The Copy Theory of Movement: Spell Out

Kyle Johnson
University of Massachusetts at Amherst
kbj@linguist.umass.edu

SICOGG 12
17 August 2010

Movement has been used to model a variety of syntactic relations that, frankly, oftentimes look quite different. Here are some examples.

(1) a. Mary kaupir ikke skó? (*Icelandic*)
Mary buys not shoes — Head Movement

b. I asked which book Mary had read — $\overline{\text{A}}$ Movement

c. A child seems to have left. — A Movement

d. every bank a different guard stood before every bank. — Quantifier Raising

Why are we tempted to see each of these cases as special cases of the same relation? Perhaps because they (sort of) share these three properties.

(2) Semantic Displacement
Some part of the meaning of the moved expression is applied to a position different from where it is spoken.

(3) Terseness
The moved item semantically occupies two positions, but is spoken in only one of them.

(4) Locality
The two positions that a moved item is related to are subject to a locality condition.

These properties don't manifest themselves in exactly the same way across these various kinds of movement, though, and so that's a challenge to seeing these as shared properties.

(5) Difference in Semantic Displacement

a. *Total Reconstruction:*
Mary kaupir ikke skó. ≡ ¬ Mary kaupir skó

b. *Variable Binding:*
Which book Mary had read ≡ The set of propositions such that $\exists x$ Mary had read x, x a book.
A guard stands before every bank ≡ $\forall x$ if x is a bank then a guard stands before x

(6) Differences in Locality

a. *Head Movement Constraint:*

* Have Mary should read a book.

b. *Ross's Islands*

Which book has Mary shown [$_{CP}$ would change your life]?

* Which book has Mary shown [$_{DP}$ the proof [$_{CP}$ would change your life]]?

c. *Tensed S Condition*

* A child seems [$_{CP}$ (that) has left].

* every bank a different guard showed [$_{CP}$ that the road stood before every bank].

(7) Differences in how Terseness is violated

(8) ngɔ̄nū ǹ wà nā ǹ kā́ ngɔ́nʊ̍ á (*Vata*)
sleep you want NA you FUT-A sleep Q
'Do you want to sleep?'

(Koopman 1984, (2a): 154)

Here "ngɔnʊ" ('sleep') has been clefted and is pronounced in both the cleft position and the position inside its VP. This is probably movement since locality conditions are satisfied.

(9) * tākā ǹ wà fòtó̍` mūmʊ̍` ǹ tá̍ká̍ ɓɔ́ àbà
show you like picture ITIT you showed REL Aba
'It's show that you like the picture you showed Aba.'

(Koopman 1984, (15): 159)

When a verb clefts in Vata, both copies must be pronounced. (8) is ungrammatical if either verb isn't pronounced. When nominal material clefts, by contrast, only the higher copy may be pronounced. (10) illustrates.

(10) ngɔ́nʊ́lì mí ǹ wà à
sleep-NOM IT you want Q
'Is it sleeping you want?' (Koopman 1984, (2b): 154)

What we have in these examples is a kind of $\overline{\text{A}}$ movement of a verb. When verbs, or predicates, move, we sometimes get violations of Terseness that involve two pronunciations of the moved predicate, and that's what Vata illustrates.

I don't know of anything similar, though, when a DP has $\overline{\text{A}}$ moved. In those cases, violations of Terseness take a different shape. They produce resumptive pronouns. For instance, in Lebanese Arabic there are resumptive pronouns that show Semantic Displacement effects when islands aren't violated. (See Aoun and Benmamoun 1998 and Aoun, Choueiri, and Hornstein 2001.)

(11) təlmiiz-a$_1$ lkəsleen ma baddna nχabbi [wala mʕallme]$_1$ ʔənno *huwwe* zaʕbar
student-her$_1$ the-bad NEG want.1P tell.1P [no teacher]$_1$ that he cheated.3SM
b-l-faħṣ
in-the-exam
'her bad student, we don't want to tell any teacher that he cheated on the exam.'

(12) * təlmiiz-a₁ lkəsleen ma badkun tχabbro [wala mʕallme] ʕan l-bənt yalli *huwwe*
student-her the-bad NEG want.2P tell.2P [no teacher] about the-girl that he
zaʕbar maʕ-a b-l-faħṣ
cheated.3sm with-her in-the-exam
(Her bad student, you don't want to tell any teacher about the girl with whom he cheated on the exam.)

(Aoun et al. 2001, (25b) & (29b): 381–2)

I think, then, that there is a difference in how Terseness is violated depending on the category of the thing being moved. There is a potential problem for this belief in the wh-copying construction that colloquial German (and other languages) display.

(13) German
a. Wen glaubt John wen Mary getroffen hat?
who thinks John who Mary met has
'Who does John think Mary has met?'
b. * Wieviel Geld meint sie wieviel Geld das kostet?
how-much money thinks she how-much money that costs
'How much does she think that costs?'

(Rett 2006, (1b): 355, (6): 358)

It's hard to say in these situations whether the intermediate copy is a resumptive pronoun or a copy. But there are certain properties of the copy construction which suggest that it is different from a "normal" movement construction with just a simple violation of Terseness. There are interesting semantic differences between the two constructions. A striking one is described in Felser (2004). She notes that in cases where a phrase has moved in across-the-board fashion out of two coördinated clauses, as in (14), the question seems to assume that the answer will provide individuals that meet the descriptions provided by both of the clauses.

(14) Wen glaubst du, dass sie getroffen hat und dass sie liebt?
who think you that she met has and that she loves
'Who do you think that she met and that she loves?'

(Felser 2004, (37a): 560)

By contrast, a parallel across-the-board movement but with the wh-phrase pronounced in the lower positions as well, as in (15), seems to assume that the answer will provide the identity of individuals that meet the descriptions provided in each of the clauses separately.

(15) Wen glaubst du, wen sie getroffen hat und wen sie liebt?
who think you who she met has and who she loves
'Who do you think that she met and that she loves?'

(Felser 2004, (37b): 560)

This difference in meaning suggests that there is a separate quantification, one for each of the lower wh phrases, in the copy construction that is absent in the non-copy construction version. If that is correct, it will require that the copy construction include more quantificational expressions than are found in the simpler, single pronunciation, movement structure.

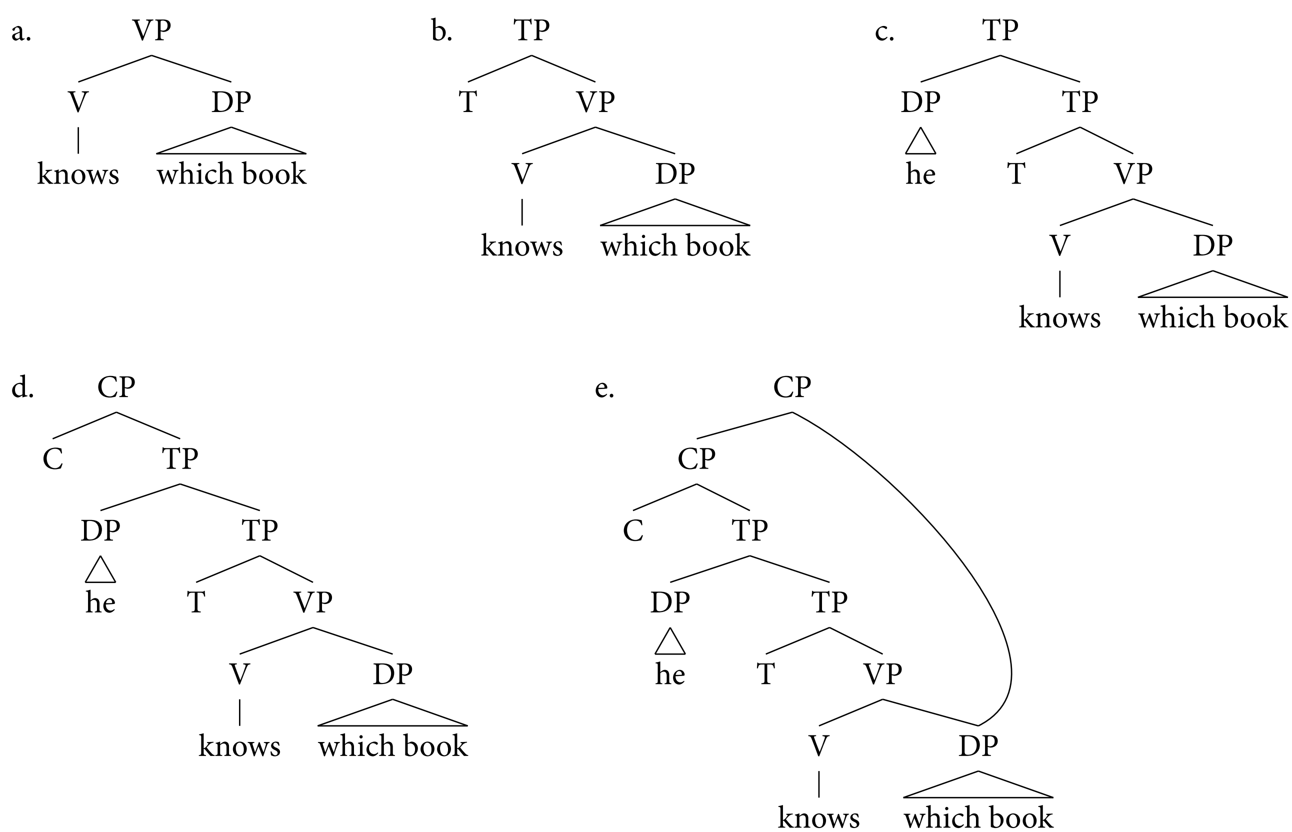
I'll assume that there is a difference between moved DPs and moved other things that is responsible for how Terseness is relaxed in them.

What we want, then, is a theory of movement that explains these three properties: Semantic Displacement, Terseness and Locality. But that theory should also be flexible enough that it gives us a handle on why these three properties manifest themselves differently depending on the particulars of the movement operation. I'm going to take a few, very small, steps in that direction, building on an idea about what movement is that was in an early unpublished manuscript by Stanley Peters and Robert Richie, carried forward by Engdahl (1980) and has now found many proponents, including Gärtner (1997), Starke (2001), Nunes (2001), Frampton (2004), Citko (2005), Kobele (2006) and de Vries (2007). That idea is that movement gives an expression two positions by re-merging it.

(16) MERGE(α,β) $=_{\text{def.}}$ $[_{\gamma}\ \alpha\ \beta]$, where the linear order of α and β is not determined.

A derivation that involves movement:

(17) (She asked) which book he knows.



You can see how (17e) provides a way of capturing Semantic Displacement. The moved term — here *who* — is syntactically in two positions and so its denotation has two positions where it can be applied. The differences in how Semantic Displacement arises are going to come about, I will claim, from the particular ways in which the expressions that are "moved" get broken up into two different positions. That is going to be the focus of most of my lectures, but we won't start that process until tomorrow.

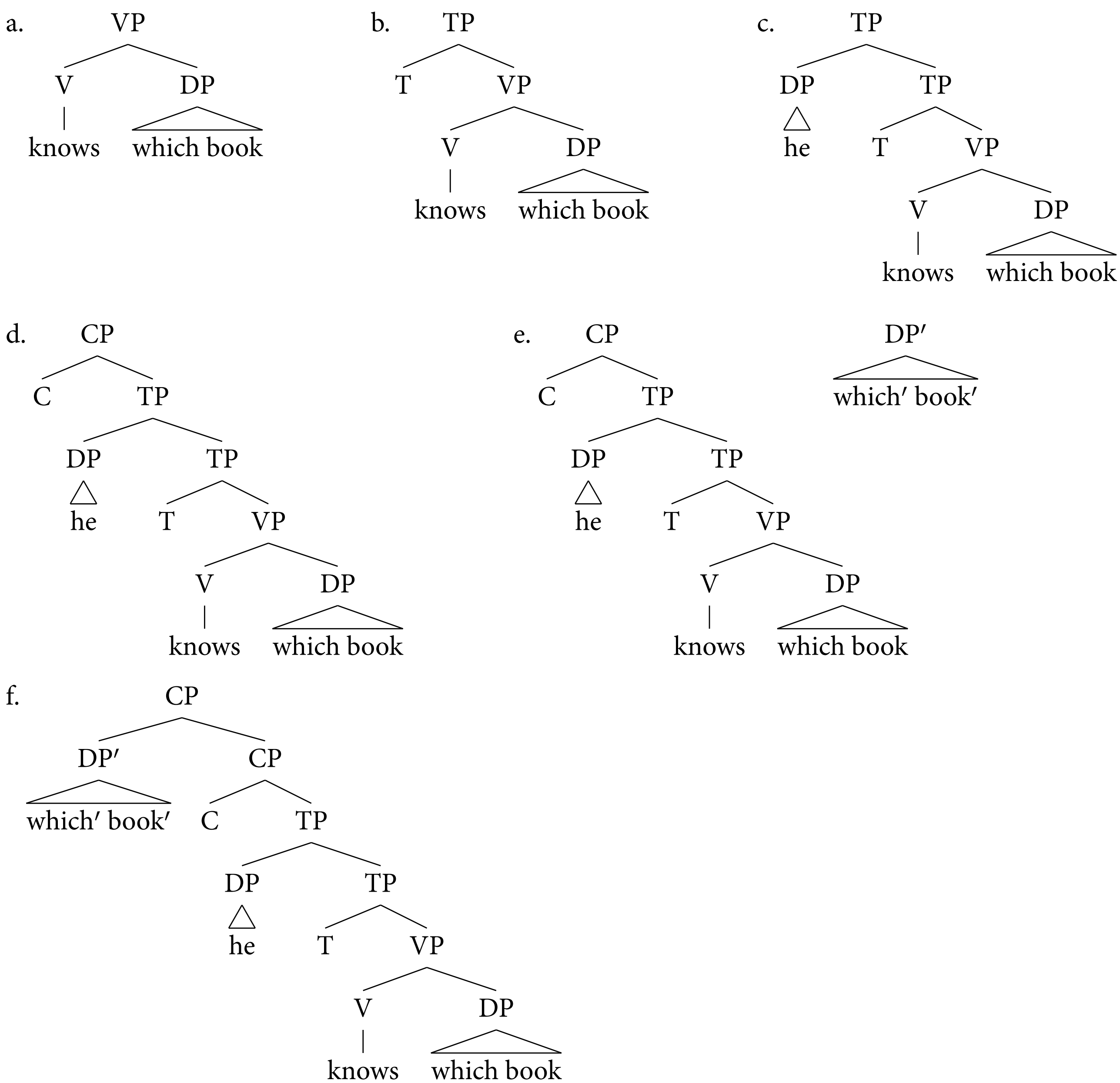
It's not obvious that these representations provide any particularly obvious explanation for why movement is subject to locality conditions. I've taken a very vague stab at trying to make that connection in a paper that was delivered in the 2009 meeting of the Chicago Linguistics Society. I'll have nothing to say about deriving Locality in these lectures.

Today I'm going to tackle how these representations derive Terseness. I'll adapt a popular explanation for Terseness that is due to Jairo Nunes. Nunes worked with a different account of movement: the "copy" theory of movement. This theory does not countenance multidominant representations, like that in (17e), but instead involves a "copy" operation.

(18) COPY(α) = α', an exact syntactic and semantic replica of α.

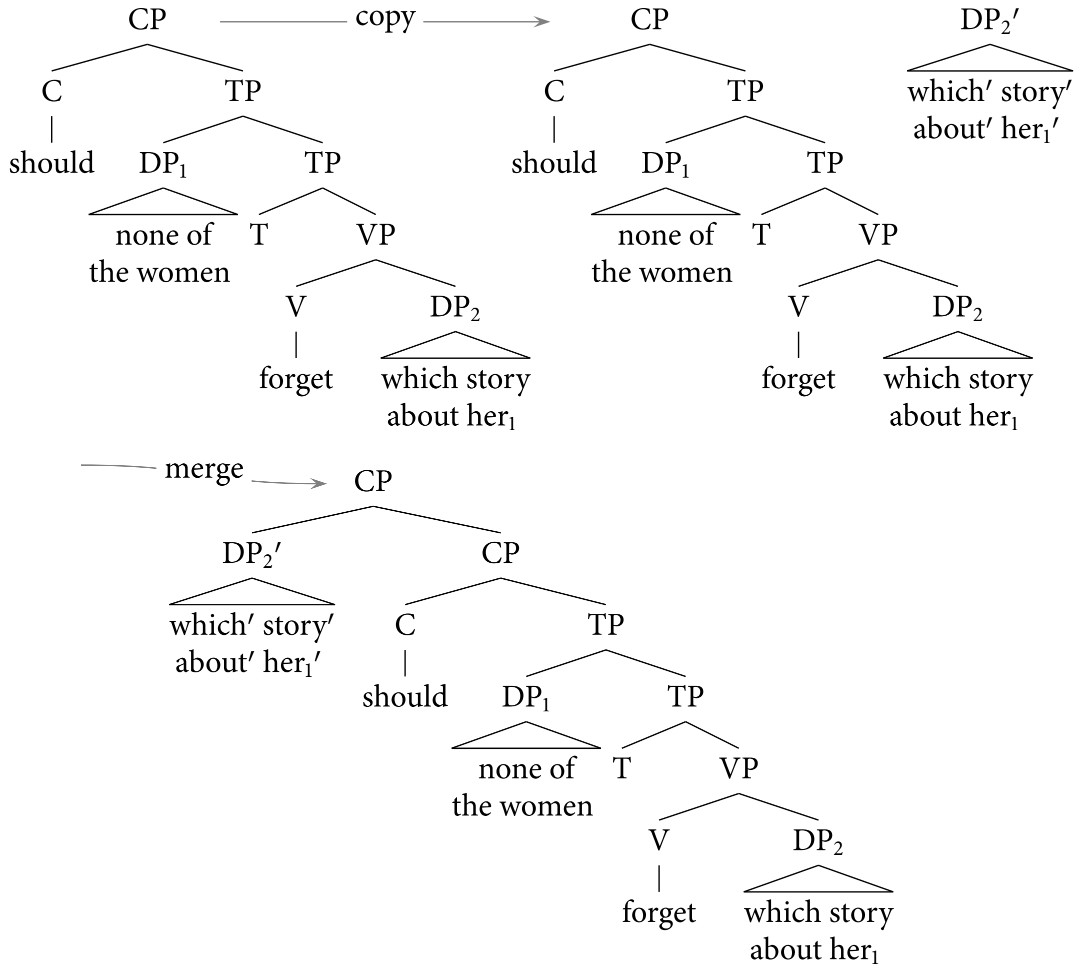
This gives us derivations like (19).

(19) (She asked) which book he knows.



This theory too is able to account for Semantic Displacement, and it does so in a way rather like that of the remerge account I will argue for. So, for instance, it gives an account of "reconstruction," a special case of Semantic Displacement that (20) illustrates.

(20) Which story about her should none of the women forget?



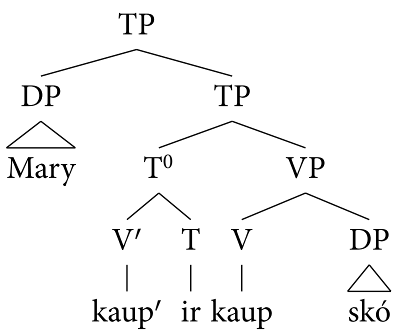
There is a copy of *her* spoken in a place different from where it is (apparently) interpreted.

These representations make Terseness arise from an operation that "deletes" one of the two phrases in the copy relation. I'll sketch the way this is done in Nunes (2004, chapter 1), which is a reworked version of his 1999 University of Connecticut dissertation, and an improved version of Nunes (1995), and then I'll modify it so it works with phrase markers with multidominance in them.

Nunes speculates that there is a deletion process that can be invoked to remove (portions of) one of the copies. He points out that adopting the simple assumption that a term and its copy cannot be distinguished by the constraints that define a well-formed linearization will cause movement structures to be unlinearizable. The deletion process he proposes could be invoked to "fix" these representations, making them linearizable and also deriving Terseness.

