idea here is that the VP 'stinks' in (11), which is not semantically underspecified, contributes its semantic content to the proposition (12), whereas the demonstrative 'that' contributes merely a constraint on how the speaker can intend to fill in the corresponding gap. We can think of the subscripted '*distal*' as a kind of propositional metadata—not as a component of the proposition itself, but as a limitation on what can be a component in the subscripted position.

We should prefer a pure constraint semantics over this kind of mixed content/constraint theory on methodological grounds. As I argued in §1.3, we should spell out the nature of linguistic meaning by first spelling out its explanatory role. Mixed content/constraint accounts find linguistic meaning playing two distinct roles: the meanings of context-insensitive expressions play the role of fixing the expressions' semantic contents, whereas the meanings of semantically underspecified expressions play the different role of constraining the intentions a speaker can have in utter-

ing the expressions. Whereas we cannot adopt a semantic theory on which meaning plays a wholly content-determining role for the reasons I've spelled out here, we can easily adopt a wholly constraint-based semantics by saying that the meanings of context-insensitive expressions also function by constraining speakers' intentions, but simply do so to a greater degree. The meaning of 'stinks' is so strict that, in uttering it, I can only have the intentions involved in referring to the property of stinking. Assuming that it is better to have meaning play a unified explanatory role than a mixed explanatory role, and that a wholly constraint-based account of linguiatic meaning is possible, I submit that these considerations give us a reason to adopt such a fully-general account.

On this view, to say that an expression *e* is semantically underspecified is to say that it the that *e*'s meaning places relatively loose constraints on the communicative intentions of someone who utters *e*. This is a powerful way to think of underspecification because speech-act types can be broadly individuated in a variety of ways, and so expressions' meanings can give rise to their semantic underspecification in a variety of ways. I will try to precisify this point in the next section in the course of spelling out the basics of the speech-act-theoretic approach to lexical and compositonal semantics.

2.3 Speech-Act-Theoretic Semantics

What would a constraint-based semantic theory look like? I want now to take an initial step toward answering this question by articulating the basic framework of speech-act-theoretic semantics, according to which semantic values are types of speech acts.

Speech-act-theoretic semantics is inspired by Grice's (1968) theory of utterance-type meaning. According to Grice, a sentence S is meaningful for a group G just in case the members of G have "a procedure in their repertoires" for uttering S if they have communicative intentions of a

certain kind (1968: 127–8).[23] Because Grice also takes the performance of a speech act to be a matter of producing an utterance with these same sorts of intentions,[24] we can interpret him as holding that linguistic meaning is a relation borne by expressions to types of speech acts. For example, it follows from Grice's view that the meaningfulness of 'Paul philosophizes' boils down to the fact that speakers have a procedure in their repertoires for uttering the sentence if they overtly intend to bring about the belief (in some addressee) that they believe that Paul philosophizes. This explication is equivalent, by Gricean lights, to the idea that the sentence's meaning is constituted by speakers' shared procedure for uttering it in order to assert that Paul philosophizes. Along similar lines, we can say that Grice takes the meaningfulness of 'does John philosophize?' to be constituted by speakers' shared procedure for uttering the sentence in order to ask whether John philosophizes. Given this theory of linguistic meaning, it makes sense to think of a sentence's semantic value as a certain type of illocutionary act—the type that speakers have a procedure to perform by uttering the sentence. Moreover, Grice argues that the procedure we have for performing illocutionary acts using sentences are constituted by our procedures for referring and predicating with the sentence's simple parts (1968: 129–37). Given this view, it is natural to say that a Gricean compositional semantic theory would be one that showed how the types of acts of referring and predicating that are the semantic values of words add up to the types of illocutionary acts that are the semantic values of sentences.[25]

In order for Grice's theory to count as a constraint semantics, we need only interpret his talk of *having a procedure in one's repertoire* in the right way. In particular, we need to say that for a speaker to have a procedure in her repertoire that links an expression E to a type of communicative intention CI entails that the speaker can have an intention of type CI only if she utters E or

[23] I will discuss this part of Grice's theory of meaning in greater detail in §4.6.

[24] Grice (1968: 117–22). For more explicit statements of this Gricean theory of speech acts, see §1.2 and the works cited there.

[25] This is just a first pass. I will have much more to say in Chapters Four and Five.

some other expression similarly linked to CI. The nature of the modal, 'can', in this statement is an issue that a full-fledged metasemantic theory would have to resolve. I will replace this statement with a more satisfactory account of linguistic metasemantics in Chapter Four.

I won't try to give a speech-act-theoretic semantics for a large fragment of English here, but I can give a simple sketch, elaborate enough to precisify the discussion of underspecification that I have outlined so far. For now, I won't attempt to show how the meanings of sentences and other clauses are composed, as that will be my goal in Chapter Two. Instead, I will assume that every clause #(φ) consists of a sentence radical φ embedded within a mood-marker #. My goal here will be to give a simple account of the compositional semantics of moodless sentence radicals, and I will assume that the semantic value of a sentence radical is a force-free speech-act type of expressing some proposition. I don't thereby intend to commit myself to the idea that speakers can perform force-free acts of expressing propositions in isolation; rather it is compatible with everything I'll say here that expressing a proposition is always part of, or an abstraction away from, an illocutionary act of some kind, just as it is compatible with what I say here that acts of referring and predicating are only ever performed as parts of, or abstractions away from, illocutionary acts. (I'll have much more to say about what I mean by this in Chapter Five.)

On my preferred formulation, an actic semantics consists of two parts: first, an assignment of types of acts to semantically primitive expressions, and, second, a collection of compositional priniciples assigning different kinds of acts of predicating to different types of binary-branching syntactic constructions. (Again, I will spend a good deal of Chapter Five justifying this formulation and comparing it to other options.) Here are the rules for a very small fragment of English:

'stinks' $\Rightarrow$ a type that may be tokened only by referring to the property of being stinky.[26]

---

[26] '$\Rightarrow$' symbolizes whatever relation holds between an expression and the type of act that serves as its semantic value.

'Fido' ⇛ a type that may be tokened by referring to Fido.

'everyone' ⇛ a type that may be tokened by referring to the property of applying to everyone

Sentence Radical Composition [$_S$ [NP][VP]]

S's semantic value is a type that can be tokened by expressing a proposition by referring to an object *x* in a way that tokens NP's semantic value, referring to a property *F* in a way that tokens VP's semantic value, and predicating *F* of *x*.

Sentence Radical Composition [$_S$ [DP][VP]]

S's semantic value is a type that can be tokened by expressing a proposition by referring to a property *F* in a way that tokens DP's semantic value, referring to a property *G* in a way that tokens VP's semantic value, and predicating *F* of *G*.

The resulting theory-fragment allows us to show how the types of acts of expressing propositions that are performable with the sentences 'Fido stinks' and 'Everyone stinks' are constituted by the types of speech acts performable with their parts. I will consider a theory capable of dealing with a larger fragment in Chapter Five.

The key to understanding semantic underspecification is that the semantic value of an underspecified sentence radical will be a broadly-individuated speech-act type—one that can be tokened by expressing a variety of propositions. In cases of lexical underspecification, this will be because the semantic value of a semantically primitive constituent of the sentence radical is a type that can be tokened by referring to different objects on different occasions. For example:

'that' ⇛ a type that may be tokened only by acts of referring to non-persons that are relatively distal from the speaker's perspective (in a perhaps metaphorical sense).

'he' ⇛ a type that may be tokened only by acts of referring to males.

'I' $\Rrightarrow$ a type that may be tokened only by acts of referring to oneself.

The basic idea is straightforward: an expression's meaning plays the role of narrowing down the kind of speech act that someone might be performing in uttering the expression, but not necessarily to the point of uniqueness. The interpreter's task—which is to non-demonstratively infer which speech act the speaker is performing—is made easier because her sensitivity to the expression's meaning cuts down the space of possibilities about what the speaker might mean. Because each speech-act type X—whether a communicative illocutionary act or an act of referring—can be individuated in terms of the kinds of communicative intentions a speaker must have in order to perform a token of X, we could equivalently formulate the semantic values of words and sentences in terms of the kinds of communicative intentions a speaker might have in uttering them. (This point will be crucial to the account of mood that I will defend in Chapter Three.) On this account, each expression's semantic value plays the role for which I posited meaning in §1.3. The semantic values of semantically underspecified expressions play the same role, but do so in a more loosely-constraining way.

To account for DP restriction, we can formulate principles that deliver results like the following for DPs of the form 'Every NP':

'Every NP' $\Rrightarrow$ a type that may be tokened only by acts of referring to the property of *applying to every* ψ (where ψ may be any restriction of the property that a speaker can refer to in tokening NP's semantic value).[27]

[27] This treatment of quantifier phrases is a rough first pass. A proper treatment would have to take a stand on whether implicit restriction is triggered only in cases of determiners which are downward monotonic with respect to their nominals (see Barwise and Cooper 1981; Neale 2005), and would have to be more precise about what it means to say that a property ψ is a restriction of a property φ.

If Stanley and Szabó (2000) are right that nominals contain aphonic restrictor variables, then our compositional principles will have to be more complex in order to take into account the complex acts of predicating involved in referring to a restricted property with a DP's nominal. (I will have more to say about these complex acts of predicating in Chapter Five.)

The meanings of several of the other underspecified expressions that I discussed in §2.1 could be captured by compositional principles that assign them the following semantic values. Again, each of these is a rough first shot, and final versions would have to be sensitive to whatever hidden syntactic structure is present in the expressions' underlying syntax.

'It's raining' $\Rightarrow$ a type that may be tokened only by acts of expressing a proposition that it's raining at *l* at *t*, where *l* is a particular location and *t* is a particular time.

'Elmar's shark' $\Rightarrow$ a type that may be tokened only by acts of referring to a shark that bears *R* to Elmar, where *R* is a particular relation.[28]

'[$_{\text{Adj}}$ expensive [$_{\text{NP}}$ $F(x)$]]' $\Rightarrow$ a type that may be tokened only by acts of referring to the property of having a cost that is higher than some degree relevant to the cost of *F*'s, where F may be any property.[29]

[28] If Peters and Westerståhl (2013) are right, then it may be better to give a more complex, quantificational account of the possessive. Nothing about the framework I am defending here would conflict with such an account, although the quantificational approach would require showing how 'Elmar's shark' can be used to performa a complex act of referring to a property of properties.

[29] This is a simplified account based on Kennedy's (2007) defense of the idea that comparative adjectives contain hidden variables that determine a contextually salient standard on the basis of some contextually salient comparison class. I

Following Kratzer (2012), we can represent a modal claim 'must φ' as including variables for a conversational background *f* and an ordering source *g*, where each of these is taken as a sentential variable.[30] We can then give a preliminary account of modals as follows:

> must$_{f,g}$ φ $\Rrightarrow$ a type that can be tokened by expressing a proposition that some proposition p is true in the *q*-best worlds at which *r* is true, where *p*, *q*, and *r* are propositions that one could express by tokening φ, *f,* and *g,* respectively.

These could all be elaborated and improved, but they successfully illustrate my general strategy for understanding the semantics of any given underspecified expression, which is to specify the limits on the kinds of speech acts that a speaker can perform in uttering the expression. As the examples illustrate, these limits may be specified in a variety of ways, in keeping with the varieties of semantic underspecification that we encounter in natural language.

Conclusion and Mile-Marking

In Chapter One, I sketched my overall pragmatic and methodological framework for understanding communication and language use, and argued that linguistic meaning must play a specific kind of role in order to fit into that framework. In this chapter, I have argued not only that speech-act-theoretic semantics allows meaning to play that role, but also that truth-conditional semantics is internally unstable in a way that causes it to collapse into the speech-act-theoretic approach.

The argument can be summed up straightforwardly: if truth-conditional semantics is workable, then there must be some property of sentences, sentences-in-contexts, or utterances of sentences that fixes their truth conditions. But if this is so, then there must be some fact in virtue of which sentences (etc.) have these properties that can be distinguished from the facts in virtue of

[30] Like Kratzer, I'll remain neutral about where, if any, these variables are located in modal sentences' syntactic structures.

which speakers express propositions with sentences. But, given scrutiny to semantically underspecified expressions, it becomes clear that semantic content collapses into a speaker-based, intention-constituted notion, and so semantics must be the study of constraints on speakers' intentions.

I have also begun to sketch the basic components of speech-act-theoretic semantics as I understand it. Over the course of the next several chapters, I will flesh out this sketch considerably. In Chapter Three, I will expand the theory to cover sentences and embedded clauses, with a particular emphasis on mood. In Chapter Four, I will replace the notion of constraint on intentions, as well as the modal force implicit in my many statements here that an expression's semantic value is the type of speech act that a speaker can perform with it, with a more fully-articulated disposition-based metasemantics. In Chapter Five, I will explain in greater detail what it is for word-sized speech-act types to add up to sentence-sized speech-act types, and give a fuller treatment of the syntax–semantics interface and the nature of compositionality.

CHAPTER THREE: ILLOCUTIONARY ACTS AND THE SEMANTICS OF MOOD

In the last chapter, I began to show how to construct a lexical and compositional semantic theory for a language. But my account there was incomplete, because it didn't account adequately for the semantics of whole clauses. The speech acts that we perform with clauses (and sentences in particular) aren't limited to expressing propositions. Rather, in uttering a sentence, a speaker performs an illocutionary act. The semantics of whole clauses—and their mood in particular—raises a variety of difficult issues that have occupied many semanticists' attention, particularly over the past several decades. I turn to the semantics of mood here. Although I will only have space to adequately address the imperative mood in this chapter, I believe that my account generalizes to interrogatives and declaratives as well.

Introduction: Mood, Meaning, and Semantics

These three sentences differ with respect to *mood*:[1]

(1) Oliver buys Dan a drink. *declarative*

(2) Oliver, buy Dan a drink! *imperative*

(3) Does Oliver buy Dan a drink? *interrogative*

No syntactic criterion individuates the moods cross-linguistically (Han 1998), and linguists have argued that the clause-types must be individuated in terms of both syntactic structure and semantic function (Kaufmann 2012: §1.1.3; Portner 2004; Sadock and Zwicky 1985). Within a given language, mood is a syntactically realized property of clauses: each clause has a mood and no

[1] Linguists tend to call the relevant property 'clause-type', so as to disginguish it from verbal mood, which is a distinct morphosyntactic property of verbs (Portner 2004; Kaufmann 2012). Because philosophers use 'mood' almost exclusively to pick out the clause-level property that distinguishes (1)–(3), things will be simpler if I use 'mood' and 'clause-type' interchangably, stipualating that I will never mean verbal mood by 'mood'.

clause has more than one. The three clause-types on display in (1)–(3) are not the only ones, but they are the only types that are universal to all languages, and they are the only ones that I will focus on here.

(1)–(3) have a lot in common. With the apparent exception of 'does' in (3), they contain occurrences of all and only the same words, in the same order, and in closely related syntactic arrangements. The sentences seemingly have some component of their meanings in common, though the details about this are pre-theoretically fuzzy. Each of (1)–(3) would normally be used to talk about the same things—namely, Oliver, Dan, and the relation of buying a drink for someone. Seemingly—though this turns out to be a matter of contention—the three sentences can also be used to perform speech acts with the same propositional content: *that Oliver buys Dan a drink*.

Despite these similarities, the syntactic differences that distinguish (1)–(3) also signal semantic and pragmatic differences. We might stick to the idea that (1)–(3) can be used to express the same proposition, but we must also recognize that they are typically used to perform speech acts that differ in illocutionary force: whereas (1) can be used to *assert* or *suggest*, (2) can be used to *request* or *command*, and (3) can be used to *ask whether*. These differences in what we typically do with (1)–(3) reflect differences in their meanings. My primary topic here is the *semantics* of mood, although this is closely connected to the pragmatic differences as well.

My aim is to defend a Gricean, speech act theoretic account of of mood, according to which the semantic value of a clause is the type of illocutionary act with whose potential it is endowed, and according to which illocutionary acts are individuated by the kinds of intentions speakers have in performing them. The semantic values of declarative, imperative, and interrogative sentences are types that may be tokened only by assertive acts, directive acts, and acts of questioning, respectively. The result will be a semantic analysis of mood that builds on the key ideas of the act-theoretic semantic framework that I began to develop in Chapter Two.

I will focus mainly on the case of imperatives, for two reasons. First, the fact that many interrogatives are also *wh*-phrases (e.g. 'who are you?') makes comparing them with clauses of other types more complicated. Second, and more importantly, Griceans disagree with most contemporary semanticists in a fundamental way about the nature of questioning: Griceans take questions to be a genre of directive speech acts, whereas most contemporary accounts treat questions and directives as distinct categories of speech act. Because the accounts I will be comparing take the semantics of interrogatives to be intimately tied to the pragmatics of asking questions, this disagreement about the pragmatic issues makes it cumbersome to compare semantic views.

My plan is the following. In §3.1, I articulate the explanatory goals of a semantic theory of imperatives. In §3.2 I outline my own theory of imperatives and articulate the most threatening criticism of it—namely, that it fails to account for imperative clauses embedded within the scope of logical operators. In §3.3, I consider and reject several prominent theories of mood that take the semantic values of non-declaratives to be either propositions or truth conditions. In §3.4, I consider dynamic theories of the imperative, which understand the meanings of imperatives in terms of the effects that uttering them will have on conversational context. I reject dynamic theories in §3.5, arguing that they collapse into Gricean, act theoretic views like my own. In §3.6, I show how a Gricean, act theoretic view like my own can account for embedded imperatives.

3.1 What Do We Want from a Semantics for Imperatives?

A variety of explananda have been proposed for a theory of the semantics of mood. I consider six categories here, the first four of which I take to be genuine explananda, and the last two of which I look on with greater suspicion.

3.1.1 Imperative Meaning

The most basic requirement of a semantics for the imperative is to explain what is distinctive about the meanings of imperatives. We have very strong pre-theoretic intuitions that imperative sentences (e.g., 'Oliver, buy Dan a drink!') have different meanings than their declarative counterparts (e.g., 'Oliver buys Dan a drink.'). A more theory-laden reason for taking their meanings to differ arises from the fact that semantics often uses as data intuitions about the truth conditions of (utterances of) sentences. But our intuitions also suggest that (utterances of) imperatives aren't the right sorts of things to be true or false. A semantic theory of the imperative should precisify and explain all of these observations

A full semantic account of imperatives would also have to explain how imperatives' distinctive meanings arise compositionally from the meanings of their parts and their syntactic structures. This is not possible in the present state of inquiry because the syntactic features that distinguish the major clause-types in English and across languages are poorly understood.[2] I will therefore follow most other semanticists working on non-declaratives by making minimal assumptions about the grammar of clause-types, and of the imperative in particular. I will assume that clauses of each type can be factored into *mood markers* that encode their clause-type—symbolized by '⊦' for declarative, '!' for imperative, and '?' for interrogative—and moodless sentence radicals that

[2] See Han (1998) and Zanuttini, Pak, & Portner (2012) for discussions of how the syntax of imperatives bears on their semantics.

encode the constraints they place on which propositions speakers can express with them.[3] I symbolize clauses by prefixing their mood-markers:

(1*) ⊦(Oliver buy Dan a drink)

(2*) !(Oliver buy Dan a drink)

(3*) ?(Oliver buy Dan a drink)

I can therefore safely reframe the question about the semantics of imperatives as the following one: what is the semantic contribution of the mood-marker '!'?

3.1.2 Imperative Use

Closely conneted to the semantic question about imperatives is a pragmatic question, raised by the fact that imperatives are used to perform speech acts of a particular type. I will call speech acts of this type 'directives', where this label should be understood as applying to a genus that subsumes a variety of distinct types of illocutionary acts (including, for example, commands and re-

---

[3] See, for example, Charlow (2013), Davidson (1979), Grice (1968), Lewis (1970, 1975b), Sadock (1974), Schiffer (1972: 105–10), Searle (1969: 30), Starr (*ms*). For an extended defense of the view that mood-markers can be factored out at LF across languages, see Han (1998). Not all of the theories I will discuss below can usefully be taken as representing the meanings of imperatives this way; e.g., Hanks (2007, 2011) takes the semantic contribution of mood to be inseparable from the meanings of the rest of sentences (see §3.3.3). Portner (2004, 2007) and Zanuttini, Pak, & Portner (2012) take the semantic values of imperatives to be properties (see §3.4.1). This latter idea—recast as the idea that the semantic values of imperatives' mood-free radicals are acts of referring to properties—could be made to fit nicely with the version of speech-act-theoretic semantics that I defend here, particularly if the contents of intentions and directive speech acts are properties rather than propositions. On this issue, see footnote 6 of Chapter One, where I provisionally adopt the assumption that intentions and directive acts have propositions as their contents. To go along with that idea, I will here assume that directives contain mood-free sentence radicals whose semantic values are types of acts of expressing propositions. Moreover, I will continue with my awkward-sounding technical usage in schematic reports like 'A intends to *p*' and 'S directs A to *p*', which I use to report the facts that A bears the intenting relation to the proposition *p* and that the content of S's act of directing is an A-dependent proposition p, respectively. If it turns out that any of these provisional assumptions is bad, it would be unproblematic to rewrite this chapter in a way that adopts the alternative assumptions.

quests). A semantic theory of imperatives should explain, or set the stage for pragmatics to explain, how imperatives' disinctive meanings give rise to their distinctive use.[4]

We might also want our account to explain some facts about the felicity conditions of imperatives—why, for example, the following snippets of dialogue are defective in some way (Starr *ms*: §2.3):

(4) a. Unicorns don't, never have, and never will exist.

⋮

b. Bring me a unicorn!

(5) a. The door is open.

⋮

b. Open the door!

Both of these data are instances of much more general phenomena. As (5) demonstrates: for any sentence radical φ, it is (usually) infelicitous to follow an utterance of ⌜⊦φ⌝ with an utterance of ⌜! φ⌝. But the phenomenon is obviously more general than that description would suggest. What Starr here takes note of is a particular, linguistically-bound case of a general fact about directive communicative acts—namely, that it doesn't make sense to direct someone to make something the case when that thing already is the case. Suppose that (5a) had not already been uttered, for example, but that it was merely obvious to everyone involved in the conversation that the door was already open. It would be equally as problematic to utter (5b) in this situation as it would be to utter it in the wake of an utterance of (5a). The same would be true of any attempt, linguistic or nonlinguistic, to direct someone to open the door. Quite generally: it does not make sense to direct someone to do what's already done. Whatever explanation we give of cases like (4) and (5) therefore shouldn't be tied too closely to language.

---

[4] For other articulations of this desideratum, see Charlow (2013: §4.2), Davidson (1979: 116).

3.1.3 Embedded Imperatives

Imperative clauses seem to embed within the scope of at least some binary connectives. In some cases, as in (6)–(7), both clauses thus embedded are imperatives.

(6) Buy me a drink and make it a stiff one! !φ and !ψ

(7) Buy me a drink or leave me alone! !φ or !ψ

In other cases, imperatives and declaratives can apparently be mixed inside the scopes of connectives.

(8) If the bartender comes back, buy me a drink.[5] (if ⊢φ)(!ψ)

(9) a. Give me all your money or the puppy gets it. !φ or ⊢ψ

b. You mow the lawn or I'll clean the garage (Starr *ms*: §2.1)

(10) a. Say that again and I'll scream. !φ and ⊢ψ

b. Stay here and I'll go to the store.

(11) If he comes back, buy me a drink and I'll pay you back. (if ⊢φ)(!ψ and ⊢σ)

There is less than full agreement about how the logical forms of sentences like these should be understood, and this lack of agreement is connected to the fact that the grammar of mood is itself a matter of dispute. Even given our simplifying assumption that a clause's mood can be factored out and represented as a prefixed mood-marker, several questions remain about the grammar of (6)–(11).

For example: how do the scopes of mood-markers interact with the scopes of connectives? Should (6) be represented as (6*a), with the mood-marker taking wide scope, or as (6*b), with mood-markers taking narrow scope?

---

[5] It might be best not to think of 'if' as a binary connective; I group conditional imperatives in with mixed-mood conjunctions and disjunctions because many of the same issues arise for both.

(6*) a. !(A buy S a drink and A make it a stiff one) !(φ and ψ)

b. !(A buy S a drink) and !(A make it a stiff one) !φ and !ψ

If mood-markers always take wide scope over connectives, as in (6*a), then things are relatively simple. For this would mean that there aren't any embedded imperatives after all: what appear to be embedded imperatives would actually be complex sentence radicals embedded under a single imperative mood-marker.

One prima facie reason to think that this strategy won't work, and that mood-markers at least sometimes take narrow scope with respect to logical connectives, is the existence of sentences like (8)–(10), which embed clauses of two different types. Mixed-mood sentences like these seemingly cannot be treated as complex sentence radicals embedded under single mood-markers, because they contain too many mood-markers, and the resulting sentences are apparently neither simply declaratives nor simply imperatives but instead complex combinations of the two. It seems that the meanings of imperative clauses must make it possible for them to be conjoined and disjoined with, and conditionalized on, declarative clauses.

One way out of this conclusion would be to suppose that sentences like (8)–(10) are not as they appear. One way of implementing this idea, as it applies to (9) and (10), would be to hold that they are actually the conditionals (9*) and (10*) in disguise.

(9*) a. If you don't give me all your money, the puppy gets it. ⊦(If not-φ)(ψ)

b. If you don't mow the lawn, I'll clean the garage.

(10*) a. If you say that again, I'll scream. ⊦(If φ)(ψ)

b. If you stay here, I'll go to the store.

Charlow argues that we should apply this strategy to at least some mixed-mood conjunctions and disjunctions. He motivates it by pointing out that it is plausible that (9a) and (10a) can be used to perform the same speech acts as (9a*) and (10a*). Nonetheless, Charlow acknowledges that "there

is a mystery about how to account for" these disguised conditional readings: "the pressure is to come up with an analysis of the relevant connectives ('or' and 'and') that allows them to transform imperatives into subordinate clauses functioning to restrict the domain of quantificational modals like 'will'" (2013: *fn*.45). In other words: even in cases where conditional readings are plausible, it wouldn't be trivial to show how these readings could arise grammatically. Moreover, if we interpret (9a) and (10b) as cases of genuine disjunctions and conjunctions, rather than as disguised conditionals, it seems that we can give a pragmatic explanation of their conditional readings. Declarative disjunctions sometimes "sound like" conditionals, after all ('Either the money is in my bank account or the puppy gets it'), as do declarative conjunctions ('John says that one more time and I'll scream'), and whatever pragmatic explanation we give of the conditional readings of these cases can likely be extended to cases like (9a) and (10a).

What about conditional imperatives, such as (8)? Some authors have argued, on the grounds that mood signals force and we don't perform illocutionary acts with the consequents of conditionals, that mood-markers always take wide-scope in conditionals, so that the logical form of a conditional imperative is !((if φ)(ψ)).[6] But there are excellent reasons to disagree. One is the existence of conditionals like (11), whose consequents conjoin an imperative and a declarative. The grammaticality of (11) counts against taking the mood-markers in conditional imperatives to have wide scope, because, once again, it seems that we have too many mood-markers (Starr *ms*: §2.1). But suppose that we stick to the idea of reinterpreting mixed-mood conjunctions as disguised conditionals and give the consequent of (11) the same treatment, thereby yielding (11*):

(11*) If he comes back, then if you buy me a drink, I'll pay you back.

⊦(if φ)((if ψ)(σ))

[6] E.g., Dummett (1973: 338–48).

Again, it is intuitively plausible to read (11) as (11*), although, again, the grammatical implementation of this proposal is not straightforward, and a pragmatic explanation might turn out to be preferable.

But we shouldn't hold out much hope that all of this wide-scopery will work out. The strategy hinges on our ability to read mixed-mood conjunctions and disjunctions as indicative conditionals, and, as Starr (*ms*: §2.1) shows, we can't always do this. First, consider (9b) and its conditional counterpart, (9*b):

(9) b. You mow the lawn or I'll clean the garage !φ or ⊦ψ

(9*) b. If you don't mow the lawn, I'll clean the garage. ⊦(If not-φ)(ψ)

Starr persuasively argues that (9b) has a purely disjunctive reading that cannot be assimilated to (9*b):

> (Context: It's Friday and we have a list of chores to do over the week- end. We've decided to that one of your chores is mowing the lawn and one of mine is cleaning the garage, though there are others. We have to do one chore total today. Right now, we are considering which one.)
>
> [X] a. Me: I don't know, you mow the lawn or I'll clean the garage.
> b. You: I'm fine with either.
>
> It is often said that disjunctions which mix imperatives and declaratives only have a negative conditional meaning (Dummett 1973). For example, [Ya] means that if you don't mow the lawn, I will take away your cell phone. But unlike [X], a response like [Yb] only evidences that you did not understand what I said.
>
> [Y] a. Me: You mow the lawn or I'll take away your cell phone!
> # b. You: I'm fine with either.
>
> This seems to show that imperatives can scope under or. (Starr *ms*: §2.1)

Some mixed-mood conjunctions likewise can't be read as conditionals. Consider (10b), which is similar to one of Starr's examples.

(10) b. Stay here and I'll go to the store. !φ and ⊦ψ

(10*) b. If you stay here, I'll go to the store. ⊦(If φ)(ψ)

On at least one natural reading of (10b), the speaker is not saying that her going to the store is conditional upon her addressee staying where they are.

These readings threaten our attempt to do away with narrow-scoped mood. Moreover, there is no problem with embedding these true mixed-mood conjunctions and disjunctions in conditionals, and so these readings also to put an end to our ambition of understaning mood as taking wide scope over them as well.

Perhaps there is a way to explain these cases away, and to retreat once again to the position that mood always scopes over the connectives, but I can't see how that would work. And yet, some prominent semantic theories of mood seem to hang on the wide-scope analysis. For example, a central tenet of Portner's (2004, 2007) influential theory is that imperatives' semantic values differ in type from those of declaratives (this is how he explains the fact that imperatives are not truth-apt). But, assuming that Portner also accepts the standard wisdom that conjunction and disjunction range only over pairs of arguments of the same semantic type (Partee and Rooth 1983), it follows that there are no mixed-mood conjunctions or disjunctions, and that apparent examples like (9) and (10) must be explained away. This constitutes a serious problem for Portner, whose account I will return to in §3.1.4.

The framework I will defend here is compatible with either hypothesis about the relative scopes of mood-markers and connectives. I show how it could account for the possibility that mood-markers always take wide scope in §3.2.4, and how it can be adjusted to accomodate the view that mood-markers take narrow scope in §3.6.

3.1.4 Illocutionary Underspecification

Imperatives can be used to perform speech acts with a variety of different kinds of illocutionary force. Perhaps the most obvious examples are commands and requests, but ordinary language includes many verbs that can be used to talk about acts that are literally and directly performable

with imperatives. There is very little agreement about the most perspicuous and explanatorily fruitful way of classifying illocutionary acts into categories, and I won't attempt a comprehensive scheme here, but here are a few example sentences together with the types of directive acts that we might characteristically perform with them:[7]

(12)	a. Have a cookie!	*suggestion/permission*
	b. Have a nice day!/Drop dead!	*wish/curse*
	c. Take the number 9 bus.	advice
	d. Leave your name at the tone.	*instruction*
	e. Don't eat before you swim!	*warning*
	f. Go ahead, use heroin (and see what happens)!	*malefactive/dare*
	g. Go ahead, eat my lunch.	*concession*
	h. Don't sit down	*prohibition*
	i. Sit down!	*command*
	j. Please sign this form.	*request*

It is crucial that we are not mislead by examples like these into thinking that something about the meanings of sentences like (12a–i) encodes the specific illocutionary forces listed with them. In the right circumstances and with the right tones of voice, most of the sentences on the left could be used to perform speech acts of most of the types listed on the right. The type of illocutionary force listed to the right of each example sentence merely labels a reading of the sentence that we are likely to jump to when encountering the sentence in isolation. This point is somewhat complicated by the presence of expressions like 'go ahead' in (12f) and (12g) and 'please' in (12j), which clearly push us, in virtue of their meanings, toward particular ways of interpreting the sentences.

[7] Some of these examples are borrowed from Starr (*ms*: §5.1). I don't think all of the kinds of illocutionary acts he mentions there are really directives, however; some (such as what he calls "passive advertisements", as in, 'win a cruise to Jamaica') are best understood as kinds of indirect directive acts (see §3.2.3).

But we should recognize two important points with respect to expressions like these, whose de facto function seems at least sometimes to involve indicating illocutionary force. First, it is possible to concede or dare without uttering 'go ahead' and to request without uttering 'please'. In their force-indicating role, these expressions serve as linguistic substitutes for information that could otherwise be supplied extralinguistically. For example: if I utter 'eat my lunch' in a defeated tone while throwing my hands in the air, and immediately after my addressee has asked to eat my lunch, it will be obvious enough that I am conceding. Second, even when an expression that normally indicates a certain illocutionaty force is uttered, extra-linguistic facts may make it possible (and, in some cases, inevitable) that the speaker performs some other illocutionary act. For example: if a police officer who has just pulled me over says, 'please blow into this tube', it will be clear to me, notwithstanding the officer's odd gesture at politeness, that I am being given a command.

So one explanandum represented by examples like (12a–o) consists of the fact that most imperative sentences can, in varying circumstances, be used to perform illocutionary acts with a variety of different forces. Semantics should contribute to an explanation of this open-endedness by assigning meanings to imperatives that help explain why they can be used in so many ways. On the other hand, all of the types of illocutionary force listed in (12) are species of a single genus: they are all types of *directive* force. So a semantics for imperatives should also explain why the open-endedness of imperative meaning has this particular limit.

Finally, we should recognize that this point generalizes in two ways.

(i) This kind of open-endedness-within-limits is also a feature of other clause-types. Declaratives can be used to perform assertions, suggestions, assurances, and so on. Our vocabulary for the acts we perform using interrogatives is less rich, but we may wish to distinguish between (for example) questioning, interrogating (Schiffer), begging for information, and wondering out loud.

It would therefore be nice if our explanation of the open-endedness of imperative meaning were part of a general theory of illocutionary underspecification.

(ii) Illocutionary underspecification resembles other kinds of semantic underspecification, such as the kinds of underspecification that are displayed by indexicals, demonstratives, quantifier phrases, and so on. In all such cases, the meaning of a word, phrase, or clause underspecifies some property of the speech acts that can literally and directly be performed with it. Indexicals' meanings underspecify the propositional contents that can be expressed with them. Clauses' meanings underspecify the illocutionary acts that can be performed with them. It is therefore reasonable to expect that a fully general account would explain all of these kinds of underspecification in similar ways.

3.1.5 Imperative Logic?

Now I reach the first of two proposed explananda that deserve a more critical look than the previous four. This is the idea that imperative sentences can stand in consequence relations. This claim is common to many recent attempts to give a semantic theory of imperatives—particularly the emerging dynamic approaches that I will discuss in §3.4. The idea of imperative consequence is usually motivated by examples, such as these:

(13) a. Attack at dawn if the weather is fine.
b. The weather is fine.
c. So, attack at dawn! Parsons (2013: 61)

(14) a. Take out the trash and mow the lawn!
b. So, take out the trash! Charlow (2013: §2.2)

It might be tempting to conclude from examples like these that imperatives are closed under the same consequence relation as declaratives—that, in other words, the following generalization holds:

Simple Imperative Consequence

For any set of imperative sentences (closed under logical connectives) !Σ, and where ⊦Σ is a set of declarative sentences obtained by substituting a '⊦' for each '!' in every sentence in ⊦Σ, and for any sentence radical φ: $!\Sigma \vDash !\varphi$ iff $\vdash\Sigma \vDash \vdash\varphi$.

But as both Charlow (2013: §2.2) and Starr (*ms*: §2.3) both stress, there are counterexamples to this principle. The counterexample both discuss most carefully is the imperative version of Ross's paradox:

(15) a. Post the letter!

⊭ b. Post the letter or burn the letter![8]

Charlow argues that although counterexamples to Simple Imperative Consequence like (15) shouldn't dissuade us from attempting to give a logic of imperatives, they complicate the project by necessitating the development of a nonclassical imperative logic. Starr (*ms*) makes accounting for (15) one of the central explananda of his semantics, arguing that we should strive for both a univocal semantics for the connectives and a single consequence relation that applies to both imperatives and declaratives, and that attempting to account for counterexamples like (15) gives us good reason to adopt broadly revisionist, dynamic notions of both consequence and truth.

---

[8] It is worth noting that, while Starr represents the logical form of (15b) as ⌜!φ and !ψ⌝, Charlow represents it as ⌜!(φ and ψ)⌝. So although both take the invalidity of (15) to be a counterexample to a precisification of my Simple Imperative Consequence principle, in fact they have different precisifications in mind, depending on how closure under logical connectives is understood when it comes to imperatives.

Rather than taking all of this at face value, I think it's crucial to ask what an account of imperative consequence would be a theory *of*. What are we saying when we say that an imperative sentence can follow from another sentence, or vice versa? To put things in a more overtly metaphysical regiseter: in virtue of what does a consequence relation between imperative sentences hold? In my view, the semanticists who appeal to intuitions about imperative consequence fail to adequately address these questions. Insofar as they do provide answers, or are committed to answers, those answers often turn out to conflict, so that "imperative consequence" winds up picking out rather different kinds of relations in the context of different theories.[9]

Even before we get into the details of particular theories, there are excellent reasons to think that any consequence relation for imperatives will be revisionary, not just in the sense that it will be nonclassical, but also in the sense that its subject matter will be different than the subject matter traditionally aimed at by logicians. Josh Parsons gives us a nice way of approaching this issue by pointing out that the idea of imperative consequence gives rise to the following inconsistent triad:

(16) (i) There are non-trivially valid arguments containing imperatives.

(ii) Imperatives are not truth-apt.

(iii) Validity is truth-preservation.

Following Parsons, we can classify approaches to imperative consequence as type-1, type-2, or type-3, depending on whether they reject (16i), (16ii), or (16iii), respectively.

As we'll see in §3.4, several influential semantic approaches to imperatives claim, partly on grounds independent of consequence, that imperatives are truth-apt. This makes a type-2 theory of imperative consequence seem attractive, for the obivous reason that it would allow imperative consequence to be just the same sort of relation as classical consequence—namely, necessary

[9] Charlow (*ms*) makes this point nicely in arguing for pluralism about both the subject matter and the mechanics of imperative consequence.

truth preservation, standardly formalized in possible worlds semantics by saying that the semantic value $[\![\sigma]\!]$ of each sentence σ is a set of possible worlds, and that, for any set of sentences $\Sigma = \{\varphi_1,...,\varphi_n\}$ and sentence ψ, $\Sigma \vDash \psi$ iff $\cap([\![\varphi_1]\!],...,[\![\varphi_n]\!]) \subseteq [\![\psi]\!]$.

This way of thinking about consequence—even as it applies to declarative sentences—is incompatible with commitments I've already incurred in the last two chapters. Specifically: I don't think that declarative natural language sentences are the right sorts of things to be the bearers of truth or the relata of consequence relation—only propositions and psychological states with propositional content are. But even if declaratives did stand in consequence relations, in the traditional sense of necessary truth preservation, there would still be significant further obstacles to extending that account to imperatives. The most obvious obstacle to type-2 accounts is that of explaining away our strong intuitive sense that imperatives just aren't the right sorts of things to be called true or false. Moreover, this sense can be backed up with reasons. As several philosophers have argued, there are foundational reasons to be suspicious of the idea that imperatives are in the business of being true or false. For one thing, we don't use imperatives to make true or false claims, and so it's hard to see why imperatives should have truth conditions, or how they got them (Charlow 2013; Parsons 2013). And, it's not merely our fuzzy intuitions that need to be explained away; it also seems that truth cannot play its usual explanatory roles when it comes to imperatives:

> It is not just manners or style that lead us to blanch when we are asked to ascribe truth or falsity to an imperative. Such ascriptions are ungrammatical. How is an explanation of the inconsistency of, say, !φ and ¡¬φ supposed to get going, if we cannot grammatically express the reductio assumption in our metalanguage?
>
> ?? Suppose it is true that go to your room and that you may stay up. Then...
>
> What good are truth conditions for imperatives if they do not enable us to explain, in terms that are, at a minimum, intelligible to us, the things we are interested in explaining? (Charlow 2013: §4.1)

So it's not enough for a type-2 account of imperative consequence to assign truth conditions to imperatives that vindicate all of our intuitions about imperatives standing in consequence relations; the account must also explain why imperatives have those truth conditions, why they don't seem to have them, and why their truth conditions don't seem to play some of the usual explanatory roles.

Seemingly invalid arguments like (15), which suggest that imperative consequence is not classical, also give us a reason to think that developing the technical aspects of a type-2 account would be difficult. Seemingly, it would involve articulating a nonclassical consequence relation and explaining why that nonclassical notion of consequence applies to imperatives but not to declaratives. In practice, this has been a factor that has pushed theorists away from type-2 accounts and toward type-3 (or type-2+3) accounts, on which consequence does not boil down to truth preservation, or at least not when imperatives are involved.[10]

Several type-3 theories have been offered in the literature (see, e.g., Parsons 2013; Vranas 2011, 2013). One way of going type-3 that will become significant later is to adopt a dynamic theory of consequence, according to which a theory of consequence is not about truth preservation (although some defined notion of truth preservation may coincide with consequence in some cases). Rather, dynamic theories of consequence say that a sentence ψ follows from a series of sentences $\varphi_1,...,\varphi_n$ just in case updating a context with successful utterances of $\varphi_1,...,\varphi_n$ (in order) would make a successful utterance of ψ in the newly-updated context redundant—i.e., just in case any context updated with $\varphi_1,...,\varphi_n$ (in order) is a fixed point for ψ. Formally: $\varphi_1,...,\varphi_n \models \psi$ iff, for

[10] For a persuasive summary of some of the factors that cause the search for a type-2 theory to collapse into a type-3 theory, see Parsons (2012, 2013). For some other arguments that we should think outside the box when it comes to the question of what a theory of imperative consequence would be a theory of, see Charlow (2013: §2.2, *fn*.8; *ms*).

any context C, $C[\varphi_1],...,[\varphi_n][\psi] = C[\varphi_1],...,[\varphi_n]$.[11] It should be clear that adopting a dynamic notion of consequence forces a significant departure from traditional theories, not only of the technical details of the consequence relation, but also of the nature of logic itself. So, although various independent motivations for dynamic semantic and dynamic logic have been articulated,[12] any particular such motivation is not to be taken lightly, and it is therefore worth exploring all other options before a massively revisionary conception of both the nature and content of the consequence relation is adopted. Moreover, as I will argue in §3.5, dynamic semantic theories are premised on misguided theories of communication.

Both type-2 and type-3 theories start from the assumption that declarative sentences stand in consequence relations, and attempt to show how these relations can be extended to imperative sentences as well. My view is that propositions are the relata of consequence, and that natural language sentences do not stand in consequence relations, even indirectly by expressing propositions relative to contexts. On my view, our intuitions about consequence, which may seem to be prompted by collections of delcarative sentences, are actually prompted by the propositions that we easily and involuntarily imagine speakers asserting with those sentences. Given these views, it should come as no surprise that I don't think that imperative sentences stand in consequence relations either. In other words, I hold a type-1 theory of "imperative consequence"—a theory on which there is no such thing.

---

[11] This notation, which is adopted by Starr (*ms*), is due to Veltman (1996). I will explain it in greater detail below. The basic idea is that $[\varphi]$ is the context change potential (CCP) of a sentence φ, where a CCP is a function mapping contexts into contexts, and $C[\varphi]$ is the result of applying $[\varphi]$ to a context C. As Veltman explains, his reason for adopting the postfix notation is that it makes it easier to represent the result of applying a series of CCPs to a single context, as in the example in the text.

[12] An influential recent defense of dynamic consequence is van Benthem (2011). For formally precise criteria of when dynamic semantics and dynamic consequence are required, see Rothschild and Yalcin (*ms*) and Bonnay and Westerståhl (*ms*).

Type-2 theories incur the burden of explaining why imperatives have truth conditions but seem not to, as well as the burden of specifying and motivating an imperative consequence relation that deviates from classical logic. Type-3 theories also incur the latter burden, as well as the foundational burden of motivating a substantial paradigm shift away from a tried-and-tested view about the nature of consequence and the subject matter of logic. Going type-1 incurs none of these burdens. Since the only reasons for believing in imperative consequence are our intuitions that there are valid arguments whose premises or conclusions contain imperative clauses, the only explanatory burden incurred by type-1 theories is to show why these intuitions are misguided. I will articulate my own type-1 theory, as well as my explanation of the intuitions that give rise to the notion of imperative consequence, in §3.2.2.

3.1.6 Imperative–Modal Connections?

Another purported datum about which I am suspicious is the idea that the meanings of imperatives are in some way intimately related to the meanings of deontic modals. The most extreme form of this idea is manifested in theories that take the meaning of an imperative clause !φ to be identical to that of a modal clause □φ, with '□' expressing deontic necessity (e.g., Aloni 2007; Han 1998; Lewis 1975b; Kaufmann and Schwager 2009; Schwager 2006a,b). Other theories posit weaker, but still quite prominent, relationships between the semantics of imperative sentences and their corresponding modal sentences (e.g., Charlow 2013; Portner 2007; Starr *ms*). I will discuss these idea in §3.3 and §3.4. For now, I want to consider the data that has led some theorists to posit these semantically-relevant relationships between imperatives and modals in the first place.

One source of support for an imperative–modal connection arises from the fact that arguments involving deontic modals exhibit some of the same peculiarities as arguments involving

imperatives. As originally noted by Ross (1941, 1944), and as remarked upon by many contemporary semanticists (Charlow 2013: §2.2; Starr *ms*: 2.3), (15) is strikingly reminiscient of (17):

(15) a. Post the letter!

⊭ b. Post the letter or burn the letter!

(17) a. You ought to post the letter.

⊭ b. You ought to post or burn the letter.

Before drawing any sweeping conclusions from this single datapoint, however, it is a good idea to recognize it as an instance of a broader pattern.

(18) a. You intend to post the letter.

⊭ b. You intend to post or burn the letter.

(19) a. You desire to post the letter.

⊭ b. You desire to post or burn the letter.

(20) a. You fear posting the letter.

⊭ b. You fear posting or burning the letter.

(21) a. You have reason to post the letter.

⊭ b. You have reason to post or burn the letter.

(22) a. Posting the letter is the thing to do.

⊭ b. Posting or burning the letter is the thing to do.

The following generalization is plausible: taking '☺' to schematize either the imperative mood-marker, a deontic modal expression, or operators that express normative, connative, or otherwise action-guiding states about some agent:

(23) a. ☺φ

⊭ b. ☺(φ or ψ)

Whereas many approaches to imperative meaning make hay of the similarity of (15) and (17), the theory I will present in §3.2 places greater emphasis on the relationship between (15) and (18), in that it grounds the performance of directive speech acts in the intention to produce intentions, and in that it traces the etiology of our intuitions about "imperative inference" back to our sensitivity to the norms governing the rationality of jointly-held intentions. Might there be a way to bring our explanations of (15) and (17)–(23) under the same explanatory roof? One plausible option for how to do this would be to hold that the semantic function of all of the operators schematized by '☺' is to affect one's interlocutors' noncognitive, action-guiding attitudes. This is part of the agenda of contemporary expressivism.[13] It would go beyond the intended scope of this chapter to discuss or evaluate expressivism in any detail. Suffice it to say that my own account of the semantics and pragmatics of imperatives—but not necessarily modals or moral language—could be construed as expressivist in something like the contemporary sense, since I ground the performance of directive speech acts (and therefore the semantics of imperatives) in the intention to produce action-guiding mental states—specifically, intentions—in one's addressee. Of course, if this makes me an expressivist, it also makes Grice, Strawson, and Schiffer expressivists.

Perhaps the most intuitive source of support for an imperative–modal connection is the idea that the act of telling someone to do something can sometimes constitute a reason why they ought to do it. Portner defends a broad generalization of this idea, arguing that part of the purpose of uttering an imperative !φ in a world w is to update the ordering source g relative to which deontic modals are interpreted, such that if g* is the updated ordering source, $g^*(w) = g(w) \cup \{[\![\varphi]\!]\}$.

[13] See, for example, Blackburn (1984), Gibbard (1990, 2003).

Some of the predictions of this view seem right. For example, it nicely explains the felicity of the following snippet of discourse:

(25) a. A: Go present this proposal to our bankers today!

b. B: I should take the 7 a.m. flight to New York then. (Portner 2007: 353)

But there are plenty of uses of imperatives that seem recalcitrant here. For example, Portner's account makes the wrong prediction about (26):

(26) a. A: Have some tea!

? b. B: I should boil some water then.

The problem with (26) seems to be that it would be most natural to use (26a) to *invite* or *permit* someone to have some tea, and it follows from an accepted invitation or permission only that one may do what one has been invited to do, not that one should do it. Another of Portner's theory's strange predictions happens to feature in an argument he gives against a Gricean account of the kind I will defend here.

> One might think that we should simply assign imperatives a general directive interpretation, and allow the subvarieties to emerge from pragmatic reasoning of the communicative-intentional sort. That is, the speaker counts on the addressee to be able to determine his/her intention in uttering an imperative, and this will involve figuring out what subvariety of directive force must be intended. While this type of Gricean reasoning certainly plays a role, it cannot be the whole story. A key piece of evidence for this point concerns the "psycho boss". You're being sent out of town to a meeting, and your boss says:
>
> [27] Be there at least two hours early.
>
> By itself, [27] can be taken as an order, and if your boss is that kind, you might well understand it that way. Next the boss says:
>
> [28] Then, have a bite to eat.
>
> If [27] was taken as an order, [28] is bizarre. You have to take it as an order too, and your boss must be crazy to order you around at that level of detail. Of course it makes sense to interpret [28] as a suggestion, but then you have to interpret [27] as a suggestion too. Why can't [27] be understood as an order and [28] as a suggestion? You might think it's just socially impossible to switch from being the ordering-boss to being the suggesting-boss so abruptly, but a parallel sequence with overt operators wouldn't be odd in the same way as [27]–[28]:

[29] a. You are ordered to be there at least two hours early.
b. Then, I suggest you have a bite to eat. (Portner 2007: 356)

For reasons I needn't fully articulate, but which are intimately related to Portner's account of "the connections between the semantics of imperatives and the semantics of modals" (2007: 356), his account predicts that we can't go back and forth between different kinds of directive speech acts willy nilly. Portner thinks that a Gricean theory along the lines of the one I will propose below makes the opposite prediction. (I'm not so sure that it makes any such prediction, but nevermind.) In the quoted passage, Portner claims that the data bears his prediction out. I emphatically disagree, and, indeed, I think that Portner's objection to a Gricean account works better as an objection to his own account. There is nothing bizarre about the same speaker following an order/command with a suggestion. Indeed, Portner's examples (27)–(28) and (29) strike me as perfectly natural, and I fail to see why Portner thinks otherwise.

In short, I think that the idea of an imperative–modal connection is considerably overblown, and that the degree to which many contemporary theories are designed to accommodate it is not among their strengths. Of course, there is an obvious sense in which different kinds of imperatives and modals are connected, but the Gricean theory I will offer here can account for that. I will argue that imperatives are used to perform directive speech acts, and that particular forms of directive speech acts are individuated in terms of the reasons on the basis of which their speaker intends her addressee to follow her directions. In addressing to you a commanding that *p*, I intend to bring about your intention to bring about *p* on the basis of my authority to require you bring about *p*. If I really do have this authority, then there is an obvious sense in which you should then intend to bring about *p*—namely, you should do *p* because I told you to. Likewise, if I advise you to bring about *p*, I intend you to recognize that *p* is in your interest. If *p* really is in your interest, then you should do *p* because it is in your interest. But I needn't *really* have authority in order to command you to do something, and *p* needn't really be in your interest in order for me to ad-

vise you to make *p* the case, and a theory of imperatives should also recognize this as well. This point bears emphasizing: suppose that I utter 'take the number 9 bus', thereby giving you *bad* advice—suppose, for example that taking the number 8 bus would be in your interest, but taking the number 9 bus wouldn't. And suppose that you accept my advice. The sentence 'you should take the number 9 bus' does not become true, even if we interpret the deontic modal 'should' relative to a teleological ordering source. The imperative–modal connection is therefore a weak one that depends not only on the content of a speaker's intentions, but, at least in some cases, on their aptness. We shouldn't try to account for this sort of imperative-modal connection using the tools of semantics.

3.2 Act-Theoretic Semantics for Mood, Part 1

Now that we have a better idea of what a good theory of the semantics and pragmatics of imperatives should look like, I want to begin arguing that a speech act-theoretic semantics built on the foundation of a Gricean theory of communication is what we need. I won't finish articulating the theory until §3.6, but I will lay out the basics here.

3.2.1 A Speech Act-Theoretic Semantics for Imperatives

The speech-act-theoretic account of clause types that I will offer here is built on top of the Gricean account of communicative illocutionary acts that I outlined in §1.2. The basic idea behind the theory is that the different clause-types are used to perform different types of illocutionary acts: we perform assertives with declaratives, directives with imperatives, and questions with interrogatives. What's novel to my theory is how this platitudinous-seeming idea is fleshed out within the semantic framework that I began to outline at the end of Chapter Two.

I begin with the stipulation that the component of the compositional semantic theory that I began to outline at the end of Chapter Two is actually a fragment that deals only with the seman-

tics of *sentence radicals*. Taking φ as a schematic variable ranging over sentence radicals, we add the following three compositional principles to our semantics:

Declarative Composition [$_S$ ⊢φ]

S's semantic value is a type that is tokenable by an assertive act whose content could be expressed by tokening the semantic value of φ.

Imperative Composition [$_S$!φ]

S's semantic value is a type that is tokenable by a directive act whose content could be expressed by tokening the semantic value of φ.

Polar Interrogative Composition [$_S$?φ]

S's semantic value is a type that is tokenable by an act of asking-whether whose content could be expressed by tokening the semantic value of φ.

This first draft will come in for significant revision in §3.6, but the main ideas of the view are already intact here, and it is already easy to see how this account satisfies most of the criteria outlined in §3.1. By assigning them different types of illocutionary acts as their semantic values, I both explain how and in what way the three clause-types differ in meaning (§3.1.1) and I explain how this difference in meaning is connected to a difference in use (§3.1.2). At the end of Chapter Two, I argued that it makes sense to think of the semantic value of an expression as a constraint on the kinds of communicative intentions that a speaker can have in uttering it, and that, as Griceans, we can represent these constraints as types of speech acts. The semantics for clause-types that I've just sketched fits nicely into this picture. The semantic value of a sentence is a type that can be tokened by performing a certain range of illocutionary acts.[14] The meaning of the clause's

[14] As I said in Chapter Two, my use of 'can' in this sentence is sloppy and under-developed. I will replace it with a fuller account of the metasemantic relation borne by expressions to their semantic values in Chapter Five.

sentence radical constrains the act of expressing a proposition that can be performed in uttering the clause (in a way that is compositionally determined in accordance with the kinds of principles outlined in Chapter Two), and the meaning of the clause's mood-marker constrains the force of acts that can be tokened in uttering it.

The account of underspecification that I defended in Chapter Two also yields a parsimonious explanation of the fact that the meaning of each clause-type underspecifies the particular types of illocutionary act that are performable with sentences of that type (§3.1.4). In Chapter Two, I argued that the semantic values of semantically underspecified expressions are act-types that are individuated more broadly than are the act-types serving as the semantic values of other expressions. For example: whereas the semantic value of 'loves' may be tokened only by referring to the love relation, the semantic value of 'he' may be tokened by referring to any male. In the same vein, the semantic value of an imperative sentence ⌜!φ⌝ is a type that can be tokened by any directive act—including a command, a request, an invitation, etc.—whose content can be expressed by tokening the semantic value of ⌜φ⌝. Of course, not every situation lends itself to every kind of directive act, just as not every situation lends itself to referring to any male. Aside from the constraints imposed by the meanings of the expressions we utter, other constraints on how we can reasonably intend to affect our addressees are imposed by a potentially limitless (or, at least, impossible to spell out in advance) array of features of the situation in which we speak.

3.2.2 Intuitions about "Imperative Consequence"

In §3.1.5, I pointed out that one source of data that has animated the semantic literature on imperatives is the intuitive sense that there are some valid arguments whose premises or conclusions contain imperative clauses. Since most semanticists take their job description to include the prediction and explanation of how consequence plays out in the languages they study, data about

imperative consequence naturally seem to fall within the explanatory scope of a semantic treatment of imperatives. I argued that there are three ways of handling this intuitive data:

type-1 theories reject the idea that imperatives stand in a genuine consequence relations, and incur the burden of explaining away our intuitions that they do.

type-2 theories accept imperative consequence, hang onto the idea that consequence is truth preservation, and incur the burdens of (a) explaining how and why imperative sentences are truth-apt depsite seeming not to be, and (b) articulating and motibating a nonclassical logic of imperatives.

type-3 theories accept imperative consequence, take our intuitions that imperative are not truth-apt at face value, and incur the burdens of (a) motivating a move away from the idea that consequence is truth preservation, and (b) articulating and motibating a nonclassical logic of imperatives.

Most of the recent literature on imperatives has revolved around type-2 and type-3 theories. In §3.1.5, I began to advocate a type-1 approach on which our apparent intutions about sentences standing in consequence relations are actually intuitions about the properties of speech acts that we reflexively take the sentences to be evidence of. I am now in a better position to spell out the details of that account.

In performing a directive speech act, one intends to bring about certain intentions in an addressee. Possession of any given intention rationally commits an agent to having certain other intentions and rationally prohibits her from having still others. For example: intending to buy Oliver and Rachel drinks rationally commits me to possessing an intention of buying Oliver a drink, and also rationally prohibits me from intending to avoid buying Rachel a drink. Our sensitivity to these facts about the norms governing rational intention possession are the ultimate source of our

intuitions about apparent instances of imperative consequence. Consider the intuition that (14) is a valid argument, for example:

(14) a. Take out the trash and mow the lawn!

b. So, take out the trash!

The explanation is easiest to grok if we label the following four intentions:

(i1) The intention to take out the trash and mow the lawn.

(i2) The intention to produce (i1) in an addressee A.

(i3) The intention to take out the trash.

(i4) The intention to bring about (i3) in A.

In addressing a literal and direct speech act to A in uttering (14a), a speaker S would normally possess (i2), and so intends to bring about (i1) in A.[15] Possession of (i1) would rationally commit A to also possess (i3). I assume the following general principle:

Constraint Principle (CP)

If being in mental state *x* rationally commits one to being in mental state *y*, then intending to produce *x* in an agent rationally commits one to intending to produce *y* in the same agent.

It follows that intending to produce (i1) in A rationally commits S to intending to produce (i3) in A. But this is just to say that S's possession of (i2) rationally commits S to possessing (i4) as well. And (i4) is an intention that S would typically have in performing a literal and direct speech act with (14b). So, any speaker with intentions of the type that uttering (14a) typically signals is ra-

[15] This is somewhat misleading, since 'the trash' and 'the lawn' are incomlete descriptions that could be intended to pick out different trash and different lawns on different occasions of use. I will ignore this detail for simplicity. This example also glosses over problems about embedded imperatives that I will return to in §3.2.4 and §3.6.

tionally committed to having intentions of the type that (14b) typically signals. This sort of fact is what our intuitions about "imperative consequence" track.

In this respect, these intuitions are much the same as the ones that give rise to the idea that declarative sentences stand in consequence relations. The meaning of 'Paul plays bass and Ringo drums' is such that someone who performs a literal and direct assertion with it must intend an addressee to to believe that Paul plays bass and Ringo drums.[16] But, assuming that the principles governing what it is rational to believe validate conjunction elimination, anyone who believes that Paul plays bass and Ringo drums is also rationally committed to believing that Paul plays bass. It follows by (CP) that anyone who performs an assertion with a conjunction has intentions that rationally commit them to the intentions that would be involved in performing assertions with each of the conjuncts.[17] In this way, our intuitions that declarative sentences stand in consequence relations will mirror our sensitivity to the norms governing rational belief possession, which are, plausibly, imposed (at least in part) by the consequence relation governing beliefs' propositional contents.

Couldn't we define consequence in the following way? For a set of sentences $\Sigma = \{\varphi_1,...,\varphi_n\}$ and sentence ψ, $\Sigma \vDash \psi$ iff possessing the intentions required to performing literal and direct speech acts with all of $\varphi_1,...,\varphi_n$ rationally commits one to possessing the intentions required in performing a literal and direct speech act with ψ. There is a good reason not to draw this conclusion. Namely: sentences' meanings underspecify the intentions involved in performing a literal and direct speech act with them; they thus constrain our communicative intentions in ways that are

---

[16] Again, this description glosses over several potential sources of semantic underspecification in the sentence arising from both proper names (which Paul and which Ringo?) and generics (how much and in what way do they have to play bass and drum in order for the proposition to be true?).

[17] Does this mean that anyone who asserts *p* also asserts all of the logical consequences of *p*? No: being rationally committed to intending *q* is not the same thing as intending *q*. Moreover, we may want to delimit rational commitment to some subset of the logical consequences of our beliefs—say, just the ones that are *obvious* in some sense. All of this is consistent with my points here.

much weaker than might initially seem to be the case. For example: once we face up to the reality of semantic underspecification, the following three inferences are invalidated by the proposed criterion:

(14) a. Take out the trash and mow the lawn!

b. So, take out the trash!

(30) a. Elmars shark bit me and drew blood.

b. So, Elmar's shark bit me.

(31) a. One plus one equals two.

b. So, one plus one equals two.

By the proposed criterion, (14b) doesn't follow from (14a) because 'the trash' underspecifies the intentions that would determine which trash is being talked about, and so, one could produce literal and direct utterances of (14a) and (14b), tokening their respective semantic values, such that one's intentions in uttering the former do not rationally commit one to having the intentions one might have in uttering the latter. Likewise in the case of the possessive 'Elmar's shark' in (30), which could be used to pick out different relations between Elmar and a shark depending on the intentions with which it is uttered. In both cases, it is only if we imagine these (and other) intentions to be held fixed that the criterion proposed above predicts that (14) and (30) are valid arguments. But this is just to say that it is not the norms governing (14) and (30) themselves that our intuitions track, but the norms governing particular speech acts that we reflexively imagine (14) and (30) to be evidence of. Our intuitions seem to be about the sentences themselves only because we involuntarily fill in a back story that holds a speaker's referential intentions constant in an intuitive way across the sentences, thereby treating them as evidence of particular speech

acts.[18] Given the theory of illocutionary force proposed in §1.2 and the account of illocutionary underspecification proposed in §3.2.1, the proposed criterion of validity also invalidates (31). This is because (31a) could be used to perform a suggestion and (31b) could be used to perform an assertion. According the account I gave earlier—(GI2) in §1.2—suggesting that *p* involves intending merely to bring about an act of considering whether *p*, and having this intention does not rationally commit an agent to also possessing an intention to bring about belief that *p*. Even more broadly: there is nothing written into the meanings of any of the sentences above that guarantees that they be addressed to the same agent, and this is another form of constancy that we must assume in order for the kind of consequence I've just spelled out to make sense. I therefore conclude that our intuitions about sentences standing in consequence relations issue, quite generally, from our tendency to see sentences as evidence of speech acts, and to make a variety of assumptions about the nature of those speech acts that are unwarranted on the basis of the meanings of the sentences alone.[19]

Aside from explaining away our intuitions about imperative consequence, the type-1 account I have offered also has several other virtues. One is that it explains the apparent invalidity of (15), which is a motivation for Starr's dynamic account of consequence.

(15) a. Post the letter!

⊭ b. Post the letter or burn the letter!

---

[18] It is this artificial constancy of referential intentions that indexical logics attempt to mimic. But, as I argued in Chapter Two, those logics are highly misleading for this very reason.

[19] Does this position threaten the whole enterprise of formal logic? No! Logics are built around formal languages that lack mood, semantically underspecified expressions, and other features that prevent natural language sentences from being mapped directly onto propositions. The sentences of intepreted formal languages can therefore be treated as symbolic proxies for propositions in a way that natural language sentences cannot. Moreover, if we treat propositions as either sets of worlds or as structured entities, we can simply define consequence as a relation on the propositions themselves.

Someone who intends to post a letter is not rationally committed to intending to either post or burn the letter, and so intending to produce the former intention does not rationally commit one to intending to produce the latter.

So-called imperative inconsistency gets a similar explanation:

(18) a. Buy me a drink!

b. Don't buy me a drink!

An agent who intends to buy me a drink is rationally prohibited from intending not to buy me a drink, and so, in intending to produce the former intention rationally prohibits me from intending to produce the latter intention in the same agent.

One nice feature of this account is that it can be marshalled to explain some other data that have interested contemporary semanticists, but which they typically don't take to be a matter of logical consequence. For example, we can explain the infelicity of (4), which Starr uses to motivate his dynamic semantics.

(4) a. Unicorns don't, never have, and never will exist.

⋮

b. Bring me a unicorn!

The infelicity of uttering (4b) after (4a) results from the wider principle governing the interaction of beliefs and intentions that one can't rationally intend to do something that is ruled out by one's beliefs.[20] If a speaker has asserted that unicorns don't exist by uttering (4a), then directing the same addressee to bring him a unicorn by uttering (4b) would involve intending to bring about an intention that is incompatible with the belief intended earlier. But knowingly intending to bring about an irrational or impossible intention would itself be irrational or impossible. Moreo-

[20] For endorsements of various versions of this principle, see Bratman (1984), Donnellen (1968: 212), Grice (1968: 125), and Neale (2004: §2.6).

ver, as I pointed out in §3.1.2, this is an instance of a phenomenon so general that it would not make sense to explain it semantically. Rather, the infelicity of (4) should be explained on the basis of constraint on the rationality or possibility of intention possession, given one's other beliefs and intentions. This is the same kind of explanation that I have given for our intuitions about apparent instances of imperative consequence.

In appealing to the notions of rational commitment and rational prohibition, I needn't settle on a particularly precise theory of those notions. Based on what I've said so far, my commitments paint at best a partial picture of what those notions entail, including at least the following principles:

- Agents are rationally committed to believe propositions that are (obvious) logical consequences of the contents of their current beliefs, and rationally prohibited from believing propositions that are (obviously) logically contradicted by the contents of their current beliefs.
- Agents are rationally prohibitted from simultaneously intending two (obviously) contradictory propositions.
- Agents who intend p and q are rationally committed to intending p and to intending q.
- Agents whose beliefs rule out p are rationally prohibitted from intending p.
- Agents are rationally prohibited from intending to do something that they believe has already been done.

Each of these principles is plausible, independently motivated, and outstrips the explanatory purview of semantic and pragmatic theory. There are various ways in which we might want to elaborate and systematize them. For example, we might wish to generalize the first principle by closing the contents of rational beliefs not only under logical entailment but also under a probabilistic extension thereof. Moreover, it may be possible to develop a much more general and systematic account of the relations that must obtain between intentions' contents in order for agents to be

rationally committed to or prohibited from jointly possessing intentions.[21] We might also wish to add further principles concerning desire, and perhaps also other intentional states. It is at least possible that we could cook up a formal system to model rational commitment and prohibition quite generally. Would this formal system be a logic? Not in the narrow sense in which I use the term—as a label for the theory of the consequence relation, which is narrowly a matter of truth preservation. Of course, there are well-attested, broader uses of 'logic', according to which a logic is any old formal system used to track relations between representations. And, to some extent, which of these uses of 'logic' we adopt is a verbal matter. All that matters for my purposes—and what makes my theory of imperative consequence a type-1 account—is that we recognize that a "logic" of rational commitment, if we're even able to give one, will be a formal system of this broader kind, and won't be a theory of truth preservation.

3.2.3 Making as if to Illocute

We sometimes utter imperative sentences without intending to produce an intention in anyone, just as we sometimes utter declarative sentences without intending to produce a belief in anyone. Utterances of these kinds may seem to be counterexamples to semantic theories, like the one I've constructed here, that understand the meaning of mood in terms of illocution, which is in turn understood in terms of communicative intentions. Several authors have claimed that these putative counterexamples are genuine counterexamples.[22] Here are some of the cases used by Wilson and Sperber (1988) in their argument:

---

[21] Much work in philosophy and artificial intelligence has gone into systematizing and formalizing the relations that must obtain between a rational agent's beliefs and intentions—e.g. Bratman (1984, 1987), Broome (2004), Cohen and Levesque (1990a,b), Grice (1971), Setiya (2010).

[22] E.g., Davidson (1979), Starr (*forthcoming a*), Wilson and Sperber (1988).

(32) a. Peter: I'm not going, but, hypthetically, how would one get to the station?

b. Mary: Take a number 3 bus.

—Mary doesn't intend to get Peter to intend to take the bus.[23]

(33) Go on, throw it. Just you dare.

—The speaker doesn't really intend for Peter to throw the snowball.

(34) Get well soon.

—Peter has no control over if/when he will get well, and so can't intend to.

(35) Start, damn you!

—Mary utters this alone in her car to her car, which can't have intentions.

(36) Please don't be late again.

—Peter can't hear Mary say this, and the result is predetermined.

The semantic value of each of (32b)–(36) is a type of speech act that can be tokened only by directing someone to bring about *p*, where this requires communicatively intending to produce an intention that *p* in that someone. Moreover, I have suggested that the role of a sentence's semantic value is to constrain the speech acts that speakers can perform with the sentence. But, in the relevant circumstances, each of (32b)–(36) is uttered without intentions of the relevant kind. What gives?

---

[23] This is slightly different from Wilson and Sperber's example, in that it makes the hypothetical nature of Peter's question. Wilson and Sperber argue that all or most acts of advising and permitting with imperatives can't be understood as "attempt[s] to get the hearer to perform the action described" because "there is no reason to think that [the speaker] cares whether [the hearer] follows her advice" (1988: 212–3). In general, I think argument is flawed—on par with saying that I don't intend you to believe that I believe *p* in asserting *p* because I ultimately don't care one way or the other. The mere fact that I have produced the utterance is good evidence that I care, at least a little bit. Moreover, there is nothing incoherent about communicatively intending to produce an effect about which I have no preferences (perhaps beyond the most basic desire to be cooperative).

The short answer is that the content and force of the illocutionary act a speaker peforms in producing an utterance—and the fact that she performs a speech act at all—are wholly constituted by the speaker's intentions. Although a sentence's meaning constrains those intentions by constraining how an addressee will normally interpret the speaker, the relevant kind of constraint is not binding.[24] This is because there are other ways in which an addressee can work out a speaker's intentions. A speaker can be sensitive to these other ways and exploit them in order to communicate non-literally. All of this should sound familiar to anyone who knows about Grice's (1975) theory of implicature or Searle's (1975) theory of indirect speech acts.[25]

Notice that each of (32b)–(36) *could*, in other circumstances, be used to perform directive speech acts that token the types that my theory identifies with their semantic values. Mary could utter (32b) in order to command or advise Paul to take the bus, for example, and she could utter (35) in order to rudely command her masseuse to begin massaging her feet. I reserve the notion of a *literal speech act* for cases of this kind, wherein the speech act performed by uttering an expression tokens the expression's semantic value. The speech acts described by (32b)–(36) do not count as literal by this definition. Instead, each of (32b)–(36) describes an act of *pretending to perform*—or, to borrow Grice's turn of phrase, *making as if to perform*—a directive illocutionary act in order to perform another, indirect act.

If I am right about this, then the question we should ask about cases like (32b)–(36) is this: why would the speaker pretend to perform a directive speech act? What are they thereby trying (intending) to accomplish, and why would pretending to perform a directive act be a way to ac-

---

[24] This raises to salience the question of what I mean by 'constrain'. I will replace this vague notion with something more precise in Chapter Five.

[25] Searle focuses on cases of performing one illocutionary act indirectly by performing another one directly, and neglects cases in which we indirectly perform an act by *pretending* to perform another one. Grice rectifies this with his discussion of *making as if to say*, but focuses his discussion of implicature and the maxims of conversation narrowly on what amount to direct and indirect *assertive* acts, neglecting indirect directive and interrogative acts.

complish that? The answers to these questions vary from case to case. In some cases, such as when acting in a play, pretending to perform a speech act may be an end in itself, or, insofar as it does serve some further goal (such as, for example, entertaining an audience), intending to accomplish this goal does not contstitute the performance a further illocutionary act. I am tempted to assimilate some audienceless utterances, including perhaps (35) and (36), to this paradigm. In uttering (35) to her car, Mary makes as if to command her car to start, but unless her intentional stance is running wild (which, admittedly, is a possibility in the heat of the moment), she does not thereby intend to produce any sort of cognitive response in the car. And, unlike self-directed utterances like writing in one's diary, where it is natural and unproblematic to construe the addressee of an illocutionary act as being one's future self, this is not plausible in many cases of acting or talking to one's car or under one's breath. Why, if not to perform an illocutionary act, does Mary produce (35) and (36)? A plausible answer, I think, is that she pretends to issue commands to her car and to Peter, respectively, in order to vent her feelings of frustration or powerlessness at them. Why would pretending to command something or someone be a way to vent one's frustration at that thing or person? A plausible, if rough-hewn answer, is that to perform a command is to presuppose that one has a certain amount of power or authority over whatever one commands, and that pretending to have this power or authority feels good when one is powerless and frustrated. Obviously, this explanation is folksy and incomplete, and to render it more satisfying would require saying more about the psychology of frustration. I am optimistic that the explanation could be fleshed out in this way.

In other cases, we pretend to directly perform an illocutionary act in order to indirectly perform an illocutionary act of another kind. In Grice's reference letter case, for example, a letter-writer makes as if to say (i.e., pretends to directly assert) that a student is punctual and has good handwriting in order to implicate (i.e., indirectly assert) that the student is not a very good philosopher (1961; 1975: 33). Examples of implicature typically focus on assertive acts—performing

or making as if to perform one such act in order to indirectly perform another—but this reflects only the limited interests of theorists rather than the limited scope of the phenomenon. We frequently make as if to ask questions in order to request things other than information—for example, by uttering, 'Could you pass the salt?' Likewise, we sometimes pretend perform assertive acts in order to perform directive acts—as, for example, when we say, 'you left the light on' in order to request that it be turned off.[26] As (32b) illustrates, we sometimes make as if to perform a directive act in order to perform an assertive one: Mary doesn't intend to produce an intention in Peter to take the number 3 bus, but she does intend to produce a belief that taking the number 3 bus is a way that he could get to to the station. It's tempting to interpret this as a run-of-the-mill quality implicature: Peter asked a question—he requested (true) information. A cooperative response from Mary would require that she provide that information, most likely by performing an assertive act. Since directive acts aren't aimed at providing information, and aren't truth apt, to peform one would be to do something uncooperative. So, in seeming to perform one, she flouts the maxim of quality,[27] and the best explanation of why she did so is that she was intending to inform Peter that *p* was an answer to his question, rather than directing him to bring about *p*.

---

[26] Searle argues that indirect directives are particularly common "because ordinary conversational requirements of politeness normally make it awkward to issue flat imperative sentences (e.g., 'Leave the room') or explicit performatives (e.g., 'I order you to leave the room'), and we therefore seek to find indirect means to our illocutionary ends (e.g. 'I wonder if you would mind leaving the room'). In directives, politeness is the chief motivation for indirectness" (Searle 1975a: 36). Searle is undeniably right that the rudeness of a directive act can often be defrayed by performing it indirectly. It is far less clear why this might be.

[27] Grice's formulation of the maxims of conversation is designed only for occasions when providing information—i.e., performing an assertive speech act—is the way to make a cooperative contribution to a conversation. This is most obvious in the case of the maxim of quality, which requires making one's contribution a true one that is supported by evidence (1975: 27). Since some cooperative contributions to a conversation, including questions and directives, are not merely untrue but not apt to be true, the maxim as stated covers only certain kinds of contribution, perhaps because it is a special case of a broader maxim of quality that governs questions and directives as well. Some revision may also be required for the maxim of quantity, since it is stated in terms of providing the right amount of information. I don't attempt to articulate these revisions or generalizations here.

In the same vein, Mary should be understood as performing indirect speech acts with (33) and (34). In uttering (33), she makes as if to dare Peter to throw a snowball in order to indirectly command him not to. This is akin to an ironic assertion whereby a speaker makes as if to assert a proposition in order to indirectly assert the proposition's negation.[28] There are various things that Mary could be trying to do in uttering (34), but one plausible candidate is that she is indirectly reporting her desire that Peter get well soon by making as if to request that he do so. Again, it makes sense that pretending to request would be a good way to report desire, because one of the intentions that constitutes a request that *p* is the intention to produce an intention that *p, partly on the basis of the speaker's desire that p*. A request, in other words, is always in part the signaling of a desire; in the nonliteral and indirect case of (34), that desire is *all* Mary communicates.

3.2.4 The Problem of Embedded Imperatives

The most serious potential roadblock for the Gricean, act-theoretic picture I've been advocating is raised by the fact that clauses—mood-markers and all—seem to embed within the scope of logical connectives. This is illustrated by (6)–(11), each of which contains at least one imperative clause embedded inside the scope of a connective.

(6) Buy me a drink and make it a stiff one!	!φ and !ψ
(7) Buy me a drink or leave me alone!	!φ or !ψ
(8) If the bartender comes back, buy me a drink.	(if ⊦φ)(!ψ)
(9) a. Give me all your money or the puppy gets it.	!φ or ⊦ψ
b. You mow the lawn or I'll clean the garage	(Starr *ms*: §2.1)

[28] Wilson and Sperber seem to suggest that dares are generally quasi-ironic in this way. But as most of my childhood friends could have told you, many dares are meant to be taken literally.

(10) a. Say that again and I'll scream. !φ and ⊢ψ

b. Stay here and I'll go to the store.

(11) If he comes back, buy me a drink and I'll pay you back. (if ⊢φ)(!ψ and ⊢σ)

Act-theoretic semantics aims to explain the meanings of declaratives and imperatives by saying that they can be used to perform assertive and directive speech acts, respectively. Sentences like (6)–(11) raise a serious problem for this view. One way of stating this problem is to note that each of (6)–(11) contains more than one clause, but isn't used to perform more than one illocutionary act. Moreover, (7) contains two imperative clauses, but, in uttering the sentence, one wouldn't be performing a directive act that would token either clause's semantic value. (The analogous point can be made about disjoined declaratives and assertion.) So the problem is this: if a clause's meaning is a matter of the speech act it is used to perform, how can we account for its semantic contribution to sentences in which it is embedded, given that they aren't used to perform the relevant kind of speech act?

We could avoid this problem if it were possible to treat all sentences, including (6)–(11), as consisting of a single mood-marker scoping over a (perhaps complex) sentence radical. In that case, most of compositional semantics would consist of a recursive specification of the types of acts of expressing propositions that we can perform with sentence radicals, and the semantics of mood could be handled with a single composition rule for each mood-marker, along the lines presented in §3.2.1:

Imperative Composition $[_S\ !\varphi\ ]$

S's semantic value is a type that is tokenable by a directive act whose content could be expressed by tokening the semantic value of φ.

But, for reasons I discussed in §3.1.3, the prospects for this strategy aren't good. Sentences like (9b) and (10b) seemingly can't be reinterpreted as complex sentence radicals flanked by a single mood-marker. Moreover, this problem ramifies, because sentences like (9b) and (10b) can themselves be embedded inside the scope of logical operators and in the consequents of conditionals. This means that the kind of composition represented by (6)–(11) is genuinely semantic and productive, and it follows that our problem can't be solved simply by specifying a handful of further speech act-types to serve as the semantic values of genuine, mixed-mood sentences. Rather, we need a way of recursively specifying the types of speech acts performable with complex, mixed-mood sentences on the basis of specifications of the types of speech acts performable with their parts. This solution will also have to explain how the semantic value of a sentence can be a type of speech act even though we don't perform a speech act of that type when uttering the sentence in a variety of embedded positions.

This is an instance of what is often called the *Frege–Geach problem*, which, in its various guises, has often been seen as devastating to views that cash out the semantic contributions of some expressions in terms of their speech act potentials.[29] Below, I will argue that the problem has a solution that draws on distinctively Gricean resources. Because this solution also takes some inspiration from dynamic approaches to imperatives, which I discuss in §§3.4–3.5, I will leave my discussion of it until §3.6.

3.3 Content-Theoretic Accounts of Imperatives

The theory of mood that I began to outline in the last section is one way of articulating and expanding upon idea that the semantic function of a sentence's mood is to encode the force of the speech acts that can be literally and directly performed with the sentence. Other ways of working out the same idea can be found in the work of Hare (1970) and Searle (1968, 1969). I think that

[29] See, e.g., Dummett (1973: *ch.*10), Geach (1965), Schroeder (2008), Starr (forthcoming a).

my way of filling out this idea is preferable to the alternatives for two reasons. First, it is built on the foundation of a Gricean speech act theory that is preferable to (e.g.) Searle's (1969) version of speech act theory for the reasons spelled out in §1.2. Second, the speech-act theoretic semantic framework in which my theory is framed is independently motivated (see Chapter One) and takes seriously the demands that a semantic theory be a finitely statable account of how sentence meanings are built up from word meanings (see Chapters Two and Six). These points notwithstanding, I agree with the basic idea behind Hare's and Searle's view: that the meaning of mood should force us away from the idea of sentence meanings as truth conditions or propositions, and into the realm of speech act theory.

In this section, I want to consider three ways of resisting this basic idea by identifying the semantic values of imperatives with semantic contents—truth conditions or propositions of some kind. Views of this kind are motivated by the astounding successes enjoyed by truth-conditional semantics over the last half century. As I suggested in Chapter Two, one could have been forgiven during much of this period for taking compositional semantics to be tantamount to either truth-theoretic semantics in the style of Davidson, model-theoretic semantics in the style of Montague, or structured proposition-theoretic semantics of the kind defended by Soames (1986). The dominance of these ideas, which were developed primarily with declarative sentences in mind, has led semanticists to attempt to account for non-declaratives within them as well. For Davidson (1979), this meant developing a version of the performative-reductionist analysis discussed in §3.3.1. Although Montague ignored nondeclaratives,[30] Lewis (1970) quickly developed the best-known ver-

[30] Given Montague's brief career as a semanticist, this statement may not be entirely fair. But there is some textual evidence that Montague's lack of engagement with nondeclaratives was not due to mere lack of time. Montague's only allusion to mood comes in the context of his claim that the explanatory goals of syntax are entirely subordinate to those of semantics. Montague makes this claim after first saying that syntax should be interested primarily in characterizing "the set of declarative sentences" (1970b: 223, *n*.2). Although it might be claimed that this passage is compatible both declarative-reductionism and modal-reductionism, Montague's claim here at least suggests that he holds the even stronger view that non-declaratives fall outside the scope of semantic theory.

sion of performative-reductionism as an approach to non-declaratives within a broadly Montagovian framework, and went on to develop a version of modal-reductionism (Lewis 1975b), which is the kind of theory I'll discuss in §3.3.2. In §3.3.3, I'll discuss Hanks' recent attempt to extend a structured proposition-theoretic semantic framework to non-declaratives by rethinking the nature of propositions themselves. What all these account share is the idea that the core of a semantic theory of imperatives should take the form of a theory of the special features of the semantic contents of imperatives.

3.3.1 Performative-Reductionism

Performative-reductionism, which was the earliest-to-develop content-theoretic approach to nondeclaratives, is so-called because its central idea is that either the logical forms or semantic values of sentences that look to be non-declarative on the surface are actually akin to those of declarative, explicit performative sentences. For example, this sort of approach might try to construe (37a) as having a logical form akin to that of (37b):

(37) a. Buy me a drink.

b. I (hereby) direct that you buy me a drink.

Views of this general kind have been defended by a variety of authors,[31] including David Lewis:

> I propose that these non-declaratives ought to be treated as paraphrases of the corresponding performatives, having the same base structure, meaning, intension, and truth-value at an index or on an occasion. And I propose that there is no difference in kind between the meanings of these performatives and non-declaratives and the meanings of the ordinary declarative sentences considered previously. (Lewis 1970: 222).

---

[31] See also Ross (1970) and Cresswell (1973: ch.14). Davidson (1979) considers, but ultimately rejects, a closely-related proposal on which the logical form of a directive !φ is akin to that of the pair of sentences, 'I (hereby) direct that. φ'. (This is an instance of Davidson's broader paratactic theory of attitude reports and indirect speech—see Davidson (1968).) Most of the arguments against Lewis's version of performative-reductionism also apply to Davidson's version, and Davidson's version also suffers from the more general flaws of his paratactic theory of indirect speech (see, e.g., Burge 1986; Schiffer 1987; Starr *forthcoming a*: §2.3.1–2.3.2.

Aside from reducing away those pesky non-declaratives, the big attraction of performative-reductionism is that explicit performatives share many of the features of non-declaratives that demand explanation. For example, the explicit performative (37b) can seemingly be used to perform the same range of speech acts as the imperative sentence (37a).[32] Most notably: explicit performatives themselves seem not to be truth apt, thus mirroring one of biggest apparent roadblocks to a truth-conditional analysis of nondeclaratives. In order to overcome this roadblock, Lewis claims that performatives, and so non-declaratives, really are truth apt, but that they seem to lack truth conditions for pragmatic reasons.

> I would wish to say that 'I bet you sixpence it will rain tomorrow' is true on an occasion of utterance iff the utterer does then bet his audience sixpence that it will rain on the following day; and, if the occasion is normal in certain respects, the utterer does so bet; therefore his utterance is true. Austin says it is obviously neither true nor false, apparently because to utter the sentence (in normal circumstances) is to bet. Granted; but why is that a reason to deny that the utterance is true? To utter 'I am speaking' is to speak, but it is also to speak the truth. This much can be said in Austin's defense: the truth-values (and truth conditions, that is intensions) of performatives and their [non-declarative] paraphrases are easily ignored just because it is hard for a performative to be anything but true on an occasion of utterance. Hard but possible: you can be play-acting, practicing elocution, or impersonating an officer and say 'I command that you be late' falsely, that is, say it without thereby commanding your audience to be late. I claim that those are the very circumstances in which you could falsely say 'Be late!'; otherwise it, like the performative, is truly uttered when and because it is uttered. It is no wonder if the truth conditions of sentences embedded in performatives and their non-declarative paraphrases tend to eclipse the truth conditions of the performatives and non-declaratives themselves. (Lewis 1970: 224)

However satisfying this explanation is when it comes to explicit performatives—and I find it relatively satisfying—the same line of thought is less convincing when applied directly to non-

---

[32] A frequent complaint against performative reductionism and its close relatives alleges that the approach cannot account for illocutionary underspecification, since any way of filling in the blank in ⌜I (hereby) ____ that you φ⌝ will be too specific to cover the range of acts performable with any given mood (e.g., Davidson 1979; Starr *forthcoming a*). But once we recognize illocutionary underspecification as a special case of semantic underspecification more generally (see §3.1.4), we can say that whatever expression goes in the blank merely places loose constraints on which kinds of illocutionary act can be named with it. Thus my use of 'direct' rather than 'command' or 'request' in (37b).

declaratives. When someone lacking authority utters, 'I order you to leave the park,' it makes sense to respond by saying, 'False: you don't have the authority to do that', just as this response makes sense to a non-athoritative utterance of 'I hereby place you under arrest'. But the same response seems considerably more infelicitous following the imperative sentence, 'Leave the park!' So although Lewis does a nice job of explaining away the apparent non-truth-aptness of explicit performatives, his explanation seemingly does not transfer over to non-declaratives themselves. This is particularly problematic for Lewis and other performative-reductionists, who take performatives to have semantic values identical to their non-declarative counterparts.

Moreover, although non-declaratives and their corresponding explicit performatives (such as (37a) and (37b)) can be used to perform a similar range of illocutionary acts, their other semantically relevant properties don't match. As Charlow (2013: §3.1) and Parsons (2012) point out, non-declaratives and their performative counterparts have intuitively very different logical profiles. And as Starr (*forthcoming a*: §2.3.1) points out, they make different semantic contributions to sentences in which they're embedded.[33] All of this suggests that non-declaratives cannot be reduced to performatives.