To see this, we'll need to spell out what those constraints are and what linearizations are. I will assume that syntactic representations are converted into phonological representations (PFs) by matching vocabulary items to terminals in the syntactic representations and linearizing those vocabulary items. I shall adopt the formalism, made popular by Kayne (1994), of expressing a linearization as a set of ordered pairs. A linearization results from an algorithm which evaluates a syntactic structure and computes from the information in that structure how each vocabulary item in the structure is ordered relative to every other vocabulary item in the structure. So, for instance, the structure in (21) would map onto the ordered pairs in (22).

(21)



(22) $\left\{\begin{array}{lll} \text{Mary} < \text{T}^0 & \text{T}^0 < \text{V} & \text{V} < \text{skó} \\ \text{Mary} < \text{V} & \text{T}^0 < \text{skó} & \\ \text{Mary} < \text{skó} & & \end{array}\right\}$

Note that I've represented the stem, or root, of the verb *kaupir* here with *kaup*. The set of ordered pairs in (22) involves the ordering relation "<," precedence. (22) corresponds to the string in (23).

(23) Mary kaupir kaup skó

This isn't quite right. We'll come back to its correction a bit later.

One thing to highlight about (22) is that the elements in the ordered pairs are words, or vocabulary items. They are not the terminals that make up those words. This differs from Kayne (1994), but it is what Nunes, and I, need. Thus, for instance, *kaup* and *ir* are not arranged according to the linearization algorithm.

In Kayne's work, the linearization procedure produced linearizations which were then subjected to well-formedness conditions. These conditions require that everything in the sentence be linearized and that the linearization be consistent. I will formulate those conditions as (24).

(24) a. All vocabulary items in the phrase marker p must be in the linearization of p.
(Totality)

b. For all vocabulary items, a and b in p, the linearization of p cannot include both $a < b$ and $b < a$.
(Antisymmetry)

c. For all vocabulary items, a, b, c in p, if the linearization of p includes $a < b$ and $b < c$ then it must include $a < c$.
(Transitivity)

He then builds a linearization algorithm that has a variety of interesting consequences for the shapes that phrase markers may have.

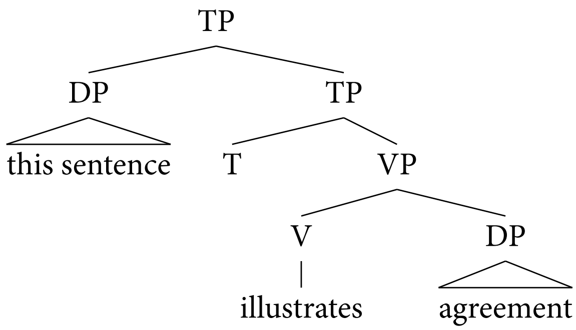
Kayne's version of Totality has the consequence that multidominant phrase markers are blocked. I've changed them so that this consequence is removed, but we will want to put in place something that generally has the effect of mapping phrases onto contiguous strings. I will therefore add to (24) a violable constraint, I'll call it "Contiguity," following Fanselow and Ćavar (2001, (56): 130), that has this consequence.

(25) Contiguity

Let A be the set of vocabulary items dominated by A and b be a vocabulary item not in A. If b precedes something in A, then b precedes everything in A, and if b follows something in A, then b follows everything in A.

These constraints – Contiguity, Totality, Antisymmetry and Transitivity – are sufficiently draconian that they manage to constrain the structure-to-string mapping almost enough to ensure reasonably accurate outcomes. Imagine that the linearization algorithm did nothing more than generate all possible orderings of vocabulary items and submit them to the constraints. The strings produced would include the correct one and a small number of alternatives. For instance, a linearization algorithm of this sort would produce from (26) a collection of sets that, once filtered through the constraints, would result in those listed in (27). (I will indicate the linearizations with the (more compact) strings they correspond to, rather than with the full sets of ordered pairs.)

(26)



(27)
a. this sentence T illustrates agreement
b. sentence this T illustrates agreement
c. this sentence illustrates agreement T
d. sentence this illustrates agreement T
e. this sentence T agreement illustrates
f. this sentence agreement illustrates T
g. sentence this T agreement illustrates
h. sentence this agreement illustrates T
i. T illustrates agreement this sentence
j. T illustrates agreement sentence this
k. illustrates agreement T this sentence
l. illustrates agreement T sentence this
m. T agreement illustrates this sentence
n. T agreement illustrates sentence this
o. agreement illustrates T this sentence
p. agreement illustrates T sentence this

The ill-formed linearizations in (27) are, many of them, well formed in other languages. For instance, (27f) corresponds roughly to how German would linearize this structure, and (27i) corresponds roughly to how Nuiean would. While not all of these outcomes are ones that we might want to permit cross-linguistically,[1] I will nonetheless treat them all as language-particular possibilities. The step from this range of linearizations to the one that is correct for English, then, engages that component of the theory which models word order variation. There are a variety of proposals in the literature on how to model word order variation. One of those is built into Kayne's linearization scheme. We don't need to choose among them, though, and it will be convenient (and harmless) to avoid engaging the details. In what follows, therefore, I will leave open how the choice from the possibilities allowed by the constraints to the one appropriate for English is made. I will call that portion of the linearization procedure that makes the language particular choice, the "language particular component."

[1] (27n,o) are vanishingly rare according to Dryer (1996).

The linearization algorithm will have four parts, then: a function that produces orderings among the vocabulary items in a sentence, a set of constraints, a procedure that steers how those constraints choose the linguistically viable ones from that set, and then a final component — the language particular component — that picks the language particular best ordering. The function that produces all the possible sets of orderings among vocabulary items in a phrase I will call *lin*.

(28) Let L(P) be a set of ordered pairs, $x < y$, where x and y are vocabulary items dominated by P, and "<" means "precedes."
lin(P) $=_{\text{def.}}$ the set consisting of every L(P).

Nunes's method of deriving Terseness hinges on the proposal that Antisymmetry, and the other constraints on a linearization, cannot distinguish a copy from the thing it is copied from. Moreover, the way Nunes executes his idea relies not on my version of Totality — designed for multidominant representations – buy Kayne's. Nunes system, then, invokes (29) and (30).

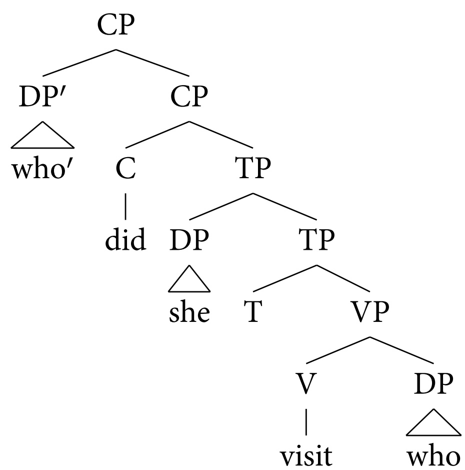
(29) *lin* applies to every X^0 in a phrase marker. *Totatlity*

(30) For Antisymmetry and the other constraints on linearization, α and Copy(α) are the same thing.

When these are coupled with the copy theory of movement, they will produce unlinearizable results.

Consider, for instance, what *lin* and the language particular component will together produce for a phrase marker created by movement.

(31)



The linearization of (31) that satisfies the language particular component as well as Totality and Transitivity is (32).

(32) = who′ did she visit who

$$\left\{\begin{array}{lllll} \text{who}' < \text{did} & \text{did} < \text{she} & \text{she} < \text{T} & \text{T} < \text{visit} & \text{visit} < \text{who} \\ \text{who}' < \text{she} & \text{did} < \text{T} & \text{she} < \text{visit} & \text{T} < \text{who} & \\ \text{who}' < \text{T} & \text{did} < \text{visit} & \text{she} < \text{who} & & \\ \text{who}' < \text{visit} & \text{did} < \text{who} & & & \\ \text{who}' < \text{who} & & & & \end{array}\right\}$$

This linearization has pairs like *who′ < visit* and *visit < who* in it, and under Nunes's proposal, these will be violations of Antisymmetry: *who′* and *who* are indistinguishable for Antisymmetry, and these pairs amount to saying, then, that *who* both precedes and follows *visit*.

To produce the correct outputs, Nunes suggests that there is a deletion process which removes the items introduced by movement that cause the violation of Antisymmetry. That deletion process is called "chain reduction." It can be formulated as (33).

(33) Chain Reduction

Delete α, α a syntactic constituent that has been put into the Copy relation.

Chain Reduction should be seen as an operation that is part of PF, the component of the grammar that converts syntactic representations to phonological ones. We want Chain Reduction to remove material from the phonological representation of the sentence, but leave unaffected how that material is semantically interpreted. In the case of (31), it could remove DP or DP′ and thereby produce representations from which the linearization algorithm produces the strings in (34).

(34) a. = who did she visit

$$\left\{\begin{array}{llll} \text{who}' < \text{did} & \text{did} < \text{she} & \text{she} < \text{T} & \text{T} < \text{visit} \\ \text{who}' < \text{she} & \text{did} < \text{T} & \text{she} < \text{visit} & \\ \text{who}' < \text{T} & \text{did} < \text{visit} & & \\ \text{who}' < \text{visit} & & & \end{array}\right\}$$

b. = did she visit who

$$\left\{\begin{array}{llll} \text{did} < \text{she} & \text{she} < \text{T} & \text{T} < \text{visit} & \text{visit} < \text{who} \\ \text{did} < \text{T} & \text{she} < \text{visit} & \text{T} < \text{who} & \\ \text{did} < \text{visit} & \text{she} < \text{who} & & \\ \text{did} < \text{who} & & & \end{array}\right\}$$

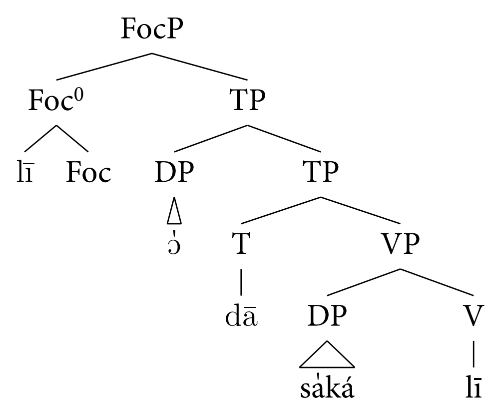
Both these representations satisfy Antisymmetry, and they also satisfy Totality, if Chain Reduction is seen as removing the terminals that Totality requires be in the linearization. Of these, only (34a) is the correct one. We can credit this to the language particular component.

This is how Terseness is derived.

Moreover, this method provides an interest account for those examples where Terseness seems to be relaxed. For instance, Vata predicate cleft constructions involve a structure like (35), in which the verbal root has moved and joined with a functional head that encodes focus.

(35) lī ɔ̍ dā sa̍ká lī
eat she/he perf rice eat
'she/he has eaten rice.'

(36)



Nunes argues that because the result of clefting a verb in Vata puts it within an X^0, Antisymmetry is able to be satisfied without invoking Chain Reduction. If constraints like Antisymmetry make reference to vocabulary items, and not the terminals from which those vocabulary items are composed, then putting a copy into a vocabulary item will effectively "hide" it from Antisymmetry. The representation in (36), for instance, can be assigned the linearization in (37).

(37) $\left\{ \begin{array}{llll} \text{Foc}^0 < \text{ɔ́} & \text{ɔ́} < \text{dā} & \text{dā} < \text{sȧká} & \text{sȧká} < \text{lī} \\ \text{Foc}^0 < \text{dā} & \text{ɔ́} < \text{sȧká} & \text{dā} < \text{lī} & \\ \text{Foc}^0 < \text{sȧká} & \text{ɔ́} < \text{lī} & & \\ \text{Foc}^0 < \text{lī} & & & \end{array} \right\}$

This linearization satisfies not only Totality, but Antisymmetry as well. When Foc^0 is matched against the vocabulary item that spells out "lī+Foc," the correct string associated with (36) is produced.

A virtue of the remerge theory of movement is that it derives Nunes' stipulation that Antisymmetry treats something and its copy as the same thing. On the remerge theory of movement, there are no copies, and a moved item really is one thing. It's one thing in two positions. I suggest, then, that we take Terseness to provide an argument for the remerge theory of movement.

Because the details of Nunes' method of deriving Terseness rely on the copy theory of movement, we'll have to translate it into something that fits the remerge theory. We can't adopt Chain Reduction, for instance. Deleting the vocabulary items that have been put into two positions by MERGE will not create a representation that allows those items to be pronounced in just one of the two positions they occupy, as it did on Nunes's scheme. Indeed, invoking an operation that is tied to the existence of Chains, in the manner that Nunes's deletion operation is, also no longer has traction. There is nothing in a multidominant representation that corresponds to a chain. We will have to look elsewhere for the mechanism that brings these representations into compliance with Antisymmetry, and thereby delivers Terseness.

If we jettison the version of Totality that Nunes (and Kayne) propose, and resort to the one in (24a), then the linearization algorithm, as I presented it above already has a way of avoiding the violations of Antisymmetry that movement will create.

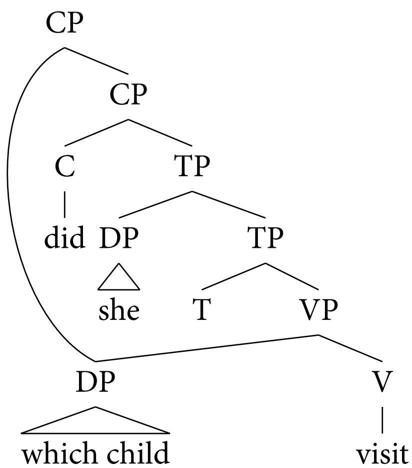
(24a) *Totality*

All vocabulary items in the phrase marker *p* must be in the linearization of *p*.

Because *lin* is defined so that it generates every possible set of orderings, including those that are incomplete, it need not produce linearizations that will violate Antisymmetry to begin with. It's the job of the constraints, and the language particular component, to determine which of the sets of orderings offered by *lin* survive. Because the version of Totality in (24a) only requires that every vocabulary item within a structure be mapped onto a position in the resulting string, it will allow orderings that do not take into account all of the positions a vocabulary item might occupy. For these reasons, then, *lin* need not produce an ordering that makes a vocabulary item that has two (or more) positions fall into more than one spot in the string. Since this is what Antisymmetry requires, this is how Terseness arises.

To see this, consider some of the orderings that *lin* will produce for (38). These include those in (39).

(38) Which child did she visit?



(39) a. = which child did she visit

$$\left\{\begin{array}{lllll} \text{which} < \text{child} & \text{child} < \text{did} & \text{did} < \text{she} & \text{she} < \text{T} & \text{T} < \text{visit} \\ \text{which} < \text{did} & \text{child} < \text{she} & \text{did} < \text{T} & \text{she} < \text{visit} & \\ \text{which} < \text{she} & \text{child} < \text{T} & \text{did} < \text{visit} & & \\ \text{which} < \text{T} & \text{child} < \text{visit} & & & \\ \text{which} < \text{visit} & & & & \end{array}\right\}$$

b. = did she visit which child

$$\left\{\begin{array}{lllll} \text{did} < \text{she} & \text{she} < \text{T} & \text{T} < \text{visit} & \text{visit} < \text{which} & \text{which} < \text{child} \\ \text{did} < \text{T} & \text{she} < \text{visit} & \text{T} < \text{which} & \text{visit} < \text{child} & \\ \text{did} < \text{visit} & \text{she} < \text{which} & \text{T} < \text{child} & & \\ \text{did} < \text{which} & \text{she} < \text{child} & & & \\ \text{did} < \text{child} & & & & \end{array}\right\}$$

Both of these sets of orderings obey Totality, since every vocabulary item shows up in the resulting strings. They also obey Antisymmetry and Transitivity. However, they differ with respect to Contiguity, which is repeated in (40), and that part of the language particular of English that requires wh-phrases to be spoken in their higher position. I will call that the Wh Criterion; we can formulate it with (41).

(40) Contiguity

Let *A* be the set of vocabulary items dominated by A and *b* be a vocabulary item not in *A*. If *b* precedes something in *A*, then *b* precedes everything in *A*, and if *b* follows something in *A*, then *b* follows everything in *A*.

(41) The Wh Criterion

If just one wh-phrase is merged to CP, then *lin*(CP) must position that wh-phrase so that it precedes everything else in that CP.

(39a) obeys the Wh Criterion, but violates Contiguity. (It violates Contiguity because *she* (for example) precedes *visit* but not the other vocabulary items in the VP.) (39b), by contrast, obeys Contiguity but violates the Wh Criterion. The language particular component is, by definition, inviolable and therefore of these two orderings, (39a) is the better. Further, there is no way of avoiding a violation of Contiguity if the Wh Criterion is to be satisfied. In particular, there are no elements of *lin*((38)) that by virtue of violating Totality, Antisymmetry or Transitivity manage to satisfy the Wh Criterion and also avoid violating Contiguity. There are no candidates that beat out (39a) by relying on violations of one of the other of our constraints on linearization because these will all involve either additional violations of the Wh Criterion or violations of Contiguity. We must understand Contiguity to be violable.

(42) Of the constraints on *lin*, only Contiguity is violable.

(43) Every element of *lin*(P) which incurs *n* violations of Contiguity is ungrammatical if there is an element of *lin*(P) that incurs fewer than *n* violations Contiguity.

Consider next (44), which is also a member of the set of orderings that *lin* produces when applied to (38).

(44) = which did she visit child

$$\left\{\begin{array}{lllll} \text{which} < \text{did} & \text{did} < \text{she} & \text{she} < \text{T} & \text{T} < \text{visit} & \text{visit} < \text{child} \\ \text{which} < \text{she} & \text{did} < \text{T} & \text{she} < \text{visit} & \text{T} < \text{child} & \\ \text{which} < \text{T} & \text{did} < \text{visit} & \text{she} < \text{child} & & \\ \text{which} < \text{visit} & \text{she} < \text{child} & & & \\ \text{which} < \text{child} & & & & \end{array}\right\}$$

This satisfies Totality, Antisymmetry and Transitivity, and so threatens to be a successful linearization. Under the system proposed here, there are two potential reasons that (44) is blocked. One is that it may not satisfy the Wh Criterion. This depends on how the underlying causes of the Wh Criterion pan out. It could be that it requires the entire DP headed by a wh-determiner to be spelled out in the merged-to-CP position. Irrespective of the Wh Criterion, however, Contiguity will disfavor (44) relative to (39a). Contiguity is violated in both (44) and (39a), but it is violated less in (39a). In (44), CP, both TPs, VP and DP violate Contiguity; in (39a), CP, both TPs and VP violate Contiguity, but DP doesn't. There is therefore one fewer violation of Contiguity in the case of (39a). Here, then, is the reason for letting the evaluation procedure be sensitive to the numbers of times that Contiguity is violated.

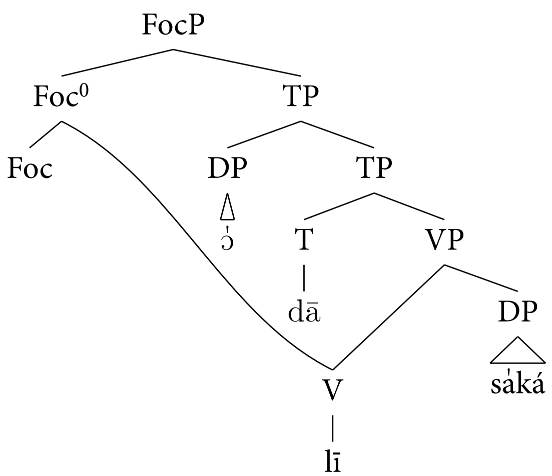
The Wh Criterion makes unavoidable a violation of Contiguity, and of those candidates that satisfy the Wh Criterion (and violate Contiguity), (39a) is the best: it violates none of the other constraints on linearizations and it violates Contiguity the fewest number of times required by the Wh Criterion. This correctly derives, then, that (38) maps onto the string in (39a). Without the Wh Criterion, (39b) becomes the winning linearization, because it uniquely violates none of the constraints on linearizations. What we see, then, is that our linearization algorithm, in concert with a multidominant model of movement, allows overt movement only if there is a language particular constraint that forces a phrase to be spoken in the higher of its two positions. Further, when there is a language particular constraint forcing a phrase to be spelled out in its higher position, then Contiguity will favor spelling out all of that phrase in the higher position.

The account Nunes gives of those cases where Terseness is lifted is preserved on my recasting of his system. Consider how my system will apply to the case of the Vata clefted predicates, for instance.

(35) lī ɔ̍ dā sa̍ká lī
eat she/he perf rice eat
'she/he has eaten rice.'

On the remerge definition of movement, this will now have the representation in (45).

(45)



Because *lin* cannot see inside Foc^0, the orderings it will produce do not make reference to the verbal root that is part of Foc^0. Just as in Nunes's scheme, the best output *lin* will produce is (37).

(37) = lī ɔ̍ dā sa̍ká lī

$$\left\{\begin{array}{llll} \text{Foc}^0 < \text{ɔ̍} & \text{ɔ̍} < \text{dā} & \text{dā} < \text{sa̍ká} & \text{sa̍ká} < \text{lī} \\ \text{Foc}^0 < \text{dā} & \text{ɔ̍} < \text{sa̍ká} & \text{dā} < \text{lī} & \\ \text{Foc}^0 < \text{sa̍ká} & \text{ɔ̍} < \text{lī} & & \\ \text{Foc}^0 < \text{lī} & & & \end{array}\right\}$$

Managing to avoid a violation of Antisymmetry works the same in both Nunes's and my account. Notice as well that (37) manages to avoid violating Contiguity as well.

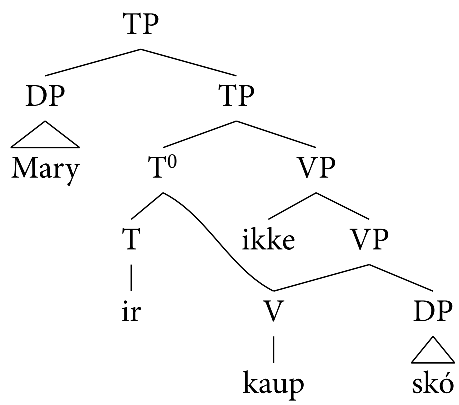
This gives us an explanation for why the clefted verb in Vata can be spelled out in both positions. But recall that the clefted verb in Vata must be spelled out in both positions. Nunes's ensured this by making the process that prevents spelling out the verb in both spots costly: Chain Reduction is employed only to the point necessary. This explanation doesn't have a correlate under the scheme I am proposing, however. Under both systems, that the verb is spelled out in the higher position is probably independently ensured by Totality. Totality requires that the vocabulary item made up of Foc^0 and the verb be part of the resulting string, and to the extent that this requires that the verb be part of that vocabulary item it will force the verb to be put in the higher position. So, what we need to figure out is why the verb must be spelled out in its lower position as well.

It's useful to compare this situation with the one involving Icelandic verb movement that we started with.

(1a) Mary kaupir ikke skó? (*Icelandic*)
Mary buys not shoes — Head Movement

On the remerge theory, this will get a representation like (46).

(46)



A winning output from *lin* — one that satisfies Totality, Transitivity, Antisymmetry, Contiguity and, to the extent known, the language particular component too — is (47)

(47) = Mary kaupir ikke kaup skó

$$\left\{\begin{array}{llll} \text{Mary} < \text{T}^0 & \text{T}^0 < \text{ikke} & \text{ikke} < \text{V} & \text{V} < \text{skó} \\ \text{Mary} < \text{ikke} & \text{T}^0 < \text{V} & \text{ikke} < \text{skó} & \\ \text{Mary} < \text{V} & \text{T}^0 < \text{skó} & & \\ \text{Mary} < \text{skó} & & & \end{array}\right\}$$

This isn't the outcome we want. Here, we want to force the verb to not be pronounced in its lower position.

Why are the Icelandic and Vata outcomes opposite? I don't know, but I speculate that it has to do with morpho-phonological requirements. Perhaps Icelandic has no vocabulary item that corresponds to a verbal root. Indeed, the citation forms of Icelandic verbs are not roots but, like English, inflected forms. If there is no vocabulary item that can be matched to the V position in (46), then this will explain why nothing is pronounced in this position.

For Vata, it may be that the lower verb is required to be spelled out so that the tones associated with the verb can be pronounced. Koopman (1984) shows that clefted verbs in Vata always appear with midtones in their cleft position, and display the tones associated with the verb in their lower copy only. This was illustrated in (8).

(8)

ngɔ̄nʊ̄	ǹ	wà	nā	ǹ	ka̍	ngɔ́nʊ̍	a̍
sleep	you	want	NA	you	FUT-A	sleep	Q

'Do you want to sleep?'

(Koopman 1984, (2a): 154)

Perhaps the tones associated with the verb are also subject to Totality in Vata. If so, whatever forces the clefted verb to be expressed with midtones will consequently force the verb to also be spoken in its lower position, where it can support the expression of those tones.

References

Aoun, Joseph, and Elabbas Benmamoun. 1998. Minimality, reconstruction, and PF movement. *Linguistic Inquiry* 29:569–597.

Aoun, Joseph, Lina Choueiri, and Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32:371–403.

Citko, Barbara. 2005. On the nature of merge: External merge, internal merge, and parallel merge. *Linguistic Inquiry* 36:475–496.

Dryer, Matthew S. 1996. Word order typology. In *Handbook on syntax*, ed. J. Jacobs, volume 2, 1050–1065. Walter de Gruyter.

Engdahl, Elisabet. 1980. The syntax and semantics of questions in Swedish. Doctoral Dissertation, University of Massachusetts, Amherst, Massachusetts.

Fanselow, Gisbert, and Damir Ćavar. 2001. Remarks on the economy of pronunciation. In *Competition in syntax*, ed. Gereon Müller and Wolfgang Sternefeld, 107–150. Berlin: Mouton de Gruyter.

Felser, Claudia. 2004. Wh-copying, phases, and successive cyclicity. *Lingua* 114:543–574.

Frampton, John. 2004. Copies, traces, occurrences, and all that: Evidence from Bulgarian multiple *wh*-phenomena. Unpublished manuscript, Northeastern University.

Gärtner, Hans-Martin. 1997. Generalized transformations and beyond. Doctoral Dissertation, University of Franfurt/Main.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.

Kobele, Gregory Michael. 2006. Generating copies: An investigation into structural identity in language and grammar. Doctoral Dissertation, University of California, Los Angeles, Los Angeles, CA.

Koopman, Hilda. 1984. *The syntax of verbs*. Dordrecht, The Netherlands: Foris Publications.

Nunes, Jairo. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral Dissertation, University of Maryland.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32:303–344.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.

Rett, Jessica. 2006. Pronominal vs. determiner *wh*-words: evidence from the copy construction. In *Empirical issues in syntax and semantics 6*, ed. Olivier Bonami and P. Cabredo Hofherr, 355–374. CNRS.

Starke, Michal. 2001. Move dissolves into merge: A theory of locality. Doctoral Dissertation, University of Geneva.

de Vries, Mark. 2007. Internal and external remerge: On movement, multidominance, and the linearization of syntactic objects. Unpublished manuscript, University of Groningen.

Trace Conversion and Wh Movement

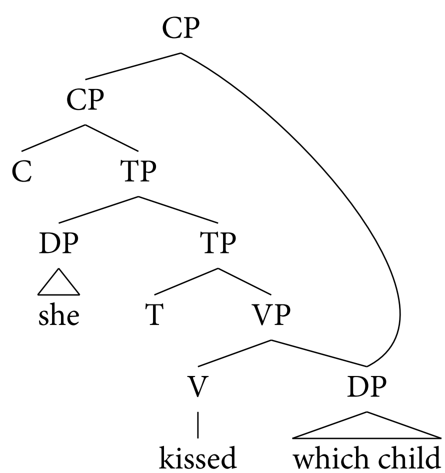
Kyle Johnson
University of Massachusetts at Amherst
kbj@linguist.umass.edu

SICOGG 12
18 August

Today we'll start to take a closer look at how movement structures are interpreted. Our focus in this lecture will be Wh Movement. Tomorrow, we'll look at Quantifier Raising.

Simple wh-questions have the shape in (1), on the remerge model of movement.

(1) (I know) which child she kissed.



A standard, simple, view of the meaning of questions is that they denote a set of propositions, each proposition offering a kind of answer in those cases where the question is answer-seeking. This is the view introduced by Hamblin and modified by Kartunnen. One way of representing a set is with the λ-operator, which can be used to represent a function.

(2) $\lambda x \mathrm{P}(x)$ = that function which, when applied to a, gives $\mathrm{P}(a)$.

A function can be equated with the set of things that that function holds of. For (1), for instance, we can give the question a denotation like:

(3) $\lambda p\ \exists x\ x$ is a child $\wedge\ p$ = she kissed x

So, the challenge is to get this kind of meaning out of (1).

The central problem a remerge definition poses is that it baldly predicts that the single meaning that is associated with the moved item should be found in both of its positions. That isn't what we want from questions. Instead, we must associate the moved wh-phrase with both a binder meaning and a variable meaning. The first person to appreciate, and try to solve, this problem is Elisabet Engdahl.

In Engdahl (1980), what she proposes is that the moved wh-phrase has two meanings, and they are introduced at their two positions. The meaning introduced in the lower position must be a variable. But it should also provide a way of explaining "reconstruction," one of the manifestations of Semantic Displacement that is found in wh-movement.

(4) Which picture of himself should no one put on his website?

Note that it is the position from which movement has occurred that matters.

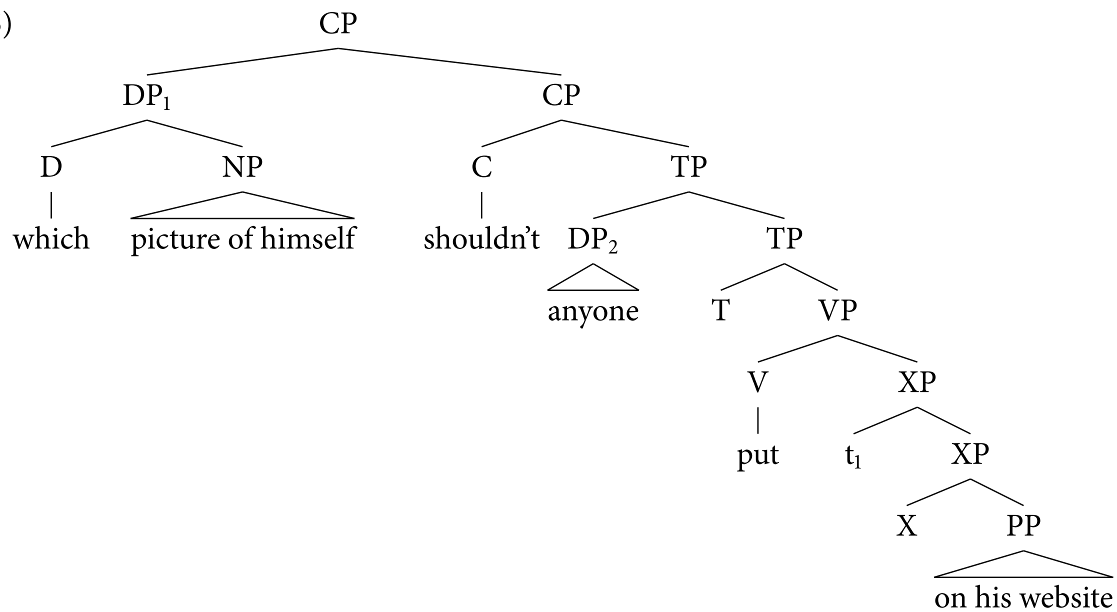
(5) a. Which picture of himself_1 does this indicate that no one_1 should bring?

b. * Which picture of himself_1 indicates that no one_1 should bring?

Engdahl (1980) suggested doing that by letting the NP part of the moved wh-phrase be interpreted in its lower position. We can speculate that the NP part is not interpreted in its higher (spoken) position however, since in this position it is neither c-commanded by its binder, nor meets the locality condition that anaphors typically impose on their antecedents.

There is another, somewhat less obvious, difficulty involved in capturing these reconstruction cases. This problem is easier to appreciate in cases involving universal quantification, like that in (6).

(6)



$[\![t_1]\!] \approx$ picture-of-himself$_2(x_1)$

If we adopt a Hamblin/Kartunnen style analysis of questions, moved in the direction we're going, then for (6) we'll get an interpretation along the lines of (7).

(7) $\lambda p\ \exists x$ p = not anyone$_2$ should put picture-of-himself$_2(x)$ on his website.

(7) characterizes the question as seeking the identity of a single picture with the expansive property of being of a bunch of guys, none of whom should put it on their website. That's not what we want. We want something that allows the pictures to vary with the variable it contains. The anaphoric connection between a moved phrase and its trace must be capable of carrying this duty. Elisabet Engdahl gave us a way of doing that.

What she suggests is that the wh-phrase in the lower position gets an interpretation like the definite description in examples of "donkey anaphora," like that in (8).

(8) Everyone who owns a donkey loves the/that donkey.

These definite descriptions also act like restricted variables. Following Cooper (1979), she adopted the view that they have buried within them a function that picks out individuals which the restrictor *donkey* tells us are donkeys. In (8), that function is something like "owned by y."

(9) Everyone$_1$ who owns a donkey loves x, $f(x) \wedge$ donkey(x)

f = owned by y_1

Her idea, then, is that in the lower position, a wh-phrase is interpreted as a function that can contain a variable within it. The values this function gives can depend on the value given to the variable it contains. That's how, (8), the donkeys vary with the values given to *everyone*. Indeed, it is this function that we can see questions as asking for the identity of.

(10) a. Which picture did you say you'd show every girl?
$\lambda p\ \exists f\ p$ = you said you'd show every girl$_1$ $f(x_1)$
f might be:
Sally → the picture of the Eiffel tower
Mary → the picture of the Milkmaid
Myrtle → the picture of hot-rods
f might be:
her favorite picture
f might be:
the picture of George Clooney

And we get from this model an account, too, of why the trace can only get a value that varies with respect to expressions that c-command it.

(11) Which picture did you show the guy every girl$_1$ knows?
$^*\lambda p\ \exists f\ p$ = you showed the guy every girl$_1$ knows $f(x_1)$
f can't be:
Sally → the picture of the Eiffel tower
Mary → the picture of the Milkmaid
Myrtle → the picture of hot-rods
f can't be:
her favorite picture
f can be:
the picture of George Clooney

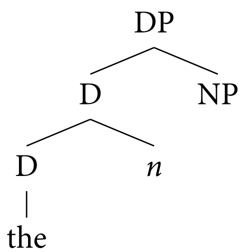
What we need to do now is put into the meaning of the phrase in the lower position the contribution that the restrictor — the NP — makes.

The first step we can take, then, is to reïnvoke Engdahl's idea that the trace left by movement is semantically like a donkey-type DPs. Let's start by considering the syntax, and semantics, of these expressions.

If we start with a model of donkey-type DPs like that offered in Cooper (1979), we will want to build in a function whose arguments can be bound. We should notice that it is not just definite descriptions that can have this interpretation in donkey anaphora sentences, but personal pronouns can as well:

(12) a. Every man who owns a donkey kisses it.
b. Every man who owns a donkey kisses the donkey.
c. Every man who owns a donkey kisses that donkey.

We should build into pronouns, traces and definite descriptions a relational meaning, then, and to the extent that this relation is the same in all these cases, we will want an explanation for why it travels in this particular pack. A commonplace idea about explaining the similarity between pronouns and definite descriptions is to adopt Postal (1969)'s proposal that pronouns are "intransitive" definite determiners. One way this can be expressed is by building the relational meaning into the determiner/pronoun. Let's look at such a way that steals shamelessly from a handout by Irene Heim used in teaching a course at the LSA institute several years ago. Its ingredients can be found in Elbourne (2005), Chierchia (1995) and Rullmann and Beck (1998). The first innovation will be to let definite determiners/pronouns take two arguments: one is an index, and the second the NP you see in the case of (most) definite descriptions.[1]

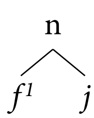
(13)


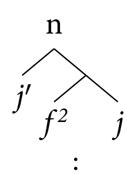
The NP will express a presupposition. "n" represents an index, which we will take to be capable of being complex. It can bundle together a function and its arguments. Let f^n be a variable ranging over functions of any adicity. I'll follow Engdahl's notation in representing the valency of f^n as follows.

(14) a. f^0 = an f with no variable in it (the constant function)
b. f^1 = an f with one variable in it
c. f^2 = an f with two variables in it
⋮

An index can have different shapes, depending on the valency of the f it contains. I'll represent the arguments of f with j.

(15) a.


b.


c.


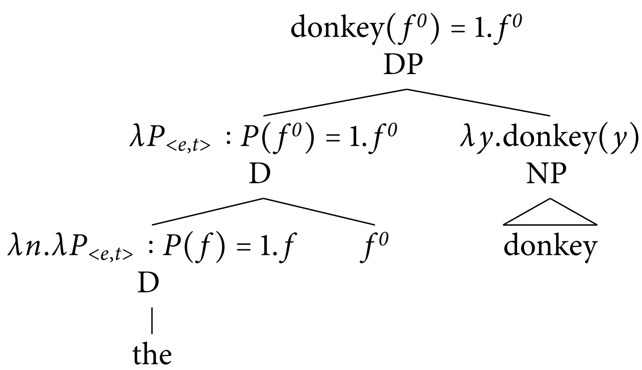
Heim's handout expresses how the index and the NP are put together with the denotation for *the* given in (16):

(16) $[\![\text{the}]\!] = \lambda n.\lambda P_{<e,t>} : P(n) = 1.n$

where the business between ":" and "." gives the conditions under which the function is well-defined, and so expresses the presupposition. When f is a constant function, as it might be in definite descriptions in out of the blue statements, we'll get something like:

[1] This comes from Elbourne 2005, chapter 3.

(17) The donkey needs a kiss.

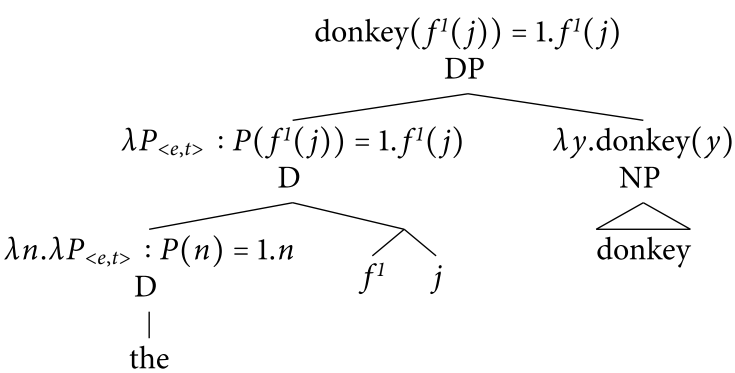


= The value given to f^0, if donkey holds of that value. Undefined otherwise.

NB: The uniqueness presupposition has been suppressed.

When f has an adicity greater than 0, it will come with silent arguments. I'll represent these with "j." This is what we have in (12b).[2]

(12b) Every man who owns a donkey kisses the donkey.

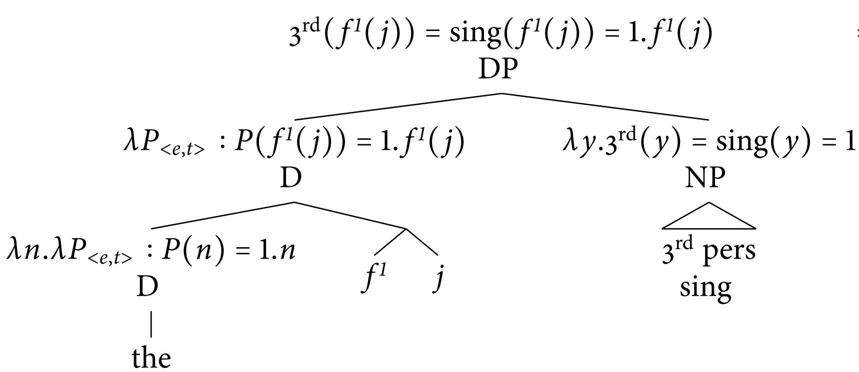


= The value given to $f^1(j)$, if donkey holds of that value. Undefined otherwise.

Because j is c-commanded by *every man* in (12b), it, and the result of applying f^1 to it, can vary in a way that depends on the values given to *every man*.