---

[33] A syntactically-loaded version of performative reductionism holds that nondeclaratives are grammatically very similar to performatives at some semantically relevant level of representation (Sadock 1974). For arguments that this idea gets the syntax of imperatives all wrong, see Han (1998) and Starr (*forthcoming a*: §2.3.1). But, since both Han and Starr also defend the idea that mood can be factored out at the semantically relevant level of syntactic representation (see §3.1.1), it seems to me that their criticisms don't rule out a version of performative reductionism on which mood markers have the semantic functions, if not the same syntactic structures, of explicit performative phrases. This syntactically un-loaded version seems to be all that Lewis had in mind in saying that non-declaratives "paraphrase" performatives.

3.3.2 Modal-Reductionism

A second proposal for reducing the meanings of imperatives to those of declaratives is *modal-reductionism*, according to which imperative clauses have semantic values that are either identical or otherwise closely related to those of corresponding deontic modals, as in (38).[35]

(38) a. Buy me a drink.

b. You [ought/must/should] buy me a drink.

It has frequently been noted that pairs of sentences like (38) can be used to perform a range of similar speech acts. It is typical to give a pragmatic explanation of this fact, according to which uttering a deontic modal of the form $\Box\varphi$ is a somewhat standardized way of indirectly performing a directive illocutionary act that one might have performed directly with an imperative of the form $!\varphi$ (Searle 1975: 37–8). Perhaps the most basic idea behind modal-reductionism is that this overcomplicates things, and that the right way to capture the close relationship between the use conditions of (38a) and (38b) is to say that they have the same semantic value.

The full-blown defenses of modal reductionism that have emerged over the past fifteen years have leaned heavily on the sorts of considerations that I discussed with less than full enthusiasm in §3.1.6 (e.g. Han 1999; Kaufmann 2012). For example, several authors have noted argued that imperatives and deontic modals have similar logical profiles (e.g., Aloni 2007; Charlow 2013: §2.3; Portner 2012). My response in §3.1.6 was that a wide variety of operators used to express action-guiding states have similar logical profiles, and so these intuitions about logical profile do no more to establish an imperative-modal connection than they do to establish an imperative-intention connection of the kind that I advocate.

Another point sometimes marshalled in favor of modal-reductionism is that many kinds of directive illocutionary act correspond to different flavors of modal, where the flavor of a modal is determined by the nature of its ordering source (e.g., Han 1999; Kaufmann 2012: ch.3; Portner

[34] Some proponents of modal reductionism include Aloni (2007), Han (1999), Kaufmann (2012), and Lewis (1975b).

2007).[35] To take just two of a variety of possible examples, the grounds on which one *commands* one's addressee corresponds naturally to a deontic ordering source—which can be thought of as a function from worlds to sets of propositions about what one ought to do from the perspective of some authority or set of rules, and the grounds on which one *advises* one's addressee correspond naturally to a teleological ordering source, which we can think of as a function from worlds to sets of propositions about an agent's goals or interests.[36] Given this pattern of correspondences, it becomes tempting to assimilate illocutionary underspecification to the semantic underspecification exhibited by modals by adopting a form of modal reductionism on which different types of directive speech act are simply modals interpreted relative to different kinds of ordering sources (see, e.g., Kaufmann 2012: chs.3–6). But again, this argument fails to achieve any interesting generalizations that aren't already present in the speech act-theoretic semantics for imperatives that I outlined in §3.2. This is because every body of information that can serve as a modal base (at a world) can also serve as a reason on the basis of which a speaker can intend her addressee to adopt a certain intention—i.e., as a value for ρ in the illocutionary intention schema (GI) that I outlined in §1.2.

(GI) Gricean Illocution

S performs an illocutionary act α in uttering *u* iff *S* utters *u* intending:

(1) to produce thereby a certain response Δ in a certain addressee A;

---

[35] On Kratzer's (1981) treatment (see also Portner 2009: §3.1), a modal □φ is evaluated relative to a world w, modal base *f*, and ordering source *g*, where $f(w)$ and $g(w)$ are sets of propositions, and such that □φ is true relative to *w*, *f*, and *g* iff φ is true in the worlds in $\cap f(w)$ that are ranked best according to $g(w)$, where a world w_1 is ranked at least as highly as w_2 by $g(w)$ iff at least as many of the propositions in $g(w)$ are true at w_1 as are true at w_2. In short: an ordering source at a world determines which propositions are best at that world; in the context of a modal-reductionist analysis of imperatives, this amounts to specifying which of the addressee's potentioal actions it would be best for her to perform, relative to a particular standard.

[36] There is much disagreement over how to carve up modal bases terminologically, but the details don't matter much here. My usage corresponds roughly to Portner's (2007).

(2) A to recognize S's intention (1);

(3) A's response Δ to be at least partly on the basis of her recognition of (1) and (in some cases) partly on the basis of some further reason ρ.

Whereas the modal reductionist identifies the source of directive underspecification as the underspecification of an ordering source (at a world) $g(w)$, I identify it as the underspecification of a speaker's intended reasons ρ for her addressee to form an intention Δ. For any given performance of a direct and literal directive act with an imperative sentence, the modal reductionist and I are (respectively) free to think that g and ρ will be fleshed out such that $g(w) = \rho$. In other words: our theories make substantially similar predictions because we hold that the things we do with imperatives are underspecified by their meanings relative to substantially similar bodies of information.[37]

Kaufmann (2012: §2.3.3) argues for modal reductionism on the grounds that imperatives can be used to answer questions that deontic modals would provide direct answers to, as in (39):

(39) a. How should I get to the station?

b. You should take the number 9 bus.

c. Take the number 9 bus.

On its own, this is a very weak argument for the view that (39b) and (39c) have the same semantic value. After all: in nearly any circumstance in which it would be felicitous to ask a question by uttering any of (40a)–(40h) (among many others), it would be felicitous to answer by uttering any of (41a)–(41h) (among many others).

(40) a. How should I get to the station?

b. How do I get to the station?

[37] One big difference between the two views is that the modal reductionist typically offers no story about what determines which ordering source is relevant on a particular occasion, beyond vague hand waving at context. For criticisms of this sort of hand waving, see Chapter Two.

c. How will I get to the station?

d. How does one get to the station?

e. I'm not going, but, hypothetically, how does one get to the station?

f. Which way is it to the station?

g. Which way to the station?

h. Which bus goes to the station?

(41) a. You should take the number 9 bus.

b. The number 9 bus.

c. The number 9 bus goes [there/to the station]..

d. Take the number 9 bus.

e. You could take the number 9 bus.

f. I recommend taking the number 9 bus.

g. I usually take the number 9 bus.

h. Most people take the number 9 bus.

So, the fact that (39b) and (39c) are both good ways to answer (39a) gives us no more reason to identify their semantic values than we have for identifying the semantic values of all of (41a)–(41h). In general, since asking a question places very specific constraints on what would count as a cooperative response from one's addressee, they create situations that are pregnant with the potential for indirect speech acts. For this reason, it is reasonable to expect that of most of the above Q&A data should be explained by appealing to indirect speech acts of one kind or another (see §3.1.5; Grice 1975; Searle 1975a). More generally, this fact about questions and their answers casts significant doubts on most arguments for semantic theses—such as Kaufmann's—that rely on data about which utterances constitute good answers to questions asked with particular sentences.

In addition to being supported by only shaky evidence (some of which I haven't reviewed here—see Kaufmann (2012) for a comprehensive defense), modal-reductionism faces significant challenges. One obvious challenge is that modals like (38b) are obviously truth apt, whereas imperatives seemingly aren't. Both van Rooy (2000) and Kaufmann (2012) attempt to navigate around this roadblock by drawing an analogy between imperatives and performative uses of modals. Intuitive examples of performatively used modals include the following sentences, uttered by an authority figure:

(42) a. You should buy me a drink.

b. You must show up on time for work.

As I've already pointed out, an obvious explanation of performative uses of modals is that they are just those utterances whereby one makes as if to assert a modal proposition in order to direct someone to bring about the modal's prejacent. As Searle points out, this explanation locates performative uses of modals within a broad pattern of other ways of indirectly performing directive speech acts—a pattern that, due to its broad linguistic diversity, is almost certainly best explained pragmatically (1975: 36–43). Nevertheless, it has been tempting for semanticists to become excited by the following symmetry: a modal is used performatively if it is used to alter what is necessary or possible rather than to assert that something is necessary or possible. This can be put in satisfyingly Kratzerian terms by saying that to use a necessity modal performatively is to use it to alter the modal base and/or ordering source so that it conforms to the prejacent, rather than to assert that the prejacent follows from the worlds in the modal base that are best with respect to the ordering source. Van Rooy (2000) goes so far as to argue that this distinction between descriptive and performative uses of modals is wholly semantic, and posits a lexical ambiguity to encapsulate this idea. If so, then perhaps imperatives are synonymous with performative modals only. Kaufmann (2012: §2.2.2) offers a host of convincing arguments against this unpromising

line of thought, but goes on to argue that performative uses of modals can be isolated on the basis of a variety of pragmatic facts about the contexts in which they occur. A use of a modal $\Box\varphi$ counts as performative, she argues, just in case it occurs in a context—here consisting of a Stalnakerian context set c, a modal base f, and an ordering source g—that meets the following conditions:[38]

AuthorityCondition
The speaker is an epistemic authority on f and g: for each w in the context set and any world w compatible with the speaker's beliefs in w: $f(w) = f(w')$ and $g(w) = g(w')$. (Kaufmann 2012: §4.2.2)

Epistemic Uncertainty
Both φ and $\neg\varphi$ are compatible with the speaker's beliefs: for each w in the context set, there are w' and w'' compatible with the speaker's beliefs in w: φ is true in w' and φ is false in w''. (Kaufmann 2012: §4.2.3)

Ordering Source Restriction
If there is a contextually salient decision problem to which $\Box\varphi$ is an answer, then g is prioritizing (deontic, bouletic or teleological) and considered by speaker and addressee to be the criterion for resolving the decision problem. Otherwise, g is speaker-bouletic. (Kaufmann 2012: §4.2.3)

I am doubtful that conditions on context like these represent the best way to distinguish performative uses of modals from descriptive uses, but I won't dwell on all of the sources of those doubts here. Kaufmann's idea about imperatives is that since the above conditions set apart performative uses, and since the disguised modals that we call imperatives can only be used performatively, we should take the imperative mood-marker to encode a semantically triggered presupposition to the effect that all of the above conditions are met, thereby assuring that imperatives (whose presuppositions are met) always function as performative modals in actual conversation. This is Kaufmann's account of the meanings of imperatives.

---

[38] These helpfully clear glosses of Kaufmann's conditions are borrowed from Starr (*ms*: §4.2), mainly because Kaufmann's own formulations are framed in technical vocabulary and notation that would take a while to set out.

Needless to say, this is all very complicated, and there is reason to worry that the degree of complexity on display in Kaufmann's account far outweighs its explanatory advantages. Starr articulates this point nicely:

> How simple and plausible is this addition to the semantics of imperatives? This presuppositional content is of unparalleled richness that exceeds imperatives' actual content. While the first two presuppositions are reasonable things to expect from an exclusively performative modal, the third is a surprising and disjunctive presupposition. It is at least controversial that the addition of these presuppositions meets the standard of simplicity. (Starr *ms*: §4.2)

Starr goes on to articulate a series of other objections to Kaufmann's account, mostly revolving around the fact that it does nothing that cannot be accomplished more perspicuously by other means. If it isn't already obvious, I agree.[39]

3.3.3 Hanks contra the Content–Force Distinction

The semantic accounts of imperatives discussed in the last two subsections revolve around different versions of the idea that we should assign familiar propositional semantic contents to imperatives, and then add on a pragmatic explanation for why imperatives can be used to perform directive acts and seem not to be truth apt. Hanks (2007, 2011) defends the bolder proposal that we should reconceive of the nature of propositional content itself by building illocutionary force into propositions and collapsing the force–content distinction. On his theory, declarative, imperative, and interrogative sentences express propositions of fundamentally different kinds, each of which is a type that can be tokened by the performance of different kinds of mental acts or speech acts. Whereas declarative sentences express *assertive propositions*, which may be tokened by speech acts of asserting or mental acts of judging, interrogative sentences express *interrogative propositions*, which may be tokened by speech acts of asking and mental acts of wondering, and impera-

[39] Other arguments against modal reductionism aboud. Hanks points out, for example, that we report utterances of modals and imperatives in different ways, so that the following sentence is not contradictory: 'Jones told smith that she ought to go to the store but did not tell Smith to go to the store' (2007: 151).

tive sentences express *imperative propositions*, which may be tokened by directive speech acts and mental acts or states of desiring, wishing, or intending (Hanks 2007: 152). Hanks argues that these distinct act-types are distinguished by the different kinds of combinatory acts that unify their parts into a structured whole—*predicating* (or *applying*), *asking*, and *wanting* (or *attempting*), respectively. These combinatory acts that unify propositions, rather than any components of propositions, are the semantic contributions of the suite of morphosyntactic features that comprise the declarative, interrogative, and imperative moods, respectively. Hanks (2011: 13,16) represents the propositional contents of clauses of different types as follows:

(43) a. George is clever.

b. ⊦⟨George, CLEVER⟩

(44) a. Is George clever?

b. ?⟨George, CLEVER⟩

(45) a. George, be clever!

b. !⟨George, CLEVER⟩

In these representations, '⊦', '?', and '!' aren't grammatical mood-markers in the way I used them above. Rather, they represent the distinct varieties of combinatory act that unifiy the propositions (43b), (44b), and (45b), according to Hanks.

Hanks' account resembles my own in several respects. First, Hanks thinks that at least many expressions' semantic values are types that may be tokened by speech acts. (His account differs from mine in that these types may also be tokened by mental acts and states—a point I'll return to.) I also agree with Hanks' suggestion that the semantic values of clauses of different types are types that may be tokened by illocutionary acts of different kinds, and that the semantic contribution of a sentence's mood is intimately related to the illocutionary force of the speech acts that can

be performed with it. As will become clear in Chapter Six, my speech act-theoretic semantic framework is similar to Hanks' act-theoretic framework in certain other ways as well.

But Hanks' views are also dissimilar from mine in several ways, all of which count, I think, in favor of my theory. I argued in Chapter Two, for example, that expressions' semantic values should not be identified with the propositional contents that they are used to express on particular occassions. As we'll see, one of Hanks' arguments for his theory takes this dubious assumption about semantic content as a key premise. For closely related reasons, I also think Hanks makes a mistake in holding that the types that serve as sentences semantic values can be tokened by thoughts as well as speech acts. If a broadly Gricean theory of speech acts is correct, then illocutionary acts are constituted by speakers' audience-directed communicative intentions, and should be typed in terms of the kinds of intentions required to perform them. But this surely isn't true of intentional psychological states such as ordinary beliefs and desires. Of course, Hanks presumably doesn't accept a Gricean theory of speech acts, but he does not offer an alternative theory of the metaphysics illocutionary acts or the mind–language interface, and I am doubtful that any such alternative could make sense of the idea that whatever it is that constitutes an act of asserting *p* is the same thing that constitutes an act of believing *p*.

Hanks gives several arguments for his theory. In two recent papers, he shows how his framework allows for solutions to persistent puzzles about the semantics of proper names (2011) and *de se* attitude ascriptions (2012). Even if we assume that these solutions are successful, however, they don't give us a reason to prefer Hanks' theory over a speech act-theoretic semantics like the one defended here, because Hanks' solutions draw on a resource that is common to both approaches—namely, the possibility of individuating the act-type that serves as an expression's semantic value either more broadly or more narrowly than we individuate the contents expressed by uttering the expression on particular occasions (see Chapter Two for my use of this technique).

Hanks' strongest positive argument in favor of his view revolves around embedded questions. He notes that most of the mainstream theories about the semantics of interrogatives hold that their semantic values are different sorts of things than the semantic values of declaratives: whereas declaratives are normally held to express propositions, interrogatives are often held to express properties of propositions or sets of propositions (Hamblin 1973; Karttunen 1977; Groenendijk and Stokhof 1982, 1997). This is problematic, says Hanks, because interrogative clauses can embed in many of the same contexts as declarative that-clauses, whose semantic values are standardly taken to be propositions. Hanks uses knowledge ascriptions as his main example:

(49) (a) Jones knows that Smith is tall.

(b) Jones knows whether Smith is tall. (Hanks 2007: 144)

As Hanks points out, (49a) and (49b) can differ in truth value, and this must be due to some difference in the semantic values of declarative that-clauses and interrogative whether-clauses. Semanticists like those cited above attempt to account for this by saying that interrogative clauses have different types of semantic values than declarative clauses, but this leads to the unwelcome consequence that 'knows' is ambiguous. Hanks gives two objections to positing this ambiguity: first, doing so does nothing to explain the fact that both that- and whether-clauses can embed in certain other positions—e.g., '…depends on…'—and so it seems inevitable that positing an ambiguity to solve our problem with (49) will force us to posit various other ambiguities as well. Second, that-clauses can be coordinated with whether-clauses in knowledge ascriptions, as in (50):

(50) Jones knows that Smith is tall and whether Wilson is Married. (2007: 145)

Hanks concludes that the semantic values of declaratives and interrogatives must be the same type of thing (so that they can be coordinated), but different flavors of that type (so that (49a) and

(49b) can have different truth conditions), and he argues that his theory of propositions best explains these data.

Although this argument is more compelling than Hanks' other arguments for his view, and although it presents a serious challenge to the other available theories of embedded interrogatives, it remains inconclusive. One problem arises from the fact that whether-clauses (and *wh*-phrases) cannot embed in all positions where that-clauses can embed:

(51) # (a) Jones believes whether Smith is tall and whether.

(b) Jones desires whether Smith is tall and whether.

(c) Jones guesses whether Smith is tall and whether

An obvious explanation of the ungrammaticality of (51a)–(51c) would be that whether-clauses don't have semantic values of the same kind as that-clauses. But, obviously, we can't accept the soundness of both this argument and Hanks' argument. Another prima facie roadblock to Hanks' argument is the fact that different attitude verbs can be coordinated with the same complement, as in (52):

(52) Jones both hopes and believes that his favorite team will win.

According to Hanks, belief reports and hope reports express propositions that relate agents to fundamentally different kinds of propositions: my beliefs relate me to assertive propositions and my hope relate me to imperative propositions. But we would have to do posit a substantial hidden grammatical structure in (52) to get a reading on which Jones' belief that his favorite team will win and his hope that his favorite team will win relate him to two distinct propositions. The most natural thing to say about (52) is that it reports two different relations that Jones bears to the same proposition—the force-neutral proposition that Jones' favorite team will win.

Although these considerations do not refute Hanks' argument, they do strongly suggest that the syntax and semantics of embedded interrogatives is much more complex than his arguments assume. As Groenendijk and Stokhof argue: "if one takes into consideration certain constructions involving interrogatives, a flexible approach to the relationship between syntactic categories and semantic types may be of great help" (1989: 421). Although I don't have anything to say about the details, it is my suspicion that some of this flexibility will need to be posited in order to account for (49)–(52), and that the result will undermine Hanks' argument from (49)–(50) to the conclusion that declaratives, imperatives, and interrogatives express fundamentally different kinds of propositions.

Another of Hanks' arguments revolves around his claim that we must adopt an act-theoretic account of propositions in order to solve the problem of the unity of the proposition (2007: §4). This is the problem of saying what it is in virtue of which a structured proposition is a representational unity greater than the mere collection of its parts. What, for example, distinguishes the proposition that Fred loves Mary from the proposition that Mary loves Fred? Assuming that propositions are structured entities, both of these propositions are made up of the same components—namely, Mary, Fred, and the binary LOVE relation—and so we need an account of what unifies these components in specific, different ways in the different propositions. Hanks argues that we must solve this problem by appealing to the combinatory acts—paradigmatically, the act of predicating—that speakers perform in speaking and thinking. It is these acts, says Hanks, that provide the compositional glue that holds propositions together. This idea is radical because it reverses the explanatory relationship that is normally held to obtain between intentional states and their propositional contents: whereas intentional states are normally held to represent things as being a certain way in virtue of having propositional content that represents things as being that way, Hanks maintains that intentional states and speech acts are the fundamental bearers of representational content and that propositions represent things as being a certain way in virtue of

being types of intentional states and speech acts that represent things as being that way. In Hanks' words, representations "are things that we produce in speaking and thinking about the world" (2007: 143).

Suppose that Hanks is right that the problem of the unity of the proposition forces us to undertake this sweeping reconceptualization of the natures of representation, intentionality, and propositional content, and to identify propositions with types of complex acts. Does this give us a reason to adopt Hanks' semantic account of mood? To see that it does not, we need only recognize that Soames (2010) has defended the same sweeping reconceptualization of the nature of representation, intentionality, and propositional content in response to the problem of the unity of the proposition, but has done so without collapsing the force–content distinction. On Soames' view, propositions are types of force-neutral mental acts, and all such mental acts are unified by neutral acts of predication. Soames argues that performing any mental act or speech act with the proposition ⟨George, CLEVER⟩ as its content requires entertaining the proposition that George is clever, and that this is accomplished by predicating cleverness of George, thereby unifying George and cleverness into a representational whole. But he also argue that taking a specific mental attitude, such as belief or desire, toward this proposition, or performing any particular illocutionary act whose content is the proposition, requires performing some further mental act or speech act, beyond the act of predicating that constitutes entertaining the proposition. So although Soames offers essentially the same solution to the problem of the unity of the proposition as Hanks, he does so without collapsing the force–content distinction, and therefore without committing himself to anything like Hanks' account of the meaning of mood.

Soames' act-theoretic, but also force-neutral, solution to the problem of the unity of the proposition should lead us to question why Hanks' theory works the way it does. For Hanks, combinatory acts like predication play two distinct roles simultaneously. First, they play the role of the glue that holds propositions together. Second, they play the role of distinguishing assertive,

interrogative, and directive speech acts (as well as those acts' mental analogues). Soames shows us how to construct a theory that is broadly similar to Hanks', but in which these two explanatory roles come apart. We should therefore ask: what are the costs and benefits of collapsing the two roles?

Hanks might respond that collapsing the two roles gives us substantial conceptual economy, but Hanks' collapse also carries significant disadvantages. To see why, we need to generalize the notion of a combinatory act in two ways. First, notice that although I don't identify semantic values with propositional contents, I still need to say what distinguishes the semantic value of 'Fred loves Mary' from the semantic value of 'Mary loves Fred'. In other words, any theory, such as my own, that takes semantic values to be non-propositional nonetheless incurs the problem of the unity of the *semantic value*. So, in talking about the problem of the unity of the semantic value, I can generalize over both my theory and Hanks' theory. Second, notice that the problem of the unity of the semantic value affects complex sub-sentential semantic values as well as the semantic values of whole sentences. Just as there is a problem of what unifies the semantic value of the sentence 'Fred loves Mary' and thereby distinguishes it from the semantic value of 'Mary loves Fred', there is a problem of what unifies the semantic value of the complex VP 'loves Fred more than Mary' and thereby distinguishes it from the semantic value of 'loves Mary more than Fred'. Whatever we take the semantic values of 'loves', 'Mary', 'Fred', and 'more than' to be, we need some account of what unifies them in two different ways as the distinct semantic values of these two different VPs.[40] A consequence is that whatever plays the role of unifying complex expressions' se-

[40] I will return to this problem and offer my solution in Chapter Six, where I argue that semantic values are unified by force-neutral combinatory acts like predication, and that the notion of predication must be generalized to include combinatory acts that unify complex subsentential semantic values. For example: we refer to the property of loving Mary by referring to Mary, referring to the loving relation, and predicating the latter of the former. We can then predicate the complex property thus referred of Fred, thereby expressing the proposition that Fred loves Mary. Chapter Six explores how to give a rigorous account of the syntax–semantics interface that is compatible with this account.

mantic values, it can't in general be the semantic contribution of mood, which is a syntactic property of clauses and not subsentential expressions. Since Hanks also has to say what unifies complex, sub-sentential semantic values, he also has to posit some other, sub-sentential combinatory acts. Moreover, these combinatory acts will have to be force-neutral, since the same VP can appear in clauses of different kinds, presumably with the same semantic value in each case, and since Hanks takes force-loaded combinatory acts to be the semantic contribution of sentences' moods. So, even Hanks is committed to there being *some* combinatory acts that can unify semantic values in a force-neutral way. It makes sense, then, to wonder why he takes the combinatory acts that unify speech acts' contents to be the same combinatory acts that constitute their illocutionary force.

In running together the semantic contribution of a clause's mood and the act that unifies the clause's semantic value, Hanks gets into other kinds of trouble as well. He argues that the semantic value of a sentence is a type of speech act/mental act whose content and force/attitude are inseparable. But, as he notes, the fact that embedded clauses aren't uttered with their usual force is in tension with this view.

> Suppose I sincerely utter "If Jones likes Smith then Smith likes Jones". Clearly,I have not asserted that Jones likes Smith. But the sentence "Jones likes Smith" has the very same content when it occurs inside this conditional as it does when it occurs alone. Hence, there cannot be an element of assertion in this content. Otherwise I could not use the sentence without asserting it. (Hanks 2007: 153)

The same problem arises for clauses embedded under negation, disjunction, and in some other positions. Hanks attempts to solve the problem by arguing that it is part of the semantic import of these operators that they cancel the force of propositions embedded in them:

> This argument assumes that what is absent from the content of a sentence when it is used as the antecedent or consequent of a conditional is always absent from that content. But another way to view these facts is to hold that assertive force is always part of the content of an indicative sentence, but this element of force is cancelled when the sentence is uttered as part of a conditional. This is one of the semantic effects of the word "if". "If" cancels the assertive force

> of the content of the sentence that follows it—this is a semantic feature of the conditional. Intuitively, "if" signals to the hearer that what follows is not being asserted. (Hanks 2007: 153–4)

Hanks compares his proposed semantic cancellation to a couple of pragmatic ways that the force of a speaker's utterance can be cancelled. One of these is retraction, as in an utterance of 'Jones likes Smith. No, I take it back.' However, Hanks argues that retraction differs from the kind of cancellation involved in conditionalization, negation, and disjunction, because cancellation-via-retraction literally happens after the fact—one first performs a speech act, and then cancels it—whereas one never performs the illocutionary act associated with a clause embedded in a context that involves semantic cancellation. Hanks next draws an analogy between his notion of semantic force-cancellation and the way in which a speech act's force gets cancelled when uttered in a fictional context, such as when performed in a play or written into a poem. Hanks quotes a passage in which Frege argues from the fact that a sentences in fictional contexts are used with their customary content but not their customary force to the conclusion that force must be distinct from content. Hanks responds that assertoric force is part of the content of a declarative sentence uttered by an actor on stage, but that "there are conventions about being on stage or writing poetry that cancel the assertive force that would normally accompany utterances of those sentences" (2007: 155).

> The general point is that an utterance of an indicative sentence counts as an assertion of the content of that sentence unless some other factor intervenes to cancel the assertive force of that content. This intervening factor can be the presence of a connective like "if…then…" or it can be the fact that the sentence is uttered by an actor on stage, or any numberof other things (e.g. intonation). (Hanks 2007: 156)

At this point, we should keep in mind that this account of cancellation is needed only because Hanks holds that combinatory acts like predication play the following two roles at once: (a) they unify clauses' semantic values, and (b) they constitute the illocutionary force of clauses' semantic values. In arguing that various linguistic and conversational contexts can cancel the customary

force with which a clause is uttered, Hanks therefore concludes that combinatory acts don't play the (b)–role in those contexts. But Hanks still needs combinatory acts to play their (a)-role: the problem of the unity of the semantic value is every bit as urgent when it comes to the semantic values of the antecedents of conditionals, for example. It follows that the (a)–role and the (b)–role can sometimes come apart, so that whatever plays one role can be present when whatever plays the other is made to be absent. In other words: we still predicate cleverness of George, even when uttering 'George is clever' in the antecedent of a conditional. But, in that context, the act of predicating is, effectively, force-neutral. So, in addition to needing force-neutral combinatory acts to unify complex sub-sentential semantic values, Hanks needs *effectively* force-neutral combinatory acts to do the job of unifying clausal semantic values in a range of special contexts. Still, he thinks that we should maintain that the combinatory acts that unify clausal acts are paradigmatically force constituting. At this point, it seems that it would be simpler if Hanks adopted a view, like-Soames' and my view, on which combinatory acts play only the (a)–role, and that the (b)–role is instead played by the features of communicative intentions that I outlined in §1.2.

Further difficulties with Hanks' account arise when we consider some of the ways in which clauses of different types interact, both grammatically and in conversation. What can his account say about the semantic values of clauses embedded under logical connectives, for example? First notice that Hanks is barred from taking the sort of approach that I explored in §3.1.3, on which mood takes scope over connectives. This is because Hanks needs the combinatory acts that are the semantic contributions of the moods of embedded clauses to play the role of unifying those clauses' semantic values. Only unified propositions can be conjuncts, disjuncts, and the antecedents and consequents of conditionals, and Hanks holds that it is the semantic contribution of mood that does the job of unifying propositions.

One problematic consequence of this view is that negation must take scope over the imperative mood in sentences like (46):

(46) (a) Don't leave me here!

(b) ¬(!φ)

The interaction between mood (and particularly imperative mood) and negation is a fraught topic (Han 1998: ch.2) and semanticists have expressed doubt that negation ever takes scope over mood (e.g., Starr *ms*: §3.3). Moreover, reading (46a) as having the structure of (46b) gives rise to a serious internal inconsistency in Hanks' view. On one hand, Hanks is committed to the idea that negation, as a matter of its semantics, creates a force-cancelling environment, but on the other hand, (46a) would normally be used to perform a directive speech act. But if this directive force isn't semantically contributed by the mood of !φ because this contribution is cancelled by being embedded under negation, then where does its directive force come from? In fact, the very same problem arises for negated declarative sentences like (47):

(47) (a) Peter isn't right.

(b) ¬(φ)

Hanks needs negation to scope over mood in sentences like (47a) in order to explain why we don't assert negated propositions. But in that case, where does an utterance of (47a) get its assertoric force? Not the mood of the embedded declarative clause, whose usual contribution of force has been cancelled. This is a genuine conundrum for Hanks.[41]

---

[41] It might be tempting to say that Hanks can solve his problems with negation by taking it to be something other than a sentence connective. For example, perhaps he could hold that the logical form of (46a) has a structure like (46c), with negation serving as part of the semantic value of a complex predicate expression:

(46) (a) !(λxλy[x doesn't leave y here]addressee$_c$, speaker$_c$)

But on this view, it's hard to see how Hanks can capture a wide variety of inference patterns—such as modus tollens or disjunctive syllogism, that assume that negation can be treated as a sentence connective. Quite generally, we should be suspicious of any view according to which negation can *never* function as a sentence connective.

The fact that Hanks' account requires mood to take narrow scope with respect to connectives also makes mixed-mood sentences like (8)–(11) problematic. To get at this problem, first consider what Hanks says about disjunctions whose disjuncts are both declaratives:

> Although a speaker asserts neither disjunct by uttering 'George is clever or Karla is foolish', she still performs an assertion and hence an act of predication. The speaker asserts that George is clever or Karla is foolish. How should we understand the act of predication contained in this assertion? Let *p* and *q* be propositions expressed by declarative sentences, that is, types of predicative actions. To assert that p or q is to predicate a disjunctive relation, expressed by 'or', of *p* and *q*. Two propositions *p* and *q* bear this disjunctive relation just in case either *p* is true or *q* is true. As types of predicative acts, the propositions *p* and *q* are true or false and hence can stand in this disjunctive relation. (2011: 20).

Here Hanks implies that the syntactic environment '…or…' cancels the assertoric force of the semantic contributions of its embedded clauses' moods, but itself has the semantic equivalent of declarative mood, in the sense that its semantic contribution is not merely the disjunction relation, but also the assertorically-loaded predication of that relation. This account is quite obviously incompatible with imperative disjunctions like (7), which we don't use to perform assertive speech acts.

(7) Buy me a drink or leave me alone! !φ or !ψ

Perhaps we could generalize Hanks' account of '…or…' by saying that it does not semantically contribute any particular kind of combinatory act, but instead contributes a combinatory act that is underspecified with respect to its force, such that its force is inherited from the semantic contributions of the clauses it embeds. This would deliver the same result as holding that mood takes scope over disjunction. But this strategy leaves us without an explanation for mixed-mood disjunctions like (9):

(9) a. Give me all your money or the puppy gets it.

b. Mow the lawn or I'll clean the garage.

What combinatory act unifies the semantic value of (9b)? The answer can be neither assertive-predication nor its directive counterpart, since we perform neither an assertive not a directive speech act with a literal utterance of (9). It is not clear what Hanks could say about this.

The problems I have just enumerated are all broadly grammatical in nature, but by running together whatever plays the semantic value-unifying role with whatever plays the force-constituting role, Hanks' view is also in tension with some conversation-level phenomena. Imagine the following dialogue:

(48) (a) Mary loves Fred. [uttered by Dan]

(b) Does Mary love Fred? [uttered by Rob]

(c) That's what I just said. [uttered by Dan]

The most natural interpretation of (48c) is that Dan uses 'that' to refer to the proposition Rob expressed with (48b) and 'what I just said' to refer to the proposition he himself had expressed in uttering (48a), and then identifies those two propositions. Moreover, this reading correctly predicts that Dan says something true in uttering (48c). But this interpretation is not available to Hanks because his theory predicts that Dan and Rob express different propositions in uttering (48a) and (48b). Specifically: Dan expresses an assertoric proposition in uttering (48a) and Rob expresses a distinct, interrogative proposition in uttering (48b). So, assuming that 'that' and 'what I just said' are being used as devices of propositional anaphora in (48c), what Dan says thereby must be false. This recalcitrant prediction is good prima facie evidence against Hanks' view.[42]

[42] It is also worth pointing out that nothing in this example hangs on the fact that (48c) involves indirect quotation, since the following dialogue makes roughly the same point: A: 'Can every even integer greater than 2 be expressed as the sum of two primes?'; B: 'That's Goldbach's conjecture.'

3.4 Dynamic Theories of Imperatives

With an array of doubts cast upon content-theoretic approaches to imperatives, I turn my attention to dynamic theories. I use the term 'dynamic' to discuss two kinds of approaches, one broadly pragmatic and the other semantic. The basic idea behind all dynamic approaches is that we should understand linguistic expressions in terms of their effects on the shared and evolving context in which a conversation takes place. Context, on this view, is a body of representations that is constructed from the shared propositional attitudes of the participants in a conversation, and that both makes possible and is updated by the speech acts that these participants perform.[43] Dynamic accounts understand speech acts in terms of their effects on context, and then give either a pragmatic or a semantic explanation for why uttering sentences of certain kinds allows speakers to have the relevant kind of contextual effect. Following Karen Lewis (2011, 2012, *forthcoming*), I will say that what distinguishes dynamic semantic theories from dynamic pragmatic theories is whether the contextual effect of uttering a sentence is built directly into the sentence's meaning or explained in terms of broad considerations about rational cooperation.

As an illustration, consider Stalnaker's (1978) theory of assertion. Stalanker argues that we should think of the context of a conversation as—or at as least including—a representation of the participants' shared presuppositions. This can be modeled either as the *common ground*—the set of propositions mutually presupposed by the speakers—or as the *context set*—the set of possible worlds compatible with all of the speakers' presuppositions. Stalnaker proposes that we understand assertion in terms of its *essential effect* on the context, which, he argues, is to add the asserted content to the common ground and thereby cull incompatible worlds from the context set. Stalnaker takes this effect to be an essential component of the illocutionary force of asserting.

---

[43] This notion of context has gone by several different names: 'common ground' and 'context set' (Stalnaker 1978), 'conversational scoreboard' (Lewis 1979a), 'conversational record' (Thomason 1990), 'information structure' (Roberts 1996/2012), 'discourse context' (Roberts 2004), 'information state' (Veltman 1996), 'conversational state' (Starr *ms*) and so on.

(The distance from this claim to the claim that illocutionary acts are constituted by their essential effects will come come under scrutiny in §3.5.)

Stalnaker counts as giving a dynamic *pragmatic* account of declarative sentences because he appeals to general pragmatic resources to explain why uttering a declarative sentence, whose content is a proposition, is a good way to add that proposition to the common ground and thereby perform an assertion. But we could just as easily go the route of dynamic semantics and build the essential effect of assertion into the meanings of declarative sentences. In a dynamic semantic theory, sentences' semantic values are *context change potentials*—functions mapping contexts to contexts. Following Veltman (1996), we can capture this idea by saying that a dynamic semantic theory for a language $\mathscr{L}$ specifies an update system $\langle \Sigma_{\mathscr{L}}, \mathscr{C}, \cdot[\cdot] \rangle$, where $\Sigma_{\mathscr{L}}$ is the set of sentences of $\mathscr{L}$, $\mathscr{C}$ is the set of possible states of the context, and $\cdot[\cdot]$ is a function that assigns to each sentence $\sigma \in \Sigma_{\mathscr{L}}$ a function $\cdot[\sigma]$ that maps context states into context states ($\cdot[\sigma] : \mathscr{C} \longmapsto \mathscr{C}$). We can construct a dynamic semantic account of declarative meaning paralleling Stalnaker's by taking each possible context-state $c \in \mathscr{C}$ to be a set of propositions (i.e., a possible state of the common ground). Then, taking the semantic value $[\![\psi]\!]$ of a sentence radical ψ to be a set of worlds (i.e., the proposition it expresses), we can make the following generalization about the semantic values of declarative sentences:[44]

For every context-state $c \in \mathscr{C}$ and every indicative sentence $\ulcorner \vdash \psi \urcorner \in \Sigma_{\mathscr{L}}$:

$c[\vdash \psi] = c \cup \{[\![\psi]\!]\}$

In English: the context change potential of a declarative sentence adds the content of its sentence radical to the common ground.

---

[44] Following Veltman's notation, I will specify the output of an expression φ's context change potential on a context c by postfixing the function to its argument: '$c[\varphi]$'. Moreover, I will assume that the semantic value of a sentence radical ψ is a proposition $[\![\psi]\!]$.

Dynamic accounts of non-assertoric speech acts typically build on Stalnaker's theory of assertion by positing richer contexts whose structures represent more of what is happening in a conversation, and by showing how the essential effects of non-assertoric speech acts can be understood in terms of the way in which they affect this added structure. For example, the most influential theory of questioning among contemporary linguists holds that the essential effect of a question is to partition the common ground into sets of possible answers (Ginzburg 1995a,b, 1996; Roberts 1996/2012, 2004). These partitions on the common ground are sometimes called a conversation's *issues* or *questions under discussion*; they can be thought of as representing the current set of unresolved topics of conversation, and further details about the relative priority of questions and the order in which they were asked or ought to be resolved can be represented by imposing further structure on the 'question unde discussion stack' (Roberts 1996/2012: §1.2). Just as part of the appeal of Stalnaker's dynamic theory of assertion is that it yields predictions about how past assertions affect what can be said later in a conversation—for example, by affecting how context-sensitive expressions are interpreted—the dynamic theory of questioning has been used to predict and explan a variety of facts about how conversations unfold—for example, how asking a question can affect the relevance of future questions and assertions (Roberts 1996/2012). And, as in Stalnaker's account of assertion, we can either write the essential effect of questioning directly into the meaning of the interrogative mood by adopting a dynamic semantic framework, or we can give a pragmatic explanation of why uttering an interrogative sentence is a good way to partition the context set. One plausible version of a pragmatic account of this kind might appeal to the idea that interrogatives' semantic values—which are typically identified with properties of propositions or sets of propositions—are of a different semantic type than declaratives' propositional semantic values, and so "belong" in a different dimension of the context. As we'll see in the next two sections, both of thse strategies have been extended to dynamic accounts of imperatives.

3.4.1 Dynamic Pragmatic Theories of Imperatives

Dynamic theories of directive force follow the same strategy as dynamic theories of asserting and questioning: they posit further structure in contexts and explicate the directive force of an act in terms of the act's effect on this further structure. Dynamic pragmatic theories of imperatives also rely on the strategy of assigning semantic values of a type different than those of declaratives and interrogatives, such that each type of semantic value is native to a particular dimension of context. Whereas declaratives are taken to express propositions and interrogatives express sets of propositions, the dynamic pragmatic theories I'll consider here both take imperatives to express properties.

Taking up an idea first proposed by David Lewis (1975b), dynamic accounts take directive speech acts to update an aspect of the context that represents agents' mutually agreed-upon practical commitments—their obligations, preferences, or plans. Perhaps the most influential such account due to Portner (2004, 2007, 2012). In addition to the Common Ground (CG) and Questions Under Discussion (QUD), Portner posits a contextual *To-Do List* (TDL), which he models as a function from each participant in a conversation α to lists of properties of the form $\ulcorner \lambda x : x = \alpha \,.\, Fx \urcorner$ (2007: 358). The content of an imperative sentence, on Portner's account, is also a property of this form, and the essential effect of a directive act is to add such a property to the addressee's To-Do List. As the name would suggest, the To-Do List is a representation of the actions—or, more broadly, states of the world—that are taken, by participants in a conversation and for the purposes of the conversation, to guide agents' rational behavior.

> We may think of the Common Ground and To-Do List as being the public, or interactional, counterparts of the individual agent's beliefs and desires. That is, as far as the participants in an interaction go, an agent's actions will be judged rational to the extent that, if undertaken in any world compatible with the Common Ground, they would tend to make this world maximally highly ranked according to that agent's To-Do List. (Portner 2004: 8)

Rather than defining the publicly accepted criteria on rational behavior directly in terms of the To-Do List, Portner uses it to define a partial ordering on worlds compatible with the Common Ground (CG) for each agent α (2004: 8):

Partial Ordering of Worlds
$(\forall w_1, w_2 \in \cap CG)((w_1 <_\alpha w_2)$ *iff*
$(\exists \varphi \in TDL_\alpha)(\varphi w_2 \alpha = 1 \ \& \ \varphi w_2 \alpha = 0 \ \& \ (\forall \psi \in TDL_\alpha)(\psi w_1 \alpha = 1 \rightarrow \psi w_2 \alpha = 1)))$

In something closer to English: a world w_1 is better than a world w_2 from the perspective of an agent's To-Do list just in case the agent has more of the properties on her list in w_1 than in w_2. Portner argues that this partial ordering figures in the following principle governing the rational behavior of the participants in a conversation (2004: 9):

Agent's Commitment

For any agent α, the participants in the conversation mutually agree to deem α's actions rational and cooperative to the extent that those actions in any world $w_1 \in \cap CG$ tend to make it more likely that there is no $w_2 \in \cap CG$ such that $w_1 <_\alpha w_2$.

To act rationally for the purposes of the conversation, according to Portner, is to act so as to gain as many of the properties on one's To-Do List as possible. On this view, the essential effect of a directive speech act is to indirectly alter what the participants in the conversation mutually agree that the addressee ought, rationally, to do.[45]

Portner explains the connection between imperatives and directive force pragmatically. Following Stalnaker (1978) and Roberts (1996/2012), he holds that declaratives express propositions and interrogatives express sets of propositions. He adds the new claim that imperatives express

[45] Portner (2007) also argues that the To-Do List helps determine the ordering source relative to which deontic modals are interpreted. I am skeptical that this is a feature, rather than a bug, of Portner's approach (see §3.1.6), but that aspect of his view is separable from the rest.

properties whose arguments are restricted to the addressee at a context of utterance. For example, (49a) expresses the property (49b) in a context *c*:

(49) (a) Leave!

(b) $\lambda x : x = \text{addressee}_c \,.\, x$ leaves (Portner 2004: 10)

All of this furnishes Portner with an elegant explanation of several of the central explananda of a semantics for imperatives. Each clause-type correspond to a dimension of context by virtue of the fact that clauses of that type have as their semantic values entities of the sort that belong in the relevant dimension of context. Since properties of speakers are the sorts of things that go on To-Do Lists, proferring such a property by uttering an imperative sentence makes it obvious that one is trying to add it to someone's To-Do List, and the fact that the property can only be possessed by one's addressee makes it obvious whose To-Do List it belongs on. Both the distinctive meaning and the distinctive use of the imperative mood are thus explained. The apparent non-truth-aptness of imperatives is explained by the fact that their contents—properties instead of propositions—aren't truth-apt.