The case in (12a) could now be given an identical treatment; the chief difference between the two being that the presupposition is expressed as an NP in (12b), but as ϕ features in (12a).

(12a) Every man who owns a donkey kisses it.

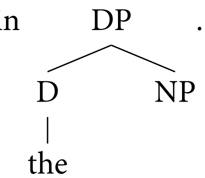


= The value given to $f^1(j)$, if third person and singular holds of that value. Undefined otherwise.

[2] Giving determiners relational indices like this is in Chierchia (1995).

Or we could adopt Elbourne's view that this kind of pronoun involves NP ellipsis, and so its semantic computation will look precisely as in (12b). We'll need a rule of Spell Out that makes "the" get pronounced as a pronoun, rather than a determiner, under these conditions. It might look something like (18).

(18) Assume that ϕ-features are on NP in

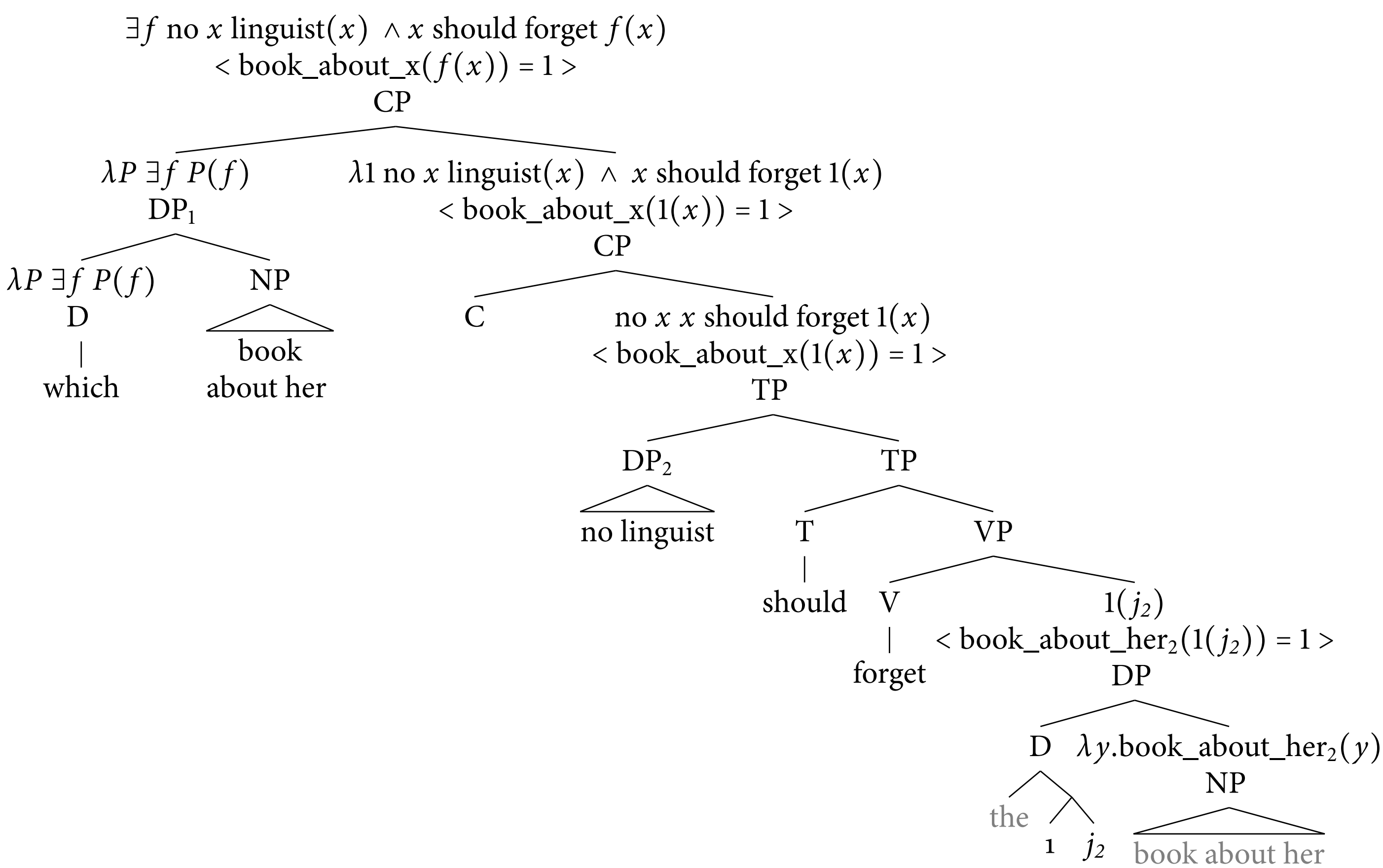


.

If SPELL-OUT(NP) does not include SPELL-OUT(ϕ), then SPELL-OUT(the) must.
Otherwise SPELL-OUT(the) = *the*.

So, this is how a normal donkey-type DP looks. We can imagine that something like this is what Engdahl would assign to the meaning of a moved wh-phrase in its lower position. In its higher position, we'll want the wh-phrase to introduce a binder that will quantify over the functions. This will make the questions about the identity of those functions. We can do that by assigning a denotation to *which* that makes it a binder, and then apply that denotation in the higher position. If we were to revert to the copy theory of movement for a moment, and represent syntactically these two denotations, we'd get something like (19).

(19) Which book about her should no linguist forget?



Note that the NP *book about her* in the higher copy is not semantically interpreted there. That's something we saw to be necessary for such cases. I've put the presupposition introduced by ⟦the⟧ and ⟦book about her⟧ in "< >." What's missing from (19) is the set of propositions part. All we've got here is a sentence that existentially quantifies over functions — functions that pick out books about x that no x, x a linguist, should forget. What we want is the set of such propositions. One way of doing that would be to give the C that heads questions a meaning that introduces the proposition part and then enrich the denotation of *which* so that it formed a set from that proposition. That is a standard method, but I'll go a different direction in a moment, so let's hold off on completing this picture for a little while.

The standard account of the semantics of movement does precisely what we see in (19). That account is due to Danny Fox,[3] who adopts a copy theory of movement and assumes that there is a rule which converts the lower copy into something that matches what we have in the lower position of (19). He calls that rule "Trace Conversion," and the way he formulates it is in (20), which is slightly different from how I've built the meaning of the lower copy, but close enough to be roughly equivalent.

(20) Trace Conversion

In $[_{\phi'}\ \mathrm{DP}_n\ [_{\phi}\ \ldots \mathrm{DP}_n \ldots]]$ interpret ϕ as a function that maps, x, to the meaning of ϕ[x/n].

ϕ[x/n] is the result of replacing the head of every constituent bearing the index n in ϕ with the head *the*$_x$, whose interpretation, $[\![\mathrm{the}_x]\!]$, is: $\lambda P : P(x).x$.

(adapted from Fox 2003, (52): 111)

Fox thinks of this rule as a kind of generalized binding rule. That is, it is meant to be the rule that relates binders to their variables. To mesh with the copy theory of movement correctly, it's been imbued with the ability to change the meaning assigned to a determiner. But I think we must see his Trace Conversion rule, really, as something that is specific to movement, and not general to all variable-binding relations. If we don't restrict it to just movement contexts, we wouldn't expect the inequality in (21).

(21) Every problem$_1$ challenges us to find no problem's$_1$ solution.

≠ Every problem$_1$ challenges us to find the/that problem's$_1$ solution.

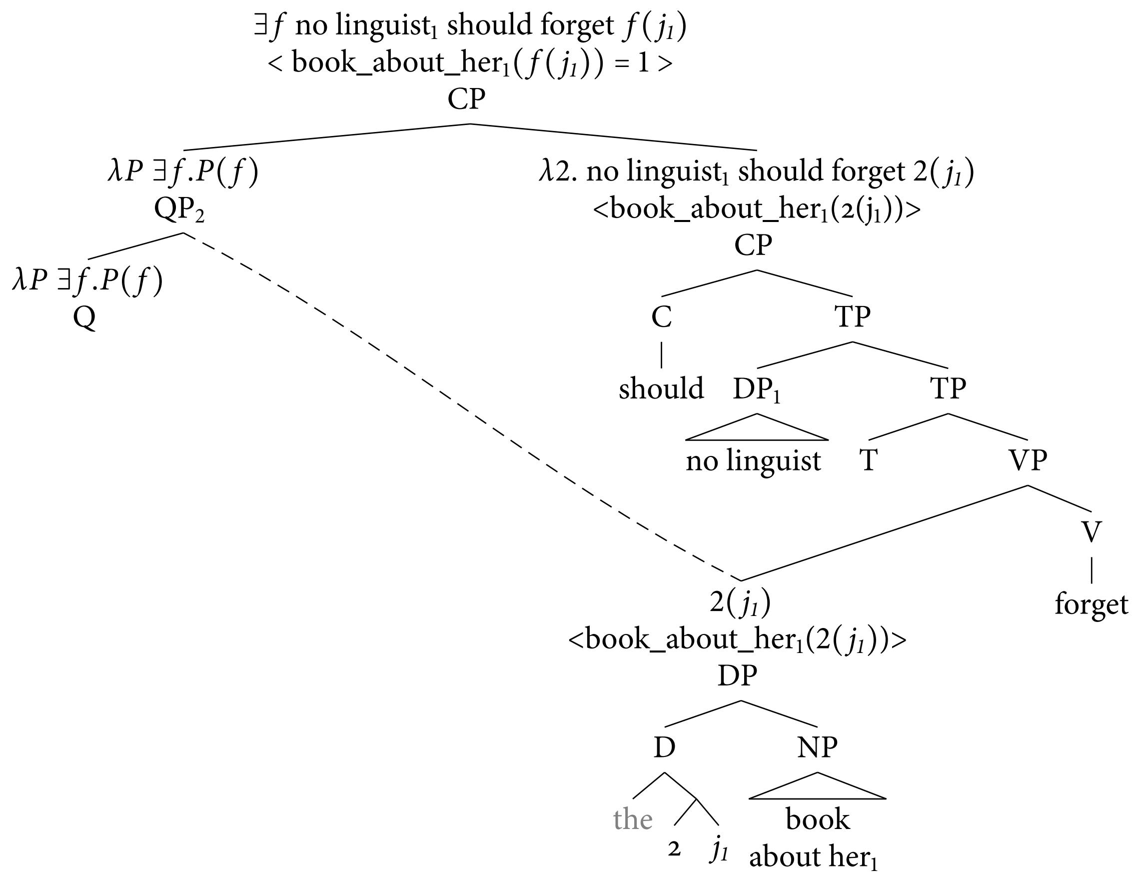
I think we should be skeptical of Trace Conversion. It says that the syntax-to-semantics mapping allows for rules that change what a determiner means. I think we can legitimately ask why that should be so, and why we don't see things like this elsewhere.

Engdahl's approach is also troubling, I feel. It says that there are expressions with two denotations, and that they are tailor made for movement relations. But this feels ad hoc to me as well. Why should there be certain expressions whose two denotations just happen to be ones that provide a binder for the other denotation?

The picture I'd like to replace these with is one that says that wh-movement involves putting together a definite description of the sort that we see in donkey anaphora, with a Q morpheme that produces the question and binds the variable in the definite description. As a first approximation, I suggest something like (22).

[3] See Fox (1999). There are others, all closely related to Fox's. Two of them are found in Sauerland (1998, 2004) and Romero (1998), as well as Engdahl (1980, 1986).

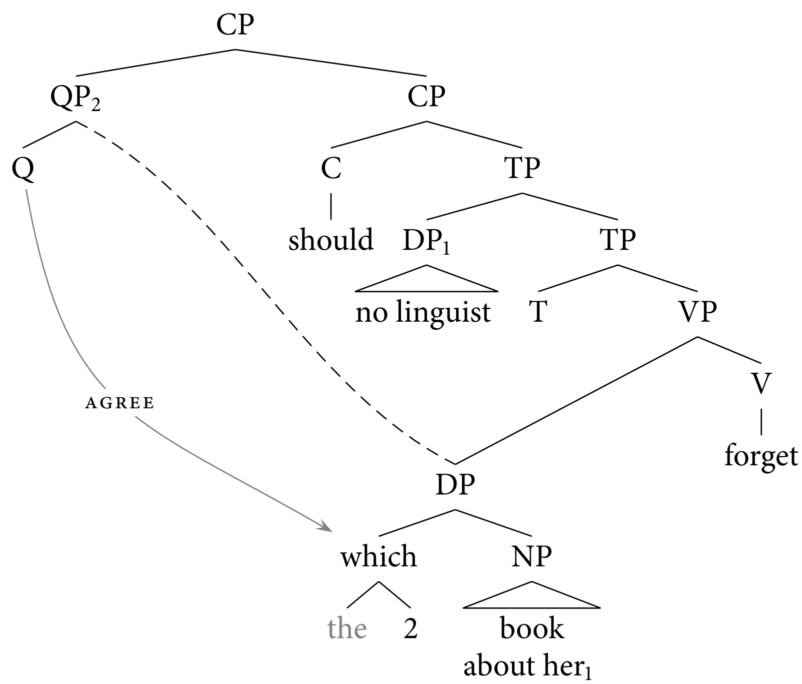
(22)



On this view, technically what has moved is just the DP portion that is interpreted as a variable. This denotation it supplies to the object position of *forget*. That DP has merged with the higher Q, which is the binder of the index within the DP in object position. It has merged with that Q, but its denotation is not computed there. I've indicated that with the dotted line. As a consequence, the QP in the higher position has the same meaning as the Q which heads it.

We need to determine where in (22) the question word *which* is inserted. I am going to assume that it is the D position of the DP that gets matched to the question word. But I want the form this D has to reflect the fact that there is the question morpheme: Q. So I suggest that we let the D get the features responsible for spelling it out as *which* from the Q morpheme under Agreement.

(23)



The DP that has moved in (23) has two different positions that *lin* can calculate its position from. The grammar of English requires that *lin* position this DP according to its higher position.

Okay, that's a start. We've got two things left to do before we'll have a complete picture.

1. Where does the set of propositions part of the meaning come from in questions?
2. What causes Q and DP to merge in the particular way indicated in (23)?

1 Alternatives

To get these remaining pieces, I'll start by looking at how questions are formed in (some) wh-in-situ languages. In these languages, the D that is found in the lower DP and the Q that binds off the variable in these lower expressions map onto separate morphemes. In Japanese, for instance, a morpheme on the verb marks the scope of a question, and in the position of the variable is an interrogative phrase.

(24) (Kimi-wa) dono-gakusei-ga nattoo-o tabe-tagatte-iru-to omoimasu-ka?
(you-top) which-student-nom natto-acc eat-desirous-be-C think-Q
(Which student do you think wants to eat natto?)

We might think of these languages as having the same syntax that I've given to English questions, but with a small difference in how the syntax-to-morphology works. In Japanese, the D and Q are mapped onto separate morphemes and, perhaps relatedly, the shared DP is spelled out in the lower of its two positions.[4] Alternatively, we could see the Q and the DP as being completely independent, and there being no remerge/movement in these examples.

Interestingly, though, in these kinds of questions there is (sometimes) a kind of intervention effect that does not arise in overt movement cases. This shows up for some dialects of Korean, according to Beck and Kim (1997). According to them, the presence of *man* ('only') in (25) is responsible for destroying the relationship between *nuku* and *ni*, thereby causing this sentence to be ill-formed.

[4] See Hagstrom (1998, 2000) and Kishimoto (2005) for proposal that would translate into my system that way.

(25) * Minsu-man nuku-lûl po-ss-ni?
Minsu-only who-acc see-Pst-Q
'Who did only Minsu see?'

Beck (2006) provides an explanation for these intervention effects — sometimes called "Beck Effects" — that gives to questions a slightly different semantics than I have adopted. Her semantics will, it turns out, provide the missing pieces to our picture so far. So I will modify what we have to bring it in line with her analysis.

Her leading idea follows Hamblin (1973) more closely than it does Kartunnen in the Hamblin/Kartunnen style account of questions. Hamblin suggested that the question word in questions introduces not a variable that gets bound off, but instead introduces "alternatives" that the set of propositions which makes a question vary on. These alternative generating terms have also been used by Rooth (1985) to model focus. What will go wrong in (25) is that the focus sensitive operator *man* will interfere with the question particle *ni*'s access to the alternatives generated by *nuku*. Let's take a brief, sketchy, look at this.

The idea in Rooth (1985) is that focused items have, in addition to their "normal" denotation another denotation that certain operators like *only* interact with. That other denotation – its focus value – is a set made up of alternatives to the term. That set is made up of objects that are of the same semantic type as the normal denotation of the focused term. In something like (26), then, *Sally* has a normal semantic value that allows it to refer to Sally, and the focus semantic value in (26b).

(26) She only visited Sally_F.
a. $[\![\text{Sally}]\!]$ = the individual named Sally
b. $[\![\text{Sally}]\!]^f$ = { Jerry, Max, Sam, Sean, Mary,…}

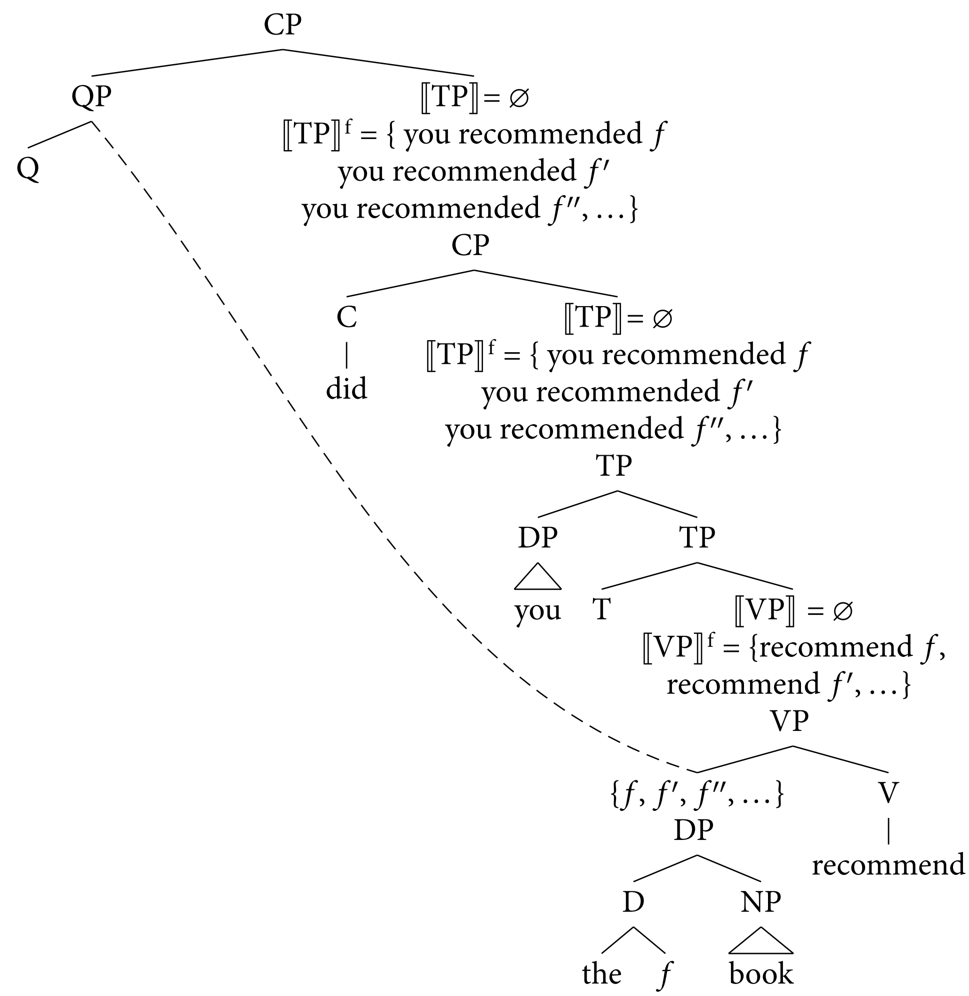
Phrases that contain terms with a focus semantic value inherit a focus semantic value by composing their normal denotation with the term in a point-wise fashion. In the case of a verb composing with its object, the verb will compose with each of the alternatives in the focus semantic value of the object by function application, and produce a set of alternative VP meanings.