If we interpret Charlow's (2013) theory of imperatives as a genuine dynamic pragmatic one, then it is similar to Portner's in several respects, although he takes imperative semantic values to be different kinds of properties than Portner.[46] Charlow refers to the context of conversation as a *representor* $\langle S, \Lambda \rangle$, with a set of possible worlds S representing *information* and a set of action-

[46] I've hedged in describing Charlow's theory as a dynamic one because it's not clear that Charlow himself conceives of his theory in that way. Charlow frequently suggests that the pragmatic function of a speech act is to update an addressee's mental state, rather than to update mutually represented states of the context. In the text, he defines the notion of representor $\langle S_a, \Lambda_a \rangle$ "for an agent *a*", and says that it represents "*a*'s information" and "*a*'s plans" (2013: §6.2). If we take this description literally, then Charlow's theory of illocutionary force is already much closer to a Gricean one than the other theories I will discuss in this section. The distinction between dynamic and Gricean theories of illocutionary force will be the main topic in §3.5, and so I'll return to this issue there. For now, I will present Charlow's theory as if it were a standardly dynamic one, with the disclaimer that this may not be what Charlow intends.

descriptors Λ representing agents' *plans* (I will follow Charlow in representing each agent *a*'s plans as Λ_a). Charlow takes imperatives to express requirement conditions on plans, and to perform a directive speech act by uttering an imperative !φ is to "tell an agent what to plan" by expressing "the property a plan has if it is decided on φ" (2013: §5.2). A plan Λ "is decided on φ, in the relevant sense, iff φ is required relative to Λ" (2013: §5.2). Charlow then defines the relevant notion of *requirement* by defining a partial ordering on worlds as follows.

> A plan for an agent *a*, Λ_a thus determines a set of propositions $P_{\Lambda a}$: the set of propositions *p* such that for some $\alpha \in \Lambda_a$, *p* is the proposition that *a* performs α. Such a set of propositions can straightforwardly determine a preorder on possibilities as follows.
>
> $w \leq_{\Lambda a} v$ iff $\{p \in P_{\Lambda a} : v \in p\} \subseteq \{p \in P_{\Lambda a} : w \in p\}$
>
> In English, *w* is at least as good as *v*, with respect to Λ_a, iff *w* involves *a*'s performance of at least those actions in Λ_a that *a* performs in *v*, and perhaps more besides. (2013: §5.2)

Charlow calls the worlds that are minimal with respect to the ordering $\leq_{\Lambda a}$ the *Λ_a–best worlds*, and refers to the set of these worlds as $min_{\Lambda a}$. This setup allows Charlow to capture his idea that the essential effect of a directive is to impose a requirement on a plan, and that a speaker can do this by uttering an imperative sentence that expresses that requirement as a property of the addressee's plan. Charlow models the semantic value of (49a) as (49c), for exmaple:

(49) (a) Leave!

(c) $\lambda\Lambda_a$. 1 only if $\forall w \in min_{\Lambda a} : \{w\} \vDash \text{addressee}_c \text{ leaves}$

In English: (49a) "places the following as a condition on a plan: the plan must prefer [worlds in which the addressee leaves] to [worlds in which the addressee doesn't leave]; a plan receives a positive verdict (i.e., 1) only if it meets this condition. … By encoding such a property, an imperative…tells an agent what/how to plan" (2013: §5.2).

Again: this explains both the semantic and pragmatic difference between imperatives and declaratives in terms of the fact that they are used to update the context in different ways: whereas a

declarative is used to add its content—a proposition—to the context's information coordinate, an imperative is used to impose its content—a property—as a requirement on the plan coordinate. The intuition that imperatives aren't truth-apt is likewise explained by saying that they express properties rather than propositions.[47]

So, despite their differences, Portner's and Charlow's accounts share most of their obvious strengths. They also share some weaknesses. One peculiarity of both accounts is that they apply, as stated, only to *requirement directives* and not to *permission directives*. Both Portner and Charlow need to tell an additional story about how I can permit you to have some tea by saying 'have some tea!', for example. Portner (2012) gives a fairly elaborate pragmatic explanation of these cases. Charlow gives distinct semantic accounts of requirement and permissive readings of imperatives, suggesting that he takes imperatives to be systematically ambiguous between the two readings. The need to explain away permission directives in these ways is not a damning problem for Portner and Charlow, but it would be nicer if they didn't have to.

A much more serious problem for both Portner and Charlow arises from embedded imperatives. Portner does not attempt to give an account of conditional imperatives, although he makes it clear in a recent paper that doing so is on his To-Do List (2012: §5). Charlow takes conditional imperatives to express restricted properties of plans—a notion that requires only a simple extension of his account of imperatives (2013: 5.5). But Charlow's account seems not to be general enough, because he treats conditional declaratives and conditional imperatives as two distinct phenomena, requiring distinct clauses in his semantics. But this means that he has no way of accounting for conditionals, such as (11), whose consequents are mixed-mood conjunctions or disjunctions.

---

[47] Charlow also goes on to give an account of imperative consequence by defining a relationship between imperatives and closely related modals, thus preserving the supposed virtues of modal-reductionism in this respect (2013: 5.4).

(11) If he comes back, buy me a drink and I'll pay you back. (if ⊢φ)(!ψ and ⊢σ)

In order to account for the meaningfulness of (11), what we need is an account of conditionals that generalizes across consequents with different moods and combinations of moods, and Charlow does not offer this. Although there may be some way of accomplishing this in his system, it is not obvious how it would work.

Serious problems for both Portner and Charlow arise when we consider mixed-mood conjunctions and disjunctions such as (9) and (10):

(9) a. Give me all your money or the puppy gets it. !φ or ⊢ψ

b. You mow the lawn or I'll clean the garage (Starr *ms*: §2.1)

(10) a. Say that again and I'll scream. !φ and ⊢ψ

b. Stay here and I'll go to the store.

These sentences pose problems for Portner and Charlow because standard wisdom in contemporary semantics says that conjunction and disjunction range only over pairs of semantic values of the same type (e.g., Partee and Rooth 1983), but both Portner and Charlow take the semantic values of imperatives and declaratives to be of different types (cf. Starr *ms*: §4.1). Charlow considers treating (9a) and (10a) as conditionals in disguise, but as we saw in §3.1.3, not all mixed-mood conjunctions and disjunctions can be treated in this way. Charlow's response is to redefine conjunction and disjunction in terms of a general notion of what it is for a sentence φ to *hold at a representor* ⟨S, Λ⟩ (represented as ⌜⟨S, Λ⟩ ⊨ φ⌝). Here are Charlow's definitions (2013: §6.2):[48]

⟨S, Λ⟩ ⊨ ⊢φ iff ∀*w* ∈ S : φ is true at *w*

⟨S, Λ⟩ ⊨ ¬⊢φ iff ∀*w* ∈ S : φ is not true at *w*

[48] I have made some adjustments to Charlow's definitions in order to avoid introducing otherwise unnecessary terminology and notation. As in Charlow's version, uppercase Greek letters range over both declarative and imperative clauses.

$\langle S, \Lambda \rangle \vDash !\varphi$ only if $\forall w \in min_{\Lambda a} : \varphi$ is true at w

$\langle S, \Lambda \rangle \vDash (\text{if } \varphi)(!\psi)$ only if $\forall w \in min_{\Lambda a} \cap \{w' : \varphi \text{ is true at } w'\} : \psi$ is true at w

$\langle S, \Lambda \rangle \vDash (\Phi \wedge \Psi)$ iff $\langle S, \Lambda \rangle \vDash \Phi$ and $\langle S, \Lambda \rangle \vDash \Psi$

As Charlow explains, this treatment allows him to make sense of mixed-mood conjunctions, while also allowing him to capture some high-level generalizations about how the semantics of sentences connects up to their pragmatics:

> …there is no problem understanding a conjunction of the form $(!\varphi \wedge \psi)$. This expresses the property a representor has iff it satisfies the property expressed by $!\varphi$ as well as that expressed by ψ. Effectively, then, we are claiming that sentences of any type express the very same sort of semantic object: in general, sentences express properties of representors. Imperatives express properties that bear on the planning parameter of a representor, while declaratives express properties that bear on its informational parameter. Sentences are not typed according to the semantic type of their denotation, since all sentences express characteristic functions of representors. Instead, they are typed according to the kind of property of a representor they express. (Charlow 2013: §6.2)

This result is very elegant, and my own account of mood-mixing—which I'll outline in §3.6—shows its influence. But we should also take note of the fact that the definition of holding-at given above does not include disjunction. Instead, Charlow defines disjunction in terms of conjunction in a footnote:

> Disjunction is defined in terms of negation and conjunction as follows. $S \vDash (\varphi \vee \psi)$ iff $S \vDash \neg(\neg\varphi \wedge \neg\psi)$. This holds iff $\forall w \in S : \{w\} \vDash \neg\varphi \wedge \neg\psi$ iff $\forall w \in S : \{w\} \vDash \neg\varphi$ or $\{w\} \vDash \neg\psi$ iff $\forall w \in S : \{w\} \vDash \varphi$ or $\{w\} \vDash \psi$. This is the correct notion of disjunction for the sake of our cognitively oriented semantics: for a disjunction $(\varphi \vee \psi)$ to hold ("be known") at S, it is required only that, for every world in S, at least one of φ or ψ holds at that world. This is crucially (and correctly) distinct from a notion of disjunction on which $(\varphi \vee \psi)$ holds at S iff $S \vDash \varphi$ or $S \vDash \psi$. Knowing $(\varphi \vee \psi)$ does not require knowing at least one of φ or ψ; one can know a disjunction without knowing which disjunct is true. (Charlow 2013: §5.1, *fn.*30)

But, as Starr explains, this account of disjunction is seriously flawed, in that it seemingly can't account for disjoined imperatives.

> However [Charlow's] analysis of disjunction is problematic. For declaratives, disjunction cannot be defined as $\langle S, \Lambda\rangle \vDash \Phi \vee \Psi \iff \langle S, \Lambda\rangle \vDash \Phi$ or $\langle S, \Lambda\rangle \vDash \Psi$. That would imply that if an agent accepts a disjunction then their information supports one of the disjuncts. But you can accept *Gabe is sleeping or he's awake* without knowing that he's sleeping or that he's awake. So disjunction is instead defined in terms of conjunction and negation: $\langle S, \Lambda\rangle \vDash \Phi \vee \Psi \iff \langle S, \Lambda\rangle \vDash \neg(\neg\Phi \wedge \neg\Psi)$ (Charlow 2013: §5.1). Since imperatives do not embed under negation, and it is unclear what such an embedding would should mean, this semantics does not capture imperatives scoping under disjunction or mixed disjunctions…. (Starr *ms*: §4.3)

So although Charlow accounts for mixed-mood conjunction, he fails to give a sufficiently general account of the semantics of conditionals, and he fails to account for imperatives embedded under disjunction. Portner's theory fares even worse in this regard: he has not yet attempted to account for any of the data embedded imperatives, and there are good reasons to think that he cannot do so. Dynamic pragmatic theories of mood thus show a good deal of promise, but also leave much to be desired.

3.4.2 Dynamic Semantic Theories of Imperatives

Dynamic semantic theories build the essential effects of illocutionary acts directly into clauses' meanings by identifying their semantic values with context change potentials. Theories of this kind promise many of the same explanatory advantages as dynamic pragmatic theories, but without the challenge of giving a pragmatic explanation of the connection between clause-type and illocutionary force. Since this challenege is part of what led to Portner's and Charlow's troubles with embeddings, it's worth seeing if a dynamic semantic account can do better.

Although dynamic semantic variants of many of the accounts I've looked at so far are potential options, the most fully articulated dynamic semantic theory of imperatives in the literature is due to Starr (*ms*).[49] His account is built around the idea that contexts are *information relations*—sets of ordered pair of propositions (*ms*: §3.1). The ordering of these pairs represent mu-

[49] Lascarides and Asher (2004) offer an alternative dynamic semantic treatment of imperatives that I won't consider here.

tual preferences that the participants in a conversation have (for the purposes of the conversation) for the first coordinate of each pair over the second coordinate. A preference for the proposition that Oliver buys Dan a drink (p) over its negation could be represented by the pair $\langle p, \neg p \rangle$, for example, and a context's information relation r might contain many such pairs. (In order to handle disjunction, Starr officially identifies a conversation's context with a set R of information relations. I'll return to this point below.) Starr next goes on to show how other dimensions of context can be defined in terms of information relations. He defines the *information* in a context (cf. Stalnaker's Context Set) as the set i of worlds that are related by information relations, and the *contextual questions* (or *issues*) of a context (cf. Roberts' Questions Under Discussion) as the set of propositions p such that $\{\langle p,\varnothing\rangle,\langle i\text{–}p,\varnothing\rangle\}$ is among the context's information relations (*ms*: §3.1).[50] Whereas asserting adds information to the context and questioning add issues, the essential effect of directive acts, according to Starr, is to change "the preferences to which an agent is publicly committed" by introducing new (non-trivial) preferences to the context's information relation(s) (*ms*: §3.2–3.3).

Starr implements these ideas formally as follows (*ms*: §3.3):

> …!A changes R by adding a preference for live A-worlds over live non-A-worlds. Since R is a set of info relations, this involves adding to a new preference each info relation r ∈ R. This semantics, applied to an atomic imperative, can be stated in update semantics format (Veltman 1996):
>
> **Basic Imperative Semantics**
> Where α is an atomic radical and $\bar{a} = i_r - a$, $R[!\alpha] = \{r \cup \{\langle a, \bar{a}\rangle\} \mid r \in R \;\; a = i_r[\alpha]\}$.
>
> What are A and i_r[A]? A is the original sentence minus its mood; a **propositional radical** (Wittgenstein 1953: §22n). It is not a well-formed sentence, but as a sub-sentential constituent it still has a semantics. It operates just on information i_r because it is the informational core of

[50] One might wonder whether this way of modeling questions does justice to the phenomenon: does it really make sense to identify the propositions we're mutually wondering about as just those propositions whose truth we would prefer to the empty proposition (which, by the way, is also identified with every necessarily false propisition)? This certainly sounds strange to me: why can't we recognize an alternative as open while preferring some necessary falsehood to it? Perhaps this is just one of those unsatisfying spandrels of possible worlds semantics that we just have to pretend not to notice.

the sentence. But it is not a 'move in the language game' because it doesn't operate on R. Information is modeled in terms of possible worlds, and for convenience possible worlds are treated as functions that map every radical to a truth-value. A radical's semantic role is to eliminate worlds where it's false:

Basic Radical Semantics
For any atomic radical α, $i[\alpha] = \{w \in i \mid w(\alpha) = 1\}$

We can simplify this picture somewhat by saying that a propositional radical φ expresses a proposision $[\![\varphi]\!]$, defined as sets of possible worlds. Taking i_r to be the information contained in an information relation r, we can then say that the semantic function of an imperative sentence !φ is to add the preference $\langle p, \neg p \rangle$ to every $r \in R$, where $p = i_r \cap [\![\varphi]\!]$ and $\neg p = i_r - [\![\varphi]\!]$. In something closer to English: the semantic function of an imperative is to add a preference to each information relation in the context for the worlds that are compatible with both that information relation's information and the imperative's radical over the worlds that aren't compatible.

The pragmatic upshot of this account resembles those of Portner's and Charlow's accounts: the speech acts that we perform in uttering declaratives and imperatives are understood in terms of the way they shape mutual information and action-guiding states, respectively. Starr then defines a choice function whose role is to take "a body of preferences and returns the set of alternatives a rational agent should pursue given those preferences" (ms: §3.2). Starr then formulates a principle governing the behavior it is rational to pursue (for the purposes of the conversation) in terms of these good alternatives, which he sums up as follows:

> …adding to the preferences that an agent is publicly committed to can change which alternatives they can be expected, for the purposes of the conversation, to pursue *if they are rational*. This fact shows that understanding the conversational dynamics of imperatives in terms of preference change explains how they serve their conversational function: to promote some alternative(s) that the speaker might be publicly expected, at least for the purposes of the conversation, to pursue. (Starr ms: §3.2)

This account of the pragmatics of imperatives closely resembles Portner's in several respects. Like Portner, Starr takes the pragmatic aim of an imperative to be that of updating *public* action-

guiding representations in terms of which rational behavior is defined *for the purposes of the conversation*. The two italicized features—that Portner and Starr take the essential effects of directive acts to be public and as-if attitudes—will loom large in §3.5.

Starr goes on to give an analogous semantics for declaratives, followed by mood-neutral treatments of conjunction, disjunction, and conditionals. Declaratives update the context's information in roughly the way Stalnaker envisioned, but, because of the complexity of Starr's framework, the definition looks more complicated than the dynamic semantic translation of Stalnaker that I outlined at the start of this section:

Declarative Semantics
Where $a_\rho = \cup(\{a\}[\rho])$, $R[\vdash\rho] = \{\{\langle a_\rho, a'_\rho \rangle r \in R : \langle a_\rho, a_\rho \rangle \in r \ \ a_\rho \neq \varnothing\}$
- Filter the alternatives in each info relation with ρ, throwing out pairs that would promote ∅ and eliminating any info relations that would contain only these pairs. (Starr *ms*: §3.3)

Starr's clause for conjunction is straightforward, but notable because it is neutral with respect to the moods of the clauses conjoined:

$$R[\Phi \wedge \Psi] = (R[\Phi])[\Psi]$$

In English: the result of updating the context with a conjunction is the same as the result of updating the context with the first conjunct followed by updating the resulting context with the second conjunct. Disjunction is more complicated. As Starr points out, there is a sense in which both conjunctions of imperatives and disjunctions of imperatives introduce preferences for both of their coordinated clauses, but we need a way of capturing this that also captures the fact that disjunction is weaker than conjunction. Borrowing an idea from recent dynamic work on disjunctions of declaratives, Starr argues that we can treat disjunctions as introducing preferences in distinct information relations, which are then unioned.

$$R[\Phi \vee \Psi] = R[\Phi] \cup R[\Psi]$$

In English: the result of updating the context with a disjunction is the union of the two results of updating the conject with each of the disjuncts. An example is helpful here. Take (7):

(7) Buy me a drink or leave me alone! !φ or !ψ

Suppose that we're in a simplified context R={r} in which the only things we're wondering about are whether you'll buy me a drink (p) and whether you'll leave me alone (q), and in which we don't yet have any mutual preferences regarding these alternatives. Starr's theory predicts that (7) will have the following effect on this context: first, it will generate a context R_1 such that R_1 = R[!(you buy me a drink)] = $\{r_0 \cup \{p \cap i_r, i_r\text{–}p \cap i_r\}\}$, next it will generate a context R_2 = R[!(you leave me alone)] = $\{r_0 \cup \{q \cap i_r, i_r\text{–}q \cap i_r\}\}$, and then it will yield the output context $R_+ = R_1 \cup R_2$. In other words: disjunctions create new information relations—one including a preference for each disjunct—that live on alongside the old information relations in the context.

> My proposal…is that disjunctions of imperatives put in play a preference for each disjunct, either of which can be acted on, while both of the preferences put in play by a conjunction must be acted on. To capture this, I model a conversational state not as a simple info relation, but a set of them. For a simple case where $R = \{r\}$, the idea is that a conjunction of imperatives introduces two new preferences into r, while the disjunction of two imperatives will spawn two info relations: one where the preference of the first disjunct has been added to r and one where the second disjunct has been added to r. (Starr *ms*: §3.2)

The justification of this account of disjunction ultimately comes through in the way that Starr's choice function generates a condition on rational behavior from sets of information relations, but the intuitive idea is fairly clear: disjoined imperatives introduce two different options for what to prefer, only one of which we must act in accordance with in order to be rational. We can sum up Starr's predictions about mixed-mood conjunctions and disjunctions in similar terms. In uttering,'buy me a drink and I'll save your seat', I introduce both a preference that you buy me a beer and the information that I'll save your seat. In uttering 'buy me a beer or I'm leaving', I split the context into two distinct options, at least one of which (but not necessarily both) we must behave

in accordance with in order to count as rational: the first now includes a preference that you buy me a beer, and the second now includes the information that I'm leaving.

Starr next shows how to give an account of conditionals that is neutral between declarative, imperative, and interrogative consequents.[51]

Definition 12 (Conditional Semantics)

$$R[(\text{if } \phi)\, \psi] = \left\{ \begin{array}{ll} \{r \cup \{r\}[\phi][\psi] \mid r \in R\} & \text{if } i_{R[\phi][\psi]} = i_{R[\phi]} \\ \varnothing & \text{otherwise} \end{array} \right\}$$

From what I can tell, there is no natural way to elucidate this account clearly in nontechnical English, but we can get a bit closer. As long as updating R with Ψ after updating it with Φ doesn't eliminate any worlds (i.e., as long as R[Φ] is a fixed point for ·[Ψ]), then (if Φ)Ψ updates each information relation r in the context to produce a new information relation r_+, as follows: for each information relation r in R, r+ is the union of r and r^*, where r^* is what you would get by updating r with (Φ and Ψ). Although this account is not straightforward, it does have a significant advantage that isn't enjoyed by any other theory of conditionals on the market—namely, it is defined over consequents of any mood or any grammatical mixture of moods. Starr argues that his account makes the right predictions about how uttering a conditional affects the context. For example, uttering a conditional imperative (if ⊦φ)!ψ effectively adds a conditional preference to the context—a preference for the content of ψ only relative to information at the context that is compatible with the content of φ. Given Starr's further assumptions about how these conditional preferences guide what is rational for agents to do, this is an elegant result. As we've seen, giving a

[51] Starr's theory of conditionals generalizes from a dynamic strict conditional analysis of the kind that has recently been defeded by, e.g., Gillies (2007, 2009, 2010). Because Starr's theory hangs on the viability of a dynamic strict account of conditionals, it is worth noting that this analysis is quite controversial.

similarly general account of conditionals represents a serious hurdle for Portner and Starr, and so Starr's theory bests theirs in this regard.

Starr also offers a novel explanation of our intuition that imperatives are non truth evaluable:

> a sentence type is truth-evaluable just in case the question of truth can arise for its instances. The idea is that a sentence type is truth-evaluable when learning about its truth or falsity of one of its tokens is informative. This amounts to the following definition. A sentence type *T* is truth-evaluable just in case there exists a *w* where an instance of type *T* is true and there exists a *w* where is false. (Starr *ms*: §3.7)

This account relies on Starr's dynamic take on truth: "φ is true in w just in case accepting φ, when you have the information that {*w*}, returns the same information" (ms: §3.7). The result does a nice job of predicting that imperatives (and interrogatives) aren't truth evaluable, with only the following small cost: in defining truth-aptness in this way, Starr gives up the option of saying that some well-formed declarative clauses—such as category mistakes or, for example, 'colorless green ideas sleep furiously'—aren't truth-apt. Assuming that Starr has something else to say about declarative sentences that, as it were, aren't even false, this seems not to be a huge price to pay.

In these and other ways, then, Starr's view does the best job of predicting the empirical data about imperatives. The most serious problems for Starr's account are, I think, foundational in nature. I will raise some of these problems—most of which are also problems for dynamic pragmatic views—in §3.5. A problem that is more specific to the dynamic semantic approach is raised by Charlow, who argues that dynamic semantics embodies a misunderstanding of the division of explanatory labor by requiring that semantic theories include a full theory of belief and preference update.

> …the change a preference state undergoes when it shifts to requiring φ is very different from the simple shift involved in simply adding φ to your stock of preferences. The former requires an account of how prior preferences—ones that conflict with the preference for φ so as to block the global preference state from requiring φ—are revised so that conflict is ameliorated. Such an account would appear to be a matter of proper concern for epistemology, rather than semantics; we should not build a substantive theory of (rational) preference-revision into any part of our semantics.

> I say this as a proponent of a picture of semantics on which a semantic theory for a sentence is a theory about what sort of cognitive instruction that sentence proffers (Section 5.1). However, cognitive instructions, as I understand them, are relatively thin things. An assertion that φ might accurately be described as a cognitive instruction whose force is to propose to get the addressee to accept that φ. When φ conflicts with an agent's prior information, then, in order to fully accept the assertion, the agent cannot simply add the proposition that φ to her stock of beliefs. She must revise her prior beliefs to eliminate the source of conflict. Not just any old revision will do; the revision must be, for example, conservative with respect to those priors. How rational agents revise priors conservatively is the topic of a vast amount of work in formal and informal epistemology, rather than natural language semantics.
>
> That seems to me like the correct division of theoretical labor. The job of natural language semantics is to associate sentences with general, and relatively skeletal, cognitive instructions. These have a plausible claim to being what is meant by utterances in natural language. It is the job of epistemologists (and theorists interested in rational changes in attitudes, more generally) to give a substantive account of how rational agents can comply with these instructions.
> …
>
> In a slogan: semantics furnishes a theory of cognitive directives. Epistemology furnishes a substantive theory of diachronic compliance for cognitive directives. What is left to the semantics? It would seem to be just the idea that imperatives propose to make their prejacents required. (Charlow 2013: §5.6.2)

Charlow's point strikes me as quite forceful: because dynamic semantics identifies sentences' semantic with *functions* that map each state of the context into another state, theories in the genre burden themselves with the task of fully specifying the updates associated with particular sen-

tences, thereby plunging foolhardily into the fraught territory of epistemology and artificial intelligence.[52]

Dynamic accounts have mostly succeeded in ignoring Charlow's problem by focusing on highly idealized conceptions of the nature of conversation. For example, Stalnaker (1978) and Roberts (1996/2012) premise their respective accounts of asserting and questioning on the assumption that the sole goal of conversation is to come to a mutual agreement about how the world is by adding information to the common ground and eliminating worlds from the context set. This idealized model is incompatible with the datum that we sometimes assert propositions that contradict what we've previously presupposed. All of the simplified examples in Starr's essay embody an analogous idealization away from conflicting preferences. Of course, there is nothing wrong with idealizations in general, but this one obscures a serious problem, for it is only once the the idealization away from contradictory information or preferences is lifted that drawing the right boundary between semantics and epistemology becomes crucial and problematic.

---

[52] In response to Charlow's worry, Starr offers the following reply:

> Calculating the ideally rational preferences that should result from combining a new imperative and a body of preferences is indeed a challenging feat that requires much more than grasping a linguistic convention. But it is not this feat I claim competent speakers routinely carry out once they become fluent with imperatives. The preferences that result from the semantic process of updating with an imperative may not even be rational, they needn't reflect all of the consequences an ideally rational agent would draw and they certainly do not guarantee that the agent will choose to bring about the alternative that is best according to them. Words are vessels by which we steer from mental state to mental state. But they are just one of the tools we rely on to set and convey our tack. Rationality—our general ability to make sense of the world and each other—serves as our compass, weaving a coherent route from mere waypoints. The approach to imperatives developed below not only preserves this role for rationality (pragmatics) in language use, it explains some of the rational processes involved in a precise and independently motivated way. (Starr *ms*: §1)

I'm not sure what to make of this reply. Starr is quite explicit that the semantic value of an imperative, together with the prior context, jointly determine the state of the posterior context, and that the context (with the help of a choice function) determines "which alternatives [participants] can be expected, for the purposes of the conversation, to pursue *if they are rational*" (*ms*: §3.2). I don't see how to square that with Starr's comment in the quoted passage that the "preferences that result from the semantic process of updating with an imperative may not even be rational" and "needn't reflect all of the consequences an ideally rational agent would draw".

3.5 From Dynamic Pragmatics to Griceanism

Dynamic theories of imperatives seem to be the most promising on offer. But dynamic theories collapse into Gricean, speech act theoretic accounts when sufficient attention is paid to the theories of communication underlying the two approaches.

3.5.1 Updating, Proposing, and Intending

The first stage in my argument concerns the nature of the essential effects in terms of which dynamic accounts define illocutionary force. I will argue that if these effects can serve as the basis for an account of illocutionary acts, then they must be regarded as the effect that a speaker communicatively intends to have in performing the illocutionary acts, rather than as the effects that performances of illocutionary acts actually have.

In arguing that the essential effect of assertion is to add information to the common ground, Stalnaker goes most of the way toward agreeing with this conclusion. To see this, we can first get clear about Stalnaker's definition by reconstructing it as (CEA).[53]

Characteristic Effect of Assertion (CEA)

A speaker asserts $[\![\vdash\varphi]\!]$ by uttering $\vdash\varphi$ at a stage t of a conversation only if either (a) it updates the context so that that $\mathcal{C}_{t+1} = \mathcal{C}_t \cup \{[\![\vdash\varphi]\!]\}$ and $c_{t+1} = c_t \cap [\![\varphi]\!]$ or (b) her assertion is rejected but would have updated the context in this way had it not been rejected.

---

[53] This reconstruction is based on Stalnaker's assumption that the semantic value $[\![\vdash\varphi]\!]$ of a (declarative) sentence $\vdash\varphi$ is a set of possible worlds. The reconstruction is also based on the idealized assumption that conversations can be broken up into stages, such that a conversation moves from stage t to stage $t+1$ when a participant produces an utterance at t, and such that C_t and c_t are the states of the common ground and the context set at t, respectively.

Stalnaker makes it clear that he does not intend (CEA) as an attempt to define, analyze, or reduce assertion to its characteristic effect, and he gives several reasons for thinking that any such attempt will be hopeless.[54]

> I should emphasize that I do not propose this as a *definition* of assertion, but only as a claim about one effect which assertions have, and are intended to have—an effect that should be a component, or a consequence, of an adequate definition. There are several reasons why one cannot define assertion in terms of this effect alone. One reason is that other speech acts, like making suppositions, have and are intended to have the same effect. A second reason is that there may be various indirect, even nonlinguistic, means of accomplishing the same effect which I would not want to call assertions. A third reason is that the proposed essential effect makes reference to another speech act—the rejection of an assertion, which presumably cannot be explained independently of assertion. (1978: 87)

Stalnaker also considers cases which seem to tell against his characteristic effect's being necessary for assertion—cases in which a speaker produces an assertion but fails to produce its essential effect—including case in which the speaker knows that her assertion will fail to have its essential effect. He rejects these putative counterexamples on the following grounds:

> My suggestion about the essential effect of assertion does not imply that speakers *intend* to succeed in getting the addressee accept the content of the assertion, or that they believe they will, or even might succeed. A person may make an assertion knowing it will be vetoed, a labor negotiator may make a proposal knowing it will be met by a counterproposal, or a poker player may place a bet knowing it will cause all the other players to fold. Such actions will not be pointless, since they all have secondary effects, and there is no reason why achieving the secondary effects cannot be the primary intention of the agent performing the action. The essential effects will still be relevant even when it is a foregone conclusion that the assertion, legisla-

---

[54] Stalnaker himself sometimes dispenses with these subtleties in later articles, making his account seem like a proposal about what makes a speech act an assertion. This has led some commentators to interpret him as offering a constitutive account of assertion. For example, MacFarlane (2011: §3) quotes the following passage when presenting Stalnaker's account of assertion as definitional: "I suggested that an assertion should be understood as a proposal to change the context by adding the content to the information presupposed. This is an account of the force of an assertion, and it respects the traditional distinction between the content and the force of a speech act. Propositional content is represented by a (possibly partial) function from possible worlds to truth-values; assertive force is represented by the way in which any such function is used to change the context that the speaker shares with those to whom he is speaking" (Stalnaker 1999: 10–11). It is tempting to understand Stalnaker's talk of "a proposal to change the context" as an account of assertion as an act performed with the intention to change the context.

> tive act, proposal, or bet will be rejected, since one generally explains why the action has the secondary effects it has partly in terms of the fact that it would have had certain essential effects had it not been rejected. (1978: 87)

The key point here, as I understand it, is that the clause about rejection in (CEA) is crucially important. Stalnaker takes the characteristic effect of assertion to be a necessary condition only for *unrejected* assertions. A speech act is an assertion, according Stalnaker, only if either it has the its essential effect or *would have had* that effect were it not been rejected.

This reading is not entirely satisfactory, since it ignores the fact that the two passages I've quoted seemingly contradict one another over the issue of whether essential effects are intended. In the first passage, Stalnaker says that the characteristic effect is one that assertions "are intended to have", but in the second passage he says that his "suggestion about assertion does not imply that speakers intend to succeed in getting the addressee accept the content of the assertion." We might try to reconcile these remarks by taking Stalnaker's use of 'accept' in the second quoted passage to pick out an attitude stronger than presupposition, but this reading does not square with the rest of what Stalnaker says in the second quoted passage. It is therefore not entirely clear how to read Stalnaker on this issue.

Assuming that we want to endorse the spirit of Stalnaker's account of assertion, what *should* we think about whether essential effects must be intended? Must a speaker intend to add p to the common ground in order to assert p? The answer, I think, is *yes*. Consider Stalnaker's reason, in the second passage, for suggesing otherwise—that we sometimes assert p knowing full well that our assertion will be rejected, and that we typically do this in order to have another, indirect effect. This should call to mind Grice's distinction between *saying* and *making as if to say*, and my distinction between *performing a speech act* and *making as if to* (or *pretending to*) *perform a speech act* (§3.2.4). One way to motivate these distinctions is by means of considerations like the following: (i) for any acts X, Y, and Z, if I am pretending to do X in order to do Y, then I am not really doing X but *merely* pretending to do X, and (ii) if I know that I cannot do X by doing Z, then I

cannot really do X by doing Z, but I can at best (at least in some cases) merely pretend to do X by doing Z. If these principles are right, then a Gricean cannot hold that I assert that *p* by uttering *u* if I know that I won't get my addressee to believe that I believe that *p* by doing so. Stalnaker should say something similar: if I know that I won't have a certain effect *e* on the context by uttering *u*, then, by uttering *u*, all I can really be doing is making as if to produce *e* in order to indirectly have some other effect. We can preserve the idea that adding a proposition to the common ground is an essential effect of assertion by saying that, in cases where the speaker knows they won't get *p* into the common ground, they merely make as if to assert *p*.

With Stalnaker's worrying cases explained away, it makes sense to reconstrue the *essential effect* of assertion as the *essentially intended effect* of assertion: a speech act counts as asserting *p* only if the speaker intends to add *p* to the common ground. An excellent positive reason for reconstruing Stalnaker in this way is that it gives him a way of distinguishing illocutionary acts from perlocutionary acts: to successfully assert that *p* requires only intending to add *p* to the common ground, and to successfully communicate via one's assertion requires only adding the proposition that one had this intention to the common ground; to add *p* to the common ground goes a step further, and into the realm of perlocution. This takes us one step closer to Grice and allows us to drop the notion of rejection from the definition.

I am far from the first person to make these observations. They are most influentially articulated in the form of Thomason's (1990) dynamic account of speaker meaning, which aims to combine the best aspects of Stalnaker to Grice:

> To mean *p* is to intentionally reveal an intention to make *p* asserted through the hearer's recognition of the status of an intention or plan of the speaker's. (Thomason 1990: 345)

At first glance, this doesn't look like a promising start toward an account of the nature of assertion: because it defines communicative intentions as intentions "to make *p* asserted", it looks to be

giving an account of speaker meaning in terms of assertion, and not—as in the Gricean account presented in §3.2—vice versa. But this appearance is superficial. Thomason goes on to say that "to make p asserted is to add *p* to the presumptions of the conversational record" (345). In other words, to make *p* asserted is merely to have the essential effect of assertion. It would be less confusing to rewrite Thomason's definition as follows.

> To mean *p* is to intentionally reveal an intention to [add *p* to the common ground] through the hearer's recognition of the status of an intention or plan of the speaker's.

But this is clearly an account of assertive force, not an account of speaker meaning in general. We need similar definitions, with different essential effects plugged into the brackets, to account for other kinds of illocutionary force. Suppose that we adopt Portner's account of the essential effect of directive force, for example. Plugging it in, we get the following account of directive force:

> To direct *A* to *F* is to intentionally reveal an intention to add F to the A's To-Do List through the A's recognition of the status of an intention or plan of the speaker's.

Undoubtedly, there is plenty of room to argue about the precise formulation of these definitions. For example: we could plug in different essential effects for directive force, and we could flesh out those effects formally in ways that look like Portner's, Charlow's, or Starr's theories. We could also—if we were so inclined—go through a series of counterexamples, remedies, and epicycles parallel to those explored by Griceans in their quest to define speaker meaning. But my goal has been merely the following: to establish that the move from Stalnaker's account to Thomason's, via the addition of intending and intention-recognition to Stalnaker's definition, is warranted, and to establish that Thomason's account would have to be reworded and reconstrued to account for different kinds of illocutionary force.

3.5.2 Collapse into Griceanism

Immediately after articulating the above-quoted definition of speaker meaning, Thomason offers a puzzling comparison of his view to Grice's:

> Notice that this is a considerable departure from Grice's 1957 definition; and it makes plan recognition central to meaning. … It is also much less convoluted than Grice's definition; nevertheless, it gives much the same results as Grice's, over a healthy spread of cases. (1990: 345)

There is much to disagree with here. First, I doubt that Thomason's definition is really "less convoluted" than Grice's. Rather, I suspect that he has merely elided complexity by relying on the notions of *revealing a plan* and *recognizing the status of a plan*; making those notions fully explicit would bring the wording of Thomason's definition closer to Grice's.[55] Second, Thomason implies that the fact that his definition "makes plan recognition central to meaning" sets it apart from Grice's explication of speaker meaning. But speaker meaning is all about intention-recognition for Grice, and plans are just complex intentions or structures of intentions; so Thomason's definition is, in this respect, somewhere between a tweak and a notational variant of Grice's. Third—as I am about to show—there are interesting cases in which Thomason's definition does not give "the same result as Grice's", and these differences reflect well on Grice's account and poorly on Thomason's.

Most broadly, I am skeptical about the extent to which Thomason's definition "is a considerable departure from Grice's 1957 definition". I think it would be better to call it a slight but significant departure, and a departure for the worse. Once dynamic approaches to illocutionary force are framed in terms of Thomasonian plan/intention-recognition, the distance from Thomason's dynamic approach to a Gricean approach is the distance from thinking that our aim in speaking is to affect the shared conversational context to thinking that our aim in speaking is to affect the

---

[55] I suspect that another reason that Thomason's definition may seem simpler is that it hasn't yet been subjected to a barrage of counterexamples at the hands of philosophers like Stephen Schiffer.

psychological states of our addressees. This is the short distance from a bad theory of communication to a good one.

One kind of reason for preferring the Gricean approach is pretheoretic and intuitive: in conversing with another agent, it just *seems* wrong to say that my aim in speaking is not to change their beliefs, but to change the context. Normal speakers don't seem to even have the relevant concept of context, and even if they did, I doubt that it would figure in their intentions in speaking to other people. Yet, dynamic theorists are quite explicit about their commitment to the idea that what constitutes the speaker's half of communication is an intention to affect the conversation's context rather than to affect the addressee's mental state.

> On the present view, it is the common ground, not the speaker, that's "informed", and it is mutual belief behavior, and not knowledge, that's sought. This permits a generalization over rhetorical questions, quiz questions, etc., which are problems for more solipsistic views of information in discourse. (Roberts 1996/2012: 6, *fn7*)

> A communicative intention is "a plan to move the conversation from one state to another, where conversational state is an abstract record of what's 'going on' in the conversation." (Murray and Starr 2013: slide 16)

The intuitive point is that this just seems wrong: in asserting that I am from Canada, what I'm really aiming to do is to get you to believe that (I believe that) I am from Canada. It seems to me dubious to say that my aim is to change the state of the context.

So far, this point is *merely* intuitive, and far from decisive. To develop the intuition into an argument, we need to think more carefully about what dynamic theorists mean by 'context'. Most often, dynamic pragmatic and semantic theorists holds that context is somehow constructed from the mutual attitudes of the participants in a conversation. This view is most naturally applied to the part of the context that represents participants' shared information—often called common ground—and various authors have identified this part of the context with the set of propositions mutually believed (e.g., Bach and Harnish 1979) or mutually known (e.g., Clark and Marshall

1981) by the participants in a conversation.[56] Most influentially, Stalnaker defines the common ground as the set of propositions presupposed by all of the participants in a conversation:

> Roughly speaking, the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of a conversation. A proposition is presupposed if the speaker is disposed to act as if he assumes or believes that the proposition is true, and as if he assumes or believes that his audience assumes or believes that it is true as well. Presuppositions are what is taken by the speaker to be the *common ground* of the participants in a conversation, what is treated as their *common ground* or *mutual knowledge*. The propositions presupposed in the intended sense need not really be common or mutual knowledge; the speaker need not even believe them. He may presuppose any proposition that he finds it convenient to assume for the purpose of the conversation, provided he is prepared to assume that his audience will assume it along with him. (Stalnaker 1978: 84)

Dynamic theorists use the state of a conversation's context to explain what can happen at a given stage of the conversation—which presupposition triggers can be felicitously uttered, for example, or which entities are picked out by anaphora. It therefore makes sense to think of common ground as determined by what speakers present themselves as knowing, rather than by what speakers actually know. The attitudes in which the state of the context is grounded should therefore be *public* attitudes (cf. Thomason 1990: §§6–7) and *as-if* attitudes. Stalnaker's notion of presupposition satisfies both constraints. In arguing that we should understand an act's directive force in terms of its (intended) effect on an action-guiding dimension of context, Portner and Starr make it clear that they take this action-guiding aspect of context to be constituted by public, as-if attitudes as well, albeit action-guiding attitudes such as preferences, plans, or practical judgments.[57]

---

[56] Mutual knowledge is also sometimes called 'common knowledge'. I go into greater detail about these and other related mutual attitudes in Chapter Four.

[57] As I pointed out in §3.4.1, Charlow diverges on this issue. In arguing that an act's force is constituted by the type of effect it characteristically has on a representor, he defines a representor as modeling the attitudes of particular addressees, rather than public, as-if attidudes. If this interpretation is correct, then he already agrees with my point in this subsection.

> We may think of the Common Ground and To-Do List as being the public, or interactional, counterparts of the individual agent's beliefs and desires. That is, as far as the participants in an interaction go, an agent's actions will be judged rational to the extent that, if undertaken in any world compatible with the Common Ground, they would tend to make this world maximally highly ranked according to that agent's To-Do List. (Portner 2004: 8)
>
> The characteristic effects associated with each sentence type are not understood as effects on the agents' private information, questions or preferences. Building on Stalnaker 1973, 2002, they are effects on what is mutually taken for granted for the purposes of the agents' linguistic activities. So the basic idea of a preference semantics for imperatives amounts to the idea that *Dance Frank!* makes Frank's dancing mutually preferable to Frank's not-dancing (for the purposes of the conversation). (Starr *ms*: §1)

Stalnaker spells out his notion of presupposition in terms of speakers' dispositions to behave as if they have certain private attitudes. One serious challenge for Portner and Starr is to give a similar account of how the public, as-if, action-guiding attitudes they posit are grounded in agents' private attitudes (and perhaps dispositions). What individual states of mind do a group of agents have to be in, in order to have a certain mutual preference for the purposes of a conversation? This challenge is non-trivial, and it is far from clear how it can be met. Lewis gives the most thorough treatment of the issue of what determines dimensions of context. In the context of his metaphorical treatment of the context as a conversational scoreboard, he distinguishes each conversational participant's "mental scoreboard" from the public "conversational score", and gives a functionalist account of the relationship between the two notions and the private psychological states of speakers:

> Conversational score is, by definition, whatever the mental scoreboards say it is; but we refrain from trying to say just what the conversationalists' mental scoreboards are. We assume that some or other mental representations are present that play the role of a scoreboard, in the following sense: what they register depends on the history of the conversation in the way that score should according to the rules. The rules specifying the kinematics of score thereby specify the role of a scoreboard; the scoreboard is whatever best fills this role; and the score is whatever the scoreboard registers. (1979a: 239)

But even if we assume that *some* psychological state or other can play the role of the private realizer of the public, as-if, action-guiding dimension of the context (which is by no means obvious),

this would not be enough for Portner and Starr. They also need the psychological state in question to be fundamentally non-cognitive. And this is a tough aim to achieve. Following Stalnaker's model, the most natural way of glossing the idea of public, as-if preferences (to take Starr's account as an example), would be as those preferences that each participant "takes for granted as part of the background of a conversation", where a speaker takes a preference for granted in this way if he is "disposed to act as if he assumes or believes that [the preference exists], and as if he assumes or believes that his audience assumes or believes that it [exists] as well." But this gloss is not available to Starr, because it would make public preferences a special case of presuppositions, and that, in turn, would make directive force a special case of assertive force. So, Portner and Starr not only owe us an explanation of how their proposed action-guiding dimensions of context are grounded in agents' private psychological states; these psychological states also mustn't boil down to agents' presuppositions about action-guiding states.

Supposing that such an account can be given, further problems lurk. Now that we have a better idea of what it would be for a speaker to intend to update the context, how plausible is it that this is what performing a speech act consists in? For example: does performing an assertion require intending to produce mutual presuppositions in all of the participants in a conversation, or does it require intending to produce a belief in a particular addressee? There are excellent reasons to prefer the latter answer to the former. The former answer would imply that, in order to assert *p*, I have to intend to bring it about that I will presuppose that my addressee presupposes *p*. But this simply isn't the case. To take just one obvious and straightforward counterexample: I could assert that *p* in a suicide note. Although this would presumably involve intending that my addressee believe (that I believe) *p*, it would not involve intending that there be any point at which my addressee and I mutually presuppose *p*. Indeed, if I'm really serious about the note, then my beliefs rule out the possibility that such a mutual presupposition will ever obtain. The same goes for directive speech acts: I can direct my executor to do various things via my will, in which case I in-

tend for the effect of my speech act to obtain only after I won't be around to participate in the mutual attitudes that would constitute the public, as-if, action-guiding dimension of context. This gives us good reason, I think, to deny that the intended effects of speech acts are public.

There are also problems with the idea that the intended effects are as-if attitudes rather than genuine attitudes. This is particularly clear in the case of directives: what I ultimately care about in directing you to do something is that you do it; producing an intention (or other action-guiding state) is ultimately a means to that end. But whereas there are clear connections between agents' private action-guiding states and their actions, the connection between as-if action-guiding states and action is far less clear. Starr, for example, uses his notion of public, as-if preferences to explain which actions are preferable "for the purposes of the conversation". But, even if we can make sense of something's being preferred for the purposes of a conversation, what is the connection between this notion and speakers' actions? One possible answer might be Portner's: "as far as the participants in an interaction go, an agent's actions will be judged rational" only if they bear some relation to the action-guiding dimension of the context. But contexts, like conversations, are fleeting, and we often direct our addressees to do something long after the conversation will come to an end. In this case, what we hold our addressee responsible for is actually planning (or preferring, etc.) to do what we directed them to do, and for following through on that plan (preference). It's hard to see what bearing the mutual, as-if plans (or preferences, etc.) of a past conversation could have at that point.

In defending his conception of assertion in terms of intended effects on the common ground ("conversational record"), Thomason attempts to address this disconnect between our public, as-if attitudes and our actions:

> In saying that assertion does not aim intrinsically at belief, I have driven a wedge between communication and belief. Of course, this also divorces communication from action—and perhaps one motivation of the belief-based approaches was to connect the two. On the view I'm advocating, what is the connection between assertion and belief? Well, the conversational

> record is like a courtroom record; it is evidence, but we may not choose to believe all of it. What we choose to believe from what we are told is a special case (but a maximally complex special case) of what to make of evidence. (Thomason 1990: 340)

If the contents of the common ground are evidence, what are they evidence of? They might count as evidence of the other speakers' private attitudes—albeit defeasible evidence, given that presupposing *p* is neither necessary nor sufficient for believing *p*. In fact, the question of how public, as-if attitudes can count as evidence about agents' private attitudes merely recapitulates the problem of how agents' public, as-if attitudes are connected to their actions. Boiled down to its essentials, the problem is this: what reliable connection is there between agents' as-if attitudes and their real attitudes, such that their as-if attitudes can constitute either partial reasons for action or evidence of their real attitudes? This is an open and unanswered question—and one that becomes more confusing in the case of public, as-if action-guiding states, since plans and preferences just don't seem to be the sorts of things that could count as evidence anyway.

In §3.5.1, I argued that the essential effects of illocutionary acts must be reconceived as essentially intended effects. In this subsection, I have argued that these essentially intended effects must be effects that we intend to have on our addressees, and not on contexts constituted by public, as-if attitudes. Together, these two arguments effect the collapse of dynamic accounts of illocutionary force into Gricean accounts.

3.5.3 Objections and Replies

How might dynamic theorists reply to my arguments? One way would be to point out cases in which we do seem to speak with the intention of updating a public, or as-if representation rather than the private attitudes of an addressee. For example, Roberts argues that dynamic views allow us to avoid a common kind of putative counterexample to Gricean accounts:

> On the present view, it is the common ground, not the speaker, that's "informed," and it is mutual-belief behavior, and not knowledge, that's sought. This permits a generalization over

> rhetorical questions, quiz questions, etc., which are problems for more solipsistic views of information in discourse. (Roberts 1996/2012: 6, *fn*.7)

Rhetorical questions and quiz questions are sometimes seen as problems for Gricean speech act theory because they involve uttering interrogative sentences without the usual sorts of intentions: in asking a rhetorical question, one doesn't intend one's addressee to answer, and, in asking a quiz question, one doesn't intend them to answer on the basis of a belief that the speaker desires to know the answer. This problem is similar to the one about audienceless utterances, advice, and similar cases considered in §3.2.4, and my response is similar as well. Rhetorical questions are obvious cases of making as if to ask a question in order to indirectly perform an assertive speech act. Quiz questions, on the other hand, may be perfectly literal and direct acts of questioning—attempts to get one's addressee to assert an answer—but they differ from typical questions because of the grounds on which the questioner expects her addressee to intend to answer. In asking a prototypical question, we intend that our addressee form an intention to answer partly because she recognizes that we desire to know the answer. In asking a quiz question, we intend that our addressee form an intention to answer partly because she recognizes that we want her to demonstrate knowledge of the answer. This account draws on the theory of illocutionary underspecification outlined in §3.2.1, and mirrors my explanation of how acts of commanding, requesting, and advising can all be performed literally and directly by uttering an imperative.

Thomason sketches another case that seems better handled by his theory than by Grice's—that of "a congressman interviewing a witness at a hearing. What the congressman has to assume in order to keep track of the conversation and what he believes are two different things" (1990: 341). Moreover, as with other cases in which the participants in a conversation know that what matters is what is preserved for posterity, rather than what their addressee comes away thinking, it is plausible that a witness being interviewed by a congressman would intend her speech acts primarily to change the public record in a very real sense. But this just shows how different testifying

before congress is from regular conversation. As opposed from the kind of "testimony" that interests epistemologists, congressional testimony is a kind of conventional speech act in Bach and Harnish's (1978) sense, which is to say that its performance is partly grounded in extra-linguistic social institutions or conventions, such as, for example, the congressional record. Regular conversation differs from cogressional testimony at least this regard, and so it would be a bad idea to use this fact to shed light on communication or linguistic meaning in general.

The most interesting argument against Griceanism and in favor of a dynamic approach rests on the putative explanatory power of dynamic approaches to context. By positing shared and evolving contexts that are constituted by speakers' mutual, as-if attiudes, dynamic theorists have attempted to explain and predict a wide range of linguistic phenomena. For example, Stalnaker's notion of common ground (and closely related notions), has been used to give semantic accounts of many different kinds of context-sensitive expressions.[58] A dynamic notion of context has often appealed to in explanations of the felicity of uttering presupposition triggers.[59] Portner (2007) uses facts about the To-Do List to make predictions about the felicity conditions and semantic values of deontic modals. And Roberts (1996/2012) uses features of the Questions Under Discussion to make predictions about which contributions to a conversation will be relevant. If we adopt a Gricean theory of speech acts rather than a dynamic one, do we have to give up these and similar dynamic explanations of context-sensitive linguistic phenomena?

The short answer is no. Context plays two roles in dynamic theories: it is what is updated by speech acts, and it determines the properties of context-sensitive expressions. What my arguments in this section have shown is that we should abandon the idea that context plays the first

---

[58] Expressions that have been given semantic treatments that appeal to dynamic notions of context include indicative conditionals (e.g., Stalnaker 1975), definite descriptions (e.g., Heim 1981; Lewis 1979a; Roberts 2003, 2009), demonstratives (e.g., Roberts 2002), anaophoric pronouns (e.g., Heim 1981; Kamp 1981, 1995), proper names (Cumming 2008), and so on.

[59] E.g., Beaver (2001), Chierchia (1995), Heim (1983). For an overview, see Beaver and Geurts (2011).

role. It does not follow that there is *no* role for context or mutual, as-if attitudes in a theory of language use, and it is consistent with everything I've said here that these attitudes play some or all of the roles alluded to in the last paragraph. All that my arguments here commit me to is the conclusion that updating the context is not the constitutive aim of performing a speech act; we can still say that context update is a common kind of downstream perlocutionary effect of performing a speech act that is normally expected by the speaker. On my view, the problem with dynamic theories is that they collapse the distinction between these downstream perlocutionary effects and intended effects that are actually constitutive of speech acts. And it may be that semantic theories should pick up on how these downstream effects license future speech acts. (That said, I must admit that I am skeptical about many appeals to context in semantics and pragmatics, including most of the appeals alluded to above; I gave some reasons for this skepticism in Chapter Two.)

3.6 Griceanism and Embedded Imperatives

Just one problem for my Gricean account of imperatives still remains. Assuming that we cannot maintain that mood always takes wide scope (which it seems we can't—see §3.1.3 and §3.2.4), how can the Gricean account for genuine cases of imperative embedded under logical operators? Consider again our stock of examples:

(6) Buy me a drink and make it a stiff one!	!φ and !ψ
(7) Buy me a drink or leave me alone!	!φ or !ψ
(8) If the bartender comes back, buy me a drink.	(if ⊢φ)(!ψ)
(9) a. Give me all your money or the puppy gets it.	!φ or ⊢ψ
b. You mow the lawn or I'll clean the garage	
(10) a. Say that again and I'll scream.	!φ and ⊢ψ
b. Stay here and I'll go to the store.	

(11) If he comes back, buy me a drink and I'll pay you back. (if ⊦φ)(!ψ and ⊦σ)

In constructing an act-theoretic semantic account for sentences like these, we can draw on some of the successes of dynamic accounts. First, notice that many of these sentences aren't used to perform illocutionary acts of any of the usual types. Whenever indicatives and imperatives are mixed under a connective, for example, the result is neither an assertion or a directive but a speech act of a different, sui generis kind. For example, a hypothetical imperetive sentence (if ⊦φ)(!ψ) is not used to perform an assertive or a directive act, but some third kind of act that is systematically related to both. We could call acts of this kind *hypothetical directives*. Intuitively, hypothetical directives are constituted by a communicative intention to produce a contingency plan in an addressee. For example: performing a literal and direct speech act with (8) intuitively involves intending one's addressee to enter a mental state such that if she comes to believe that the bartended has come back (together with certain other beliefs, such as the belief that I am still thirsty), she will form an intention to buy me a drink.

We need to enrich the Gricean theory of speech acts in order to account for these hybrid speech acts. The natural way of doing this would be to match up hybrid sentences with hybrid illocutionary act-types via the different kinds of effects that a speaker who utters such a hybrid sentence thereby intends to have on an addressee. Because dynamic theories understand the semantic values of multiple-clause sentences directly in terms of their essential effects on context, dynamic theorists of mood have already developed many of the theoretical tools that we need to do this.

Charlow's account can provide some inspiration. Recall that he winds up identifying the semantic values of sentences with properties of representors, where a representor $\langle S_a, \Lambda_a \rangle$ is a model-theoretic representation of an agent *a*'s information S_a and the agent's plan Λ_a. Charlow spells out the semantic values of declarative and imperative clauses, as well as several of their pos-

sible combinations, by giving a recursive specification of the conditions under which a sentence φ holds at a representor ⟨S, Λ⟩ (represented as ⌜⟨S, Λ⟩ ⊨ φ⌝).

⟨S, Λ⟩ ⊨ ⊦φ iff ∀w ∈ S : φ is true at w

⟨S, Λ⟩ ⊨ ¬⊦φ iff ∀w ∈ S : φ is not true at w

⟨S, Λ⟩ ⊨ !φ only if ∀w ∈ $min_{\Lambda a}$: φ is true at w

⟨S, Λ⟩ ⊨ (if φ)(!ψ) only if ∀w ∈ $min_{\Lambda a}$ ∩ {w' : φ is true at w'} : ψ is true at w

⟨S, Λ⟩ ⊨ (Φ∧Ψ) iff ⟨S, Λ⟩ ⊨ Φ and ⟨S, Λ⟩ ⊨ Ψ

A sentence holds at a representor just in case the representor has the property that the sentence denotes, and representors are just abstract representations of agents' beliefs and plans (intentions). Moreover, Charlow thinks of speech acts as proposals for updating an addressee's mental state. So, Charlow's account is not far away from a Gricean, act-theoretic one: the distance is roughly that which I covered in §3.5.1. In other words, we can bring Charlow's insights into a speech act-theoretic semantics of the sort that I outlined in §3.2 by construing the properties Charlow identifies with the denotations of clauses as the effects that a speaker would have to communicatively intend to have on an addressee in order to token the clauses' semantic values.

The implementation of this kind of idea within the framework outlined in Chapter Two and §3.2 is relatively straightforward, but leaves open many options. One option would be to model speakers' overall psychological states using model-theoretic structures of the same kind that dynamic theorists use model the context, and to specify the communicatively intended effect associated with each clause type in terms of the communicatively intended effect associated a given sentence radical. This could proceed in a way that is precisely structurally analogous to the any of the ways in which dynamic semanticists specify the context-change potentials of clauses in terms of the context-change potentials of sentence radicals. The fact that this way of proceeding is possible

makes the task of giving a speech-act-theoretic semantics for mood as trivial as shopping around to the different dynamic views and combining their best qualities.

I will take a different approach here—one that draws on the same notion of rational requirement that I used in §3.2.3 to explain intuitions that purport to be about imperative consequence. We begin by giving compositional principles for each of the three major clause-types:

Declarative Composition [$_S \vdash \varphi$]

S's semantic value is a type that is tokenable by an act constituted by the communicative intention to produce a cognitive state whose content is a particular proposition that can be expressed by tokening φ's semantic value.

Imperative Composition [$_S !\varphi$]

S's semantic value is a type that is tokenable by an act constituted by the communicative intention to produce an intention whose content is any proposition that can be expressed by tokening φ's semantic value.

Next we give generalized act-theoretic compositional principles for each of the sentential connectives. The simplest case is conjunction, which I'll use to illustrate the overall strategy

(∧C) *Conjunction Composition* [$_S$ S_1 and S_2]

S's semantic value is a type that is tokenable by any speech acts constituted by communicatively intending to produce a mental state *e* in one's addressee, where *e* meets the following condition: for some mental states e_1 and e_2 that one can communicatively intend to produce in an addressee by uttering S_1 and S_2, respectively, to be in *e* is to be in a state such that one is rationally committed to being in both e_1 and e_2 and if one is in both e_1 and e_2 then one is rationally committed to being in *e*.

The key notion here is that of rational commitment, which I appealed to in §3.2.2 in order to explain the intuition that some arguments that involve imperatives can be valid. If the mental states we're considering are just beliefs, then logic places strict limits on our rational commitments: we're rationally committed to believing whichever propositions are entailed by the propositions we believe and we're rationally prohibited from believing whichever propositions are logically inconsistent with the propositions we believe.[60] From this conception of rational commitment, it follows that if we plug in declarative clauses for S_1 and S_2 in (∧C), the semantic value of S will be a type that we can token by asserting a conjunctive proposition each of whose conjuncts could have been asserted by uttering S_1 and S_2, respectively.[61]

What if we plug in an imperative for S_1 and a declarative for S_2, as in (10)?

(10) Stay here and I'll go to the store! !φ and ⊢ψ

In this case, (∧C) predicts that one way to token (10)'s semantic value would be by performing the sort of hybrid speech act that would be constituted by communicatively intending that my addressee both intend to stay here and believe that I am going to the store. This is the correct prediction. Moreover, since norms of rational commitment govern intentions and combinations of

[60] It should be clear that these rules place only hard upper and lower bounds on rational commitment, and we should expect that, for example, probability theory will spell out more finely-grained norms.

[61] (∧C) leaves other interesting possibilities open, even when we're just dealing with declaratives. For example, since I could assert that Frank eats with 'Frank eats' and merely suggest that Linda drinks with 'Linda drinks', (∧C) leaves open the possibility that I could utter 'Frank eats and Linda drinks' with the communicative intention to produce an effect in an addressee that would combine *a belief that Frank eats* and *an act of considering whether Linda drinks*. The resulting speech act would neither be an assertion nor a suggestion, but a hybrid of the two. Can we really do this? I think so: all we need to get the right reading is a combination of the right circumstances and, perhaps, a bit of hesitant intonation or an uncertain facial expression while uttering the second conjunct. If I am right about this possibility, then it constitutes a reason for preferring my view over any view that takes assertive force to scope over conjunction.

beliefs, (∧C) gives us the resources to explain the intuition that certain conjunctions that involve imperatives are either infelicitous or inconsistent (see §3.2.2).

Things play out in more or less the same way in the case of disjunction:

(∨C) *Disjunction Composition* [$_S$ S_1 or S_2]

S's semantic value is a type that is tokenable by any speech acts constituted by communicatively intending to produce a mental state *e* in one's addressee, where *e* meets the following condition: for some mental states e_1 and e_2 that one can communicatively intend to produce in an addressee by uttering S_1 and S_2, respectively, to be in *e* is to be in a state such that if one is also rationally prohibited from being in e_1, then one is rationally committed to being in e_2, and vice versa.

Charlow offers no composition principle for disjunction, and Starr argues that the obvious way for Charlow to derive one—i.e., saying that $[\![S_1 \text{ or } S_2]\!] = [\![\neg(\neg S_1 \text{ and } \neg S_2)]\!]$—won't work because imperatives don't scope under negation (see §3.4.1). The principle (∨C) avoids this problem. In essence, (∨C) gives a functional specification of the kind of mental state that we can intend to produce by uttering a disjunction. Any state we can intend to produce with a disjunction of two declaratives can be identified with a belief in the disjunction of two propositions. Any state we can intend to produce in uttering a disjunction of two imperatives can be thought of as an intention to do at least one of two things. And any state we can intend to produce by uttering a mixed declarative–imperative disjunction will be a hybrid cognitive–practical dilemma: a state that rationally compells one to intend one thing if one disbelieves a second thing, and to believe the second thing if one does not intend the first.

As usual, things are most complicated with conditionals—particularly because we need an account that generalizes across conditionals with both declarative and imperative consequents, as well as combinations of the two. Rather than attempt to account for all conditionals at once, I will

offer an account of strict conditionals first, in the hope that I will be able to generalize the account to indicative and counterfactual (declarative) conditionals later. The following principle is a rough first shot at what we need.

(if-C) Conditional Composition [$_S$ (if S_1)(S_2)]

S's semantic value is a type that is tokenable by a speech act constituted by communicatively intending to produce a mental state *e* in one's addressee, where *e* meets the following condition: for some belief e_1 and some mental state e_2 that one can communicatively intend to produce in an addressee by uttering S_1 and S_2, respectively, to be in *e* is to be in a state such that also being in e_1 rationally commits one to be in e_2 as well.

If S_2 is a declarative clause and e_2 is the belief that *p*, then (if-C) amounts to a functional specification of a mental state of believing a conditional proposition, which we can think of as the proposition that *p* is entailed by the contents of e_1. If S2 is an imperative clause and e_2 is the intention to bring about *p*, then (if-C) gives a functional specification of a different kind of mental state—namely, a contingency plan to intend that *p* should one come to have the belief e_1. Again, this is roughly the right result, though it may require further fine-tuning, and though it will have to be extended or refined in order to deal with indicative and counterfactual conditionals. But, despite its sketchiness, the account also holds the promise of capturing the meanings of more complex conditionals with mixed-mood consequents, such as (11):

(11) If he comes back, buy me a drink and I'll pay you back. (if ⊢φ)(!ψ and ⊢σ)

Applying both (∧C) and (if-C) to (11), we get a functional specification of a mental state *e* such that (roughly) if one is in *e* and one also believes that [an intended individual] has come back [from an intended place to an intended place], then one is rationally committed to intend to buy [the speaker of (11)] a drink and to believe that [the speaker of (11)] will pay one back. Notice

that this is not the functional specification of either a belief or an intention, but a complex and hypothetical hybrid of the two. This is plausibly the right result.

These compositional principles differ from the ones outlined in Chapter Two because they tie the semantics of expressions directly to the types of communicative intentions that a speaker can have in uttering the expressions, rather than taking a detour through the type of speech act they can perform in uttering a sentence. In simple cases, this is avoidable by introducing vocabulary for new types of hybrid illocutionary acts. For example, if we know that a *hypothetical directive to p if q* is constituted by an intention to bring about a contingency plan to bring about *p* only if one believes *q*, then we can give a compositional principle for conditional imperatives in the following way:

> *Conditional Directive Composition* $[_S\ (\text{if} \vdash \varphi)(!\psi)]$
>
> S's semantic value is a type that is tokenable by a hypothetical directive that *p* if *q*, where *p* is a proposition that could be expressed by tokening ψ's semantic value and *q* is a proposition that could be expressed by tokening φ's semantic value.

But this strategy doesn't generalize, because mixed-mood clauses can be embedded within one another, and so it isn't enough to posit hypothetical directives, conjoined and disjoined assertive–directive hybrids, and so on; we would also have to posit conditional assertive–directive hybrids, assertive–(directive–assertive) hybrids, and so on. This means that our composition rules for expressions with recursive potential that scope over mood can't be given in terms of a finite list of speech act types; instead, they must recursively specify speech act types by recursively specifying their underlying communicative intentions themselves. This is no problem: on a Gricean account, each new type of communicative intention thus specified will constitute a new type of illocutionary act, even if we don't have a name ready for it. The composition rules given above are general enough to accomplish this.

It's worth noticing a couple of things about this solution to the problem of embedded non-declaratives. First, it takes inspiration from the playbook of the dynamic theorists like Charlow and Starr. In fact, I think it improves on the current menu of dynamic options, particularly when it comes to disjunction. But in learning from dynamic pragmatics and semantics, I have avoided taking on board the misguided dynamic theory of communication. Second, this account of embedded non-declaratives depends intimately on the Gricean foundations of my speech act-theoretic semantics. Not just any theory of illocutionary acts will do. This is interesting because it shows how the nuts and bolts of compositional semantics are intimately tied to foundational issues about the nature of communication.

3.7 Conclusion and Mile-Marking

In this chapter I have argued that the best account of the meaning of mood is a speech act-theoretic semantics built on a Gricean pragmatic foundation. The semantic value of a clause is a type of speech act, such that this type is individuated in terms of the types of communicative intentions that a speaker can have in performing a literal and direct speech act with the clause. All of the other major semantic accounts of mood are flawed in some way or other, and the best of the lot—dynamic pragmatic and semantic theories—collapse into the Gricean account once they are combined with a good (i.e., Gricean) theory of communication.

I have not attempted to show how this account extends to interrogative clauses here, but I believe that such an extension is both possible and warranted. Indeed, because the best contemporary semantic accounts of interrogatives are framed in dynamic semantics (e.g. Roberts 1996/2012), the arguments I gave in §3.5 should apply to them as well. Because of the vastness of the literature on interrogatives—it is even larger than the literature on imperatives—I don't have space to address it adequately here.

My conclusions here lend further support to the overall goal of the dissertation, which is to establish that meaning is a relation borne by utterance-types to speech act-types, and that the semantic values of expressions for the purposes of compositional semantics are the types of speech acts to which they're thus related. Following on my argument from semantic underspecification in the last chapter, I take this chapter to constitute a second broad and persuasive argument for this view. But much remains to be done. In Chapter Four, I will investigate the nature of the relation that a type of speech act must bear to a type of utterance in order to count as its semantic value. In Chapter Five I will investigate the ways in which this conception of semantics should both shape and be shaped by our understanding of the syntax–semantics interface.

CHAPTER FOUR: METASEMANTICS WITHOUT LINGUISTIC CONVENTION

In virtue of what do expressions have meaning? Most contemporary answers to this question are forms of *conventionalism*, according to which an expression has meaning for a group of speakers in virtue of linguistic conventions in which the speakers participate. One of my goals in this chapter is to show that conventionalism is false—to show, in other words, that the facts that constitute linguistic meaning aren't facts about convention. Although a theory of convention might have some role to play in *causal* explanations of how linguistic meaning comes about and persists, convention is not necessary for linguistic meaning and can't feature in the kind of *constitutive* explanations aimed at by metasemantics.

My objections to conventionalism are based on the idea that the adequacy of a theory of linguistic meaning should be measured, at least in part, by its ability to explain the efficiency of linguistic communication. In §4.1, I draw out the consequences of the methodological framework introduced in §1.3 for metasemantics. I then deploy this premise in three arguments. The first two, which appear in §4.3 and §4.4, target *Lewisian conventionalism*, which is named after David Lewis's (1969, 1975a) influential theory of convention, and which I'll explore in §4.2. My third argument, which appears in §4.5, is more general: it aims to show that all forms of conventionalism depend on the conflation of causal and constitutive forms of explanation.

The debate into which I am wading takes place in the context of a genre of metasemantic theory that was inaugurated by Grice with the following brief proposal:

> "*x* means$_{NN}$ (timeless) that so-and-so" might as a first shot be equated with some statement or disjunction of statements about what "people" (vague) intend (with qualifications about "recognition") to effect by *x*. (1957: 220)

This explication is almost comically rife with tentativeness and hedging. Nonetheless, it should be regarded as establishing the framework for ensuing debates about metasemantics. Lewis, Stephen

Schiffer (1972, 1993, 2006), and most other conventionalists hold, like Grice, that utterance-types (including linguistic expressions such as words, phrases, and sentences) are meaningful for a community of speakers in virtue of facts about community members' mental states. These conventionalist theories, as well as Grice's own mature metasemantic views, should be thought of as attempts to improve on Grice's early sketch by better articulating the natures of these psychological states. On this broad construal of the debate, Grice and the conventionalists agree about what they're up to. The difference is that Grice, in his later (1968) work, formulates the skeleton of a metasemantic theory from which all convention-like notions can be expunged without losing the core of his view. In this regard, I take Grice's mature approach to be more promising, and I will use it as a jumping-off point for the positive view that I develop in §4.6.

The Grice-inspired metasemantic theory that I defend also supports the broader goals of this dissertation, because it is a feature of the theory that meaning is a ternary relation borne by utterance types and speech communities to types of speech acts. The types of speech act thus related to an utterance is its meaning for a given community—its semantic value for the purposes of compositional semantics. Chapter Five will be concerned with the implications of this view for compositional semantics and the syntax–semantics interface.

4.1 Metasemantic Methodology

What is the goal of metasemantics, and how should it be pursued? The metasemantic theories of Grice, Lewis, and Schiffer all take the form of statements giving necessary and sufficient conditions for expressions' meaningfulness. I will refer to these statements as 'explications'. I take the goal of explication to be the articulation of a constitutive (non-causal) metaphysical relation stronger than supervenience—either *metaphysical reduction* or *grounding*.

This conception of the goals of metasemantics probably isn't compatible with the intentions behind every instance of '=*df*' and '*iff*' in the work of Grice, Lewis, and Schiffer. In particular,

Grice's original aim seems to have been conceptual analysis rather than metaphysical reduction. I therefore adopt an anachronistic interpretation of Grice. I do so for two reasons. First, I am more optimistic about the possibility of successful explications when they are read as metaphysical claims. Second, this anachronistic reading creates dialectical friction between Grice, on one hand, and Lewis and Schiffer, on the other, since the latter philosophers' goals were clearly metaphysical rather than conceptual.[1]

Much of the vocabulary deployed by metasemanticists in describing their project—including my own use of the phrase 'in virtue of'—is now typically thought by metaphysicians to express the grounding relation rather than metaphysical reduction.[2] It is therefore plausible that what metasemantics should be concerned with is the question of what grounds linguistic meaning, rather than the question of what meaning reduces to. Grounding is often understood as a weaker metaphysical relation than reduction, in the sense that a fact F1 may ground a fact F2 even if F1 is merely sufficient, but not necessary, for F2.[3] If metasemantics should aim at grounding, rather than reducing, facts about linguistic meaning, then we mustn't rule out the possibility that no single condition is both necessary and sufficient for meaning. Since the question of whether metasemantics should aim at reduction or grounding likely can't be worked out in advance of substantial first-order metasemantic theorizing, I will consider both options.

---

[1] The idea that metasemantics aims at metaphysical reduction is most explicit in the work of Schiffer (1981, 1982, 1987).

[2] See also Neale's (2004, 2005, forthcoming a) use of phrases like 'is constitutive of', 'makes it the case that', and 'in virtue of', in discussing both Grice's and Schiffer's metasemantics.

[3] I follow Rosen's (2010) discussion of grounding. In particular, I will follow Rosen in holding that if the fact that *p* reduces to the fact that *q*, then the fact that *p* is grounded in the fact that *q* (2010: 123). On this view, grounding and reduction are both stronger than supervenience, since a fact F1 may supervene on a fact F2 despite not being grounded in F2—for example, where F1 is the fact that 1+1=2 and F2 is the fact that Stephen Schiffer is not Claudia Schiffer (*cf*. Rosen 2010: 113–4; and Schaffer 2009: 364). For other influential discussions of grounding, which diverge from Rosen in various ways, see Fine (2001, 2012) and Schaffer (2009).

How can we know whether a metasemantic theory succeeds? The answer is tied to the fact that meaning is a theoretical posit. Like our knowedge of facts about bosons and unconscious mental states, our knowledge of facts about meaning must derive from the role that we give to the concept of meaning in a theory that explains something else. In the course of reconstructing Schiffer's (1981: 68) metasemantic project, Neale makes essentially this point by saying that explications in metasemantics

> earn their keep not by yielding results that conform precisely to native speaker's intuitions—though these are where we start, of course—but in virtue of their interlocking roles in an explanatory theory of the "semantic" properties of marks and sounds… (forthcoming a)

I take a similar methodological outlook to be embodied in Lewis's dictum:

> In order to say what a meaning is, we may first ask what a meaning does, and then find something that does that. (1970: 195)

These remarks suggest a way of judging the merits of metasemantic proposals in keeping with the methodological outlook that I outlined in §1.3. The goal of a metasemantic theory is to give an account of the facts that constitute linguistic meaning—facts to which meaning reduces, or facts that ground it. The success of a metasemantic theory is to be measured, at least in part, on the basis of how well it explains the role of meaning in efficient communication. In particular: an adequate metasemantic theory must predict that the utterance types involved in efficient linguistic communication are meaningful. The key thrust of my arguments in what follows is that conventionalist metasemantic theories fail to meet this constraint.

4.2 Lewisian Conventionalism

By 'conventionalism', I mean any metasemantic theory that gives convention an explanatory role to play. I will eventually object to all forms of conventionalism, but I will begin by using Lewisian

conventionalism as my primary case study. I do this because Lewisian conventionalism is the best-worked-out and most popular exemplar of the genre.

Lewisian conventionalism can be summed up by the innocent-sounding statement that the meaning of an expression is constituted by the group's convention to use a language. For Lewisians, however, 'language' and 'convention' are technical terms. A language, according to Lewis, is

> a function, a set of ordered pairs of strings and meanings. The entities in the domain of the function are certain finite sequences of types of vocal sounds, or of types of inscribable marks; if σ is in the domain of a language $\mathcal{L}$, let us call σ a *sentence* of $\mathcal{L}$. The entities in the range of the function are meanings: if σ is a sentence of $\mathcal{L}$, let us call $\mathcal{L}(\sigma)$ the *meaning* of σ in $\mathcal{L}$. (1975a: 163)[4]

A convention, according to Lewis, is a regularity R in some community G that meets the following conditions:

(i) everyone in G believes that everyone else in G acts in accordance with R;

(ii) there is a general preference among individuals in G for conformity to R;

(iii) the belief that R exists gives everyone in G a reason (together with their other goals) to continue conforming to R;

(iv) there is at least one other possible regularity R′ that would serve the purposes of the members of G just as well if they were to conform to it instead; and

(v) it is *mutual knowledge* among members of G that (i)–(iv) obtain.

These two technical concepts, which Lewis sees as precisifications of our colloquial notions, are the building blocks of Lewisian conventionalism. Lewis tells us that a sentence S has a proposition

---

[4] Lewis goes on to suggest that sentence meanings can be thought of as sets of possible worlds (i.e., propositions), but later adjusts that idea to account for various forms of context-sensitivity (1980). I will mostly speak of expressions *being meaningful* or *having meaning*, where this should be understood as neutral on the issue of which entities, if any, should be identified with expressions' meanings. For a helpful discussion of this issue, see Schiffer (2003: ch. 3).

P as its meaning for a group G just in case, for some language L that maps S to P, there prevails in G a convention to utter sentences only if L maps them to true propositions (and to expect others to do the same).[5]

There are several well-known objections to the particulars of Lewis's own theory, my plan is to whittle away the aspects of it that give rise to the best-known problems, revealing a core that is as plausible as possible and that is shared by all adherents to Lewis's overall metasemantic strategy. I will then argue in §4.5 that all theories sharing this common Lewisian core are false.[6]

One way to object to Lewis's theory is to argue that one or more of the clauses of his definition of convention fails to obtain in cases of human language use. For example, Burge (1975) imagines isolated speech communities whose religious beliefs dictate that their language is the only possible one, or that they ought to continue to speak their language even if they could no longer communicate with each other. These communities would be counterexamples to Lewis's condition (v) as it applies to his conditions (iii) and (iv), and perhaps also to (iii) and (iv) themselves. In order to sidestep Burge's objections and others like them, I will abstract away from both the particular kind of regularity that forms the basis of linguistic conventions and from the particular clauses (i)–(iv) of Lewis's definition of convention. Given this first way of broadening the view, a metasemantic theory counts as a form of Lewisian Conventionalism if it requires for the meaningfulness of an expression *e* in a group G that the members of G have mutual knowledge that *p*, where *p* is *some* proposition about how *e* is used by the members of G. This condition on a Lewisian theory is met by Lewis's own theory, which requires that language users have mutual knowledge of con-

---

[5] This statement most clearly echoes Lewis's (1975) theory. In place of Lewis's convention of "truthfulness and trust", Schiffer argues that the relevant convention is one of uttering a sentence S only if one speaker-means that *p*, where *p* = L(S) (1972, 1993, 2006). Williamson suggests that the relevant regularity is that one utter S only if one knows that that *p*, where *p* = L(S) (2000). None of my criticisms is specific to any of these views.

[6] For some other objections to Lewisian Conventionalism, see Devitt (2006: ch.10), Hawthorne (1990, 1993), Gilbert (1989), Kölbel (1998), Laurence (1996), Magidor and Hawthorne (2009), Petit (2002), and Schiffer (1993, 2006).

ditions (i)–(iv) as they apply to a regularity of truthfulness and trust in some language, but it would also subsume any adjustments to Lewis's theory that are designed to avoid objections of the kind raised by Burge.

Next, I want to consider mutual knowledge and some of its weaker relatives.[7] According to the best-known, *iterative* definition, A and B mutually know that *p* if and only if:

(1a) A knows that *p*.

(1b) B knows that *p*.

(2a) A knows that B knows that *p*.

(2b) B knows that A knows that *p*.

(3a) A knows that B knows that A knows that *p*.

(3b) B knows that A knows that B knows that *p*.

And so on, *ad infinitum*.

A useful generalization of this notion, which is spelled out by Clark and Marshall (1981), is *shared knowledge*$_n$, which a pair of agents possess just in case they satisfy conditions (1a)–(*n*b) of the definition of mutual knowledge. Mutual knowledge is shared knowledge$_\infty$.

Lewis (1969) and Schiffer (1972) offer similar reasons for thinking that mutual knowledge should be a necessary condition on the meaningfulness of an expression in a community, and their reasoning shows that they want their metasemantic theories to explain the efficiency of communication in the way that I outlined in §1.3 and §4.1. Suppose that A utters 'the sunset is beautiful' to B, thereby intending to communicate that the sunset is beautiful. We might try to explain B's success in interpreting A by supposing that A and B both participate in a regularity of

[7] The term 'mutual knowledge' is used by Schiffer (1972), Clark and Marshall (1981), and others for what Lewis (1969, 1975) and Barwise (1989: Ch.9) call 'common knowledge'.

truthfulness and trust in some language L such that L('the sunset is beautiful') = the proposition that the sunset is beautiful. But, as Lewis and Schiffer point out, this condition is not sufficient to explain the efficiency of linguistic communication, since B might mistakenly believe that A participates in a regularity of truthfulness and trust in L′, such that L′('the sunset is beautiful') = the proposition that the sunrise is ugly. This sort of false higher-order belief could lead B to interpret A as meaning by his utterance that the sunrise is ugly, or could lead A, for similar reasons, to change his speech behavior to put it in conformity with what he thinks B will expect. In order to rule out this kind of miscommunication, A and B would have to have shared knowledge$_1$ that they both participate in a regularity of truthfulness and trust in L (rather than, for example, L′). But this condition is also insufficient to ward off miscommunication, since A might know that B participates in the regularity of truthfulness and trust in L but mistakenly believe that B believes that A participates in a regularity of truthfulness and trust in L′, and this higher-order belief would short-circuit the potential for communication in much the same way as the previous case. This kind of miscommunication would be warded off if A and B had shared knowledge$_2$ that they they both participate in a regularity of truthfulness and trust in L. Lewis and Schiffer argue that this kind of problem can iterate indefinitely, and conclude that the only way to explain the reliable efficiency of linguistic communication is to require that A and B possess shared knowledge$_\infty$—mutual knowledge—of the the linguistic regularity in which they both participate. The fact that this argument is based on the need to ensure the efficiency of linguistic communication, and that it supplies crucial support for Lewisian conventionalism, will become significant in §4.5.

I also want to consider three proposals for weakening of the standard notion of mutual knowledge. The first is due to Schiffer, who argues that, for the purposes of defining mutual knowledge, an agent can count as knowing that *p* despite the fact that "the thought that [*p*] never once entered [his] head", and also despite the fact that he "might have to be 'convinced' or 'brought to see' that he is entitled to claim to know" that *p* (1972: 36). One reason to go along

with Schiffer on this weakening is to avoid commitment to the idea that agents with mutual knowledge have an infinite number of warranted true beliefs, which might seem psychologically dubious. For reasons in this ballpark, Lewisians have generally argued that agents can be counted as possessing mutual knowledge on the basis of shared situations that furnish them with independent justificatory "premises" (Lewis 1969), "conditions" (Schiffer 1972), or "heuristics" (Clark and Marshall 1981) from which mutual knowledge could, in principle, be derived.[8] In effect, Schiffer's idea is to reduce the psychological demands of mutual knowledge by replacing it with *entitlement to mutual knowledge*—an epistemic state that a pair of agents are in if they would be justified were they to make any of the judgements corresponding to the clauses of the definition of mutual knowledge.

Bach and Harnish (1979) propose a different notion in place of mutual knowledge. They first suggest replacing knowledge with mere belief because it seems that the extra ingredients in knowledge don't actually do any work: if a group of speakers believe (and believe that they believe, etc.) that a certain sentence is a reliable way to get a certain point across, then those beliefs become self-fulfilling. Second, Bach and Harnish point out that since humans can't entertain higher-order thoughts beyond a certain degree of complexity, there is no need to posit an infinite

[8] See McCarthy et al (1977) and Barwise (1989: Ch.9) for model-theoretic arguments that the iterated definition of mutual knowledge given earlier is unsatisfactory because it does not entail that two agents with mutual knowledge will mutually know that they have mutual knowledge. Barwise considers two other definitions, including Harman's (1977) "fixed point" definition and a "shared situation" definition that Barwise attributes to Lewis (1969), Schiffer (1972), and Clark and Marshall (1981). The last of these is what I call "entitlement to mutual knowledge". Barwise argues that although the fixed point definition is useful for information-theoretic purposes, the shared-situation definition is preferable for theorizing about actual agents (1989: 218).

number of higher-order beliefs. What they wind up suggesting in place of mutual knowledge is *shared belief*$_3$ plus (just to be safe) a ban on countervailing higher-order beliefs (1979: 309, fn.1).[9]

Finally, I want to consider an alternative to mutual knowledge that is inspired by Stalnaker's notion of *presupposition*. According to Stalnaker, "a proposition is presupposed if the speaker is disposed to act as if he assumes or believes that the proposition is true, and as if he assumes or believes that his audience assumes or believes that it is true as well" (1978: 84). Whereas entitlement to mutual knowledge and shared belief are true weakenings of mutual knowledge, in the sense that any group of agents who mutually know that p will also possess entitlement to mutual knowledge that p and shared belief$_n$ that p (for any n), shared presupposition$_n$ is not a proper weakening of mutual knowledge (for any n). As Stalnaker points out, presupposing a proposition is neither necessary nor sufficient for either knowledge of or belief in it, since we may presuppose something contrary to our beliefs for the sake of a conversation (1978: 84). Instead, what matters for shared presupposition$_\mathbf{n}$ is that agents are disposed to act as if they possess shared belief$_\mathbf{n}$ for the purposes of the conversation.

Say that A and B possess some *shared attitude* that p just in case they possess either the *entitlement to shared knowledge*$_n$, *shared belief*$_n$, or *shared presupposition*$_n$ that p, for some $n \geq 1$. In that case, say that a theory is an instance of Generalized Lewisian Conventionalism just in case it requires for the meaningfulness of an expression e for a group G that the members of G possess some shared attitude that p, where p is a proposition about how e is used by the members of G.

I turn now to two arguments that, even on this very broad construal, no form of Lewisian Conventionalism can be correct.

---

[9] Bach and Harnish use the term 'mutual belief' for their proposed weakening of mutual knowledge. I deviate from their terminology for the sake of terminological consistency. My conventions are straightforward: for any propositional attitude Ψ which can be reported using a verb whose third person singular conjugation is φ, and for any proposition p, two agents posses shared Ψ_n that p just in case they satisfy clauses 1a–nb of the result of replacing every instance of 'knows' with φ in the iterated definition of mutual knowledge given above. Two agents possess mutual Ψ that p just in case they possess shared Ψ_∞ that p.

4.3 Against Shared Attitudes

The strategy of Generalized Lewisian Conventionalism is to explain the efficiency of linguistic communication within a speech community by saying that the community's members have shared attitudes about the linguistic regularities in which they participate. This idea, even in its watered-down state, becomes problematic when we consider situations in which agents communicate efficiently with language despite being in an epistemically poor position with respect to their linguistic regularities. Situations of this kind can result from the differences in vocabulary, dialect, and register that separate us all to varying degrees.

As an illustration of this point, consider the following scenario. A and B are diplomats at the United Nations who have plenty of reasons to think that they don't speak the same language. In particular, each has plenty of evidence that the other doesn't speak English. In fact, the last time they encountered one another, they really didn't speak a common language. In the mean time, however, they've both been secretly taking English lessons. Suppose that A has learned of a terrorist plot and believes that it is urgently important that he inform the UN Security Council. On the way to deliver the news, however, he gets trapped in the UN elevator and begins to have a heart attack. The only other person in the elevator is B. In a desperate attempt to get his message to the Security Council, A produces a frantic succession of utterances in all of the languages he speaks, one of which, drawing on his secret English lessons, is of the sentence,

(S) Tell the Security Council that I've uncovered a terrorist plot!