(27) $[\![\text{visited Sally}]\!]^f$ = { visited Jerry, visited Max, visited Sam, visited Sean, visited Mary,…}
$[\![\text{visited Sally}]\!]$ = visited(Sally)

Rooth then gives *only* a meaning that, when combined with the VP, returns the same ordinary semantic value that the VP has, but adds that all the members of the focus semantic value of the VP are false.

The idea in Beck (2006) is to let the wh-words have the same focus semantic value that focussed items do. But she suggests that they have no regular semantic value. This will cause the phrases they are contained in to have only focus semantic values: they will be sets of alternatives that vary only with respect to the value given to the wh-word. We are letting the part of a wh-word that the question abstracts over be a function. If we leave all the rest of our system the same, but import Beck's idea that the "variable" in the wh-word is an alternative generator, we'll get representations like that indicated in (28).

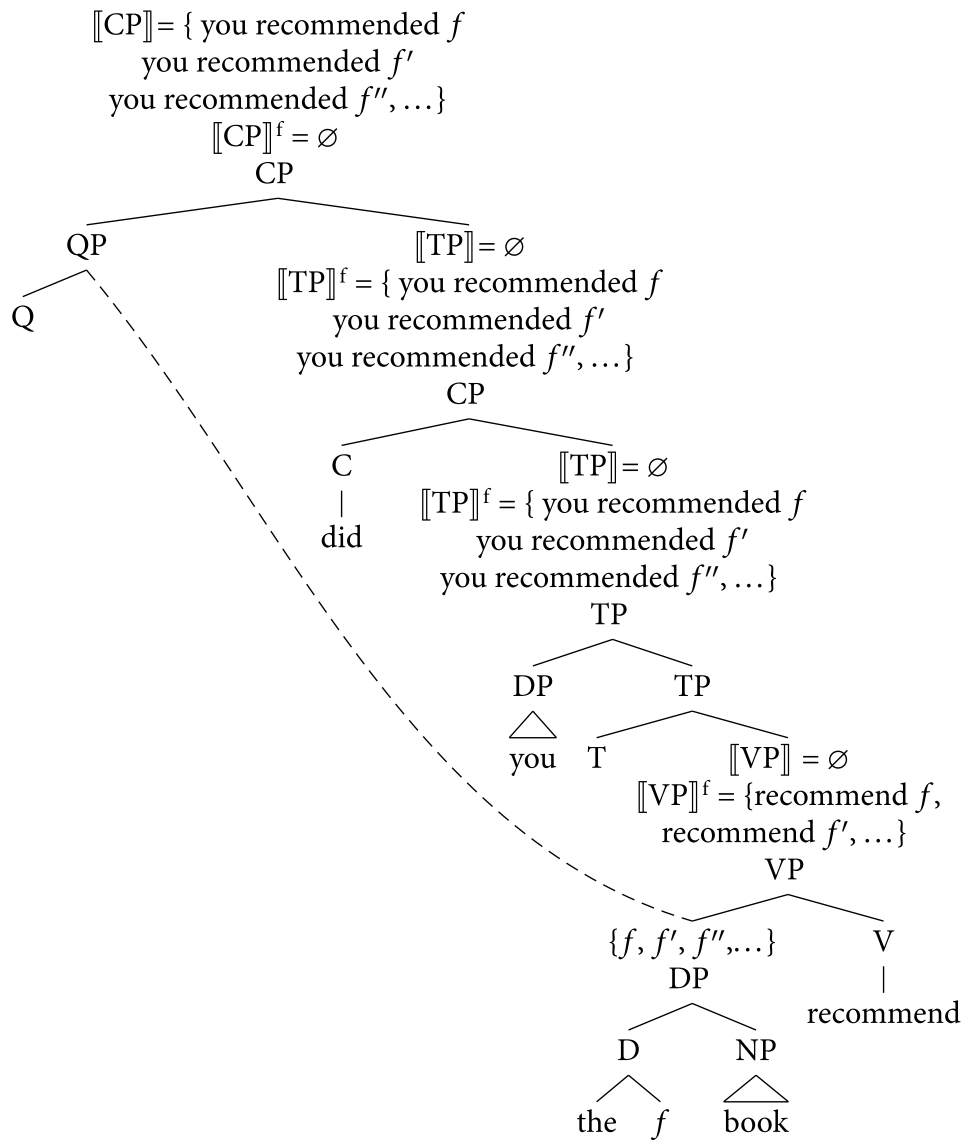
(28)



The denotation of *the* triggers the presupposition that the individual f picks out is a book.

Now, Beck's proposal is that the Q morpheme, among perhaps other things, converts the focus semantic value of its sister into a normal semantic value. So, we'd get something like (29).

(29)



Here, then, is the set of propositions component to the meaning of questions that we were search for earlier. Indeed, this is equivalent to the meaning for questions laid out at the beginning this talk.

What goes wrong in (25), then, is that *man* ('only') manipulates the focus semantic value of the clause its in before the question morpheme can get its hands on it. Interestingly, English doesn't have these kinds of intervention effects. Something parallel to (25) is perfectly grammatical.

(30) Who did only Minsu see?

We need to make it matter where the wh-phrase gets spelled out — that is what distinguishes the Korean example from the English one. I'll come back to this problem in the last lecture.

What we need now is an answer to the question why Q merges with the DP in English questions.

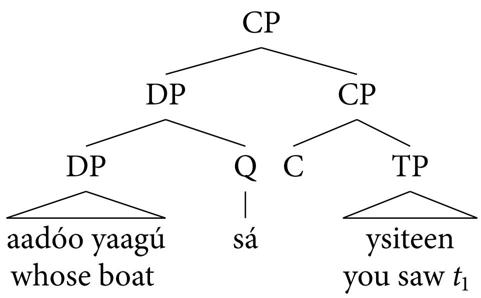
For this, I need to turn to work by Seth Cable. In Cable (2007), he studies questions of the sort that Korean and Japanese illustrate, but his object of study is Tlingit, a Na-Dené language spoken in Western Canada and Southeastern Alaska. Like Korean, Tlingit has a wh-determiner and another morpheme — I'll call it Q — in its questions. Like English, the wh-phrase moves overtly to the left edge of the question sentence. But, interestingly, unlike Korean or Japanese, the Q morpheme does not show up in what we might associate with the C position. Instead, it is merged with some phrase that contains the wh-phrase. (31) illustrates.

(31) Aadóo yaagú sá ysiteen?
who boat Q you.saw.it?
'Whose boat did you see'

(Cable 2007, (212)–(213): 155-6)

In (31), the Q particle, *sá*, has merged with a DP, inside of which lies the wh-word: *aadóo*. The whole thing has moved to the left edge of the sentence.

(32)



Moreover, Cable argues that the Q morpheme is in an Agreement relationship with the wh-word, and that there are locality conditions on that agreement relation that determine where the Q particle can be merged. Very roughly, that locality condition can be described with (33).

(33) Q can agree with D only if there is no lexical item that c-commands D but not Q.

The Q morpheme in our system is also in an agreement relationship with the wh-determiner, and so we should expect that, like Tlingit, it will have to merge in a position that does not take it too far from the wh-determiner it is agreeing with. This is what forces Q to merge to a phrase that it does not semantically combine with.

And, indeed, as Cable emphasizes, the range of phrases that the Tlingit Q morpheme can merge with are very close to the range of phrases that Pied-Pipe in wh-questions in English.

(34) a. Aadóo yaagú sá ysiteen?
who boat Q you.saw.it?
'Whose boat did you see' (=(31))

b. Aadóo teen sá yigoot?
who with Q you.went
'With whom did you go?'

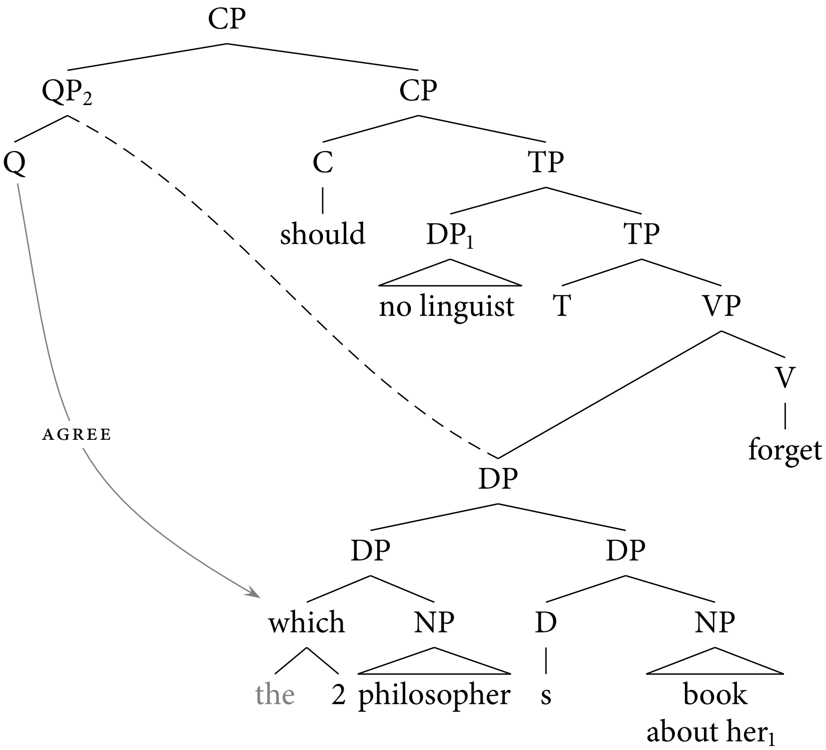
c. Daa sá ax éesh aawaxá
what Q my father ate
'What did my father eat?'

d. * Daa aawaxáa sá ax éesh
what ate Q my father
'what ate my father?'

He argues, therefore, that Pied-Piping in English arises because the phrase that moves in English has a Q morpheme merged with it in just the way that Tlingit *sá* does.[5] (35) illustrates.

[5] See also Cable (2008).

(35) Which philosopher's book about her should no linguist forget?



Now, Tlingit *sá* has to have a different semantics than we need for our English/Japanese/Korean questions, because it can show up in non-questions as well.

(36) Tléil aadóo yaagú sá xwsateen.
not whose boat Q I.saw.it
'I didn't see anyone's boat.'

(Cable 2007, (187): 141)

Cable gives it a semantics in Tlingit that, like what we need, operates on the focus semantic values of its complement and converts them into regular semantic values. In Tlingit, though, its existential force is derived from a higher, silent, operator. My suggestion, then, is that English Q has the semantics of the Q found in Korean, but the syntax of that found in Tlingit.

References

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.

Beck, Sigrid, and Shin-Sook Kim. 1997. On wh- and operator scope in Korean. *Journal of East Asian Linguistics* 6:339–384.

Cable, Seth. 2007. The grammar of Q: Q-particles and the nature of wh-fronting, as revealed by the wh-questions of Tlingit. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Cable, Seth. 2008. There is no such thing as pied-piping. Unpublished paper, University of Massachusetts at Amherst.

Chierchia, Gennaro. 1995. *Dynamics of meaning*. Chicago: University of Chicago Press.

Cooper, Robin. 1979. The interpretation of pronouns. In *Syntax and Semantics 10: Selections from the Third Groningen Round Table*, ed. Frank Heny and H. Schnelle, 61–92. New York: Academic Press.

Elbourne, Paul D. 2005. *Situations and individuals*. MIT Press.

Engdahl, Elisabet. 1980. The syntax and semantics of questions in Swedish. Doctoral Dissertation, University of Massachusetts, Amherst, Massachusetts.

Engdahl, Elisabet. 1986. *Constituent questions*. Dordrecht, The Netherlands: D. Reidel Publishing Company.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Fox, Danny. 2003. On logical form. In *Minimalist syntax*, ed. Randall Hendrick, 82–123. Oxford: Blackwell Publishers.

Hagstrom, Paul. 1998. Decomposing questions. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hagstrom, Paul. 2000. The movement of question particles. In *Proceedings of the North East Linguistic Society*, ed. Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-yung Kim, 275–286. Rutgers University: Graduate Linguistic Student Association.

Hamblin, Charles. 1973. Questions in Montague grammar. *Foundations of Language* 10:41–53.

Kishimoto, Hideki. 2005. Wh-in-situ and movement in Sinhala questions. *Natural Language and Linguistic Theory* 23:1–51.

Postal, Paul M. 1969. On so-called 'pronouns' in English. In *Modern studies in English*, ed. D. Reibel and Sandford Schane, 201–244. Englewood Cliffs, New Jersey: Prentice-Hall.

Romero, Maribel. 1998. Focus and reconstruction effects in wh-phrases. Doctoral Dissertation, University of Massachusetts, Amherst.

Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts, Amherst.

Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *which*-questions. In *Proceedings of SALT VIII*, ed. Devon Strolovitch and Aaron Lawson, 215–232. Ithaca, NY: CLC Publications.

Sauerland, Uli. 1998. The meaning of chains. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Sauerland, Uli. 2004. The interpretation of traces. *Natural Language Semantics* 12:63–127.

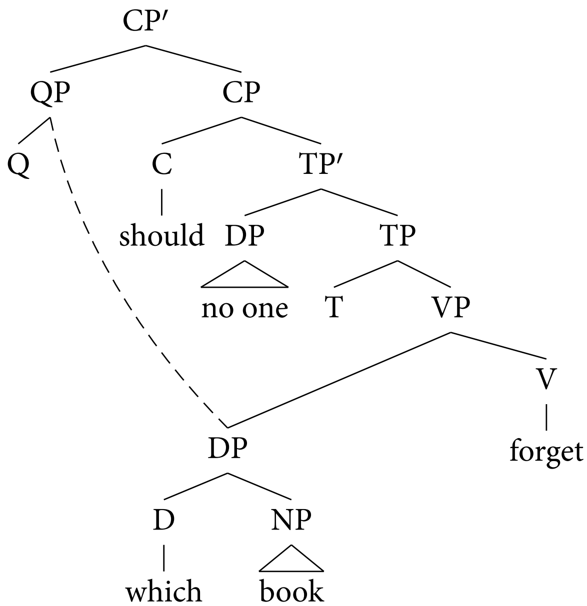
Quantifier Raising

Kyle Johnson
University of Massachusetts at Amherst
kbj@linguist.umass.edu

SICOGG 12
19 August

We left last time with a picture of wh-movement that involves representations like that in (1).

(1) Which book should no one forget?



There is an Agreement relationship between Q and D that is responsible for D getting spelled out as *which*. That Agreement relationship is responsible for ensuring that Q merges to a phrase containing D that is not too large. Moreover, there is a semantic relationship between Q and D. D introduces alternatives and makes the phrases that contain it have only a focus semantic value. Q converts those focus semantic values into regular values, and creates thereby, the question. This semantic relationship requires that Q's denotation be introduced where the question is, and this, in turn, requires that it not semantically combine with the phrase that it's merged to. Resolving these requirements triggers the multidominant structure that is seen in (1).

We get a picture, then, of the syntax-semantics mapping of questions that makes wh-movement look like a natural member of the class of question formation constructions that includes the in situ strategies of Korean and Japanese. It also allows us to dispense with ad hoc rules like Trace Conversion, while preserving the ability that Trace Conversion provided in capturing reconstruction effects.

At the same time, it gives us an explanation for why Wh Movement obeys Terseness. The syntax-phonology mapping involves, I suggested, a linearization scheme that includes an operation that generates a family of orderings (*lin*) and a set of constraints that choose from that family the best one.

(2) Let L(P) be a set of ordered pairs, $x < y$, where x and y are vocabulary items dominated by P, and "<" means "precedes."
lin(P) $=_{\text{def.}}$ the set consisting of every L(P).

(3) Linearization Constraints

a. *Totality*

All vocabulary items in the phrase marker p must be in the linearization of p.

b. *Antisymmetry*

For all vocabulary items, a and b in p, the linearization of p cannot include both $a < b$ and $b < a$.

c. *Transitivity*

For all vocabulary items, a, b, c in p, if the linearization of p includes $a < b$ and $b < c$ then it must include $a < c$.

d. *Contingiuty*

Let A be the set of vocabulary items dominated by A and b be a vocabulary item not in A. If b precedes something in A, then b precedes everything in A, and if b follows something in A, then b follows everything in A.

e. *Language Particular Component*

A set of constraints that correspond to the language particular word order choices.

(4) Constraint Evaluation

a. Only Contiguity is violable.

b. Every element of *lin*(P) which incurs n violations of Contiguity is ungrammatical if there is an element of *lin*(P) that incurs fewer than n violations Contiguity.

For (1), this system will deliver, among others, the two linearizations in (5).

(5) a. = should no one forget which book

$$\left\{ \begin{array}{llllll}
\text{Q} < \text{should} & \text{should} < \text{no one} & \text{no one} < \text{T} & \text{T} < \text{forget} & \text{forget} < \text{which} & \text{which} < \text{book} \\
\text{Q} < \text{no one} & \text{should} < \text{T} & \text{no one} < \text{forget} & \text{T} < \text{which} & \text{forget} < \text{book} & \\
\text{Q} < \text{T} & \text{should} < \text{forget} & \text{no one} < \text{which} & \text{T} < \text{book} & & \\
\text{Q} < \text{forget} & \text{should} < \text{which} & \text{no one} < \text{book} & & & \\
\text{Q} < \text{which} & \text{should} < \text{book} & & & & \\
\text{Q} < \text{book} & & & & &
\end{array} \right\}$$

b. = which book should no one forget

$$\left\{ \begin{array}{llllll}
\text{Q} < \text{which} & \text{which} < \text{book} & \text{book} < \text{should} & \text{should} < \text{no one} & \text{no one} < \text{T} & \text{T} < \text{forget} \\
\text{Q} < \text{book} & \text{which} < \text{should} & \text{book} < \text{no one} & \text{should} < \text{T} & \text{no one} < \text{forget} & \\
\text{Q} < \text{should} & \text{which} < \text{no one} & \text{book} < \text{T} & \text{should} < \text{forget} & & \\
\text{Q} < \text{no one} & \text{which} < \text{T} & \text{book} < \text{forget} & & & \\
\text{Q} < \text{T} & \text{which} < \text{forget} & & & & \\
\text{Q} < \text{forget} & & & & &
\end{array} \right\}$$

Both of these linearizations obey Totality, Antisymmetry and Transitivity. (5a) incurs no violations of Contiguity, and (5b) violates Contiguity four times: once each for VP, TP, TP′ and CP. The language particular component of English, however, is not obeyed in (5a); English requires that a wh-phrase merged to a constituent question CP be spoken in the merged-to-CP position. I called this the Wh Criterion. Because the Wh Criterion is inviolable, and Contiguity is violable, of these two linearizations, (5b) is better. Indeed, when all other linearizations produced by *lin* are considered, there are no other that have fewer violations of Contiguity and still manage to obey the language particular component, Antisymmetry, Totality and Transitivity. (5b), therefore, is how (1) gets linearized.

A problem I posed in the first lecture is how to give an account of movement that both captured the family resemblance that different forms of movement have, and yet allowed for various forms of movement to differ slightly in how they express that resemblance. The two properties that all movement operations share, I declared, are Terseness and Semantic Displacement. We've now seen how those two properties are captured for Wh Movement. Today I want to look at how the rule that moves quantificational DPs, QR, fits into this picture.

(6) every hole a marble rolled into every hole.

I choose QR because it is differs in, perhaps, the most ways from Wh movement and so considering these two movement operations gives us a kind of view of the limits of variation we want movement to be allowed.

Like Wh movement, QR produces a representation that is interpreted as a variable-binder pair. It's semantic outcome is similar, then, to Wh movement, though, as we shall see tomorrow, there is a subtle difference in the variables involved. One difference in their semantics concerns where the NP part of the moved expression is interpreted. As we've seen, in questions we want to allow the NP to be semantically interpreted in the position it moves from but not in the position it is spoken. That is, for instance, necessary to get the right interpretation out of (7).

(7) Which book about her_1 should no linguist_1 forget?

The system I've presented so far not only allows the NP part of a moved wh-phrase to be interpreted only in its lower position, it forces that. QR is different. The NP part of QR must be interpreted in the position it is spoken. One way that can be seen is by considering the disjoint reference effect in (8).

(8) every book about Julie she likes every book about Julie.