As a result of B's secret English lessons, he understands A's utterance of (S). That is: B comes to recognize, as a result of hearing A's utterance of (S), that A meant to request that B inform the Security Council that A has uncovered a terrorist plot. Call the moment at which B understands A time *t*.

The datum here is that A succeeded in getting specific and complex information across to B quickly, and in a context in which he likely couldn't have gotten the same point across non-

linguistically. In other words, this is a case of efficient linguistic communication in the sense outlined in §4.1. Given that we posit meaning in order to explain just this sort of datum, it's quite clear that we should say that (S) has the same meaning for A and B. But this fact cannot be explained by Lewisian Conventionalism: at no time up to *t* did A and B have *any* shared attitude whose content was a proposition about a linguistic regularity in which they both participate. A uttered (S) in desperation, with a very low credence that B would understand, and with good reasons to think that B wouldn't understand. Unless B is a particularly emotive interpreter, A still has no reason to think that B has understood him at *t*, any more than he has reason to think that B understood any of his other utterances. Given their past experiences, A still has many reasons to believe that B probably hasn't understood at *t*, and if he continues to produce further utterances in other languages, A can't even be counted as presupposing that B has understood.

Depending on how the situation plays out after *t*, A and B might come to have shared attitudes as a result of their exchange. It may be tempting to predict, for example, that A and B will come to believe or presuppose that they both speak English at some time after *t*. But it is not enough for the Lewisian to point to a shared attitude that *arises from* A's utterance; the Lewisian must show that the relevant shared attitude constituted the potential for efficient communication between A and B, and thereby explains how their communication at *t* took place. Even if A and B come to possess a shared attitude about the way they use (S) as a result of A's utterance, this shared fact will be explanatorily posterior to their episode of successful communication. We can suppose, moreover, that it is with his dying breath that A produces his utterance of (S), and that he then dies so quickly that by the time B understands what A said—by time *t*, that is—A is already dead. In this case, there is no moment at which the requisite shared attitude could be in place, such that the meaning of (S) for A and B could be constituted by it. Nonetheless, A and B will have communicated efficiently, and this fact demands just the sort of explanation that a metasemantic theory should provide.

One tempting reply to this objection is that (S) is already meaningful for A and B before *t*, even if they don't realize it, since A and B are both members of the wider English-speaking community, each of whose members knows that everyone in the English-speaking community speaks English, knows that everyone in the community knows this, and so on. In other words, perhaps a fact of the following kind constitutes the meaningfulness of (S) for A and B:

(4) (a) A is a member of the English-speaking community (a.k.a. ESC).

(b) B is a member of ESC.

(i) Everyone in ESC uses (S) in thus-and-such way.

(ii) Everyone in ESC [is entitled to know/believes/presupposes] (i).

(iii) Everyone in ESC [is entitled to know/believes/presupposes] (ii).

And so on, *ad infinitum*.

On some ways of construing what it would be for groups to have shared attitudes, the conjunction of conditions (i)–(*n*) would be sufficient for the English-speaking community to count as possessing a shared attitude$_n$ that they use (S) in thus-and-such way. Could this, together with (a) and (b), constitute the meaningfulness of (S) for A and B?

To see why not, first notice that (4) is not a suitable basis, by Lewis's and Schiffer's own lights, for explaining how A and B can communicate efficiently using (S). The basic problem is that B can satisfy (4) while also having false beliefs about how A uses (S) (and vice versa). Recall Lewis and Schiffer's reason for requiring that speakers of L have shared knowledge$_1$ that they conform to a regularity in L: if they didn't have shared knowledge$_1$ that they conform to a regularity in L, then one might mistakenly believe that the other conforms to a regularity in another language, L′, which could lead to systematic miscommunication. But exactly the same kind of communication-disrupting false belief is left open by (4), since (4) is compatible with B mistakenly believing that

A is a member of the English′-speaking community, where English′ differs from English in that it assigns a different meaning to (S). So, by parity of reasoning with Lewis's and Schiffer's own argument for grounding meaning in shared attitudes, (4) can't explain the meaningfulness of (S) for A and B; we would have to add a further condition to the effect that each of A and B knows that the other is a member of the English-speaking community, which is precisely what they fail to know in the UN case.

This raises the issue of what it is for mutual knowledge to be possessed by a group (as opposed to a pair of individuals). Schiffer articulates the dilemma:

> How is one to understand the phrase 'mutual knowlege* amongst the members of G'? Does this mean that each member of G knows of each member of G that he knows…? Or does it merely mean that that every member of G knows that anyone who is a member of G knows that…? The mutual knowledge* condition must be taken in the second of these two senses. (Schiffer 1972: 131)

We can think of Schiffer's two glosses as the *de dicto* and *de re* readings of 'mutual knowledge in G', respectively. Schiffer prefers the de dicto reading because it is more realistic to think that it could be satified by an actual speech community, but he does not delve into the consequences of this choice for his theory of linguistic convention. One of the implications of my argument is that nothing short of the *de re* reading of 'mutual knowledge in G' is sufficient for Lewis's and Schiffer's purposes. And since actual speech communities sometimes don't satisfy the *de re* reading, this is a serious problem for Lewis and Schiffer.

We needn't agree with Lewis's and Schiffer's motivations for positing mutual knowledge to see that a fact with the shape of (4) can't explain how A and B are able to communicate efficiently with (S). The crux of the problem is simple: even if I know that a certain group G uses an expression *e* in a certain way, that knowledge can't help me to interpret *your* utterance of *e* unless I also know (or believe, or presuppose, etc.) that *you* are a member of G. And this brings us back to where we were four paragraphs ago: although A's utterance of (S) might *bring about* B's knowledge

that A is a member of the English-speaking community, this newfound knowledge can't be used to explain the meaningfulness of (S) for A and B, since we needed the meaningfulness of (S) to explain how B was able to understand what A said in the first place.

My UN example is hypothetical, but relevantly similar situations arise in real life all the time. The essential ingredients for one familiar type of case are a piece of vocabulary that is technical, obscure, or otherwise native to a particular dialect or register, together with a set of social conditions that, for reasons having to do with politeness, snobbishness, or whatever else, might lead someone to use such a word or phrase despite justified doubts that their audience will understand. Conversations of this sort should be very familiar from academic contexts, with their constantly shifting and only partially overlapping technical vocabularies. They also arise frequently when traveling in a country where one's language is not commonly spoken but in which it would be rude or annoying to ask each interlocutor whether she speaks it before using it to engage with them.[10] The phenomenon is an unsurprising by-product of the fact that no two speakers have identical vocabularies, together with the fact that it is often socially inappropriate, inefficient, or impossible to ensure, in advance of deploying a given expression, that one's addressee will understand it.

This line of argument shows that Lewisian conventionalism fails on its own terms—as an attempt to reduce the semantic facts to facts about shared attitudes. Since a fact may ground another without being necessary for it, this line of reasoning still leaves open the possibility that Lewisian shared attitudes ground linguistic meaning some of the time, but I don't regard this view as attractive, both because it fails to tell us anything about what grounds many actual instances of linguistic meaning, and because I think that we can find a more general and explanatory account of what constitutes meaning. I will expand on these considerations in the next two sections.

---

[10] Petit (2002) makes a closely related point using similar examples.

4.4 The Meaning Without Use Problem

Hawthorne (1990, 1993), Lewis (1992), and Schiffer (1993, 2006) discuss a different sort of problem for Lewisian conventionalism under the heading of *the meaning-without-use problem*. Although I think that my argument in §4.3 to be decisive, I wish to consider the meaning-without-use problem because it can help us to diagnose a different sort of problem with the Lewisian metasemantic strategy.

The problem can be laid out as follows. Due to the productivity of natural languages and the finitude of human discourse, there are many sentences that will be meaningful for a group despite never having been uttered by any of the group's members. Indeed, some meaningful sentences will never be used. Lewisian conventionalism says that a language L is used by a community of speakers just in case they participate in a convention of uttering sentences of L only if a certain condition is met—only if L maps the sentences to true propositions, for example. The problem arises because community members will vacuously satisfy any such condition with respect to hitherto unuttered sentences. Take a language L* that differs from L only in the meanings that it assigns to sentences that haven't been uttered by the members of a certain community G. On the Lewisian account as it was originally laid out, and as I've laid it out above, there is no fact that could determine which of L or L* is the actual language of G. L and L* may assign radically different meanings to every sentence that has so far gone unused, and so, on Lewisian conventionalism there can be no fact of the matter about the meanings of unused sentences. This is a disastrous result: not only are many unused sentences intuitively meaningful; we also need them to be determinately meaningful in order to explain the fact that we regularly communicate efficiently with novel sentences.

Lewis attempts to deal with this problem by arguing that the language determined by any particular spoken fragment of a language is specified by the unique *extrapolation* of whatever grammar applies to the fragment. The idea of an extrapolation of a grammar, according to Lewis, "pre-

supposes the distinction between straight and bent" grammars (1992: 151). Lewis doesn't offer a definition of what counts as a straight grammar, but he suggests that "any grammar that any linguist would actually propose" would be straight in the relevant sense (150). This distinction between straight and bent grammars would rule out grammars specifying languages following straightforward compositional principles that break down only for sentences above one thousand words in length, for example.

Presumably, a grammar counts as straight insofar as it exemplifies an appropriately weighted cluster of the principles that guide theory choice—predictiveness, consistency, parsimony, projectability, etc. Still, straightness looks to be a vague criterion, and there is no guarantee that all of the straight grammars of of a language fragment will agree about how to extend it. Even if we ignore these defects of the straightness criterion, however, Schiffer (1993, 2006) has demonstrated other good grounds on which to reject it. His criticism boils down to the fact that we might sometimes have good reasons to think that a community's language use is guided by a bent grammar. This might happen, for example, if the community learns the language in a classroom where they are explicitly taught the bent grammar. In this case, it is conceivable that that there is a grammar for the community's spoken fragment that is straighter than the one guiding their language use, and that might even diverge from it in the meaning it would assign to possible or future utterances.

Another kind of counterexample stems from the hypothesis that language use relies on a grammar that is internally represented by or embodied in the language-processing mechanisms of the mind. This view is accepted by many linguists and philosophers of language. But if it is true, then there is no reason to assume that the grammar guiding any particular speaker's language use is a straight one. Lewis recognized this possibility, but argued that we should still rely on "straight extrapolation" in order to settle questions of meaning, since otherwise, "whenever we resort to extrapolation to answer questions of syntax and semantics, we are engaged in risky speculation

about the workings of the brain" (1992). Schiffer rightly finds this position wanting. Surely, if given a choice between a grammar chosen on the basis of general principles of theory choice and one that actually plays a crucial causal role in determining how a speaker uses and understands a sentence, the latter is the better candidate for the grammar that determines what the sentence means for the speaker. The fact that a bent, psychologically real grammar would be more difficult to discover than a straight one is inconvenient, but convenience is not a valid principle of theory choice.

In response to this line of thought, Schiffer proposes that we replace Lewis's criterion with an overtly psycholinguistic one (2006: 286):

> What makes *L* a language that is used by *x* even though the used fragment of *L* is also a fragment of infinitely many languages other than *L* is that it's the language determined by the internally represented generative grammar implicated in *x*'s language-understanding processes.

Schiffer's solution to the meaning-without-use problem is preferable to Lewis's solution because it grounds the meaningfulness of unused sentences in psychological facts that play a role in explaining how speakers could communicate efficiently with the expressions, were they to deploy them. But the resulting metasemantic theory has its own defects. Lewis anticipated such an account three decades earlier, when he rejected it on the following grounds (1975a: 178):

> I am much less certain that there are internally represented grammars than I am that languages are used by populations; and I think it makes sense to say that languages might be used by populations even if there were no internally represented grammars.

Oddly enough, an earlier incarnation of Schiffer (1987a, 1993) sides with Lewis against the more recent Schiffer (2006). The earlier Schiffer argues that since language comprehension could be achieved by some means other than an internally represented grammar—his alternative is an internalized mechanism for translating public language into the language of thought—building the idea of such a grammar into our theory of meaning would make it insufficiently general.

Schiffer's (2006) psycholinguistic proposal is flawed on other grounds. Given linguistic variation within speech communities, it seems unlikely that any two members of a speech community—let alone *all* the members of the community—share just the same internally represented grammar. This point is obvious if we individuate grammars finely—in terms of the lexical items over which they are defined, for example, or in terms of states of speakers' language faculties (as in Chomsky's I-Language hypothesis (1986)). Fine-grained individuation conditions would be necessary if we want grammars to fully determine a Lewisian language. But even if we individuate grammars coarsely—in terms of compositional principles defined over semantic categories, for example, or even as a generative procedure defined over syntactic categories—we might still find two members of a single speech community with different internally represented grammars.

The crucial point for my purposes here is that it is possible for a single linguistic expression to have the same meaning for two speakers even if their internally represented grammars differ with respect to the expression's syntactic or semantic structure. Consider a speaker S who uses and understands the complex noun phrase 'gender gap' in the familiar way, but who has learned the phrase as an idiomatic unit. S does not know the words 'gender' or 'gap', at least in the sense that S is undisposed to utter either word separately, and would fail to understand an utterance of either word outside the context of the whole phrase. Although, in effect, S treats 'gender gap' as a single word, I think it is reasonable to say that sentences containing it mean the same thing for him that they do for the rest of us. In fact, we have to say this in order to explain the fact that we would have no problem communicating efficiently about gender gaps with S. Nonetheless, it is plausible that S's internally represented grammar differs from ours, at least in how it treats the expression 'gender gap'. I can think of no way of individuating grammars that would be coarse enough to get by this point, but fine enough to play the role that Schiffer needs for it in his metasemantic theory.

The problem raised by all of this is that Schiffer therefore has no way of accounting for the publicity of linguistic meaning—the fact that expressions are meaningful for groups rather than

for individual speakers. In inserting the notion of an internally represented grammar into his solution to the meaning-without-use problem, Schiffer looks *too far* into speakers' minds for the psychological facts that ground linguistic meaning, overcompensating for Lewis's (1992) anti-psychologistic solution, which fails to look far enough into speakers' minds. What's needed is an approach that grounds meaning in psychological facts taken at just the right level of abstraction—a level that emphasizes our linguistic commonalities without overstating them and without understating them.

Some other time-slices of Schiffer (1993, 2003) have recognized this need, and suggested that we define the actual language relation in terms of *whatever* mental state plays the right sort of role in the psychology of language use, whether that role is played by a grammar, a translation manual, or something else: "*knowledge of meaning needn't be any one kind of state but is rather any state that plays the 'knowledge of meaning' processing role*" (2003: 115, emphasis Schiffer's). To make the point, Schiffer adopts the language of thought hypothesis, and defines an *L-determining translator* as "a language-processing mechanism that determines a mapping of each sentence of *L* onto a Mentalese sentence that means in Mentalese what the *L* sentence means in *L*" (1993: 246). He then puts forward the following hypothesis (*ibid.*):

> A language *L* is *used by* a population *P* iff there is a practice in *P* of meaning in *L* and the processing of *L* utterances proceeds via an *L*-determining translator.

It is important to note that the language of thought hypothesis, and Schiffer's accompanying notion of a language-determining translator, are not essential to the spirit of his view. More generally, his point is merely that a language user has *some* psychological mechanism that (a) explains how he gets from an auditory perception of some sentence to a mental representation of what the speaker said, and (b) is fine-grained enough to determine the meanings of sentences that haven't

been used yet.[11] We can't be sure of the details until we do a lot more psycholinguistics, but this somewhat vague hypothesis is the best explanation of how speakers interpret utterances of new sentences, and it also gives us a way, in principle at least, of nailing down the language a speaker uses.

I think that if we interpret Schiffer in the right way, he is very nearly right. It is not obvious that he would interpret himself in this way, however. His two examples of psychological mechanisms that might play the knowledge of language processing role are an internally represented grammar and an internally represented public-language-to-mentalese translator—either way, something about a person's mind that would presumably stay stable from one conversation to the next, absent a language class.

But the meaning of a sentence is protean. 'What grounds the electric current' means something different for a group of analytic metaphysicians than it does for a group of electricians. Many sentences are meaningful only for one group but not the other. But a speaker may be a member of both groups, or a combined group. Likewise, the sentence, 'ow's she gon', baye?' is meaningful for a group of Newfoundlanders, but not for a group of non-Newfoundlanders, nor—crucially—is it meaningful for a mix of the two. 'Ponds are bigger than lakes' means something different for a group of Newfoundlanders than for a group of non-Newfoundlanders, and may not

---

[11] Schiffer gives a slightly different version of his theory in *The Things We Mean* (2003: ch.3). There, he defines the notion of a sentence's character*, which is an ordered pair, <A,P>, such that A is the type of direct and literal illocutionary act that is typically performed by uttering the sentence, and A is what one typically says in uttering it. He struggles with the question of whether to identify meanings with characters*. He ultimately explicates knowledge of a sentence S's meaning as possession of some mental state that takes the knower from an auditory perception of an utterance of S to the knowledge that the utter has performed a literal, direct speech act with a force and content that conform to S's character*. This formulation improves on the version presented in 'Actual Language Relations' (1993), mainly in that it accounts for the fact that the meaning of a sentence constrains the forces, as well as the contents, of the speech acts we can perform with it. The earlier Schiffer recognized this too, but chose to abstract away from it in stating his solution to the meaning-without-use problem.

mean anything definite or consistent for a mixed group.[12] An individual can move between groups of metaphysicians, electricians, Newfoundlanders, and many others, communicating efficiently with the other members of each. In doing so, the individual doesn't merely continue to utter the same sentences and hope for the best. Assuming he is sensitive (in some sense) to the linguistic dispositions of the group into which he moves, his own dispositions will change in various ways to increase the chances of communicative success with the new interlocutors.[13] Since whatever brings about these changes is part of what explains the unique efficiency of verbal communication—and since it is part of what explains people's abilities to move from perceptions of utterances to correct interpretations of speech acts—they should be taken into account by a theory of sentence meaning. But my internally represented grammars or translation manuals presumably aren't protean in this way. So the psychological states on which meaning supervenes must go beyond grammars and translation manuals.

Nonetheless, I think that Schiffer's 'processing role' idea is basically on the right track, as long as we throw out the bits about public languages and conventions, and take the processing role to be constituted by some mental state or combination of mental states that are less stable than an

---

[12] 'Ow's she gon', baye?' is a informal equivalent of 'how do you do today, sir?' in the dialect of English spoken in Newfoundland. Newfoundlanders also use 'pond' to refer to lakes and 'lake' to refer to ponds.

[13] I think this phenomenon is obvious enough, especially to anyone who has started conversing with a non-philosopher immediately after participating in a heated and jargon-rich philosophical discussion. But sociolinguists have also done some interesting research on it. For example: sociolinguistic 'convergence' one of the key concepts of Communication Accomodation Theory, whose proponents have amassed loads of empirical data suggesting that convergence occurs along a wide variety of linguistic axes. See Giles et. al (1991) for an overview. In a similar vein, Armstrong (2013) and Ludlow (*forthcoming*) have documented a wide variety of ways in which the meanings of expressions can vary from moment to moment in a conversation, and developed the theory of microlanguages to explain it. Although both Armstrong and Ludlow base their theories on Lewisian conventionalism, it seems that they suffer from a particularly serious version of the meaning-without-use problem, and cannot appeal to Schiffer-style internally represented grammars for the reasons mentioned in the main text.

internally represented grammar. If we do this, then Schiffer's theory winds up looking a lot like the one that I'll offer in §4.6.

4.5 Against Convention

In light of the foregoing argument, it might be tempting to go in search of a non-Lewisian theory of convention—a theory not built on a foundation of shared attitudes about linguistic regularities—in the hope that this would allow us to hold on to a form of metasemantic conventionalism while avoiding Lewisians' cognitive and epistemic overcommitments.[14] My next argument aims to show that this tactic is misguided, and that metasemantics should make no appeal to conventions of any kind.

The thrust of this objection is that it is possible for two agents to communicate efficiently with language despite not participating in any linguistic convention. Given that we must posit meaning in order to explain the efficiency of this communication, we need to suppose that the expressions involved in these conventionless episodes of communication have meaning for the agents involved. It follows that convention is not a necessary underpinning of linguistic meaning, and that facts about meaning cannot be reduced to facts about convention. I will go on to show that this argument, together with the distinction between causal and constitutive explanations, gives us reason to reject the thesis that facts about meaning are grounded in facts about convention.

I count a metasemantic theory as a form of conventionalism if it gives some role to any notion that can be regarded as a precisification of our ordinary notion of convention. I will rely on the assumption that a convention, in this ordinary sense, is a way a group has of doing something

---

[14] Margaret Gilbert (1989), Ruth Millikan (2005, 2008), Seumas Miller (2001), and Brian Skyrms (1996, 2010) have defended alternative theories of convention. Of these, Gilbert's and Miller's theories appeal to mutual knowledge in ways that make them vulnerable to variations on my argument from §4. Any theory of meaning that gives convention a prominent role is susceptible to my next objection—particularly those of Millikan and Skyrms, since they define convention in essentially diachronic and causal, rather than synchronic and constitutive, terms.

that is (in a relevant sense) arbitrary, and that results either from an explicit agreement or from a self-perpetuating regularity in the group. This statement can be thought of as placing a pre-theoretic adequacy condition on theories of convention; in order to count as a full theory, several of the key terms in it would have to be made more precise. My claim is that no notion of convention meeting even this basic condition should play a role in metasemantics.

Given these assumptions about convention, all that my argument requires is that a group of speakers could find themselves in a state that makes efficient linguistic communication possible, but that the state either doesn't qualify as arbitrary or didn't come about by means of an explicit agreement or a self-perpetuating regularity. The potential examples tend to be science fictional: consider the case of a person whose linguistic abilities were endowed by a bump on the head or a lightning strike, or a creature whose ability to converse with English speakers is innate and immutable. The fact that I must look outside everyday life to find a clear example makes this argument somewhat more speculative than my objection to Lewisian conventionalism. Nonetheless, I can see no good reason for denying that any of these scenarios is possible. If a subset of humans had evolved to speak English (and only English) from birth, we would still have to posit linguistic meaning in order to explain their ability to communicate, despite the fact that their words' meanings would be no more arbitrary or self-perpetuating than any other feature of their biology (cf. Peacocke 1976: 168–70).[15] Likewise, although it is enormously unlikely for a lightning bolt or a bump on the head to scramble an individual's brain in such a way as to endow him with the abil-

[15] There may be an evolutionary sense in which humans' number of limbs is an arbitrary and self-perpetuating regularity. Even if so, it seems to me that these are not the same senses of 'self-perpetuating' or 'arbitrary' that would make a regularity conventional. A way of doing things is a convention only if it is arbitrary in the sense that the very individuals who participate in the convention could have achieved their ends just as well by other means. Given plausible and widely-held assumptions about alethic modality, an individual person could not have evolved to have a radically differently genotype than he or she actually has. If populations can somehow be individuated independently of their members, then biological traits may be arbitrary and self-perpetuating at the population level, but they are neither arbitrary nor self-perpetuating for the individuals who possess them.

ity to speak a new language, there is a long philosophical tradition of accepting the possibility of such freak events, and I can see no reason to diverge from this tradition. In essence, the problem is this: whatever it is about our minds/brains that puts us in a position to communicate efficiently with language, it is possible that a pair of speakers could find themselves in just the same state, but by unconventional means.

There is, of course, a literature on an issue that superficially resembles this one, springing from Davidson's (1987) example of the Swampman, who springs fully-formed from the swamp as a result of an improbable lightning strike. Some deny, on grounds of externalism about mental content, that the Swampman has genuine beliefs and desires, and some may be tempted for similar reasons to deny that a lightning bolt could endow someone's words with genuine meaning. I don't know what to think about whether the Swampman has genuine intentional states, but I think that this line of objection to my own argument is mistaken, given the different explanatory roles played by mental content and linguistic meaning. As I have laid it out, a significant part of the project of a theory of linguistic meaning is to explain the efficiency of linguistic communication, and communication, as I understand it, requires that the communicators already have contentful mental states. The relevant thought experiment, for my purposes, is one in which an agent who already possesses normal, contentful mental states acquires, by means of some unlikely accident, the ability to efficiently communicate these states via language. Explaining the efficiency with which they communicate using their newfound linguistic abilities would be just as worthwhile as explaining the efficiency with which *we* communicate, and if linguistic meaning is worth positing in our case, it is also worth positing in theirs.

So convention is not necessary for meaning, and facts about meaning therefore don't reduce to facts about convention. Although the possibility of meaning without convention does not, by itself, show that facts about meaning are never grounded in facts about convention, a slightly different construal of the same considerations should push us to accept that conclusion as well. A

useful way of putting the upshot of the foregoing argument is that although the facts about groups of speakers in virtue of which utterance-types have meaning for them might *come about* and *be sustained* by conventional means, this needn't be so. This way of putting things raises an important distinction between two different kinds of explanatory relations: constitutive determination (or grounding) on one hand, and causal determination on the other. The causal dimension of convention-invoking explanations is built into the idea of convention as I spelled it out earlier: conventions must be either explicitly *agreed upon in advance* or *self-sustaining regularities*. Most theories of linguistic convention, including those of Lewis, Schiffer, and Gilbert (1989), mix causal and constitutive explanations in ways that would be difficult to pry apart.[16] In thus trying to answer a constitutive question about meaning—an 'in virtue of what…' question—by citing facts about the etiology of meaning, conventionalist metasemantic theories conflate two different kinds of explanation.[17]

Why not mix constitutive and causal explanations? Because, at least in some cases, the same constitutive explanation of some fact is compatible with different causal explanations, and linguistic meaning seems to be one such case. Suppose that C1, C2, and C3 are speech communities that are psychologically identical to each other at a time *t* in all of the respects that make efficient linguistic communication possible, despite the fact that C1 got that way be means of a self-reinforcing regularity, C2 got that way by biological evolution, and C3 got that way by means of a very unlikely series of lightning strikes. Since all three communities would be equally capable of efficient linguistic communication at *t*, and since this capability would be rooted in the same psychological facts about them at *t,* we would have the same reasons for positing the same particular cases of linguistic meaning for each. The same metasemantic theory would be true of all three

[16] By contrast, Skyrms (2010) and Millikan (2005) aim at purely causal explanations of the development and persistence of meaning that are neutral about constitutive questions.

[17] On the causation/grounding distinction, see Audi (forthcoming: §3), Clark and Liggins (2012: 812), Fine (2001: 15–6), and Rosen (2010: 118).

communities, in other words, but we would need a different causal explanation of the origins and persistence of meaning in each case. If this is correct, then we need a metasemantic theory to give an account of the facts underlying meaning that is silent as to their origins.

4.6 A Dispositional Metasemantics

Each of the three preceding sections contains a different kind of objection to conventionalism, and each section also comes with a lesson about what a good metasemantic theory would have to look like. In §4.3, I argued that metasemantics must avoid placing the kinds of cognitive or epistemic demands on speakers that gets Lewisian theories into trouble. In §4.4, I argued that metasemantics must identify the psychological facts underlying linguistic meaning at just the right level of abstraction—not so coarsely so as to fall prey to the meaning-without-use problem, but not so finely as to erase the possibility that meaning is public. I also argued that the psychological states must be flexible enough, and sensitive enough to changing circumstances, to explain the various ways in which meaning is protean. In §4.5, I argued that metasemantics must focus on a purely constitutive explanation of meaning, and should be compatible with a variety of possible causal stories about how meaning arises and persists. In this section, I will outline a metasemantic theory that meets all of these desiderata. The main influences on this theory is Grice's mature metasemantic views, though my own theory departs from Grice's in some details, and goes beyond it in others.

Grice's most detailed metasemantic proposal can be found in his 1968 article, 'Utterer's Meaning, Sentence Meaning, and Word Meaning'.

> "For group G, utterance-type X means '$*_{\psi}$p' " $=_{df.}$ "At least some (many) members of group G have in their repertoires the procedure of uttering a token of X if, for some A, they want A to $\psi^{\dagger}$ that p, the retention of this procedure being for them conditional on the assumption that at least some (other) members of G have, or have had, this procedure in their repertoires." (1989: 127)

Although it requires unpacking and elaboration, I think this explication shows promise in some of the ways that conventionalist metasemantic theories fail.

The second clause of Grice's definition—about the retention of procedures—is his nod to the idea that linguistic meaning is conventional, or at least self-reinforcing. But notice that unlike Lewis's appeal to convention, around which his whole metasemantic approach is constructed, Grice's nod to convention is easy to excise. Grice aims at a constitutive account of meaning with the first half of his explication and a causal account of meaning with the second half. In light of my argument in §4.5 for keeping these aims distinct, I will drop the second half.[18]

> "For group G, utterance-type X means '$^*_\psi$p' " $=_{df.}$ "At least some (many) members of group G have in their repertoires the procedure of uttering a token of X if, for some A, they want A to $\psi^\dagger$ that p."

Most of the metasemantic work here is being done by the notion of *having a procedure in one's repertoire*, and I will turn to it shortly. Since the explication contains several other troubling obscurities, I first want to subject it to some sympathetic reconstruction.

First, consider the form of the second relatum in Grice's explicandum.

'$^*_\psi$p'

Grice calls the schematic operator variable '$^*_\psi$' a "dummy mood indicator" (118) or "mood marker" (110). It "corresponds to the propositional attitude ψ-ing (whatever that may be), as, for

---

[18] Grice sometimes casually refers to utterance-type meaning as 'conventional meaning', and this has led many commentators to label him a conventionalist (e.g., Laurence 1996: 269). Neale is one of the few commentators to explicitly deny this interpretation: "Grice's use of the word 'conventional' in 'conventional meaning' should not be taken too literally, for it is Grice's view that linguistic meaning is not to be explicated in terms of what other philosophers might think of as convention" (Neale 1992: *fn*12). See also Chapman (2005: 76) for an insightful discussion of Grice's aversion to the notion of convention.

example, '⊢' corresponds to believing (thinking) and '!' corresponds to intending" (123). Grice's motivation here is to point out that the timeless meaning of a sentence bears an explanatory relation both to the proposition that a speaker would intend his addressee to entertain in uttering the sentence, and to the attitude the speaker would intend his addressee to take toward that proposition. Moreover, this latter dimension of the meaning of an utterance-type is tied to certain of its features—features that Grice schematically separates out as '$*_{\psi}$', but which typically take the form of mood, intonation (in speech), and punctuation (in writing). I have already outlined theories of illocutionary force and clausal mood that I take to be compatible with Grice's views in §1.2 and Chapter Three, respectively; Grice is hinting at similar accounts with his notation here.

Grice's account can be rendered more perspicuous by framing it in the idiom of speech act theory. Drawing on the equivalence between speaker meaning and the performance of illocutionary acts outlined in §1.2, I will rephrase Grice's explication as follows.

> "For group G, utterance-type X has meaning" $=_{df.}$ "At least some (many) members of group G have in their repertoires the procedure of uttering a token of X in order to perform a speech act of some type Y."

Sentence meaning, on this reformulation of Grice's view, is a ternary relation borne by complete-utterance types (e.g., sentences) to groups of speakers and types of communicative illociutionar acts.[19]

With Grice's other obscure locutions paraphrased away, it is plain that most of the work in Grice's theory is done by his notion of *having a procedure in one's repertoire*. Grice ultimately leaves this somewhat vague technical term as a primitive in his account, and so, at first glance, the notion seems to be no more than a placeholder for a satisfying explication. Nonetheless, Grice's formulation holds the promise of several advantages over conventionalist alternatives. First, everything Grice says is compatible with the possibility that two speakers can possess the same pro-

[19] For accounts of the nature and role of sentence meaning spelled out in terms of a similar ternary relation, see Alston (2000), Neale (2004, 2005), Searle (1969), and Schiffer (1993, 2003).

cedure without possessing shared attitudes that they do. In other words, Grice's explication can be construed so as to appropriately loosen the cognitive and epistemic requirements on the existence of meaning in a way that allows it to avoid my objections to Lewisian conventionalism in §4.3. Second: it is at least plausible that procedures are mental phenomena that are identified at a level of abstraction that is specific enough to solve the meaning-without-use problem by constituting the meaning of hitherto-unused expressions in a way that generalizes across speakers and is protean from one conversation to the next. Third: the notion of having a procedure in one's repertoire gives some (albeit thin) content to the idea that meaning is constituted, not by convention itself, but by some facts that could be brought about and sustained either by conventional causal chains or in some other way. If the members of a speech community have a procedure in their repertoires that allows them to communicate efficiently using a sentence, then how they came to have the procedure—whether by convention, evolution, or lightning strike—is immaterial. This modification of Grice's theory aims to offer an account of the facts constituting meaning that is neutral about their etiology, as it should.

Despite these potential virtues, it is clear from Grice's struggle to further flesh out his 'procedure'/'repertoire' formulation that even he found it to be unsatisfyingly obscure (1989: 126–128). I will now suggest that Grice could have profitably replaced his notion of having a procedure in one's repertoire with the notion of *overlapping communicative dispositions*.

In fact, Grice briefly considers substituting dispositions for procedures after outlining his proposal, but quickly rejects this modification after considering prim Aunt Matilda, who is not disposed to utter a certain vulgarity under any circumstances despite its having meaning for her (127). Based on just this example, Grice "abandon[s] the attempt to provide a definition" of having a procedure in one's repertoire, "and content[s] himself with a few informal remarks" (127).

I think Grice stopped too soon. One needn't look further than his informal remarks about Aunt Matilda to see how he might have pursued a more satisfactory dispositional account. Grice asks us to consider a situation in which some utterance-type X

> is current for some group G; that is to say, to utter X in such-and-such circumstances is part of the practice of many members of G. In any case my Aunt Matilda (a member of G) may be said to have a procedure for X, even though she herself would rather be seen dead than utter X, for she knows that some other members of G do have a readiness to utter X in such-and-such circumstances. [127–8]

Putting things in a less epistemically-loaded way: although Aunt Matilda is not disposed to utter a vulgarity under any circumstances, she is still disposed to interpret an utterance of a vulgar expression in a way that allows her to understand others who utter it. Although she lacks the *performative disposition*—the disposition to utter the vulgar expression in order to perform a certain type of speech act—she does possess a corresponding *interpretive disposition*—the disposition to interpret speakers who utter the vulgar expression as performing speech acts of the right type. We can group performative and interpretive dispositions together as *communicative dispositions*. I'll say that two speakers' communicative dispositions *overlap* when one has a performative disposition to produce utterances of some type X in order to perform speech acts of type Y, and the other has an interpretive disposition to interpret utterances of the kind X as performances of the kind Y. Given the Gricean approach to speech act theory outlined in §1.2, to have a performative disposition is to be disposed to produce a certain kind of utterance if one has a certain kind of communicative intention, and to have an interpretive disposition is to be disposed to interpret a speaker as having a certain kind of intention (component (1) of a communicative intention, as in the schema from §1.2) if the speaker produces a certain kind of utterance.

With these notions in hand, the following further modification of Grice's account becomes available:

Utterance-Type Meaning (UTM)

"For group G, utterance-type X has meaning" =*df.* "At least some (many) members of group G are disposed to utter X in order to perform a speech act of type Y, and at least some (many) members of G are disposed to interpret utterances of X as speech acts of type Y."

This account makes the potential advantages of Grice's own formulation more concrete. Groups of speakers can have overlapping communicative dispositions without possessing any shared attitudes about their language use, and so this account avoids Lewisians' epistemic and cognitive overcommitments. Of course, agents' communicative dispositions could conceivably be constituted by Lewisian shared attitudes, and so this explains the temptation to think that expressions would be meaninful in a commuty who met Lewis's definition, but there is no reason to think that actual speech communities exemplify the definition.

It's plausible that language users' communicative dispositions are partly constituted by the current states of their language faculties—on one way of talking, by their internally represented grammars (Chomsky 1986). But two agents needn't have the same internally represented grammar in order to have the same communicative disposition, or in order to have communicative dispositions that overlap in a way that makes it possible for them to communicate efficiently. Two agents could be disposed to refer to gender gaps using 'gender gap' even though one's grammar treats it as an idiomatic unit and the other's treats it as a syntactically and semantically structured noun phrase. This multiple realizability of communicative dispositions is what allows them to ground meaning in a way that avoids overcommitments about agents' fine-grained mental states.

On the other hand, communicative dispositions are perfectly suited to ground the meaningfulness of previously unused expressions because an agent can possess a communicative disposition for an expression she has never used before. Where do we get all of these communicative dispositions to use expressions that haven't been used before? I'll argue in Chapter Five that it is the job of compositional semantics to answer that question by showing how our sentence-sized

communicative dispositions are systematically constituted by our word-sized communicative dispositions. The latter are dispositions to perform acts like referring, predicating, and expressing propositions with words, phrases, types of syntactic arrangement, and sentence radicals. The right collection of primitive communicative dispositions—dispositions to use semantically primitive expressions and syntactic structures in certain ways—can add up to a very large number of communicative dispositions to produce and interpret complex expressions, including expressions that have never been uttered before.

Although the way in which word-sized dispositions add up to sentence-sized dispositions is surely guided by something like internally represented grammars, the shiftyness of meaning across contexts gives us reason to suspect that other, less stable kinds of mental states play a role in grounding agents' communicative dispositions as well. I am disposed to use the verb 'ground' one way when talking to an electrician and another way when talking to a metaphysician. It isn't plausible that this shift in my dispositions would be best explained by moment-to-moment changes in my internally represented grammar, which, we have good reasons to believe, is relatively cut off—or "informationally encapsulated"—from personal-level mental states like my belief about the profession of my addressee (Fodor 1983). But it makes perfect sense to say that an agent's communicative dispositions are partly constituted by subpersonal-level mental states, modular competencies, or tacit knowledge, and partly constituted by beliefs, desires, intentions, and other personal-level mental states. To describe the mind in terms of communicative dispositions is to describe it at the right level of abstraction to explain how the mental states distributed throughout a speech community can together constitute meaning that is both public and protean in the right sorts of ways.

One potential complaint about (UTM) is that it gives rise to its own version of the meaning-without-use problem, because it predicts that any sentences that the members of a community aren't disposed to communicate with isn't meaningful. For example: (UTM) predicts that any sen-

tence that is too long or grammatically complex for any member in a community to understand will not be meaningful for the community. This might include some sentences containing triple or quadruple center embeddings—for example (Chomsky and Miller 1963):

(5) Anyone who feels that if so-many more students whom we haven't actually admitted are sitting in on the course than ones we have that the room had to be changed, then probably auditors will have to be excluded, is likely to agree that the curriculum needs revision.

According to an influential strand of thought in contemporary linguistics (Bever 1970; Chomsky 1956; Chomsky and Miller 1963), (5) is a perfectly grammatical sentence, in the sense that a normal English speaker's internally represented grammatical competence determines a set of rules by which (5) abides. An English speaker might nevertheless intuit (5) to be ungrammatical because of a performance error—perhaps due to a lack of working memory. But the fact that the etiology of this performance error locates it outside of the speaker's syntactic competence, which is the proper object of study of syntactic theory according to Chomsky (1957, 1965, 1986), makes it irrelevant to the question of whether (5) is a well-formed sentence.

But Chomsky and Miller don't stop at syntactic well-formedness; they also claim that (5) has "a clear and unambiguous meaning" (1963: 286), and most semanticists have gone along with the idea that semantics, like syntax, is the study of a semantic competence, which must be distinguished from performance. On this view, sentences with meaningful parts and the right grammatical structure should count as meaningful even if they are too long or complex to be uttered or understood. If this is the right way to think about linguistic meaning, then (UTM) suffers from its own version of the meaning-without-use problem, since it cannot predict the meaningfulness of unusable sentences.

I don't think that this objection holds any force. Even if Chomsky is right that we should explain ungrammaticality intuitions about sentences like (5) by appealing to a syntactic competence–performance distinction, it does not follow that we should posit the same kind of distinction in the domain of semantics. This is because syntax and semantics study different kinds of properties—structure and meaning, respectively—which are posited to do different kinds of theoretical work. Chomsky argues that syntax is the study of an idealized speaker-hearer's faculty of language—a mental organ whose internal representations determine the speaker's syntactic competence (1986: ch.1). The language faculty is posited to explain the individualistic and subpersonal psychological facts that stand behind an individual's acquisition and knowledge of language. If Chomsky is right about this, then he's right to posit a syntactic competence–performance distinction. Even so, meaning is a very different kind of object of study, posited to explain the interpersonal phenomenon of efficient communication, and so it must be a public property of utterance types. Moreover, given that our reason for positing linguistic meaning is to explain facts about the efficient use of linguistic expressions in communication, we have no good reason to countenance the meaningfulness of expressions that aren't usable in this way.

Another virtue of (UTM) is that it gives a purely constitutive account of linguistic meaning that avoids commitments about how the facts underlying meaning causally came about. Whichever psychological states realize an agent's communicative dispositions, there is no reason to deny that those very states could have come about via a self-reinforcing regularity, or via a lightning strike, via biological evolution, or via any number of other processes.

Finally, and most importantly, (UTM) grounds facts about meaning in facts that are tailored to explain why meaning plays the explanatory role for which we posit it. On the assumption that linguistic communication consists of the production and successful interpretation of speech acts, then we posit the meaningfulness of an utterance-type X for a group G in order to explain the following fact: if members of G were to perform a speech act of some type Y, they would be likely to

do so by uttering X, and if members of G were to have an utterance of X addressed to them, they would interpret the utterer as performing a speech act of type Y. G-members' possession of the right sorts of overlapping communicative dispositions is tailor-made to explain just this sort of fact. By contrast, the fact that a way of speaking is a regularity, or arbitrary, or self-reinforcing is not relevant to whether it can facilitate communication.

An added bonus of (UTM) is that it gives us a way of individuating speech communities in terms of the ways in which their members' communicative dispositions overlap—at least to the degree that speech communities even need to be individuated. As stated—with "at least some (many) members of G" being required to possess the right communicative dispositions—these individuation conditions are extremely vague. We can try to make them more precise by fixing the limiting case and working backward: a totally homogeneous speech community is one in which there is some set of expressions Γ is such that every member of the community possesses both a performative and an interpretive disposition that links the same type of speech act with each element of Γ, and no expression not in Γ is meaningful for any subset of the community. Since it is doubtful that any two natural language users have exactly the same vocabulary, it is doubtful whether there is any actual homogeneous speech community. But there are many communities that fall short of this ideal in various respects that leave them capable of communicating efficiently in all sorts of ways, and so I maintain that the vagueness in (UTM) is appropriate. What matters to the theory, after all, is that we can posit meaning in all and only those cases where efficient communication is possible, and the theory does this in the right way, on a case-by-case basis. Although (UTM) cannot give us precise criteria by which to judge whether two agents, A and B, belong to the community of English speakers, for example, it *can* tell us whether they are able to communicate efficiently with any particular expression, and this is enough for a metasemantic theory to do.

Conclusion and Mile-Marking

In this chapter, I have argued that the dominant, conventionalist view about what it is for an expression to have meaning is mistaken, and I have constructed a better view. This foray into metasemantics furthers the overall goals of the dissertation in several ways. First, it tells us more about utterance-type meaning is, and does so in such a way that allows meaning to play the explanatory role outlined in Chapter One. Second, it lends further support to the idea, first argued for in Chapters Two and Three, that meaning is a relation borne by utterance types to types of speech acts. Third, it gives us a clearer picture of what compositional semantics is the study of: the structures of language-users' communicative dispositions.

Much remains to be done. In particular, although I have given a sense of how I think semantics should work (in Chapter Two), and although I have gone some way toward showing how compositional semantics will work (in Chapters Three and Four), it remains to flesh out my theory of compositionality and the syntax–semantics interface in a way that incorporates the metasemantic insights of this chapter. That will be my task in Chapter Five.

CHAPTER FIVE: COMPOSITIONALITY IN SPEECH ACT THEORETIC SEMANTICS

Introduction

I have spent the last four chapters defending an account of the nature of utterance-type menaing, and linguistic meaning in particular. I've argued that for an expression to have meaning is for a community to be disposed to communicate with the expression in a certain way, and that the expression's semantic value is the type of speech act that members of the community are disposed to perform (and interpret each other as performing) with it.