If *book about Julie* could be interpreted in just its higher position, then we should expect *Julie* and *she* to be able to corefer. But they cannot. So here is one difference in how semantic displacement works that we should explain.

(9) QR'd material must be semantically interpreted where it is spoken, but Wh moved material is able to be semantically interpreted in only its unspoken position.

This is the syntax-semantics difference that I will focus on deriving today.

A more obvious difference in Wh Movement and QR, though, is how they get spelled out. The syntax of wh movement allows the moved expression to be spelled out in its higher position, and the language particular constraints of English pick that possibility for "single" constituent questions. For multiple questions, of course, English allows, in fact forces, the wh phrase to be spelled out in its lower position, as in (10).

(10) Which story should you tell which child?

In the case of QR, however, the moved phrase is always spelled out in its lower position. So, that's one difference we would like to capture.

(11) A wh-moved phrase can be spelled out in either of its two positions, but a QR'd expression can only be spelled out in its lower position.

These are the tasks for today: explain (9) and (11).

Let me start by sketching out a theory of QR that captures some of the standard effects it has been designed for. I'll do this by employing the copy theory of movement, since that is how the present literature on QR is written.

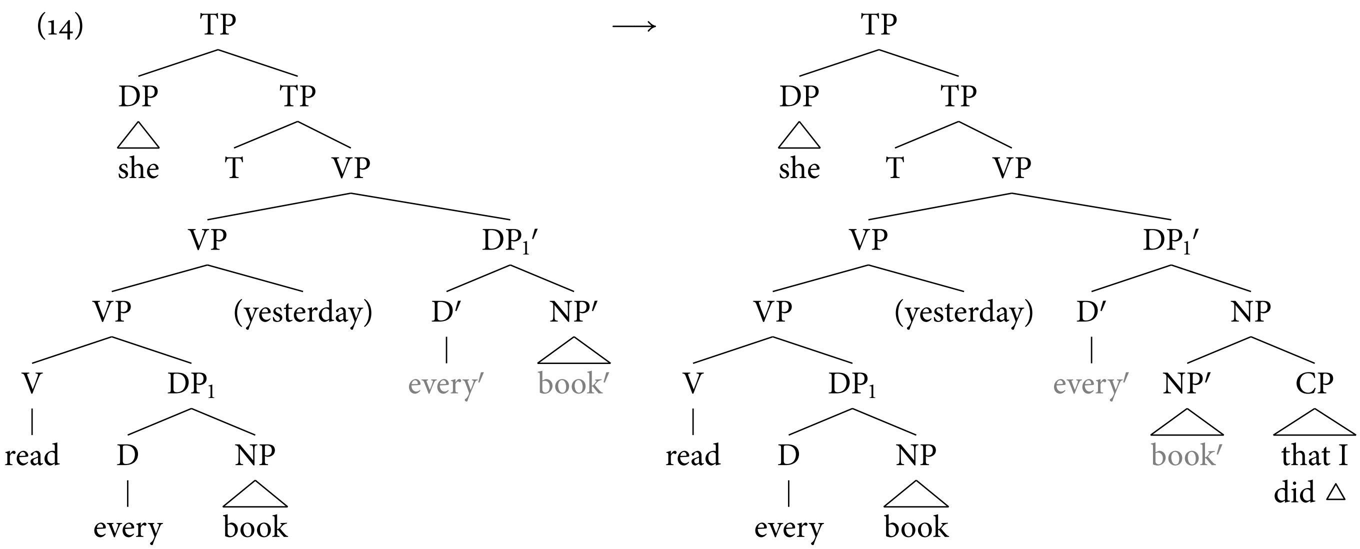
One of the situations in which QR can be seen arises when a quantificational object contains an elided VP whose antecedent is the VP the object sits in. This can happen when a quantificational object comes with a relative clause, as in (12).

(12) She read every book that I did △.

I'll the account of these so-called Antecedent Contained Deletions (ACD) in Fox (2002). On Fox's analysis, ACD is licensed when the relative clause containing the ellipsis has extraposed from the antecedent for that ellipsis.[1]

(13) She [$_{\text{VP}}$ read [$_{\text{DP}}$ every book]] (yesterday) [$_{\text{CP}}$ that I did △]

Fox (and Fox and Nissenbaum 1999) argue that these sorts of extraposition operations are the result of "late merging" a clause into a QR'd DP.

(14)


Movement produces a copy of the object DP and then merges that copy into a position outside the VP which serves as antecedent to the ellipsis. Unlike Wh Movement, this higher copy goes unpronounced and, instead, the lower copy is spoken. However, into the higher copy is merged the relative clause containing the elided VP, and this relative clause is pronounced in the position occupied by the higher copy of QR. A reason for using movement, and the copy theory of movement in particular, to model QR is that it provides a simple account of extraposition from NP and, with it, a good account of ACD. It also captures a fact about ACD that Sag (1976) established: the position where an ACD in a DP is resolved is the same position that the quantifier heading that DP is interpreted.

We can now see another difference between QR and Wh Movement: when material in the higher copy in Wh Movement is pronounced, that material gets linearized so that it precedes everything else in the clause it is dominated by. But when material in the higher copy of QR is pronounced, it follows everything else in the phrase it is dominated by. Wh Movement goes to the left, but QR goes to the right. This difference too should be derived.

(15) When Wh-moved material is spelled out in its higher position, it shows up to the left of the phrase it is merged to. When QR'd is spelled out in its higher position, it shows up to the right of the phrase it is merged to.

[1] The original argument for using extraposition structures as the source for ACD is Baltin (1987).

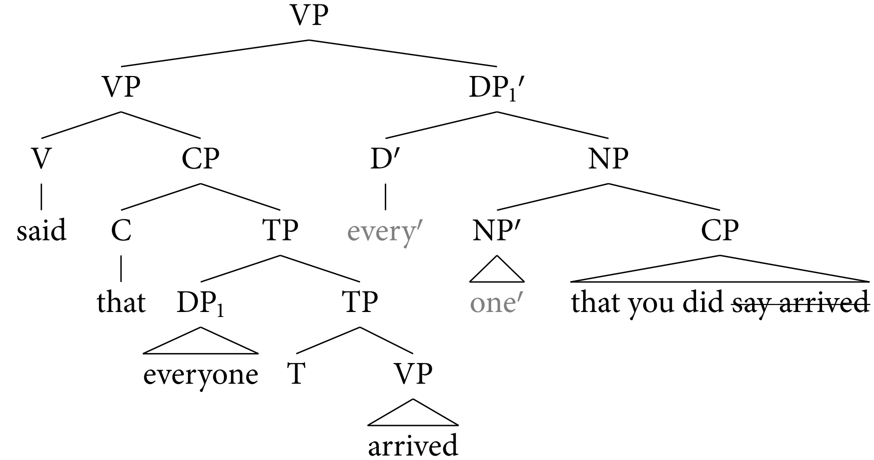
One of the most interesting pieces of support for Fox's analysis of ACD is the contrast in (16), from Tiedeman (1995).

(16) a. * I said that everyone you did △ arrived.
b. I said that everyone arrived that you did △.
△ = said that *x* arrived

(Fox 2002, (35b), (36b): 77)

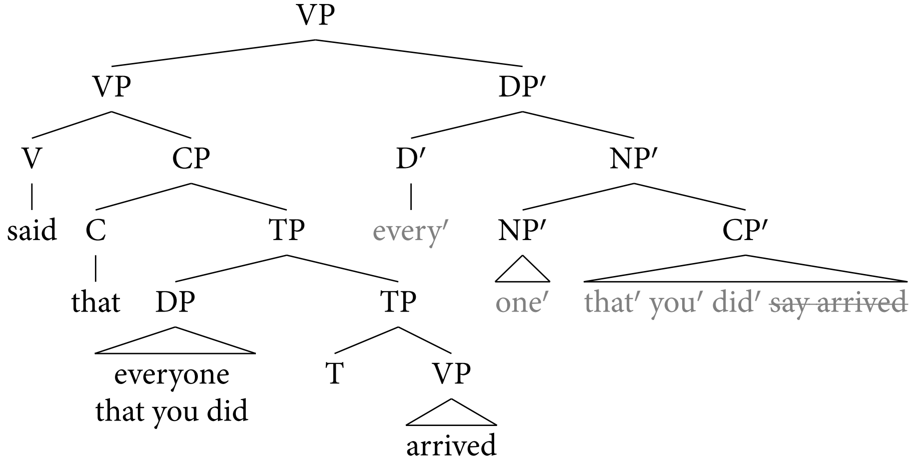
The difference is credited to extraposition being able to generate the string in (16b) but not (16a). The representation in (17) is only available for (16b). *lin* will put the material in the embedded subject between the complementizer and the embedded VP, and linearize the extraposed relative clause so that it follows everything the higher copy has merged with.

(17)



Hidden in this example, however, is yet another illustration of the difference between QR and Wh Movement semantics. To see this difference, consider a derivation in which the relative clause is part of the lower copy, and therefore maps onto the string in (16a). This derivation will QR the subject and its relative clause together into a higher position, one that puts the elided VP outside of its antecedent, as in (18).

(18)



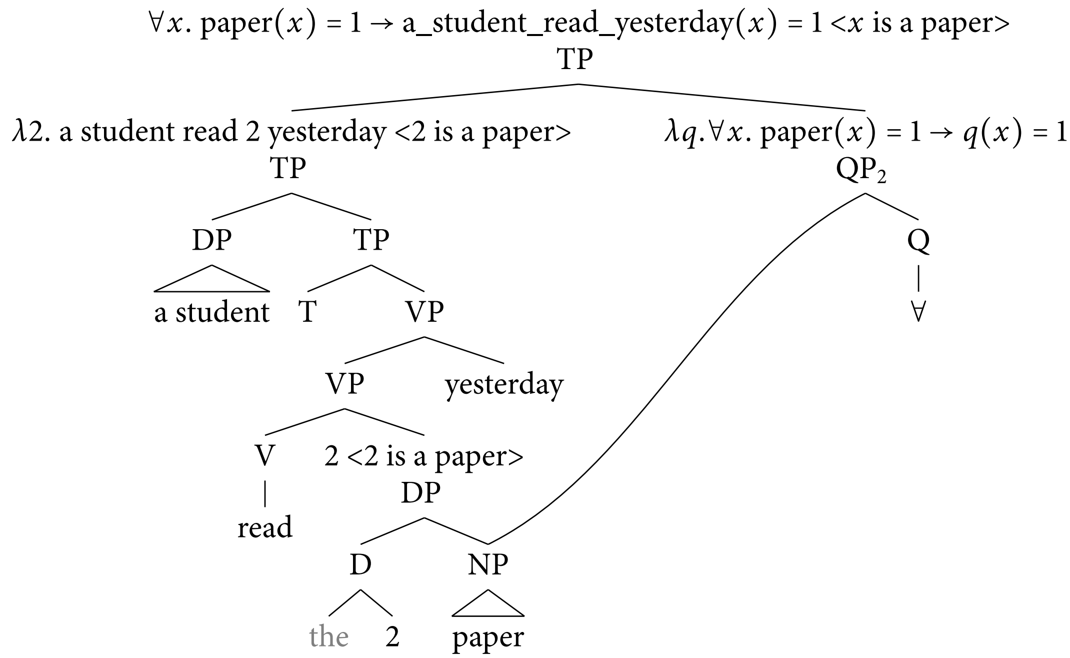
This representation is ill-formed only if the relative clause in the lower copy must be semantically interpreted. Here, then, is another illustration of (9).

Let's now consider how QR can be modeled in our remerge theory of movement. As with Wh Movement, let's let there be a definite description in the lower position that is bound by an operator in the higher position. Unlike the case with questions, however, the determiner in the lower position and the quantifier in the higher position will both combine semantically with the NP. That is because the denotation of quantifiers requires that they relate, semantically, the meaning of the NP they combine with with the meaning of the clause they are in. Quantifiers like *every*, for instance, have a meaning something like that indicated for ∀ in (19).

(19) $[\![\forall]\!] = \lambda p.\lambda q.\forall x.p(x) = 1 \rightarrow q(x) = 1$

QR will therefore produce a representation like that in (20).

(20) A student read every paper yesterday.



As with the case of questions, we want the form of D to be determined by the quantifier that is in the higher position. When the quantifier is ∀ we'll want D to be spelled-out as *every*, when it's "¬ any" we'll want D to be spelled-out as *no*, and so on. We can't do this with Agree, as Q does not c-command D and Agree only holds between things that are in a c-command relation. I suggest instead that D+Q are brought together by the morphology. Morphology is equipped with processes that allow two terminals to be mapped onto one vocabulary item. These processes show up in a variety of ways. They are responsible for mapping a preposition and determiner onto one lexical item in certain contexts in German (and other IndoEuropean languages), for instance.

(21) an dem Tisch → am Tisch
on the table → on.the table

These processes go by various names, and there seem to be slight differences in the conditions under which they may operate that depends on the case being modeled.[2] But across all these cases, there is a similar locality condition on the two terminals that combine: they must be so close together that, under normal circumstances, they would show up adjacent in the string. Let us simply adopt this descriptive consequence as a well-formedness condition on "fusion," as I will call the process that bring D and Q together into one word.

[2] See Pranka (1983), Marantz (1988, 1984), Halle and Marantz (1993), Bobaljik (1995), Embick and Noyer (2001), Matushansky (2006) and references cited therein.

(22) X and Y can fuse only if *lin* assigns them adjacent positions.

The language particular component of English causes Q and D both to come before everything else in the phrases they head.

(23) *lin*(QP) puts Q before everything else in QP
lin(DP) puts D before everything else in DP

If D and Q were to be linearized in non-QR contexts, then, they would show up in adjacent positions. They are, therefore, possible fusers.

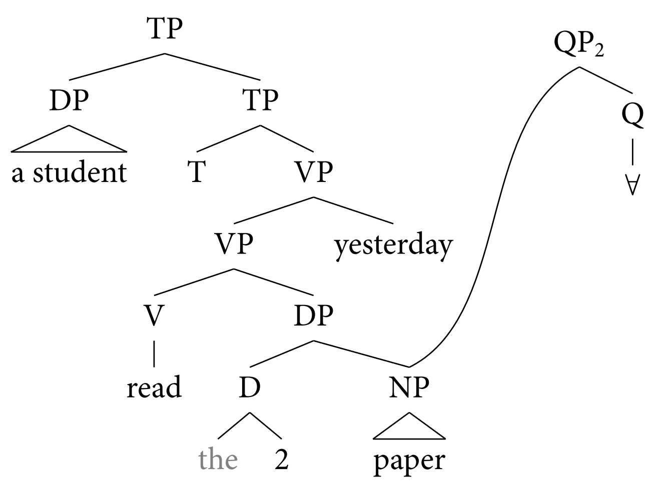
However, if D and Q are to try to fuse in structures of QR, then the condition in (22) will stand in the way. In (20), for example, *lin* will put things between D and Q and prevent adjacency. There are two strings that *lin* could produce from (20). If the NP containing *paper* is linearized in the spot assigned to QP, *lin* will deliver (24a), and if this NP is linearized in the spot assigned to DP, *lin* will deliver (24b).

(24) a. a student read D yesterday Q paper
b. a student read D paper yesterday Q

Both of these linearizations violate Contiguity, but they only violate Contiguity to the extent required by the multidominant representation. These, then, are the minimal violators of Contiguity and therefore the candidate winners. In both of them, D and Q are separated by *yesterday* and so fusion is blocked. If we assume that the English lexicon does not provide vocabulary items for the D or Q in these structures, we will have a sentence that cannot be spelled out.

I suggest, then, that fusion is achieved before the entire QR structure is built. Let's imagine the stage in the derivation that leads to (20) immediately before the QP has merged with TP. This will look like (25).

(25)



This representation has two root nodes: TP and QP. If we define *lin* so that it runs on root nodes, then in this case it will apply to TP and QP independently, and produce the ordered pairs in (26).

(26)

$$lin(\mathrm{TP}) = \left\{ \begin{array}{lllll} \text{a} < \text{student} & \text{student} < \text{T} & \text{read} < \text{D} & \text{D} < \text{paper} & \text{paper} < \text{yesterday} \\ \text{a} < \text{T} & \text{student} < \text{read} & \text{read} < \text{paper} & \text{D} < \text{yesterday} & \\ \text{a} < \text{read} & \text{student} < \text{D} & \text{read} < \text{yesterday} & & \\ \text{a} < \text{D} & \text{student} < \text{paper} & & & \\ \text{a} < \text{paper} & \text{student} < \text{yesterday} & & & \\ \text{a} < \text{yesterday} & & & & \end{array} \right\}$$

$$lin(\mathrm{QP}) = \forall < \text{paper}$$

This linearization puts nothing between D and ∀, and so they can fuse. Once they've fused and been mapped onto *every*, it is *every* that will occupy the positions assigned to D and ∀ in (26). After QP and TP have merged, no new ordering statements need to be added to meet the requirement of Totality. That's because Totality only requires that every vocabulary item in a phrase marker be assigned a position relative to every other vocabulary item, and that will be achieved by (26) (after D and Q fuse) for all the vocabulary items that will be matched to the terminals in (25). So (26) will be the linearization for the final representation (=(20)); (26) corresponds to the string in (27), which is precisely the right outcome.

(27) a student read every paper yesterday

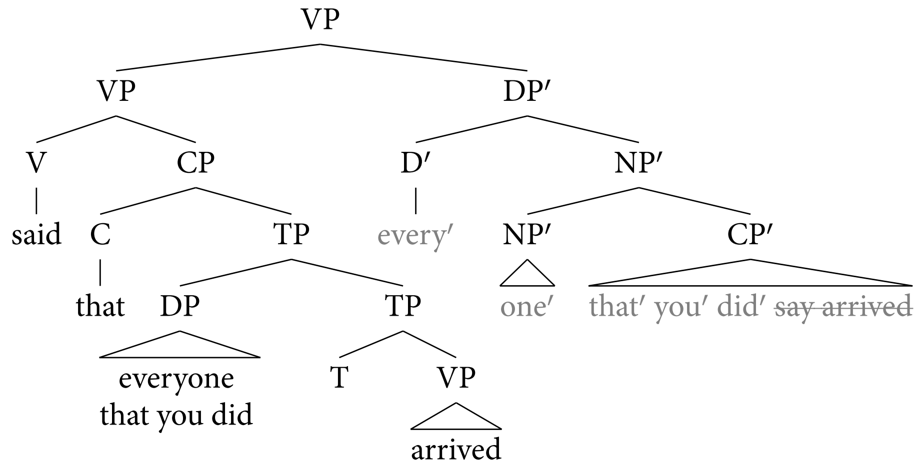
Once QP has merged with TP to form the representation that is semantically interpreted, one of two options are possible: *lin* can be run again, or the ordered pairs in (26) can simply be combined. In either case, the resulting set of ordered pairs will meet Totality, Transitivity and Antisymmetry, and it will violate Contiguity no more than is made necessary by QR. We have just derived the fact that QR'd phrases get spelled out in their lower position.

Let's consider next how this system derives (9).

(9) QR'd material must be semantically interpreted where it is spoken, but Wh moved material is able to be semantically interpreted in only its unspoken position.

We've already seen how the case of Wh Movement works. An interrogative DP can be semantically interpreted entirely in its lower position, and yet be part of a phrase that *lin* puts in a different position. Let's now consider why something parallel is not possible for QR. One of the cases I used to demonstrate (9) is (18). Under a copy theory account, this gets the representation indicated.

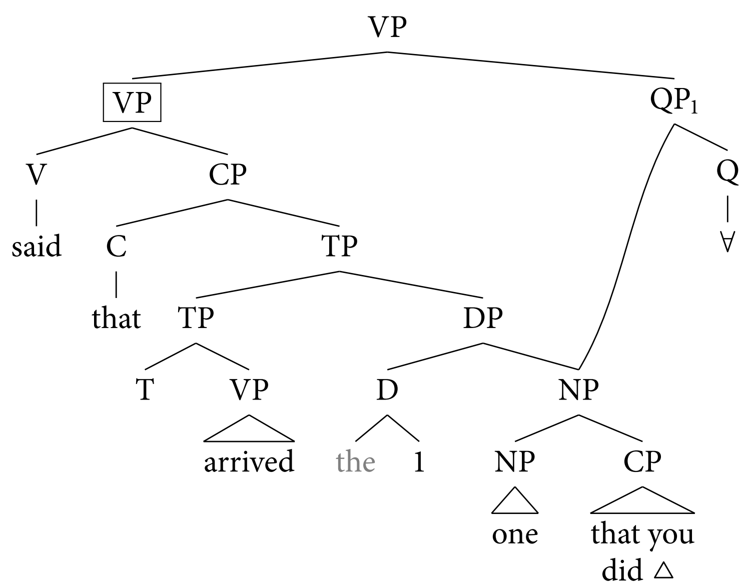
(18) * I said that everyone you did △ arrived.