'But', my truth-conditional-semanticist opponent has surely been objecting since Chapter Two, 'we need to identify sentence meanings with truth condition determining properties in order to understand how word meanings can add up to sentence meanings!' This complaint is not without motivation: compositional semantics—the study of this "adding up" of word meaning into sentence meaning—has not only been the most fruitful and precise branch of the study of linguistic meaning for the past half century; it has also been predicated on the idea that the role of meaning is to determine truth-conditions. This assumption, the truth-conditional semanticist might well complain, is not incidental to to the successes of truth-conditional semantics, but lie at their very core.

In order to fully counter this objection, I would have to show that speech-act-theoretic semantics can account for every phenomenon that truth-conditional semanticists have accounted for over the last several decades, at least as well as they have. In Chapters Two and Three, I argued that speech-act-theoretic semantics deals with some phenomena *better* than mainstream approaches, but there is a long list of expressions, constructions, and readings that I haven't begun to account for, and that extends long beyond the scope of a single dissertation. In this chapter, I will

tackle some more of what's on that list, and I'll attempt to do so in a wat that lays a solid methodological foundation so that more items to be easily ticked off in the future.

5.1 What is Compositional Semantics About?

The principal explananda of compositional semantics are the learnability, productivity, and systematicity of natural language—the facts that we can become semantically competent with indefinitely many sentences in a finite amount of time, the fact that we can be semantically competent with sentences that have never been uttered before, and the fact that our competence with some sentences will be systematically correlated to our competence with other sentences that contain the same parts or structures. Compositional semantics is the study of some facts about humans, our minds, or our behavior that must be posited to account for these phenomena. My question in this section is: *what are these facts?*

Mainstream semanticists have offered several different answers to this question, none of which I find particularly convincing. One account of the subject matter of compositional semantics, which is defended by Davidson (1965, 1967, 1973), is that a compositional semantic theory is what the user of a language (tacitly) *knows*, or is a rational reconstruction of what a language user could (tacitly) *know*, that would allow her to interpret the utterances of another user of the same language. The intuitive idea behind this formulation is that it explains how a speaker could know the meanings of an indefinitely large number of expressions in a productive and systematic way in virtue of knowing the meanings of their parts together with rules for their recursive combination. One peculiarity of Davidson's view is that he also holds that knowledge of a semantic theory would be neither necessary nor sufficient for the kind of linguistic competence that is the real target of his theory.

> Kurt utters the words 'Es regnet' and under the right conditions we know that he has said that it is raining. Having identified his utterance as intentional and linguistic, we are able to go on to interpret his words: we can say what his words, on that occasion, meant. What could we

> know that would enable us to do this? How could we come to know it? The first of these questions is not the same as the question what we *do* know that enables us to interpret the words of others. For there may easily be something we could know and don't, knowledge of which would suffice for interpretation, while on the other hand it is not altogether obvious that there is anything we actually know which plays an essential role in interpretation. The second question, how we could come to have knowledge that would serve to yield interpretations, does not, of course, concern the actual history of language acquisition. It is thus a doubly hypothetical question… (Davidson 1973: 125)

Given the double-dissociation that Davidson finds between actually knowing a semantic theory, on one hand, and being semantically competent, on the other, it is difficult to see how modeling such knowledge could explain what Davidson seeks to explain. Moreover, as Pettit (2002) has argued, *knowledge* and even *belief* seem to be the wrong categories to apply to whatever it is about speakers or their psychological states that grounds their semantic competencies or the meaningfulness of the expressions they use. After all: as I pointed out in §4.3, agents can communicate efficiently despite having justifiably very low credences about the specifics of their interlocutors' linguistic capacities.

Inspired by Chomsky's view that syntax is the study of a specialized cognitive faculty of language, some philosophers and linguists think that semantics aims to model the semantic component of a grammar that is subpersonally represented in the minds of speakers.[1] This internally represented grammar is part of the modular linguistic competence that guides speakers' language-using abilities.[2] If this is correct, then compositional semantics aims to describe—at a certain level of abstraction—an algorithm that each language user carries around in their mind, whose inputs are syntactic representations of sentences and representations of lexical meanings, and whose outputs are representations of sentences' meanings. I objected this view in §4.4, where I argued

---

[1] Versions of this view are defended or assumed by Borg (2004), Heim and Kratzer (1998), Larson and Segal (1995), and Partee (1995).

[2] On the idea of modularity, see Chomsky (1986) and Fodor (1983). It's worth noting that both Chomsky and Fodor deny that semantics could be the study of anything modular in this sense, in part on the grounds that what semantics studies is not informationally encapsulated or internal to individuals' minds in the right ways.

that the facts that ground linguistic meaning are both too protean and too sensitive to language users' personal-level mental states (such as their beliefs about whom they're talking to) to be constituted entirely by sub-personal or modular mental states.

According to Montague (1974), semantics is a branch of mathematics whose aim it is to map out the possible relationships borne by a certain class of formal structures called languages to the special classes of model-theoretic structures that "interpret" them. Taken at face value, it's hard to see how this vision of semantics could have anything to do with the language use of actual humans. In particular, it's hard to see how a purely mathematical pursuit could help us to explain the efficiency of linguistic communication, which is the explanatory role I carved out for semantics in §1.3. This is where Lewis's (1969, 1975a) ideas about conventionalism are supposed to come in, since they open the possibility that semantics could be the study of a mathematical structure which would count as the actual language of a community in virtue of the linguistic conventions in which the community members participate. One advantage of this overall strategy is that it makes compositional semantics the study of a public, interpersonal phenomenon, and this holds the promise of allowing semantic theories to play a role in explaining the efficiency of linguistic communication. But, as I argued in §§4.3–4.5, there is no notion of convention that can thus bridge the gap between Montagovian formal structures and human behavior.

On the theory I defended in §4.6, expressions are metasemantically connected to their semantic values by the overlapping communicative dispositions of the members of a speech community. To say that natural language is learnable, productive, and systematic is, on my view, just to say that speakers acquire a very large stock of communicative dispositions in a limited span of time, that they possess communicative dispositions with respect to sentences that neither they nor anybody else has ever uttered or encountered, and that the possession of a communive disposition with respect to one sentence is systematically connected to the possession of communicative dispositions with respect to other sentences with the same syntactic structures and containing the

same lexical items. The job of compositional semantics is to show how we can have communicative dispositions with these properties.

More specifically, a compositional semantic theory must show how the overlapping sentence-sized communicative dispositions of the members of a community are constituted by their overlapping word-sized communicative dispositions. To give the compositional semantics for a given sentence is to show how the communicative dispositions that link it to a certain type of communicative illocutionary act are constituted by (i) our dispositions to refer to certain kinds of objects, properties, and relations with the sentence's semantically primitive parts (and to interpret others as doing the same), (ii) our dispositions to perform certain kinds of combinatory acts of predicating with the types of binary-branching syntactic structures (and to interpret others as doing the same), and (iii) our dispositions to perform certain kinds of communicative illocutionary acts with sentences in certain moods.[3] By acquiring a finite stock of dispositions of types (i–iii), a person thereby also acquires a far larger stock of productive and systematic communicative dispositions that allow for the efficient performance and interpretation of a wide range of communicative illocutionary acts.

5.1.1 Referential and Predicative Dispositions

I want now to say more about the nature of the *referential* dispositions that underly the meanings of words and phrases and the *predicative* dispositions that underly the semantic import of certain syntactic arrangements. First off, what is it to have a disposition to refer to some object or property x with a word or phrase e? In Chapter Two, I endorsed a view of referring along the lines of

[3] There may be many other features of utterance types that may feature in how speakers are disposed to use and interpret them, and, if so, a compositional semantic theory will have to illuminate the semantics of those features as well. To name just one example: it seems clear that the intonational contour with which a sentence is uttered is semantically significant, in the sense that a speaker might be disposed to perform two different types of speech acts with two different utterance types that differ only with respect to intonation contour.

those defended by Bach (1987), Neale (*forthcoming a*), and Schiffer (1981), and according to which referring is grounded in the referential intention with which a speaker produces an utterance. Here's the partial account of referential intentions that I provisionally adopted in §2.2:

(RI) Referential Intention

S has a referential intention to refer to an object, property, or relation x with an utterance type e in producing an utterance u only if e is a proper part of an expression of which u is a token and there exists an x-dependent proposition p, and addressee A such that *S* produces u, thereby intending:

(1) to produce a state of entertaining p in A;

(2) A to recognize S's intention (1);

(3) A's act of entertaing p to be at least partly on the basis of her recognition of (1);

(4) A's act of entertaing p to be at least partly on the basis of A's sensitivity to the fact that, for some binary relation R, $R(e,x)$.

For S to refer to something with a particular expression requires that S possess a referential intention of this kind. But what is required for S to be *disposed* to refer to a certain object x with an expression e, such that this disposition can ground the meaningfulness of e? Quite generally, a *referential performative disposition* is one that would make an instance of the following schematic counterfactual true:

(RPD-X) Referential Performative Disposition for X

If S had a referential intention satisfying (RI), and, for a certain property F, F*x*, then the value of *e* in the instance of (RI) that could specify this intention might be X.[4]

The meanings of different expressions that we could plug in for X in this schema will covary with different values that might be plugged in for F. For example, if X is 'Fred', then F could be the property of being identical to Fred,[5] but if we plug 'I' in for X, then F will be the property of being identical to the speaker. A referential interpretive disposition is one that overlapped with (RPD-X) would be one that made an instance of the following schematic counterfactual true (holding the values of X and F constant):

(RID-X) Referential Interpretive Disposition for X

If A were to recognize that S had addressed an utterance of *u* to A, where u_e was a proper part of *u* that tokened X, then A would come to believe that S intended to produce an *x*-dependent thought in A, such that F(*x*).

Actual referential intentions are kinds of communicative intentions, and so are also intentions that one's addressee entertain some particular object-dependent proposition. So one can have a particular referential intention only in having a full-blown communicative intention. But one can be *disposed* to have a certain *type* of referential intention without being disposed to have any token of that type, just as a vase can be disposed to break when dropped without being disposed to break in any precise particular way. So a referential performative disposition, for example, will be one that links a certain expression to a general type of referential intention, individuated in terms

[4] I have used a 'might' counterfactual rather than a 'would' counterfactual because there will often be different expressions one could use to refer to a given object or property. For example, if someone has more than one name, then I might use one to refer to him on some occasions but another on other occasions.

[5] This assumes that names are directly referential and syntactically unstructured—both controversial assumptions.

of the value of F and *e* but not the values of *x* and *p*. I can be disposed to refer to males with 'he' without being thereby disposed to any one particular male, or to express any particular proposition.

Predicative dispositions, on my view, are the dispositions in virtue of which referential dispositions "add up" to full-blown communicative dispositions. A predicative disposition is a disposition to refer to a certain kind of complex property or express a certain kind of proposition using an expression with a certain syntactic structure. I can't see how to give a general counterfactual schema such that instances of it would be made true by all predicative dispositions, but here is an example that, hopefully, will get the general idea across:

(PPD-S) Predicative Performative Disposition for [$_S$ NP VP]

If (a) S had a referential intention r_1 to refer to an object *x* with NP, (b) S had a referential intention r_2 to refer to a property F with VP, and (c) S's utterance *u* was a token of a sentence type whose syntactic structure is [$_S$ NP VP], then r_1 and r_2 would be components of S's communicative intention to produce a thought that *x* if F in an addressee.

In §5.2, I will construct a version of speech-act-theoretic semantics on which every type of binary-branching node will be semantically significant in a different way and will therefore require its own compositional principle. Each one of these principles will be grounded in predicative dispositions like (PPD-S) and in overlapping interpretive dispositions.

Much more would have to be said about the nature of referential intentions and communicative dispositions in order for the result to constitute a full account of the subject matter of compositional semantics, but I hope to have at least given a general picture of how a speaker's stock of productive and systematic communicative dispositions could be grounded in a finite stock of referential and predicative dispositions.

5.1.2 Sentence-Sized Acts and Word-Sized Acts

One question that may be invited by my talk of word-sized speech acts *adding up* to sentence-sized speech acts, or of referential and predicative dispositions *adding up* to communicative dispositions, is the question of which of these speech acts and dispositions are more explanatorily fundamental. The answer is somewhat nuanced.

Given the Gricean picture of referring that I sketched in §2.2 and §5.1.1, a referential intention is always a kind of communicative intention, and so the notion of referring is not explanatorily prior to the notion of speaker meaning (or, equivalently, to the notion of performing a communicative illocutionary act). If word-sized speech acts aren't more basic than sentence-sized speech acts, how is it that the former could add up to the latter in any interesting sense?

The answer is that compositional semantics is not the study of how complex speech acts are built out of more basic, independently performable speech acts. Rather, a compositional semantic theory specifies the more specific features that define a given illocutionary act-type α in terms of the less specific features that define the referential and predicative act-types that one also tokens in tokening α. Equivalently: compositional semantics specifies the specific types of communicative intention that are the semantic values of complex expressions in terms of the less specific types of communicative intention that define the semantic values of the complex expressions' parts and structures.

The purpose of thus specifying specific types of communicative act in terms of less specific types of communicative act is to show how the communicative dispositions linking meaningful sentences to their semantic values are grounded in the referential and predicative dispositions linking words and syntactic structures to their semantic values, respectively.

So, to sum up: acts of referring aren't explanatorily or metaphysically prior to communicative illocutionary acts, and it is not possible to refer independently of performing an illocutionary act in the service of which one is referring. But, our referential and predicative dispositions are meta-

physically more fundamental than the full-blown communicative dispositions underlying the meaningfulness of sentences.

5.2 The Syntax–Semantics Interface

So far, I have considered only how to give the semantics for a few, syntactically very simple expressions. I want now to consider how to extend this fragment a bit further, and, in the process, develop a few more tools that will allow for more development still. All of this revolves around the syntax–semantics interface, and so I will start by reviewing the contemporary orthodoxy about that interface as it is codified by Heim and Kratzer (1998).

5.2.1 Truth-Conditional Semantics and the Minimal Interface Principle

Following the work of Klein and Sag (1985), truth-conditional semantics has come to be regulated by the ideal of what I'll call a *minimal* syntax–semantics interface. A semantic theory has a minimal interface with syntax to the extent that its compositional principles are sensitive to few aspects of sentences' phrase structures. A fully minimal interface would be one such that compositional principles would be sensitive to only the most superficial tree-geometric syntactic features. This is how Heim and Kratzer (1998) put it:

> The semantic interpretation component, then, can ignore certain features that syntactic phrase structure trees are usually assumed to have. All it has to see are the lexical items and the hierarchical structure in which they are arranged. Syntactic category labels and linear order are irrelevant. (144)

The general methodological principle can be summed up as follows:

> *Minimal Interface Principle*
>
> The mode of composition at a given node should be determined solely by how many daughters the node has and by the semantic properties of its daughters, not by any further syntactic features of the node or its daughters.

Like most other contemporary semanticists, Heim and Kratzer implement the Minimal Interface Principle by adopting a form of *type-driven interpretation*. In a type-driven semantic theory, expressions' semantic values are either entities (type e), truth values (type t), or functions built up from objects of these two types. A verb phrase's semantic value is a function from entities to truth values (type ⟨e,t⟩), for example, and a quantifier phrase's semantic value is normally taken to be a type ⟨⟨e,t⟩,t⟩ function from type ⟨e,t⟩ functions to truth values. This allows for a (mostly) minimal syntax-semantics interface because it opens up the possibility that semantic composition at binary branching nodes is a matter of mere functional application: the semantic value of any expression α with two daughters β and γ is either the value of β's semantic value given γ's semantic value as an argument or the value of γ's semantic value given β's semantic value as an argument. As Heim and Kratzer put it: "it's the semantic types of the daughter nodes that determine the procedure for calculating the mother node" (1998: 44). It is the power of this functional typing system that allows truth-conditional semantics to (mostly) get by without access to the fine-grained syntactic features of expressions.

In practical terms, the Minimal Interface Principle shows up in Heim and Kratzer's semantics as a scarcity of compositional principles. They begin with just three:

Terminal Nodes (TN)

If α is a terminal node, ⟦α⟧ is specified in the lexicon.

Non-Branching Nodes (NN)

If α is a non-branching node, and ⟦β⟧ is its daughter node, then ⟦α⟧ = ⟦β⟧.

Functional Application (FA)

If α is a branching node, {β, γ} is the set of α's daughters, and ⟦β⟧ is a function whose domain contains ⟦γ⟧, then ⟦α⟧ = ⟦β⟧(⟦γ⟧). (Heim and Kratzer 1998: 43–4)

It appears to have been the possibility of avoiding a proliferation of compositional principles, on one hand, and, perhaps more dimly, the prospect of a highly constrained syntax-semantics interface, on the other hand, that has led to the rise of the Minimal Interface Principle as a guiding methodological tenet of contemporary semantics. Klein and Sag, who originally advocated type-driven interpretation, justified their proposal by appeal to parsimony, saying that the approach "allows redundancy to be eliminated, and linguistically significant generalizations to be expressed" (1985: 163). Heim and Kratzer endorse the approach on more or less the same grounds (1998: 44–45).

I say that Heim and Kratzer *begin with* just three compositional principles and wind up with a *mostly* minimal interface because they ultimately add "one or two additional principles to the above list" (1998: 44). The principles they add violate the Minimal Interface Principle by depending for their application on fine-grained syntax. Heim and Kratzer make it clear, however, that these additions should be considered deviations from their professed methodology:

> ...we will strive to keep [the list of compositional principles] as parsimonious as possible. When we look at a new construction for which we don't yet have a semantic analysis, we always try first to accommodate it by adding only to the lexicon. (1998: 44).

This is why I say that the Minimal Interface Principle is a methodological one—because although few contemporary semantic theories ultimately exemplify the principle to the letter, most such theories exemplify it in spirit, in the sense that they are shaped by it in a large number of the theoretical choices they embody.[6]

Contemporary truth-conditional semantics, it might be said, views natural language sentences as encoding the "structure and sequence" of applications of functional application, which is the

---

[6] Some semanticists have pursued a minimal syntax-semantics interface farther than Heim and Kratzer—for example, by pushing even more complicated semantic machinery into the lexicon and by adopting type-shifting principles. I'll briefly discuss the actic analogue of type-shifting toward the end of §3. For an overview of type-shifting in contemporary truth-conditional semantics, see Pylkkänen and McElree (2006: §3).

primary way in which the semantic values of complex expressions are determined in terms of their parts.

5.2.2 The Unity of Linguistic Meaning

I will turn now to a brief comparison of speech-act-theoretic semantics and another variety of act-theoretic semantics, versions of which are defended by Hanks (2007, 2011, 2012, *forthcoming*) and Soames (2010). As in speech-act-theoretic semantics, Hanks and Soames identify the semantic values of expressions with types of acts. The most notable differences between my view and those of Hanks and Soames are that (a) both Hanks and Soames think that the act types that serve as semantic values can be tokened by mental acts rather than (Soames) or in addition to (Hanks) speech acts, and that (b) both authors' commitments to their versions of act-theoretic semantics spring from their commitments to act-theoretic accounts of the metaphysics of propositions. Specifically, both Hanks and Soames hold that a proposition is a type of complex act made up of simple acts of referring and predicating. Together with the view that sentence-sized semantic values are structured propositions whose constituents are the semantic values of the sentence's parts—a proposal influentially defended by Soames (1987, 2002), which that both Soames and Hanks accept and which I rejected in Chapter Two—an actic theory of propositions entails an actic approach to semantics.

Both Soames and Hanks embrace the semantic consequences of their theories of propositions. For Soames, the semantic values of words and phrases are types of acts of "thinking of" objects, properties, and relations, and the semantic values of sentences, which he identifies with acts of

entertaining the propositions expressed with by sentences, are types of complex acts of predicating properties and relations of one another and of objects.[7]

> The proposition expressed by ['6 cubed > 14 squared'] ... is the complex event type of (i) thinking of the cubing function f3 and the number 6, and applying the former to the latter, (ii) thinking of the squaring function f2 and the number 14, and again applying the former to the latter, and (iii) predicating the relation being greater than of the result of applying f3 to 6, followed by the result of applying f2 to 14. By contrast, the proposition expressed by ['216 > 196'] is the event type of thinking of the numbers 216 and 196, and predicating being greater than of the former, followed by the latter. Since these event types are different, the propositions are different. (Soames 2010: 114)

Hanks holds a similar view, on which names' semantic values are types of acts of referring to objects, predicates' semantic values are types of acts of expressing properties and relations, and the semantic value of a sentence is a complex combinatory act performed on the relevant objects, properties, and relations.

> The proposition expressed by the sentence 'George is clever' is, I propose, a type of action a speaker performs when she asserts that George is clever. One component of this type is a type of act of referring to George. This type of reference act is the semantic contribution of the name 'George' to the proposition expressed by 'George is clever'. (Hanks 2011: 11)

Hanks agrees with Soames that the semantic value of a declarative sentence is a complex act of predicating, although they disagree about non-declarative sentences in ways that I discussed in §3.3.3.

---

[7] After initially suggesting an actic theory, Soames (2010) argues that we should identify propositions with types of *cognitive events* rather than with types of mental acts. He does this because he thinks that it sounds strange to say that the proposition that Scott philosophizes (for example) is something I did after midnight, or that something I'm doing right now can be either true or false. As Neale (2011) has pointed out, however, a view that identifies propositions with types of cognitive events will force us to say similarly strange things, for example that a token of the proposition that Scott philosophizes just happened, or that something true just took place in my mind. For this reason, Soames has returned to act-theoretic talk (forthcoming). In any case, my arguments here apply equally well if predication is conceived of as a *combinatory event* rather than as a combinatory act.

Both Soames and Hanks argue for their act-theoretic accounts of propositions, and, by consequence, for their act-theoretic brands of semantics, on the grounds that these theories offer a solution to the problem of the unity of the proposition.[8] This problem has often been framed as the challenge of saying what it is for a structured proposition to be a genuine unity, over and above the mere collection of its components. As Soames argues, however, the real problem lies in saying what it is in virtue of which a given structured proposition represents things as being one way rather than another—what it is in virtue of which a given proposition forms a genuine *representational* unity. What relations could the angled brackets or quasi-phrase-structural trees of typical structured proposition notation represent, such that the the proposition <Fred, <LOVE, Mary>>, for example, is true just in case Fred loves Mary (and not, say, just in case Mary loves Fred)? Soames argues that the only good answer is that these notations

> encode the structure and sequence of cognitive acts of predication that are necessary and sufficient for entertaining the real propositions these abstract structures represent—where entertaining such a proposition is performing the acts of predication involved in tokening the event type that it is. (2010: 106)

A tempting, if perhaps oversimplified, way of describing Soames' and Hanks' conceptions of semantics would be to say that natural languages are just different (and presumably much richer) systems of notation for representing the structured acts we would have to perform in order to token the propositions expressed by natural language sentences. Predication thus plays a role for Soames and Hanks that is similar to the role played by functional application in truth-conditional semantics—that of the "compositional glue" that combines primitive semantic values into complex semantic values.

I can't use considerations about propositional unity to motivate speech-act-theoretic semantics because I deny that propositions and their components are the right sorts of things to serve as

---

8 For some other recent work on this problem, see King (2007, 2009, *forthcoming*) and Collins (2011).

expressions' semantic values (see §2.2). But given my view that an expression's semantic value is the type of speech act that we are disposed to perform with it, a similar concern about the unity of speech acts (and speech-act types) forces me to give predication the same sort of combinatory role that Soames and Hanks give it. The basic idea of the conception of compositional semantics that I have defended in this chapter is that the acts I perform in uttering a sentence's parts add up to the complex act I perform in uttering the sentence. But what is it that makes the act I perform with the sentence 'Rob drives' a single, complex act of saying that Rob drives, over and above the series of acts I perform in uttering 'Rob' and then 'drives'? The answer is that, in uttering 'Rob drives', I *predicate* the property of driving of Rob.

There is also a speech-act-theoretic analogue to Soames' question about why propositions form *representational* unities. Why is the act that I perform with 'Mary loves Fred' an act of saying that Mary loves Fred rather than, for example, an act of saying that Fred loves Mary? More broadly: what is it about the acts we perform with a sentence's parts in virtue of which they add up to a communicative illocutionary act of one kind rather than another?

Truth-Conditional Semantics must answer an analogous question: why is 'Mary loves Fred' true just in case Mary loves Fred and not just in case Fred loves Mary? The truth-conditional semanticist's answer has to do with how sentences' syntactic structures guide functional application up a sentence's phrase structure, step-by-step. What we normally think of as the binary *love* relation is actually bifurcated, so that it can be applied to its arguments in two steps: the semantic value of 'loves' is a function that outputs the property of loving *x* when given an object *x* as an argument. It is thus the property of loving Fred that composes with Mary. Functional application occurs independently at each node, before the result is passed up the tree. Composition is *local*.

An act-theoretic semanticist must say something similar about how word-sized acts compose into sentence-sized acts. We perform the act of expressing the proposition that Mary loves Fred by referring to Mary and predicating the property of loving Fred of her. The act of referring to the

property of loving Fred is itself a complex one that consists of acts of referring to Fred and combining the *love* relation with him. I will assume that this latter sort of combinatory act—the sort by means of which we combine relations with their second relata—is itself a kind of predication. That is: the act of referring to the complex property of loving Fred itself consists in the act of predicating the love relation of Fred.

This usage of the verb 'predicate' is a technical one. The word's colloquial usage—or the usage that philosophers and grammarians are used to, anyway—picks out acts of the kind that we perform with predicates (i.e., verb phrases) on the referents of subjects (i.e., noun phrases). Nonetheless, I think my use of the word to pick out sub-sentential combinatory acts more generally is theoretically justified. First: we don't have a nontechnical term for the kind of combinatory acts involved in performing complex subsentence-sized speech acts, and so *some* technical term is needed. Second: by saying that the combinatory acts whose structure is mapped out by the structures of natural language sentences are all acts of predicating, we can draw the sort of explicit parallel between functional application in truth-conditional semantics and predication in actic semantics that suits my purpose in this section. This is ultimately a matter of terminological convenience: all that my argument requires us to recognize is that speech-act-theoretic semantics must posit some type of combinatory act corresponding to each branching node in a sentence's phrase structure—an act by means of which tokens of the node's daughters' semantic values are compositionally combined into tokens of the node's semantic value.

The following generalization thus emerges: speech-act-theoretic semantics associates complex acts with syntactically complex expressions by associating referential acts with semantically primitive expressions and different kinds of acts of predicating with the different binary-branching structures into which the simple expressions are organized.

5.2.3 The Act Mismatch Problem and Two Solutions

I now turn to the question of whether speech-act-theoretic semantics (or, by parity of reasoning, any act-theoretic semantics) can obey the Minimal Interface Principle. To make the case, I'll attempt to construct a simple speech-act-theoretic semantic fragment that mirrors the structure of type-driven semantic theories like Heim and Kratzer's.

As I have already shown, type-driven theories succeed in obeying the Minimal Interface Principle—or come reasonably close, anyway—only because they pack most of their semantic machinery into expressions' semantic values. For a semantic theory to be type driven means that composition proceeds in a way that is predetermined by the functional types of its lexical entries in a way that depends only on a bare minimum of syntactic structure. Because predication plays the analogoue role in speech-act-theoretic semantics that functional application plays in truth-conditional semantics, the only hope of constructing an actic theory that mirrors these features is to pack predication itself into expressions' semantic values in such a way that what gets predicated of what at a given node does not depend on the node's fine-grained syntactic structure.

We are thus presented with a basic choice about how to structure an actic semantic theory:

Syntactically Driven Predication

What gets predicated of what at a given branching node α depends on α's fine-grained syntactic structure.

Lexically Driven Predication

What gets predicated of what at a given branching node α does not depend on α's fine-grained syntactic structure, but is determined by facts about α's daughters semantic values.

Trivially, Syntactically Driven Predication is incompatible with the Minimal Interface Principle, and so it seems that we must explore the possibility of building a semantic theory that implements Lexically Driven Predication. One thing that speech-act-theoretic semantics would have to do in

order to implement Lexically Driven Predication is to locate the type of act of predication that occurs at a given node in the semantic value of one of the node's daughters. This mirrors truth-conditional semantics, which locates the function that gets applied at a given node in one of the node's daughters. This means that implementing the Minimal Interface Principle would require identifying the semantic value of a verb phrase (for example) with an act of predicating.

Different act-theoretic semanticists differ on this point. Hanks takes a clear stand in favor of Syntactically Driven Predication, arguing that predication is the semantic contribution of a clause's mood (2007, 2011). It is less clear where Soames stands with regard to the two principles, but the fact that he distinguishes the act of predicating a property from the act of "thinking of" it, rather than collapsing the two into a single act, suggest that Syntactically Driven Predication is at least an open option for him, since he can hold that a verb phrase's semantic value is an act of thinking of a property, and that the act of predicating is supplied by a syntactically driven compositional principle. In their discussions the prefigure the version of speech-act-theoretic semantics I've defended here, both Neale and Grice seem to build predication into the semantic values of verb phrases in ways that suggest Lexically Driven Predication. Neale takes a verb phrase's meaning to be a matter of *its own* potential for the performance of an act of predication, thereby suggesting that verb phrases' semantic values are acts of predication (*forthcoming b*). Along similar lines, Grice argues that the meaning of a predicative adjective boils down to speakers' shared procedure for predicating some property with it (1989: 131), thereby at least suggesting the idea of identifying predicates' semantic values with types of acts of predicating.

Aside from sub-sentential semantic values rich enough to drive semantic composition in a way that isn't steered by fine-grained syntax, the other ingredient that allows contemporary semantic theories to (almost) obey the Minimal Interface Principle is a single, general purpose compositional principle for branching nodes—one that is broad enough to apply at each such node in a way that takes advantage of the rich compositional potential of the node's daughters, but

without needing any information about the node's fine-grained syntactic structure to do so. For Heim and Kratzer, this principle is (FA). A speech-act-theoretic semantics with a minimal interface needs a principle that is analogous but that replaces functional application with predication:

Daughter-Driven Predication (DDP)

For any expression α with β and γ as daughters, the semantic value of α is either (i) an act-type tokenable by performing a token of β's semantic value on the referent of a token of γ's semantic value, or (ii) an act-type tokenable by performing a token of γ's semantic value on the referent of a token of β's semantic value

Just as (FA) can deliver the semantic value of a node only if the node has one daughter whose semantic value is a function whose domain contains the other daughter's semantic value, (DDP) can deliver a node's semantic value only if one daughter's semantic value is an act-type of predication and the other daughter's semantic value is some act-type that can be tokened by referring to suitable objects of predication. For the sake of uniformity and terminological convenience, I'll call an object that a speaker refers to in performing an act of referring the act's *referent*. So, in order to compose, expressions whose semantic values are tokened by acts of predicating must have sisters whose semantic values are types that are tokened by acts of referring.

All of this works out very nicely when it comes to the composition of nodes with [$_S$ NP VP] structure. (DDP) allows us to derive the conclusion that the semantic value of 'Rob drives' is an act-type of saying that Rob drives, for example, by positing the following semantic values for 'Rob' and 'drives':

'Robert' $\Rightarrow$ referring to Robert [9]

[9] For clarity, I will leave out 'a type tokenable by performing an act of…' at the start of my specification of each expression's semantic value.

‘drive’ $\Rightarrow$ predicating the property of driving (of some object, thereby expressing the proposition that it drives)

Just as the functional types of semantic values in truth-conditional semantics specify both the sorts of things they can act on (the functional types of the entities in their domain) and the sorts of things they output (the functional types of the entities in their range), the act-type of predicating that an actic semanticist identifies with a predicative expression’s semantic value must specify the sorts of entities that can be its objects and the sort of act that an act of that type, aimed at a suitable object, adds up to. This information is given parenthetically in the above specification of the semantic value of ‘drive’.

I could go on to posit semantic values for more names, verbs, and adjectives, including transitive verbs (by appealing to the sort of considerations sketched in the previous section). The following is a good first pass at the semantic value of the transitive verb ‘love’, for example:

‘love’ $\Rightarrow$ predicating the relation of loving (of some object, thereby predicating the property of loving that object)

But now I want to cut right to a big problem case: quantifier phrases. For now, I won’t worry about what goes on within quantifier phrases; I want to focus on the semantic values of quantifier phrases themselves. Take the DP, ‘every parent’, for instance. It might at first be tempting to say that we use this phrase to refer to every parent, but it has been clear since Frege and Russell that we don’t use quantifier phrases to refer to the objects over which they quantify. That lesson isn’t completely obvious for ‘every parent’, but it is much more obvious when it comes to ‘some parent’ (*cf*. Russell’s remarks about the idea that ‘some man’ denotes “an ambiguous man” (1905)), and the lesson could not be clearer once we consider ‘no parent’. We must use such phrases to do something else.

Once again, contemporary truth-conditional semantics can serve as our guide. Following Frege, Montague, and proponents of generalized quantifier theory, it is now typical to think of quantifier phrases' semantic values as properties of properties. Truth-Conditional Semanticists, including Heim and Kratzer, identify these properties with type-⟨⟨e,t⟩,t⟩ functions—functions from type-⟨e,t⟩ functions (which serve as the semantic values of verb phrases) to truth values. The analogous actic move would be to think of the semantic values of quantifier phrases as types of acts of predicating properties of properties. For example:

> 'every parent' ⇛ predicating the property of applying to every parent (of some property, thereby expressing the proposition that it applies to every parent)

To express the proposition that every parent drives, for example, is to say of the property of driving that it applies to every parent—i.e., to predicate the property of applying to every parent of the property of driving.

But now the foreshadowed problem comes into focus. By our earlier hypothesis, the semantic value of 'drives' is itself an act of predicating the property of driving, not the act of referring to the property of driving. This gives rise to an *act mismatch*: a pair of sister expressions whose semantic values are act-types that don't compose. As reflected in our composition rule (DDP), a predicative act-type φ can compose only with a referential act-type; otherwise the tokens of φ will be infelicitous acts of predicating without objects. But in this case, neither expression's semantic value is an act of referring whose referent can serve as the target of the act of predication that is the other expression's semantic value. Instead, the daughters of 'every parent drives' are both types of acts of

predicating, neither of whose tokens has any object to be performed on. Composition breaks down.[10]

This problem is vaguely reminiscent of the well-known type mismatch that arises in contemporary truth-conditional semantics when a quantifier appears in a transitive verb's object position.[11] Despite this superficial similarity, the two problems are quite distinct in both their scope and their causes. The act mismatch turns out to pose a problem that is both ubiquitous in natural languages and open to a uniform diagnosis, and it seemingly arises for any version of speech-act-theoretic semantics designed to have a minimal syntax–semantics.

The act mismatch problem arises due to a fundamental disanalogy between functional application and predication that prevents the two from playing structurally similar roles in semantic theories. Type-driven truth-conditional semantic theories like Heim and Kratzer's depend crucially on the ability of unsaturated functions—functions that haven't themselves been given arguments—to serve as the arguments of other functions. Without this feature, type-driven semantics would break down at any node both of whose daughters' semantic values are functions. These are precisely the nodes at which act mismatches arise in act-theoretic semantics. This is because acts of predicating can't be performed without an object on which to perform them. Unless one predicates a property *of something*, one hasn't really predicated it at all. This is part of what it means to say that predication is a combinatory act. Unlike an unsaturated function, which can still serve as the argument of another function, an act of predication with no object can't serve as the object of another act of predicating—because there can be no such act. This is the fundamen-

---

[10] At first glance, it looks as though we could solve this problem by saying that if expression has an act-type of predicating a property ψ as its semantic value, then we can consider the expression's referent to be ψ. But this doesn't solve the problem because we would still need our compositional principles to tell us what gets predicated of what in act mismatch cases. As I will argue in my discussion of act shifting, this *what gets predicated of what?* problem can't be solved without appealing to fine-grained syntactic structure.

[11] See Heim and Kratzer (1998: ch.7).

tal disanalogy between functional application and predication, and it is what explains the ubiquity of act mismatches that arise in any actic theory that aspires to a minimal interface. Specifically: an act mismatch will arise in such a theory at any branching node neither of whose daughters would be given type-e or type-t semantic values by a truth-conditional semantic theory. In short: this problem has a totally different scale from the problem posed by type mismatches in truth-conditional semantics.

Can the act-mismatch problem be solved by any of the strategies that have been developed to deal with the type mismatch problem in truth-conditional semantics? Most such strategies wouldn't work even if a suitable act-theoretic analogue could be found, because they are designed to deal with type mismatches only within the much narrower scope in which they appear in truth-conditional semantics.[12] The only technique for type mismatch resolution that is both amenable to an actic reinterpretation and easy to generalize is the strategy of employing type shifting principles to coerce semantic values into playing nice with one another.[13] By analogy, we might try to resolve our act mismatch by developing a system of *act-shifting* rules, which would systematically adjust the types of acts serving as expressions' semantic values in such a way as to ensure that every act of predicating is met with an act of referring with which it can compose. What would be needed from such a system of rules seems quite simple: for every expression whose daughters' semantic values are both predicative act-types, one of the two must be transformed into a suitably corresponding referential act-type. For example: in order for (DDP) to apply to 'every parent drives', we would need our act shifting principles to transform the semantic value of

---

[12] Two examples are Montague's (1973) treatment of all NPs as quantifier phrases and the later idea of raising quantifiers at LF (May 1977, 1985, 1987; Neale 1993). Both of these strategies are designed to deal only with type mismatches involving quantifier phrases. Even if an actic analogue could be devised, it is unclear how they could be generalized to help us with act mismatches in their ubiquity. (This isn't to say that an act-theoretic semanticist would therefore have no other reasons to posit quantifier raising; it just won't help to solve the present problem.)

[13] See Partee and Rooth (1983) and Partee (1986) for the original formulations of type shifting, and Pylkkänen and McElree (2006: §3) for a contemporary overview of the technique.

'drives' from an act-type of predicating the property of driving into an act-type of referring to the property of driving.

But when we try to adopt act shifting principles that implement this strategy, we run into a difficult problem. For every expression α, both of whose daughters β and γ have act-types of predication as their semantic values, our act-shifting principle(s) would have to correctly choose which daughter's semantic value to shift. Suppose that the semantic value of β is an act of predicating the property φ^{β} and the semantic value of γ is an act of predicating the property φ^{γ}. Effectively, then, our act shifting principles must decide whether α's semantic value is constituted by an act of predicating φ^{β} of φ^{γ}, or an act of predicating φ^{γ} of φ^{β}. Moreover, they must make this choice without appealing to the fine-grained syntactic features of α, β, or γ, lest our attempt to construct a semantic theory that is blind to fine-grained syntax collapse. Since the semantic values of β and γ are both acts of predicating properties, the only non-syntactic features that could distinguish them in the eyes of our act-shifting principles would have to be either (A) features the properties themselves—something about the difference between φ^{β} and φ^{γ}—or (B) some feature, other than the properties being predicated, that distinguished the kinds of acts of predicating serving as the semantic values of β and γ. I will consider each of these possibilities in turn.

According to possibility (A), whenever our semantics encounters a node α, whose daughters β and γ have act-types of predicating the properties φ^{β} and φ^{γ}, respectively, there must be a feature of either φ^{β} or φ^{γ} that allows our act-shifting rules to decide which act-type to shift to an act-type of referring. But there is no such feature that will work. To see that this is the case, consider the fact that, for any property, we can *always* construct a verb phrase with which to predicate the property. Importantly for my purposes here, this includes properties of the sort that we normally predicate with quantifier phrases. By the lights of the theory I have been exploring in this subsection, for example, the semantic value of 'every parent' is a type of act of predicating the property of applying to every parent. But by the lights of the same theory, this same property is also the

semantic value of the verb phrase, 'applies to every parent'. What we need is a principle that, when confronted with a sentence consisting of a quantifier phrase followed by a verb phrase, will shift the semantic value of the verb phrase but not the semantic value of quantifier phrase, and that will do this solely on the basis of differences between the properties normally predicated with the two phrases. But if quantifier phrases and verb phrases can be used to predicate the same properties, then no act shifting principle can possibly accomplish this task.

The following pair of sentences illustrate the problem:

(5) Every property applies to every property.

(6) Some property applies to some property.

First, notice that these sentences are perfectly intelligible. Indeed, (5) is false while (6) is true. But, by the lights of the semantic theory I've been working with, each of these sentences is made up of a quantifier phrase and a verb phrase with the same semantic value—the act-type of predicating the property of applying to every property in (5), and the act-type of predicating the property of applying to some property in (6). Clearly, then, no act-shifting principle could tell us to shift these sentences' verb phrases rather than their quantifier phrases on the basis of information about the properties one predicates with them alone. Instead, an act shifting principle would have to discriminate these two sentences' daughters on the basis of their syntactic categories.

It might be objected that it doesn't matter what gets predicated of what in (5) and (6), since the truth conditions work out the same either way. We might therefore formulate our act shifting principle(s) so that they choose which phrase's semantic value to shift at random in such cases. But the following example illustrates that this strategy can't work.

(7) Every object is abstract.

Our theory predicts that the semantic value of 'every object' is an act-type of predicating the property of *applying to every object*, and that the semantic value of 'is abstract' is an act-type of predicating the property of *being abstract*. To deliver the correct reading of (7), on which it is false, an act shifting principle would have to transform the semantic value of 'is abstract' into an act-type of referring to the property of being abstract. If it were to transform the semantic value of 'every object' instead, then our semantic theory would predict that (7) has a reading on which we use it to say that *the property of applying to every object is abstract.* Although this reading would be perfectly coherent—indeed, it would be true—it is a reading that (7) simply does not have. It's quite clear from examples like this one that our hypothetical act shifting principle(s) would have to consistently shift the semantic values of VP's in sentences with [$_S$ DP VP] structure.

According to possibility (B), whenever our semantics encounters a node α whose daughters β and γ have act-types of predicating the properties φ^{β} and φ^{γ}, respectively, there must be a feature of the semantic values of β or γ other than features of φ^{β} or φ^{γ} that allows our act-shifting rules to decide which act-type to shift to an act-type of referring. What could these other properties be? One possibility is that acts of predicating themselves are typed, much in the same way that the functional semantic values of truth-conditional semantics are typed. On this view, we could specify an expression *e*'s semantic value by specifying both the object or property that the expression is used to refer to or predicate, but also the type of act it is as well. We could borrow our notation for these act-types from truth-conditional semantics, even if they wouldn't mean quite the same things in this case (since we aren't dealing with functional types).

'Robert' $\Rrightarrow$ referring$_{\mathbf{e}}$ to Robert

'drive' $\Rrightarrow$ predicating$_{\langle \mathbf{e,t} \rangle}$ the property of driving (of some object, thereby expressing$_{\mathbf{t}}$ the proposition that it drives)

'every parent' $\Rrightarrow$ predicating$_{\langle\langle \mathbf{e,t}\rangle,\mathbf{t}\rangle}$ the property of applying to every parent (of some property, thereby expressing$_{\mathbf{t}}$ the proposition that it applies to every parent)

We could then formulate an act-shifting principle according to which, whenever two acts of predicating meet, the act whose type is lower is shifted to become an act of referring. This allows us to deal with (5) because the semantic value of the DP 'every property' is an act-type of predicating$_{\langle\langle \mathbf{e,t}\rangle,\mathbf{t}\rangle}$ the property of applying to every property, whereas the semantic value of the VP 'applies to every property' is an act-type of predicating$_{\langle \mathbf{e,t}\rangle}$ the property of applying to every property, and so our act-shifting rule can convert the VP's semantic value into an act of referring to the property of applying to every property on this basis.

This way of typing acts of predicating not in terms of their properties opens up the possibility of giving a speech-act-theoretic semantics that is structurally isomorphic to mainstream truth-conditional semantics, but with the peculiar result that the typed functions used by truth-conditional semanticists to represent properties' semantic values would correspond, not to the properties we refer to and predicate with expressions, but to different types of acts of referring and predicating themselves.

It may, therefore, be possible to build predicating into the semantic values of non-referential expressions, and I plan to continue to explore this possibility in future work. Nonetheless, in the next section, I will continue to explore the version of speech-act-theoretic semantics that I have been assuming since Chapter Two, and according to which different types of acts of predicating are the semantic values of different kinds of binary-branching syntactic arrangements. I call this way of structuring a theory *top-down speech-act-theoretic semantics*.