This representation cannot resolve the ellipsis it contains and this indicates that the relative clause in the lower copy must be semantically interpreted.

If QR is modeled with the multidominant representations proposed here, however, the string in (18) can only get the representation in (28). Like the representation in (18), The structural requirements for resolving the ellipsis are not met in (28), either, since the ellipsis is inside its antecedent: $\boxed{\text{VP}}$.

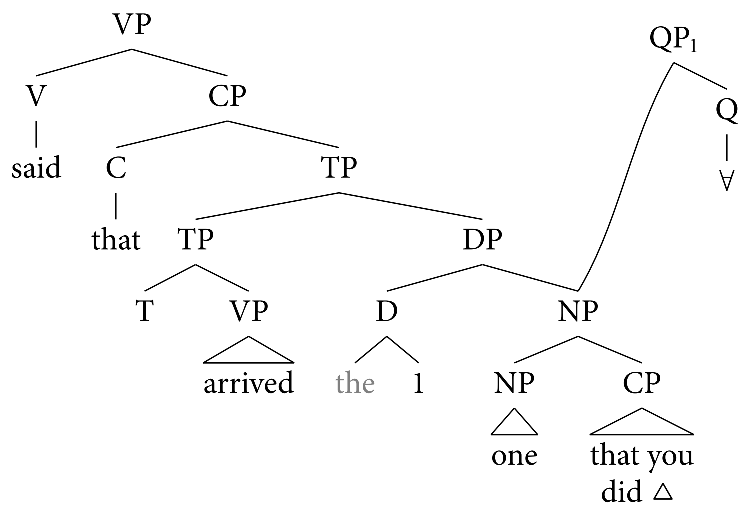
(28)



To see why it's this structure that corresponds to the string in (18), consider how *lin* will deliver the linearization of (28).

As we've seen, *lin* must run before the QP is merged to $\boxed{\text{VP}}$ in order to fuse D and Q into *every*. So, the string associated with (28) will be built upon the output *lin* produces from applying to (29).

(29)



That linearization is (30).

(30)

$$
lin(\text{VP}) = \left\{ \begin{array}{llllllll}
\text{said} < \text{that} & \text{that} < \text{T} & \text{T} < \text{arrived} & \text{arrived} < \text{D} & \text{D} < \text{one} & \text{one} < \text{that} & \text{that} < \text{you} & \text{you} < \text{did} \\
\text{said} < \text{T} & \text{that} < \text{arrived} & \text{T} < \text{D} & \text{arrived} < \text{one} & \text{D} < \text{that} & \text{one} < \text{you} & \text{that} < \text{did} & \\
\text{said} < \text{arrived} & \text{that} < \text{D} & \text{T} < \text{one} & \text{arrived} < \text{that} & \text{D} < \text{you} & \text{one} < \text{did} & & \\
\text{said} < \text{D} & \text{that} < \text{one} & \text{T} < \text{that} & \text{arrived} < \text{you} & \text{D} < \text{did} & & & \\
\text{said} < \text{one} & \text{that} < \text{that} & \text{T} < \text{you} & \text{arrived} < \text{did} & & & & \\
\text{said} < \text{that} & \text{that} < \text{you} & \text{T} < \text{did} & & & & & \\
\text{said} < \text{you} & \text{that} < \text{did} & & & & & & \\
\text{said} < \text{did} & & & & & & &
\end{array} \right\}
$$

$$
lin(\text{QP}) = \left\{ \begin{array}{llll}
\forall < \text{one} & \text{one} < \text{that} & \text{that} < \text{you} & \text{you} < \text{did} \\
\forall < \text{that} & \text{one} < \text{you} & \text{that} < \text{did} & \\
\forall < \text{you} & \text{one} < \text{did} & & \\
\forall < \text{did} & & &
\end{array} \right\}
$$

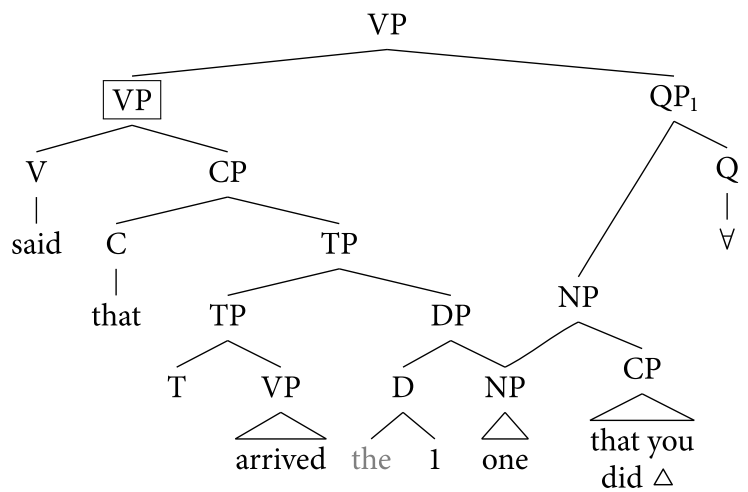
When *lin* runs again, subsequent to merging QP and VP, no new ordered pairs will be introduced, and so we will get the simple union of *lin*(VP) and *lin*(QP). That corresponds to the string in (31).

(31) said that everyone that you did arrived

This string therefore corresponds to a structure that does not resolve the ellipsis: just what we want.

To resolve the ACD, we must have a representation that involves "late merge" of the relative clause. Under the present proposal, this will look like (32).

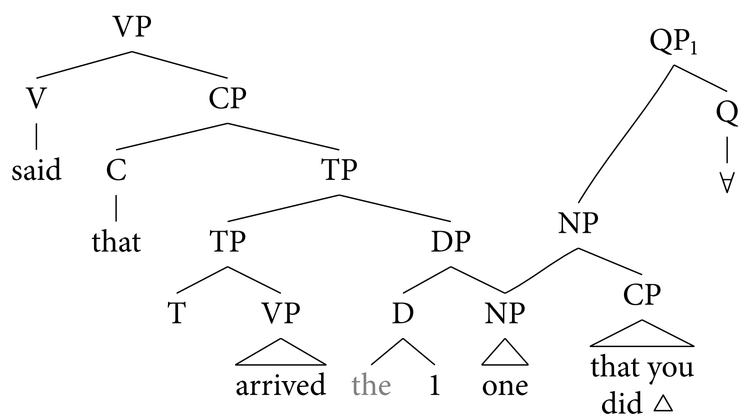
(32)



In this structure, the relative clause is not within the VP that serves as antecedent, and so the ellipsis can be resolved. To see that this structure does not correspond to the string in (18), consider how *lin* will manufacture a linearization for it.

As always, *lin* will be forced to apply before the QP has merged into the larger structure. In this case, that will look like (33).

(33)



From (33), *lin* will produce (34).

(34) = said that everyone arrived & everyone that you did

$$
lin(\text{VP}) = \left\{ \begin{array}{lllll}
\text{said} < \text{that} & \text{that} < \text{T} & \text{T} < \text{arrived} & \text{arrived} < \text{D} & \text{D} < \text{one} \\
\text{said} < \text{T} & \text{that} < \text{arrived} & \text{T} < \text{D} & \text{arrived} < \text{one} & \\
\text{said} < \text{arrived} & \text{that} < \text{D} & \text{T} < \text{one} & & \\
\text{said} < \text{D} & \text{that} < \text{one} & & & \\
\text{said} < \text{one} & & & &
\end{array} \right\}
$$

$$
lin(\text{QP}) = \left\{ \begin{array}{llll}
\forall < \text{one} & \text{one} < \text{that} & \text{that} < \text{you} & \text{you} < \text{did} \\
\forall < \text{that} & \text{one} < \text{you} & \text{that} < \text{did} & \\
\forall < \text{you} & \text{one} < \text{did} & & \\
\forall < \text{did} & & &
\end{array} \right\}
$$

Notice that because the relative clause is not yet inside VP, it is not included in the string associated with VP. As a consequence, only *lin*(QP) has information about where the relative clause will be positioned: it will follow everything else in QP. It is only after QP has merged with VP — to form (32) — that *lin* can order the material in the relative clause with the material in the VP. In order to satisfy Totality, *lin* will therefore have to apply again after (32) is formed.

The ordered pairs that this second run of *lin* will add to (34) must not only satisfy Totality, but they must also satisfy Antisymmetry and Contiguity. The best satisfaction of Contiguity will be linearizations that keep the material in the relative clause together and put it either all before the VP, or all after the VP. If *lin* puts the relative clause before the VP, however, the ordered pairs it will generate will include (35).

(35) that < every, that < one, you < every, you < one, …

If these are added to the ordered pairs in (34), however, a violation of Antisymmetry will ensue. For these reasons, the ordered pairs *lin* will generate when applied to (32) will add to (34) those in (36): ones in which the relative clause follows the VP.

(36)

$$\text{The new outputs from } \mathit{lin}(\,(32)\,) = \left\{ \begin{array}{llll} \text{said} < \text{that} & \text{that} < \text{that} & \text{T} < \text{that} & \text{arrived} < \text{that} \\ \text{said} < \text{you} & \text{that} < \text{you} & \text{T} < \text{you} & \text{arrived} < \text{you} \\ \text{said} < \text{did} & \text{that} < \text{did} & \text{T} < \text{did} & \text{arrived} < \text{did} \\ \text{said} < \forall & \text{that} < \forall & \text{T} < \forall & \text{arrived} < \forall \end{array} \right\}$$

This corresponds to the string in (37).

(37) …said that everyone arrived that you did $\triangle$

For the relative clause to be positioned outside the VP that is serving as the antecedent for the ellipsis it contains, it will necessarily be positioned linearly outside the string that corresponds to that VP. This result is perfectly general. We derive (9): QR cannot put spoken material in a position where it is not semantically interpreted.

In fact, this also sketches how (15) is manufactured.

(15) When Wh-moved material is spelled out in its higher position, it shows up to the left of the phrase it is merged to. When QR'd is spelled out in its higher position, it shows up to the right of the phrase it is merged to.

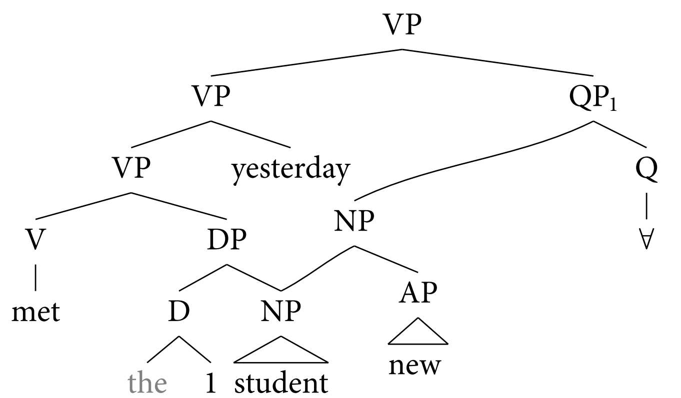
In stepping through how *lin* positions the late merged relative clause in (34), we saw not only that it must put that relative clause outside the phrase to which the QP merges, but also that it must put that relative clause to the right of the phrase to which the QP merges. This arises because the relative clause is forced to follow everything else within the QP that dominates it, and, at the same time, all that other material in the QP gets linearized within the phrase to which the QP later merges. Thus, when the QP merges with some phrase, the relative clause it contains will have to linearize itself with respect to that phrase in the same way that it is linearized with respect to the rest of the QP's material inside that phrase.

This way of deriving the rightwards direction of QR and, consequently, extraposition has another useful consequence. It correctly captures the fact that in English, the only material that can extrapose out of nominals is material that can be linearized at the right edge of those nominals. Extraposition cannot, for instance, form (38a) from (38b).

(38) a. * I met every student yesterday new.
     b. I met every new student yesterday.

(38a) would get the representation in (39).

(39)



As in every example of QR, *lin* will run before the QP has merged into the larger structure; in this case, then, *lin* will apply before QP and VP have merged. After fusion occurs, *lin* will produce from QP the ordered pairs that correspond to the string in (40), and from VP, it will produce ordered pairs that correspond to the string in (41).

(40) every new student

(41) met every student yesterday

After QP has merged with the VP to form (39), *lin* will have to run again in order to ensure that Totality is obeyed. As in the case of an extraposed relative clause, this latter run of *lin* has two choices: it may put *new* after the VP or before the VP. Unlike in the case of a relative clause, however, both of these choices violate Antisymmetry. If *new* is positioned after the VP, then it will be positioned after *student*, and that is at odds with the previous run of *lin*, which positions *new* before *student*. If, instead, the choice is to put *new* before the VP, this will cause it to precede *every*, and again this will conflict with what the previous run of *lin* did. This effect arises for any material that gets linearized between the determiner and whatever shows up at the end of the DP. Only material at the end of a DP can be extraposed.

References

Baltin, Mark. 1987. Do antecedent-contained deletions exist? *Linguistic Inquiry* 18:579–596.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: a case for overt QR. In *Proceedings of the West Coast Conference on Formal Linguistics 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 132–144. Somerville, Massachusetts: Cascadilla Press.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, Massachusetts: MIT Press.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Massachusetts: MIT Press.

Marantz, Alec. 1988. Clitics, morphological merger, and the mapping to phonological structure. In *Theoretical morphology: Approaches in modern linguistics*, ed. Michael Hammond and Michael Noonan, 253–270. San Diego, California: Academic Press.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37:69–109.

Pranka, Paula. 1983. Syntax and word formation. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Sag, Ivan. 1976. Deletion and logical form. Doctoral Dissertation, Massachusetts Institute of Technology.

Tiedeman, Robyne. 1995. Some remarks on antecedent contained deletion. In *Minimalism and Linguistic Theory*, ed. Shosuke Haraguchi and Michio Funaki, 67–103. Tokyo: Hituzi Syobo.

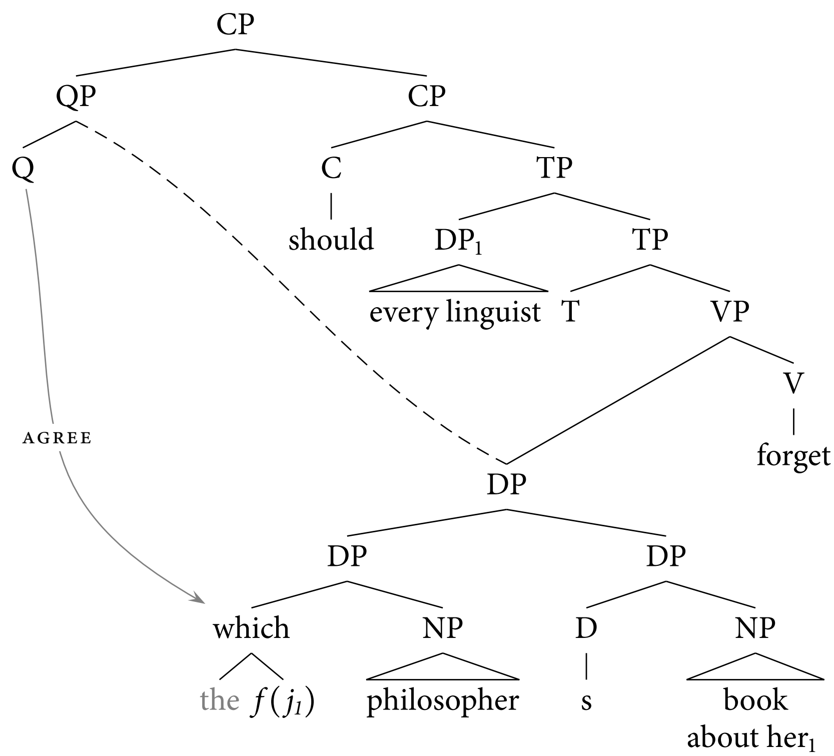
A General Theory of Movement

Kyle Johnson
University of Massachusetts at Amherst
kbj@linguist.umass.edu

SICOGG 12
20 August

In the last two lectures, I've taken a close look at two kinds of movement. One is Wh Movement, which I've argued involves putting a DP in two positions, as (1) reminds.

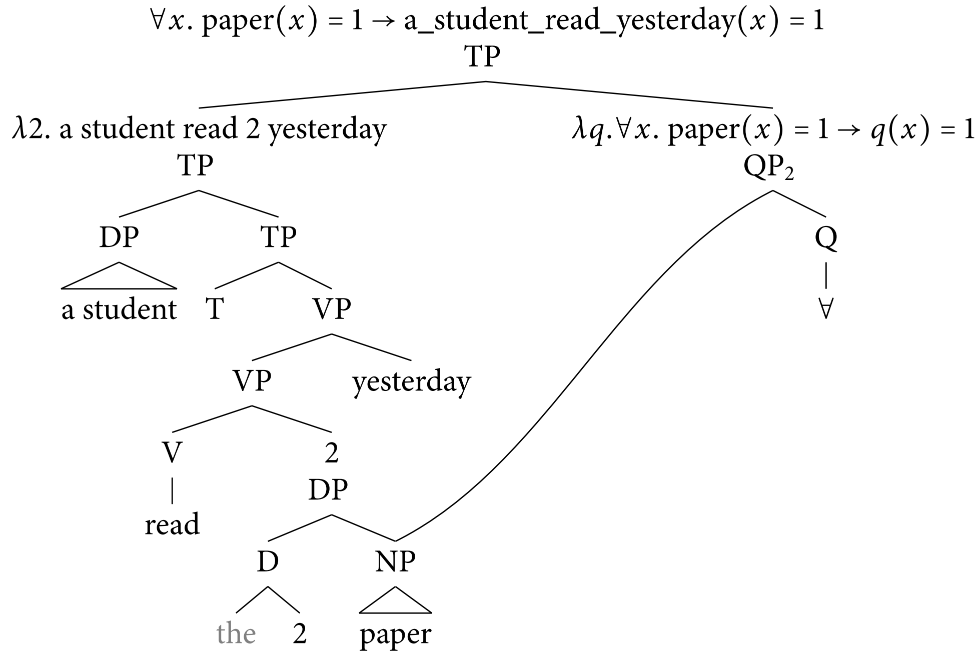
(1) Which philosopher's book about her should every linguist forget?



Presupposition: The output of $f(j_1)$ is a philosopher

And the second is QR, which causes an NP to be shared between a Determiner and a Quantifier, as (2) reminds.

(2) A student read every paper yesterday.



Presupposition: 2 is a paper

In the Wh Movement example, there is a determiner that contains a function and causes the phrases containing it to have a meaning that ranges over different values for that function. The Q morpheme that marks the scope of the question converts those alternatives into the semantic value for the question. Because the Q is in an Agreement relationship with the D, it must merge to a reasonably small phrase containing the Determiner, and it's this phrase that "moves."

In the QR example, the Q and D must both semantically combine with the same NP. Their selectional requirements force them to combine with an NP, and because they must do so in such a way that allows them to fuse, they will merge to the same NP. They must fuse because English only has vocabulary items that correspond to the D+Q combination and not to Q alone. Because of the constraints on the fusion operation, the method by which these structures are linearized will force QR'd phrases to be spoken in their lowest position.

What I've tried to illustrate with these two cases is how "movement" can be seen as the result of demands made by the morphology (Agreement and fusion) in concert with the semantics and the linearization algorithm. These demands can only be met — I claim — by allowing one term to remerge, and that gives us movement.

What I want to do in this final talk is look at how this template for a theory of movement spreads to other, sometimes more complex, cases.

I'll start by making an observation about a difference between QR and Wh Movement that I've not yet discussed. Wh Movement, recall, involves a variable over functions. That is why, for instance, a question like (3) can have the answers indicated in (4).

(3) Which picture did you say you'd show every girl?