5.2.4 Top-Down Speech-Act-Theoretic Semantics

The big difference between the version of speech-act-theoretic semantics I'll explore here and the version that I explored in the previous subsection is that this theory supplies combinatory acts of

predication by means of an array of syntactically-driven compositional principles. A helpful way to understand this distinctive feature is to think of predication as being the semantic contributions of syntactic structures rather than those of particular words or phrases.

I'll begin by tackling the sorts of expressions I've looked at so far.

'Robert' ⇛ referring to Robert

'Mary' ⇛ referring to Mary

'Fred' ⇛ referring to Fred

'drive' ⇛ referring to the property of driving

'love' ⇛ referring to the love relation

'parent' ⇛ referring to the property of being a parent

'every' ⇛ referring to the *every* relation[14]

The common theme is reference. What we do with a words and phrases, according to the theory I am proposing, is refer to things, properties, and relations.[15]

I said that words *and phrases* have act-types of referring as their semantic values. I have in mind, for example, that the semantic value of 'loves Fred'—a complex verb phrase—is an act-type of referring to the complex property of loving Fred. This interpretation of phrases gives uniformity to the theory, and it gives a consistent sort of target for compositional principles to aim at,

---

[14] By the *every* relation, I mean the relation that one property bears to another just in case every possessor of the first property possesses the second property.

[15] Some act-theoretic semanticists who more or less agree with the spirit of my proposal would object to the uniformity of this formulation. Soames talks of "thinking of" where I talk of referring, and Hanks talks of "referring" with names and "expressing" properties and relations with verbs. Although I have certainly departed from or extended the colloquial sense of 'refer' in my usage, I can see no reason to think that this results in more than a terminological disagreement with Soames and Hanks.

both in what sort of acts they combine and in what they deliver. The following four compositional principles generate semantic values for all of the sorts of sentences I've considered so far.

Predication for [$_{S}$ **[NP][VP]**]

S's semantic value is an act-type of expressing a the proposition that *x* is F that may be tokened by referring to *x* in a way that tokens NP's semantic value, referring to F in a way that tokens VP's semantic value, and predicating F of *x*.

Predication for [$_{VP}$ **[V^{t}][NP]**]

VP's semantic value is an act-type of predicating the property of bearing R to *y* that may be tokened by referring to R in a way that tokens V^{t}'s semantic value, referring to *y* in a way that tokens NP's semantic value, and predicating R of *y*.

Predication for [$_{S}$ **[DP][VP]**]

S's semantic value is an act-type of expressing the proposition that F is G that may be tokened by referring to G in a way that tokens DP's semantic value, referring to F in a way that tokens VP's semantic value, and predicating G of F.

Predication for [$_{DP}$ **[DET][VP]**]

DP's semantic value is an act-type of predicating the property of bearing F to G that may be tokened by referring to F in a way that tokens DET's semantic value, referring to G in a way that tokens VP's semantic value, and predicating F of G.

The basic idea here, as in §2.3, is that we express propositions with sentence-radicals, refer to objects, properties, and relations with words and phrases, and predicate properties and relations of each other and of objects with syntactic structures. This small fragment can be built into a larger, recursive one with the addition of compositional principles for mood and logical connectives like

those I gave at the end of Chapter Three. The scope of this theory fragment is still a far cry from Heim and Kratzer's, but it shows promise.

5.2.5 Too Many Intentions? Too Many Speech Acts?

According to the theory I've laid out, a speaker who does the usual thing in uttering the simple sentence 'Paul philosophizes' performs one speech act with 'Paul', another with 'philosophizes', another with the syntactic structure [$_S$ NP VP], another with the declarative mood, another with the whole sentence, and perhaps others with tense morphemes, intonation contours, and whatever other semantically significant bells and whistles are hidden in the sentence. Moreover, the performance of each of these speech acts will be grounded in a different kind of communicative intention. Upon hearing me say all of this, commentators have tended to respond that it is implausible that we typically perform *that many* speech acts, or that we typically have *that many* intentions, in speaking. And, of course, this objection becomes all the more intuitively forceful once we catalogue the speech acts and intentions involved in uttering more complex sentences (such as this one).[16]

Several things can be said in response to this worry. One is that it is an instance of a more general problem about individuating action, and philosophers of action have defended accounts of the individuation of action that produce similar proliferations of actions. For example, Goldman influentially defends the view that "*two act-tokens are identical if and only if they involve the same agent, the same property, and the same time*" (1970: 10). The result of this is that what may

[16] The best-articulated version of this objection that I've heard is due to Neale, who considers it in the manuscript of his forthcoming book, *Linguistic Pragmatism* (*forthcoming b*). In Neale's formulation, the initial appeal of speech-act-theoretic semantics is that it seems plausible that we do perform acts of referring and predicating in order to perform acts of asserting with sentences like 'Paul philosophizes', but this initial appeal wears off once we start identifying every meaningful component of a complex sentence with yet another act type that a speaker tokens in uttering the whole sentence.

seem like going about one's ordinary business turns out to involve the peformance of many token actions:

> If we adopt this analysis of act-tokens, we shall obtain the result sought in the previous section; that is, it will follow that John's moving his hand (at *t*), John's moving his queen to king-knight-seven (at *t*), John's checkmating his opponent (at *t*), John's giving his opponent a heart attack (at *t*), etc., are all different act-tokens. This is because the properties exemplified in these cases are different properties. The property of moving one's hand ≠ the property of moving one's queen to king-knight-seven ≠ the property of checkmating one's opponent ≠ the property of giving one's opponent a heart attack, etc. (Goldman 1970: 11)

In much the same way, it follows from my view that the many speech acts that S performs in uttering 'Paul philosophizes' are distinct actions because they exemplify distinct act types. What my account adds to this is the view that our dispositions to perform some of these actions in certain circumstances are grounded in our dispositions to perform certain others.

Moreover, I see no problem with saying that each of these speech acts is performed with certain intentions in virtue of which it counts as a token of a given act type. The opposing view, which has been a perennial worry about intention-based accounts of speaker meaning,[17] seems to be based mainly in the fact that many speakers will deny having communicative dispositions at all, or in the fact that even theorists lack introspective access to this evidence. More generally, it may simply seem implausible that naïve language users have this much stuff going on in their minds in speaking.

My response to this concern is that although it may be intuitively strange to say that speakers perform this many actions or have that many intentions in speaking, neither our intuitions, nor our folk theory of action, nor our introspective abilities should be taken as reliable guides to these matters.[18] Although the early stages of a theory might get off the ground in part due to the plausibility of its posits, those posits must ultimately be judged by the explanatory work that they do.

---

[17] See, for example, Evans and McDowell (1975: xix–xxiii).

[18] It's also worth asking ourselves how intuitive the complex set-theoretic posits of truth-conditional semantics are, by comparison.

The version of speech-act-theoretic semantics I've offerd in this and the past few chapters is ultimately justified on the theoretical grounds that it best accounts for the phenomena that a theory of linguistic meaning ought to explain.

5.3 Conclusion and Mile-Marking

My goal in this chapter has been to point the way toward a version of compositional semantics that is compatible with everything I've said in the previous chapters. Although there is much work left to do, I hope to have shown at least that the project is conceptually sound, and that there is reason to be optimistic about the possibility that speech-act-theoretic semantics stacks up well against its more orthodox competitors when it comes to their strengths as well as their weaknesses.

BIBLIOGRAPHY

Aloni, Maria (2007): 'Free choice, modals, and imperatives', *Natural Language Semantics* 15: 65–94. doi:10.1007/s11050-007-9010-2.

Alston, William P. (2000): *Illocutionary Acts and Sentence Meaning* (Cornell University Press).

Audi, Paul (forthcoming): 'Grounding: Toward a Theory of the In-Virtue-Of Relation', forthcoming in *Journal of Philosophy*.

Austin, J.L. (1962): *How to do Things with Words* (Harvard University Press).

Bach, Emmon (1976): 'An Extension of Classical Transformational Grammar', in *Problems of Linguistic Metatheory (Proceedings of the l976 Conference)* (Michigan State University Press): 183-224.

Bach, Kent (1987): *Thought and Reference* (Oxford University Press).

———. (1992): 'Intentions and Demonstrations', *Analysis* 52: 140–6.

———. (1994): 'Conversational Implicature', *Mind & Language* 9.

———. (1995): 'Standardization vs. Conventionalization' , *Linguistics and Philosophy* 18: 677–86.

———. (1999): 'The Myth of Conventional Implicature', *Linguistics and Philosophy* 22: 327–66.

———. (2004): 'On Referring and Not Referring' (unpublished manuscript).

———. (2005): 'Context ex Machina' in Zoltán Gendler Szabó (ed.): *Semantics versus Pragmatics* (Clarendon Press): 15–44.

———. (2006): 'What Does it Take to Refer', in the *Oxford Hanbook to the Philosophy of Language*, edited by Ernie Lepore and Barry Smith (Oxford University Press): 516–54.

———. (2012): 'Context Dependence', in Manuel García-Carpintero and Max Kölbel (eds.): *The Continuum Companion to the Philosophy of Language* (Continuum).

———. (*ms*): 'Why Speaker Intentions Aren't Part of Context', unpublished manuscript. http://userwww.sfsu.edu/kbach/Bach.Intentions&Context.pdf

Bach, Kent and Robert Harnish (1979): *Linguistic Communication and Speech Acts* (MIT Press).

Bar-On, Dorit (2004): 'Externalism and Self-Knowledge: Content, Use, and Expression', *Noûs* 38: 430-55.

Barker, Stephen (2004): *Renewing Meaning: A Speech-Act Theoretic Approach* (Oxford University Press).

Barwise, Jon (1989): *The Situation in Logic* (CSLI).

Barwise, Jon and Robin Cooper (1981): 'Generalized quantifiers and natural language,' *Linguistics and Philosophy* 4: 159–219.

Barwise, Jon and John Perry (1983): *Situations and Attitudes* (MIT Press).

Beaver, David (2001): *Presupposition and Assertion in Dynamic Semantics*. (CSLI Publications).

Beaver, David and Bart Geurts (2011): 'Presupposition', in Edward N. Zalta (ed.): The Stanford Encyclopedia of Philosophy (Fall 2013 Edition). URL: http://plato.stanford.edu/archives/fall2013/entries/presupposition/

Benacerraf, Paul (1965): 'What Numbers Could Not Be', *Philosophical Review* 74 (1): 47–73.

Bennett, Johnathan (1976): *Linguistic Behaviour* (Hackett).

Bever, Thomas G. (1970): 'The Influence of Speech Performance on Linguistic Structure', in W. Levelt & G. Flores d'Arcais (eds.), *Advances in Psycholinguistics* (North-Holland): 21–50.

Blackburn, Simon (1984): *Spreading the Word* (Oxford University Press).

Bonnay, Denis and Dag Westerståhl (*ms*): 'Dynamic vs. Classical Consequence', unpublished manuscript.

Borg, Emma (2004): *Minimal Semantics* (Oxford University Press).

Brandom, Robert (1998): *Making it Explicit* (Harvard University Press).

Bratman, Michael (1984): 'Two Faces of Intention', *Philosophical Review* 93 (3): 375–405.

———. (1987): *Intention, Plans, and Practical Reason* (Harvard University Press).

Braun, David (1994): 'Structured Characters and Complex Demonstratives', *Philosophical Studies*, 74: 193–219.

———. (2007): 'Indexicals', in Edward Zalta (ed.): *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition). http://plato.stanford.edu/archives/sum2012/entries/indexicals/>

Broome, John (2004) 'Reasons,' in R. J. Wallace et al. (ed.): *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. (Oxford University Press): 28–55.

Brown, Jessica and Herman Cappelen (2011): *Assertion: New Philosophical Essays* (Oxford University Press).

Buchanan, Ray (2010): 'A Puzzle about Meaning and Communication', *Noûs* 44 (10): 340–71.

Buchanan, Ray and Gary Ostertag (2005): 'Has the Problem of Incompleteness Rested on a Mistake?', *Mind* 114 (456): 889–913.

Burge, Tyler (1975): 'On Knowledge and Convention' in *Foundations of Mind* (Oxford University Press, 2007).

———. (1979): 'Sinning Against Frege' in *Truth, Thought, and Reason: Essays on Frege* (Oxford University Press, 2005).

———. (1986): 'On Davidson's "Saying That"', in E. Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (Blackwell).

Cappelen, Herman (2011): 'Against Assertion', in (Brown and Cappelen): 21–47.

Cappelen, Herman and Ernie Lepore (2005): *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism* (Blackwell).

Carston, Robyn (2002): *Thoughts and Utterances: The Pragmatics of Explicit Communication* (Blackwell).

Chalmers, David (1996): *The Conscious Mind* (Oxford University Press).

Chapman, Siobhan (2005): *Paul Grice: Philosopher and Linguist* (Palgrave Macmillan).

Charlow, Nate (2013): 'Logic and Semantics for Imperatives', *Journal of Philosophical Logic* (online first). DOI: 0.1007/s10992-013-9284-4.

———. (*ms*): 'Imperative Statics and Dynamics'. Unpublished Manuscript. URL: www.natecharlow.com/work/imperative_statics_and_dynamics.pdf.

Chierchia, Gennaro (1995): *Dynamics of Meaning* (University of Chicago Press).

Chomsky, Noam (1956): 'On the Limits of Finite-State Description', MIT Research Laboratory for Electronics, Quarterly Progress Report 41. 64–65.

———. (1957): *Syntactic Structures* (Mouton).

———. (1965). *Aspects of the Theory of Syntax* (MIT Press).

———. (1986): *Knowledge of Language: Its Nature, Origin, and Use* (Praeger).

Chomsky, Noam and G. A. Miller (1963): 'Introduction to the Formal Analysis of Natural Languages', In Luce, R. D., Bush, R. R. & Galanter, E. (eds.) (1963): *Handbook of Mathematical Psychology*, II. (John Wiley and Sons, Inc.).

Clark, Herbert (1992): *Arenas of Language Use* (University of Chicago Press and the Center for the Study of Language and Information).

Clark, Herbert and Catherine Marshall (1981): 'Definite Reference and Mutual Knowledge', in (Clark 1992): 9–59.

Clark, Michael J. and David Liggins (2012): 'Recent Work on Grounding', *Analysis* 72 (4): 812–23.

Cohen, Philip R. and Hector J. Levesque (1990a): 'Intention is Choice with Commitment', *Artificial Intelligence* 42: 213–261.

———. (1990b): 'Persistence, Intention, and Commitment', in (Cohen, Morgan, and Pollack 1990): 33–70.

Cohen, Philip R., Jerry Morgan Martha E. Pollack (eds.) (1990): *Intentions in Communication* (MIT Press).

Cole, Peter and Jerry Morgan (eds.) (1975): *Syntax and Semantics Vol. 3: Speech Acts* (Academic Press).

———. (1978): *Syntax and Semantics Vol. 9: Pragmatics* (Academic Press).

Collins, John (2011): *The Unity of Linguistic Meaning* (Oxford University Press).

Cresswell, Max (1973): *Logics and Languages* (Methuen).

Cumming, Sam (2008): 'Variabilism', *Philosophical Review* 117 (4): 525–54.

Davidson, Donald (1965): 'Theories of Meaning and Learnable Languages' in (Davidson 1984): 3–16.

———. (1967): 'Truth and Meaning' in (Davidson 1984): 17–36.

———. (1970): 'Semantics for Natural Languages' in (Davidson 1984): 55–64.

———. (1973): 'Radical Interpretation' in (Davidson 1984): 125–40.

———. (1979): 'Moods and Performances', in A. Margalit (ed.): Meaning and Use (Dordrech Reidel). Reprinted in (Davidson 1984): 109–22.

Davidson, Donald (1979): 'Moods and Performances', in (Davidson 1984): 109–122.

———. (1984): *Inquiries into Truth and Interpretation* (Oxford University Press).

———. (1987): 'Knowing One's Own Mind' in *Subjective, Intersubjective, Objective* (Oxford University Press, 2001).

Davis, Wayne (2003): *Meaning, Expression, and Thought* (Cambridge University Press).

Devitt, Michael (2006): *Ignorance of Language* (Oxford University Press).

———. (forthcoming): 'What Makes a Property "Semantic"', unpublished manuscript.

Donnellan, Keith (1966): 'Reference and Definite Descriptions', *Philosophical Review* 75: 281–304.

———. (1968): 'Putting Humpty Dumpty Together Again', *Philosophical Review* 77 (2): 203–15.

Dowell, Janice (2011): 'A Flexible Contextualist Account of Epistemic Modals', *Philosopher's Imprint*: 1–25.

———. (2013): 'Flexible Contextualism about Deontic Modals: A Puzzle about Information-Sensitivity', *Inquiry* 56 (2–3): 149–178.

Dummett, Michael (1973): *Frege: Philosophy of Language* (Duckworth).

Edgington, Dorothy (1995): 'On Conditionals', *Mind* 104 (414): 235–329.

Elbourne, Paul (2005): *Situations and Individuals* (MIT Press).

Evans, Gareth (1980): 'Pronouns', in *Collected Papers* (Oxford University Press, 1985): 214–48.

Field, Hartry (1994): 'Deflationist Views of Meaning and Content', *Mind* 103 (411): 249–84.

Fine, Kit (2001): 'The Question of Realism', *Philosophers Imprint* 1 (1): 1–30.

———. (2012): 'A Guide to Ground', in F. Correia and B. Schneider (eds.): *Because* (Cambridge University Press).

Fodor, Jerry (1983): *The Modularity of Mind* (MIT Press).

Fodor, Jerry and Ernie Lepore (2005): 'Out of Context', *Proceedings and Addresses of the APA* 78: 2.

Frege, Gottlob (1892): 'Über Sinn und Bedeutung', *Zeitschrifft für Philosophie und Philosophisches Kritik* 100: 25–50. Translated by M. Black in (Frege 1980).

———. (1980): *Translations from the Philosophical Writings of Gottlob Frege*, edited by P. Geach and M. Black (Oxford University Press).

Geach, Peter T. (1965): 'Assertion', *Philosophical Review* 74: 449–65.

Gibbard, Alan (1990): *Wise Choices, Apt Feelings* (Harvard University Press).

———. (2003): *Thinking How to Live* (Harvard University Press).

Gilbert, Margaret (1989): *On Social Facts* (Routledge).

Gillies, Anthony (2007) 'Counterfactual scorekeeping', *Linguistics Philosophy* 30 (3): 329–360.

———. (2009) 'On truth-conditions for If (but not quite only If)', *Philosophical Review* 118 (3): 325–349.

———. (2010): 'Iffiness', *Semantics and Pragmatics* 3 (4): 1–42. doi:10.3765/sp.3.4. http://dx.doi.org/10.3765/sp.3.4.

Ginzburg, Jonathan (1995a): 'Resolving Questions, I', *Linguistics and Philosophy* 18 (5): 459–527. DOI: http://dx.doi.org/10.1007/BF00985365.

———. (1995b): 'Resolving Questions, II', *Linguistics and Philosophy* 18 (6): 567–609. DOI: http://dx.doi.org/10.1007/BF00983299.

———. (1996): 'Dynamics and the Semantics of Dialogue', in Jerry Seligman Dag Westerståhl (eds.): *Logic, Language and Computation*, vol. 1 (CSLI Lecture Notes 58) (Center for the Study of Language Information): 221–37.

Green, Mitch (2007): *Self Expression* (Oxford University Press).

Grice, H.P. (1957): 'Meaning', *Philosophical Review* 66. Reprinted as Chapter 14 of (Grice 1989): 213–224.

———. (1968): 'Utterer's Meaning, Sentence Meaning, and Word Meaning', *Foundations of Language* 4. Reprinted as Chapter 6 of (Grice 1989): 117–37.

———. (1969): 'Utterer's Meaning and Intention', *Philosophical Review* 78. Reprinted as Chapter 5 of (Grice 1989): 86–116.

———. (1971): 'Intention and Uncertainty', *Proceedings of the British Academy* 57: 263–79.

———. (1975): 'Logic and Conversation', in (Cole and Morgan 1975). Reprinted as Chapter 2 of (Grice 1989): 22–40.

———. (1978): 'Further Notes on Logic and Conversation', in (Cole and Morgan 1979). Reprinted as Chapter 3 of (Grice 1989): 41–57.

———. (1989): *Studies in the Way of Words* (Harvard University Press).

Groenendijk, Jeroen and Martin Stokof (1982): 'Semantic Analysis of *WH*-Complements', *Linguistics and Philosophy* 5: 175-233.

———. (1989): 'Type-shifting Rules and the Semantics of Interrogatives', in Gennaro Chierchia, Barbara Partee, and Raymond Turner (eds.): Properties, Types, and Meanings, Vol. 2: Semantic Issues (Kluwer). Reprinted in (Portner and Partee 2002): 421–56.

———. (1997): 'Questions', in Johann van Benthem and Alice ter Meulen (eds.): *Handbook of Logic and Language* (MIT Press): 1055-124.

Han, C.-h. (1998). *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar*. Ph.D. Dissertation, University of Pennsylvania. URL: http://www.sfu.ca/~chunghye/papers/dissertation.pdf.

———. (1999): 'Deontic Modality,

Hanks, Peter (2007): 'The Content-Force Distinction', Philosophical Studies 134 (2): 141–64.

———. (2011): 'Structured Propositions as Types', *Mind* 120: 11–52.

———. (2012): 'First-Person Propositions', *Philosophy and Phenomenological Research* (online first, March 2012).

———. (forthcoming): *Propositional Content* (Oxford University Press).

Hacquard, Valentine (2006): *Aspects of Modality*, PhD Thesis, MIT.

Harman, Gilbert (1977): 'Review of *Linguistic Behavior* by Jonathan Bennett', *Language* 53: 417–24.

———. (1979): 'If and Modus Ponens', *Theory and Decision* 11: 41–53.

Hawthorne, John. (1990): 'A Note on "Languages and Language"' in *Australasian Journal of Philosophy* 68: 116–8.

———. (1993): 'Meaning and Evidence: A Reply to Lewis' in *Australasian Journal of Philosophy* 71: 206–11.

Hawthorne, John and Ofra Magidor (2009): 'Assertion, Context, and Epistemic Accessibility' in *Mind* 118 (470): 377–97.

Hawthorne, John and David Manley (2013): *The Reference Book* (Oxford University Press).

Heim, Irene (1982): *On the Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts at Amherst.

———. (1983): 'On the projection problem for presuppositions', in M. Barlow, D. D. Flickinger, and M. Westcoat (eds.): *Second Annual West Coast Conference on Formal Linguistics* (Stanford University): 114–26.

Heim, Irene and Angelika Kratzer (1998): *Semantics in Generative Grammar* (Blackwell).

Horn, Laurence R. and Gregory Ward (eds.) (2004): *The Handbook of Pragmatics* (Blackwell).

Horwich, Paul (1998): *Meaning* (Oxford University Press).

Kamp, Hans (1981): 'A theory of truth and semantic representation', in: J.A.G. Groenendijk, T.M.V. Janssen, and M.B.J. Stokhof (eds.), *Formal Methods in the Study of Language*. Mathematical Centre Tracts 135: 277–322.

Kaplan, David (1989a): 'Demonstratives' in (Almog et al. 1989): 481–563.

———. (1989b): 'Afterthoughts' in (Almog et al. 1989): 565–614.

Karttunen, Lauri (1976): 'Discourse Referents', in P. Cole (ed.): *Syntax and Semantics*, Volume 7 (Academic Press).

Kaufmann, Magdalena (2012): *Interpreting Imperatives* (Springer).

Kaufmann, S., and M. Schwager (2009): 'A Unified Analysis of Conditional Imperatives, in E. Cormany, S. Ito, D. Lutz (Eds.): *Proceedings of SALT* 19: 239–256 (CLC Publications). URL: www.semanticsarchive.net/Archive/TYzZWYyM/.

Kennedy, Christopher (2007): 'Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives', *Linguistics and Philosophy* 30: 1–45.

King, Jeffrey C. (2001): *Complex Demonstratives: A Quantificational Account* (MIT Press).

———. (2007): *The Nature and Structure of Content* (Oxford University Press).

———. (2009): 'Questions of Unity', *Proceedings of the Aristotelian Society* 109 (3): 257-277.

———. (2012): 'Speaker Intentions in Context', Noûs (online first, June 2012).

———. (forthcoming): 'Propositional Unity: What's the Problem, Who has it and Who Solves it?', to appear in *Philosophical Studies*.

King, Jeffrey C. and Jason Stanley (2005): 'Semantics, Pragmatics, and the Role of Semantic Content', In Zoltán Gendler Szabó (ed.): *Semantics vs. Pragmatics* (Oxford University Press: 111–64.

Klein, E., and I. Sag (1985): 'Type-Driven Translation', *Linguistics and Philosophy* 8: 163–201.

Kratzer, Angelika (2012): *Modals and Conditionals* (Oxford University Press).

Kripke, Saul (1977): 'Speaker's Reference and Semantic Reference'. inn Peter French, Thomas Uehling & Howard Wettstein (eds.): *Midwest Studies in Philosophy*, vol. 2 (University of Minnesota Press). Reprinted in Philosophical Troubles (Oxford University Press, 2011): 99–124.

Larson, Richard and Gabriel Segal (1995): *Knowledge of Meaning* (Oxford University Press).

Lascarides, Alex and Nicholas Asher (2004): 'Imperatives in Dialogue', in P. Kuehnlein, H. Rieser and H. Zeevat (eds.): *The Semantics and Pragmatics of Dialogue for the New Millenium* (Benjamins).

Laurence, Stephen (1996): 'A Chomskian Alternative to Convention-Based Semantics' in *Mind* 105: 269–301.

Lepore, Ernest and Kirk Ludwig (2005): *Donald Davidson: Meaning, Truth, Language, and Reality* (Oxford University Press).

Lewis, David (1969): *Convention* (Harvard University Press)

———. (1970): 'General Semantics' in (Lewis 1983).

———. (1975a): 'Languages and Language' in (Lewis 1983).

———. (1975b): 'A Problem about Permission', in E. Saarinen et ah, eds., *Essays in Honour of Jaakko Hintikka* (Dordrecht, Reidel, 1979). Reprinted in (Lewis 2000): 20–33.

———. (1976): 'Conventions: Reply to Jamieson', *Canadian Journal of Philosophy* 6: 113–20.

———. (1979a): 'Scorekeeping in a Language Game', *Journal of Philosophical Logic* 8. Reprinted in (Lewis 1983): 233–49.

———. (1979b): 'Attitudes De Dicto and De Se', *The Philosophical Review* 88. Reprinted in (Lewis 1983): 133–55.

———. (1980): 'Index, Context, and Content' in Kanger and Öhman (eds.), *Philosophy and Grammar* (Reidel).

———. (1983): *Philosophical Papers, Volume I* (Oxford University Press).

———. (1992): 'Meaning without Use: Reply to Hawthorne' in Australasian Journal of Philosophy 70: 106–10.

———. (2000): *Papers in Ethics and Social Philosophy* (Cambridge University Press).

Lewis, Karen (2011): *Understanding Dynamic Discourse* (Dissertation, Rutgers University).

———. (2012): 'Discourse Dynamics, Pragmatics, and Indefinites', Philosophical Studies 158 (2): 313–42.

———. (forthcoming): 'Do We Need Dynamic Semantics', forthcoming in Alexis Burgess and Brett Sherman (eds.): *Metasemantics* (Oxford University Press).

Loar, Brian (1981): *Mind and Meaning* (Cambridge University Press).

MacFarlane, John (2011): 'What is Assertion?', in (Brown and Cappelen 2011): 79–96.

Martinich, A.P. (1985/2001): *The Philosophy of Language* (Oxford University Press).

May, Robert (1977): *The Grammar of Quantification*, Doctoral Dissertation, MIT

———. (1985): *Logical Form: its Structure and Derivation* (MIT Press).

———. (1987): 'Logical Form as a Level of Linguistic Representation' in Ernie Lepore (ed.): *New Directions in Semantics* (Academic Press): 187–218).

McCarthy, J., M. Sato, S. Igarashi, and T. Hayashi (1977): 'On the Model Theory of Knowledge', *Proceedings of IJCAI–1977* (MIT Press).

Michaelson, Eliot (2013): 'This and That: A Theory of Reference for Names, Demonstratives, and Things in Between'. PhD Thesis, UCLA.

Miller, Seumas (2001): *Social Action: A Teleological Account* (Cambridge University Press).

Millikan, Ruth (2005): *Language: A Biological Model* (Oxford University Press).

———. (2008): "A Difference of Some Consequence Between Conventions and Rules" in *Topoi* 27: 87–99.

Montague, Richard (1970b): 'Universal Grammar', Theoria 36: 373–98. Reprinted in (Montague 1974): 222–246.

Montague, Richard (1973): 'The Proper Treatment of Quantification in Ordinary English', in J. Hinktikka, J. Moravcsik, and P. Suppes (eds.): *Approaches to Natural Language* (Dordrecht). Reprinted in (Montague 1974): 247–70.

———. (1974): Formal Philosophy (Yale University Press).

Mount, Allyson (2008): 'Intentions, Gestures, and Salience in Ordinary and Deferred Demonstrative Reference', *Mind and Language* 23 (2): 145–64.

Murray, Sarah and Will Starr (2013): 'The Semantics of Mood and Communicative Acts: From Sentential to Illocutionary Acts'. Slides from day 5 of a Workshop at ESSLI. August 9th, 2013.

Neale, Stephen (1990): *Descriptions* (MIT Press).

———. (1992): 'Paul Grice and the Philosophy of Language', *Linguistics and Philosophy* 15: 509–559.

———. (1993): 'Logical Form and LF' in Chris Otero (ed.): *Noam Chomsky: Critical Assessments* (Routledge): 788–838.

———. (1999): 'Colouring and Composition', in Rob Stainton (ed.): *Philosophy and Linguistics* (Westview Press): 35–82.

———. (2001): *Facing Facts* (Oxford University Press).

———. (2004): 'This, That, and the Other' in Marga Reimer and Anne Bezuidenhout (eds.): *Descriptions and Beyond* (Oxford University Press): 68–182.

———. (2005): 'Pragmatism and Binding' in Zoltán Gendler Szabó (ed.): *Semantics versus Pragmatics* (Oxford University Press).

———. (2007a): 'Heavy Hands, Magic, and Scene-Reading Traps', *European Journal of Analytic Philosophy* 3 (2): 77–132.

———. (2007b): 'On Location' in M. O'Rourke and C. Washington (eds.): *Situating Semantics: Essays on the Philosophy of John Perry* (MIT Press): 251–393.

———. (2008): 'Term Limits Revisited', *Philosophical Perspectives* 22: 375–442.

———. (forthcoming a): 'Implicit Reference' to be published in Gary Ostertag (ed.): *The Philosophy of Stephen Schiffer* (MIT Press).

———. (forthcoming b): *Linguistic Pragmatism* (Manuscript).

Pagin, Peter (2011): 'Information and Assertoric Force', in H. Cappelen and Jessica Brown (eds.): *Assertion* (Oxford University Press): 97–136.

Parsons, Josh (2012): 'Cognitivism about Imperatives', *Analysis* 72: 49–54. doi: 10.1093/analys/anr132.

———. (2013): 'Command and Consequence', *Philosophical Studies* (online first). doi: 10.1007/s11098-013-0094-x.

———. (ms): 'The Transformational Approach to Imperative Consequence'. Unpunlished Manuscript. URL: http://www.academia.edu/2364666/The_transformational_approach_to_imperative_consequence

Partee, Barbara (1973): 'Soe Structural Analogies between Tenses and Pronouns in English', *The Journal of Philosophy* 70 (18): 601–9.

———. (1986): 'Noun-Phrase Interpretation and Type-Shifting Principles', in (Partee 2004): 203–30.

———. (1995): 'Lexical Semantics and Compositionality', in Lila R. Gleitman and Mark Liberman (eds.): *An Invitation to Cognitive Science, Volume 1: Language*, Second Edition (MIT Press).

———. (2004): *Compositionality in Formal Semantics* (Blackwell).

Partee, Barbara and Mats Rooth (1983): 'Generalized Conjunction and Type Ambiguity', in R. Bauerle, C. Schwarze, and A. von Stechow (eds.): *Meaning, Use and Interpretation of Language* (De Gruyter). Reprinted in (Portner and Partee 2002): 332–356.

Peacocke, Christopher (1976): 'Truth Definitions and Actual Languages', in Gareth Evans and John McDowell (eds.): *Truth and Meaning: Essays in Semantics* (Oxford University Press).

———. (1986): 'Explanation in Computational Psychology: Language, Perception, and Level 1.5', *Mind and Language* 1 (2): 101–23.

Perry, John (1977): 'Frege on Demonstratives', *The Philosophical Review* 86 (4).

———. (1986): 'Thought without Representation', *Proceedings of the Aristotelian Society* 60: 137–52.

Peters, Stanley and Dag Westerståhl (2013): 'The Semantics of Possessives', *Language* 89 (4): 713–59.

Petit, Dean (2002): 'Why Knowledge is Unnecessary for Linguistic Understanding', *Mind* 111 (443): 519–50.

Portner, Paul (2004): 'The Semantics of Imperatives within a Theory of Clause-Types', in K. Watanabe R. Young (eds.): *Proceedings of SALT 14* (CLC Publications). URL: http://semanticsarchive.net/Archive/mJlZGQ4N/.

———. (2007): 'Imperatives and Modals', *Natural Language Semantics* 15: 351–83.

———. (2012): 'Permission and Choice' In G. Grewendorf T. Zimmermann (eds.): *Discourse and Grammar: From Sentence Types to Lexical Categories*, Studies in generative grammar (Mouton
de Gruyter). http://semanticsarchive.net/Archive/jI1YjMyY/

Portner, Paul and Barbara Partee (2002): *Formal Semantics: The Essential Readings* (Blackwell).

Pylkkanen, Liina and Brian McElree (2006): 'The Syntax-Semantics Interface: On-Line Composition of Sentence Meaning', in M. Traxler and M.A. Gernsbacher (eds.): *Handbook of Psycholinguistics* (2nd Ed) (Elsevier): 537-77.

Quine, Willard V. O. (1968): 'Ontological Relativity', *Journal of Philosophy* 65 (7): 185–212.

Reimer, Marga (1998): 'Quantification and Context', *Linguistics and Philosophy*, 21: 96–115.

Roberts, Craige (1996/2012): 'Information Structure in Discourse: Toward an Integrated Formal Theory of Pragmatics', in Jae-Hak Yoon Andreas Kathol (eds.), *Papers in Semantics* (Working Papers in Linguistics 49) (The Ohio State University, 1996). Revised and reprinted in *Semantics and Pragmatics* 5: 1–69. DOI: http://dx.doi.org/10.3765/sp.5.6.

———. (2002): 'Demonstratives as Definites', in Kees van Deemter & Roger Kibble (eds.): *Information Sharing: Reference and Presupposition in Language Generation and Interpretation* (CSLI): 89–196.

———. (2003): 'Uniqueness in Definite Noun Phrases', *Linguistics and Philosophy* 26: 287–350.

———. (2004): 'Context in Dynamic Interpretation', in (Horn and Ward 2004): 197–220.

Rosen, Gideon (2010): 'Metaphysical Dependence: Grounding and Reduction', in *Modality: Metaphysics, Logic, and Epistemology*, edited by Bob Hale and Aviv Hoffman (Oxford University Press): 109–136.

Ross, Alf (1941): 'Imperatives and Logic', *Theoria* 7: 53–71. doi:10.1111/j.1755-2567.1941.tb00034.x.

———. (1944) 'Imperatives and logic', *Philosophy of Science* 11 (1). 30–46.

Ross, J. R. (1970): 'On Declarative Sentences'. In Jacobs, R. & Rosen- baum, P. S., (eds.): 'Readings in English Transformational Grammar', (Ginn): 222–72.

Rothschild, Daniel and Seth Yalcin (*ms*): 'On the Dynamics of Conversation', unpublished manuscript.

Russell, Bertrand (1903): The Principles of Mathematics (Cambridge University Press).

———. (1905): 'On Denoting', *Mind* 14: 479–93

van Rooy, R. (2000): 'Permission to Change', *Journal of Semantics* 17: 119–45.

Sadock, Jerrold (1974): *Toward a Linguistic Theory of Speech Acts* (Academic Press).

Sadock, J. M. and A. M. Zwicky (1985): 'Speech Act Distinctions in Syntax', in: T. Shopen (ed.): *Language Typology and Syntactic Description*, Volume 1 (Cambridge University Press): 155–96.

Schaffer, Jonathan (2009): 'On What Grounds What', in Metametaphysics: New Essays on the Foundations of Ontology, edited by David Manley, David Chalmers, and Ryan Wasserman (Oxford University Press): 347–83.

Schiffer, Stephen (1972): *Meaning* (Oxford University Press).

———. (1981): 'Indexicals and the Theory of Reference', *Synthese* 49 (1): 43–100.

———. (1982): 'Intention-Based Semantics', *Notre Dame Journal of Formal Logic* 23 (2): 119–56.

———. (1987a): *Remnants of Meaning* (MIT Press).

———. (1987b): 'Extensionalist Semantics and Sentential Theories of Belief', in E. Lepore (ed.): *New Directions in Semantics* (Academic Press): 113–42.

———. (1993): 'Actual Language Relations' in *Philosophical Perspectives* 7: 231–58.

———. (2003): *The Things We Mean* (Oxford University Press).

———. (2005): 'Russell's Theory of Definite Descriptions', *Mind* 114 (456): 1135–1183.

———. (2006): 'Two Perspectives on Knowledge of Language' in *Philosophical Issues* 16: 275–287.

Schroeder, Mark (2008): 'What is the Frege–Geach Problem?', *Philosophy Compass* 3/4: 703–720.

Schwager, M. (2006a): *Conditionalized Imperatives*, in M. Gibson J. Howell (Eds.): *Proceedings of SALT* 16 (CLC Publications). url: http://user.uni-frankfurt.de/~scheiner/papers/schwagerFEB07.pdf.

———. (2006b): *Interpreting Imperatives*. Ph.D. Dissertation, University of Frankfurt.

Searle, John (1968): 'Austin on Locutionary and Illocutionary Acts', *Philosophical Review* 57.

———. (1969): *Speech Acts* (Cambridge University Press).

———. (1975a): 'Indirect Speech Acts', in (Cole and Morgan 1975). Reprinted in (Searle 1979): 30–57.

———. (1975b): 'A Taxonomy of Illocutionary Acts', in Keith Gunderson (ed.), Minnesota Studies in the Philosophy of Science: Language, Mind, and Knowledge (University of Minnesota Press). Reprinted in (Searle 1979): 1–29.

———. (1979): *Expression and Meaning: Studies in the Theories of Speech Acts* (Cambridge University Press).

Sellars, Wilfrid (1969): 'Language as Thought and Communication,' *Philosophy and Phenomenological Research* 29: 506–27

Setiya, Kieran (2011): 'Intention', in Ed Zalta (ed): *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition). url: http://plato.stanford.edu/archives/spr2011/entries/intention/

Siegel, Susanna (2002): 'The Role of Perception in Demonstrative Reference', *Philosopher's Imprint* 2 (1).

Skyrms, Brian (1996): *Evolution of the Social Contract* (Cambridge University Press).

———. (2010): *Signals* (Oxford University Press).

Soames, Scott (1986): 'Incomplete Definite Descriptions', Notre Dame Journal of Formal Logic 27: 349–75.

———. (1987): 'Direct Reference, Propositional Attitudes, and Semantic Content', *Philosophical Topics* 15: 47–87. Reprinted as Chapter 1 of (Soames 2009): 33–71.

———. (2002): *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity* (Oxford University Press).

———. (2009): *Philosophical Essays, Volume II: The Philosophical Significance of Language* (Princeton University Press).

———. (2010): *What is Meaning?* (Princeton University Press).

———. (2010b): *Philosophy of Language* (Princeton University Press).

Sperber, Dan and Dierdre Wilson (1986/94): *Relevance: Communication and Cognition* (Blackwell).

Stalnaker, Robert (1973): 'Presuppositions', *Journal of Philosophical Logic* 2: 447–57.

———. (1974): 'Pragmatic Presuppositions', in Milton Munitz and Peter Unger (eds.): *Semantics and Philosophy* (New York University Press). Reprinted in (Stalnaker 1999): 47–62.

———. (1975): 'Indicative Conditionals', *Philosophia* 5. Reprinted in (Stalnaker 1999): 63–77.

———. (1978): 'Assertion', in P. Cole (ed.): *Syntax and Semantics* 9 (New York: Academic Press). Reprinted in (Stalnaker 1999): 78–95.

———. (1999): *Context and Content* (Oxford University Press).

———. (2008): *Our Knowledge of the Internal World* (Oxford University Press).

Stanley, Jason (2002): 'Nominal Restriction', in Gerhard Preyer and George Peter (eds.): *Logical Form and Language* (Oxford University Press, 2002). Reprinted as Chapter 3 of (Stanley 2007): 111–132.

———. (2005): Knowledge and Practical Interests (Oxford University Press).

———. (2007): *Language in Context* (Oxford University Press).

———. (2010): 'Assertion and Intentionality', *Philosophical Studies* 151 (1): 87–113.

———. (2011): *Know How* (Oxford University Press).

Stanley, Jason and Zoltán Gendler Szabó (2000): 'On Quantifier Domain Restriction', in (Stanley 2007): 69–110.

Starr, Will (2010): *Conditionals, Meaning, and Mood* (Dissertation, Rutgers University). Online at http://williamstarr.net/research/dissertation.pdf.

———. (*forthcoming a*): 'Mood, Force, and Truth', in E. Lepore and Y. Jiang (eds.): *Language and Value* (Oxford University Press).

———. (*forthcoming b*): 'What If?', forthcoming in Philosophers Imprint

———. (*ms*): 'A Preference Semantics for Imperatives', unpublished manuscript. URL: http://williamstarr.net/research/a_preference_semantics_for_imperatives.pdf.

Stine, Gail C. (1978): 'Meaning Other Than What We Say and Referring', *Philosophical Studies* 33 (4): 319–37.

Strawson, P. F. (1950): 'On Referring', *Mind* 59: 320–44.

———. (1964): 'Intention and Convention in Speech Acts', *Philosophical Review* 76: 439–60.

———. (1969): 'Meaning and Truth', in *Proceedings of the British Academy*, (Oxford University Press).

Thomason, Richmond H. (1990): 'Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics', in (Cohen, Morgan, and Pollack 1990): 325–63.

Travis, Charles (2008): *Occasion Sensitivity: Selected Essays* (Oxford University Press).

van Benthem, Johan (2011): *Logical Dynamics of Information and Interaction* (Cambridge University Press).

von Fintel, Kai and Irene Heim (2011): *Intensional Semantics*, Spring 2011 edition (MIT Course Notes).

Vendler, Zeno (1972): *Res Cogitans* (Cornell University Press).

Veltman, Frank (1996): 'Defaults in Update Semantics', *Journal of Philosophical Logic* 25 (3): 221–61.

Vranas, P. B. M. (2009): 'In Defense of Imperative Inference', *Journal of Philosophical Logic* 39: 59–71.

———. (2011): 'New Foundations for Imperative Logic: Pure Imperative Inference', *Mind* 120 (478): 369–446.

Westerståhl, Dag (1985): 'Determiners and Context Sets', in J. van Bentham and A. ter Meulen (eds.): *Generalized Quantifiers in Natural Language* (Foris Publications Holland).

Wettstein, Howard K. (1984): 'How to Bridge the Gap Between Meaning and Reference', *Synthese* 58 (1): 63–84.

Williamson, Timothy (2000): *Knowledge and its Limits* (Oxford University Press).

Wilson, Deirdre and Dan Sperber (1988): 'Mood and the Analysis of Non-Declarative Sentences', in (Wilson and Sperber 2012): 210–229.

———. (2012): *Meaning and Relevance* (Cambridge University Press).