= { you said you'd show every girl $f(j_1)$ picture, you said you'd show every girl $f'(j_1)$ picture,...}

(4) a. f might be:
Sally → the picture of the Eiffel tower
Mary → the picture of the Milkmaid
Myrtle → the picture of hot-rods
b. f might be: "her favorite"

The variable within this function — j_1 — is bound by *every girl* and that allows the pictures that the function picks out to vary in a way that depends on the values given to *every girl*. We can make overt this dependence of the function on the values given to some c-commanding quantifier by putting within the moved phrase an overt bound variable, as in (5).

(5) Which picture of herself did you say you'd show every girl?

The DP that is part of the moved phrase — the DP that is interpreted in the lower position — is responsible for introducing this function.

But the DP that is part of the quantificational expressions that QR moves are different. Those DPs introduce a variable over individuals — things — and not functions. We can see that by considering an example like (6).

(6) A different student read every book of hers.

In this example, *a different student* can fall within the scope of *every*. The values given to *a different student* can vary with the values given to *each book of hers*. And also in this example, *hers* can be bound to *a different student*. But these two things cannot happen at the same time. Suppose, for instance, that the context of utterance for (6) includes just the students Mary and Sally, and that, moreover, each of these students has two books: Sally has *Independent People* and *The Wild Sheep Chase* and Mary has *Infinite Jest* and *Between Silk and Cyanide*. If *her* refers to neither Sally nor Mary, then (6) can describe a situation with the relations indicated in (7).

(7) With *her* = someone other than Mary or Sally (let's say, Cherlon)
a. Mary read every book of Cherlon's, and
b. Sally read every book of Cherlon's.

But if *her* is bound to *a different student*, then *different* does not fall within the scope of *every* and we get a reading in which *different* indicates that the student we are describing is different from the one we have been talking about. With *her* bound to *a different student*, (6) describes a situation with the relations indicated in (8).

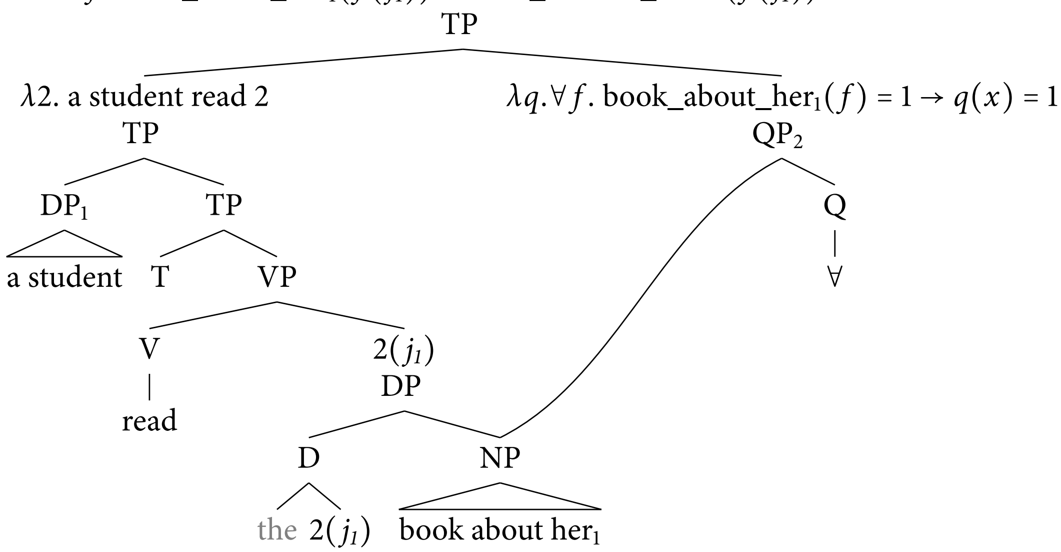
(8) With *her* bound to *a different student* = someone other than Mary or Sally (let's say: Cherlon).
a. Cherlon read every book of Cherlon's

What (6) can't describe is something with the relations indicated in (9).

(9) a. Mary read *Infinite Jest* and *Between Silk and Cyanide*, and
b. Sally read *Independent People* and *The Wild Sheep Chase*.

(9) is an interpretation we should expect, if the "trace" left by QR could introduce a variable over functions. (6) could, on that view, get the representation in (10).

(10) $\forall f.\ \text{book\_about\_her}_1(f(j_1)) = 1 \rightarrow \text{a\_student\_reads}(f(j_1)) = 1$



Presupposition: The output of $f(j_1)$ is a book about her_1.

Our conclusion is that quantifiers quantify over individuals, not functions. This, I will assume, is something lexical about quantifiers. They specify that the things they quantify over must be the semantic type of individuals, and not functions.

There is another difference in the lower DPs of questions and the lower DPs of QR. In the case of questions, the determiner of the lower DP introduces alternatives. These alternatives propagate up through the phrase marker and are used, finally, by the Q operator to form a question. There is no binding, then, between the D and Q in questions. But there is in the DPs that QR. The quantifier in the higher position binds the variable introduced by the determiner in the lower position.

So we have these two differences in the DPs that occupy the lowest position in QR and Wh Movement.

(11) a. The Q in a QR'd phrase binds the variable in the lower DP. The Q in a Wh moved phrase does not bind the lower DP, but instead operates on the meaning of the clause it combines with.
b. The lower DP in a QR'd phrase introduces a variable over individuals. The lower DP in a wh moved phrase introduces a variable over functions.

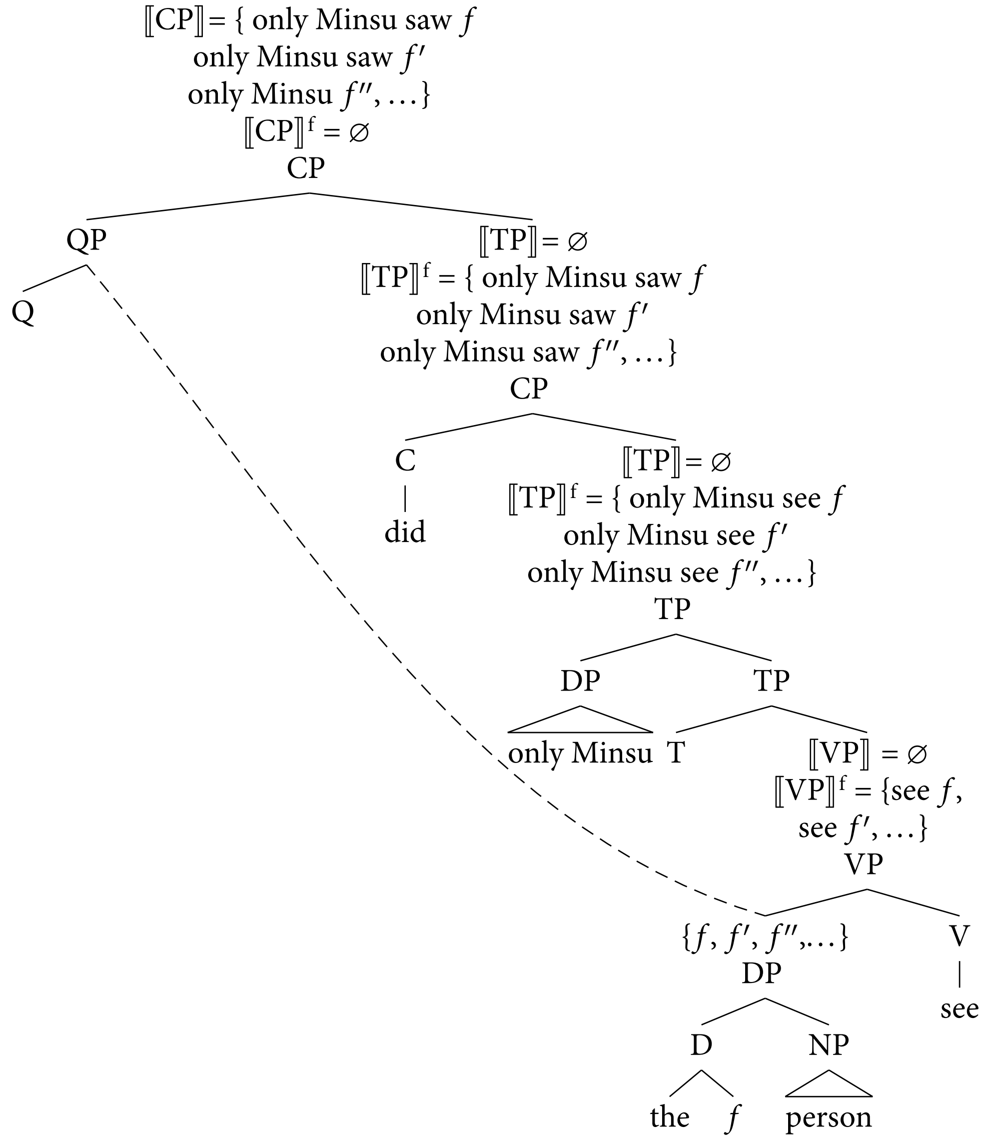
These differences in QR and Wh Movement play an important role in cases of successive cyclic derivations. The evidence for these derivations is strongest in the case of Wh Movement, and so my discussion will be confined to cases of Wh Movement. Successive cyclic derivations are ones in which two sequential applications of movement apply to reposition a term. The evidence for these derivations came first from a consideration of Wh Movement out of embedded clauses, which can occur with the two operations of movement indicated in (12).

(12) Which book did she say [$_{\text{CP}}$ that you should read]?

The proposal I will make is that the first movements in a successive cyclic derivation of Wh Movement are in fact QR. Only the last movement is Wh Movement.

There are two reasons for this. One is that it gives us an account for why overt Wh Movement in English does not show Beck effects. If the syntax for sentences like (13) is as indicated, then we should expect these sentences to have the same degraded status that they do in certain wh in situ languages.

(13) Which person should only Minsu see?

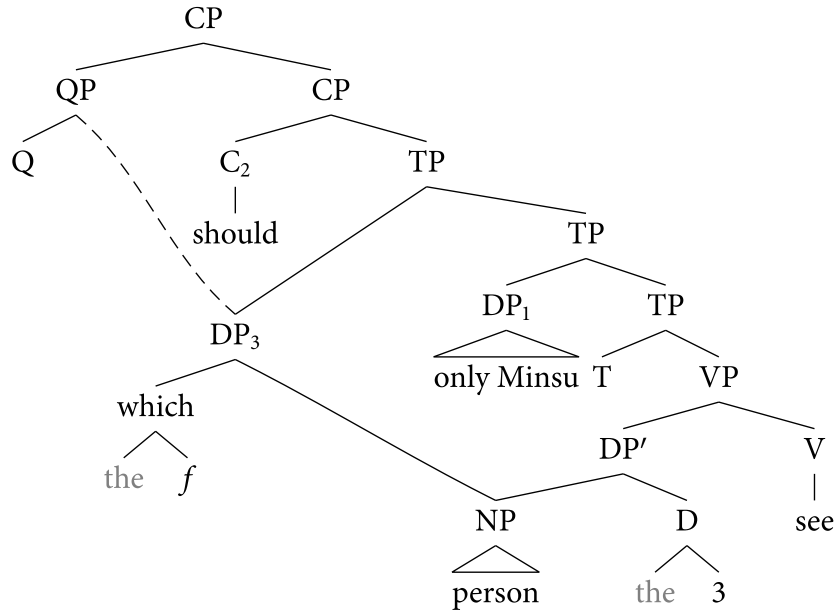


Presupposition: The output of f, f', f'', … is a person.

Beck's account of Beck effects is that focus sensitive operators, like *only*, use the focus semantic values of the things they combine with, and this disturbs the use of these focus semantic values that a higher question morpheme makes. The syntax in (13) predicts that this problem should arise in English questions as well, but it doesn't.

However if we let this sentence have a successive cyclic derivation, and allow the first step in such a derivation to be QR, we'll have a representation like (14).

(14) Which person should only Minsu see?



In this example, there is a DP headed by a determiner that will get spelled out as *which*, under Agreement with Q. These are the DPs that we've seen before function as the "traces" of Wh Movement. In this example, however, this DP shares its NP with another DP, DP′, whose head also introduces variables over functions. (This lowest DP must contain a variable over functions if we are to capture the functional question phenomena. That requires that the function over which a question abstracts be within the scope of quantificational expressions it has moved past.) These are the kinds of DPs that we've seen functioning as the traces of QR, except that the variable they introduce can be over functions and not just individuals. That I would relate to the fact that the "quantifier" in this example is not drawn from the class of lexical items that we've illustrated QR with up to this point. Those quantifiers — things like *every* and *no* — specify that they quantify over individuals only. In this case, however, we are dealing with *which*, and we've already determined that this expression can involve quantification over functions. So what we have in (14) is QR, but with a trace that can range over functions, followed by Wh Movement.

Notice that because the higher determiner is getting matched to a vocabulary item by way of Agreement with Q, it need not fuse with the determiner of the lower DP. One of the consequences of fusion is that the QR'd phrase must get spelled out in its lower position. When fusion is lifted, as it is in this example, so also is that consequence for where the expression gets spelled out. As a result, where the *which* phrase gets linearized will be determined by what the language particular component says about wh-phrases in general. For English, this means it will show up at the left edge of the question.

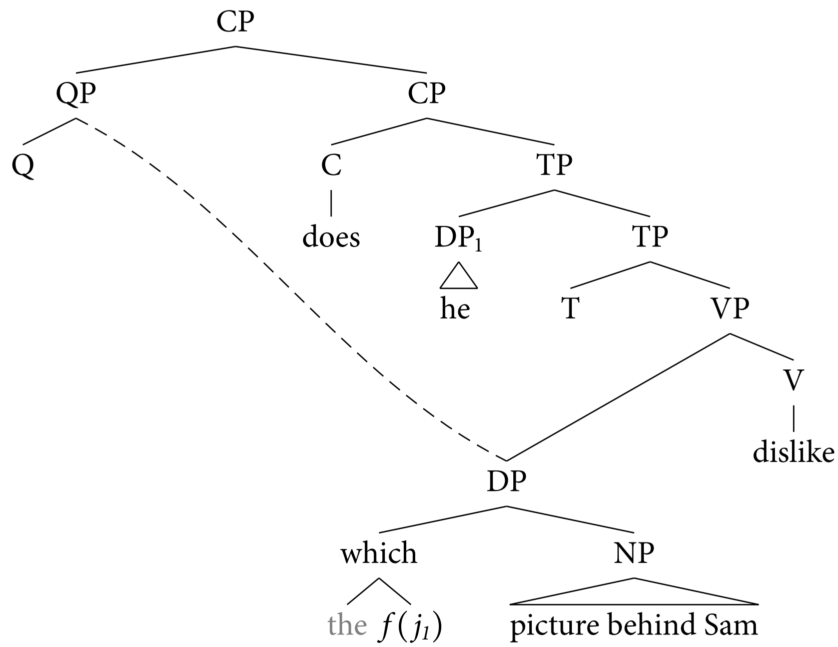
The structure in (14) does not lead us to expect Beck effects. The relationship between DP_3 and DP′ is one of variable binding, and that can happen over focus sensitive operators like *only*. The focus semantic values introduced by *which* don't commence until a larger portion of the phrase marker is encountered. As a consequence, there are no focus sensitive operators between Q and the DP headed by *which*, and therefore no Beck effect is expected.

That is one reason for believing that the first step in a successive cyclic derivation of Wh Movement is QR. Another reason has to do with the ability of Wh Movement to bleed disjoint reference effects. (15) is an example of that sort.

(15) Which picture behind Sam_1 does he_1 dislike?

It is possible to understand *Sam* and *he* to corefer in this example. Without the representations made available by successive cyclic derivations, this is unexpected. (15) should get the representation in (16), which places *Sam* within the scope of *he*.

(16)



That should make (15) parallel to (17) with regard to coreference, and that is wrong.

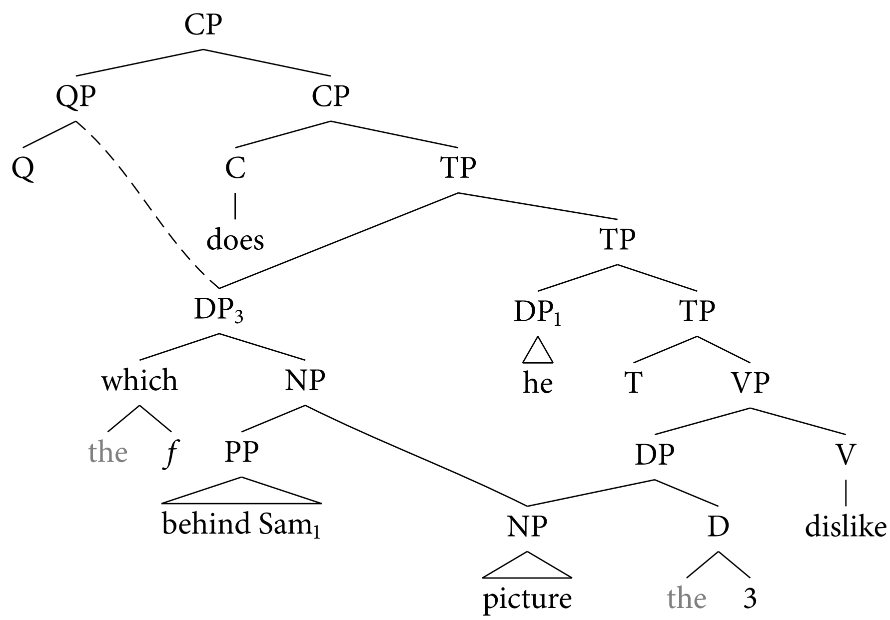
(17) * He$_1$ dislikes a picture behind Sam$_1$.

The standard solution to examples like (15) is to let the PP containing *Sam* “late merge” into a higher NP, and therefore not be in the lower NP where it would incur a disjoint reference effect.[1] If we are to adopt that solution, it means that there must be a higher position in which this PP can be semantically interpreted. That higher position cannot be the merged-to-Q position, however, since the question Q does not semantically combine with the phrase it has merged to. Without successive cyclic movement, our representations only offer one position for the material within a Wh-moved phrase to be interpreted, and that is the lowest position. But as these examples show, there needs to be some higher position that the material can be interpreted in.

If successive cyclic movement can involve QR steps, then those needed higher positions can be produced. Letting QR feed Wh movement in (15) will produce a representation like (18).

[1] The idea goes back to Lebeaux (1988).

(18)



Presupposition: 3 is a picture

Presupposition: The output of f is a picture behind Sam

QR merges the DP *which picture behind Sam* to TP in this exmaple, and this DP binds *the picture*. Because the higher DP is outside the scope of *he*, *Sam* and *he* can corefer.

That successive cyclic Wh movement can involve steps that are QR has independent support from cases where the wh-phrase contains a quantifier. A well-studied example of that kind is found in *how many* questions, like (19).

(19) How many examples of Wh movement will we have to endure?

There is in this question both the variable introduced by *how* and the variable introduced by *many*. The variable introduced by *how* ranges over amounts, or numerical degrees. The variable introduced by *many* ranges over individuals that are examples. The variable introduced by *how* gets resolved at the CP that has the question meaning. But the meaning introduced by *many* can be evaluated in different positions. Where *many* is evaluated will control where the NP is semantically interpreted as well. That is what we learn from examples like those in (20) and (21), introduced by Heycock (1995) and discussed in Fox (1999).

(20) a. * How many stories about Diana$_1$ is she$_1$ likely to invent?
b. How many stories about Diana$_1$ is she$_1$ likely to reinvent?

(fashioned after Fox 1999, (19): 167)

(21) a. * How many houses in John's$_1$ city does he$_1$ think you should build?
b. How many houses in John's$_1$ city does he$_1$ think you should rebuild?

(Fox 1999, 20: 167)

The creation verbs *invent* and *build* favor an interpretation in which *many stories* is interpreted in the embedded clause. This forces *Diana* to be interpreted in this position, and a disjoint reference effect arises as a consequence.

Let's take a closer look at how the system here captures these facts.

We should begin by getting a fix on what the structure for a *how many* DP is. We should begin with the simpler, non-interrogative, cases of *many* DPs. Hackl (2000) has argued that these expressions involve degree phrases in which the degrees being measured are numerical amounts. That permits a unified analysis of the expressions in (22).

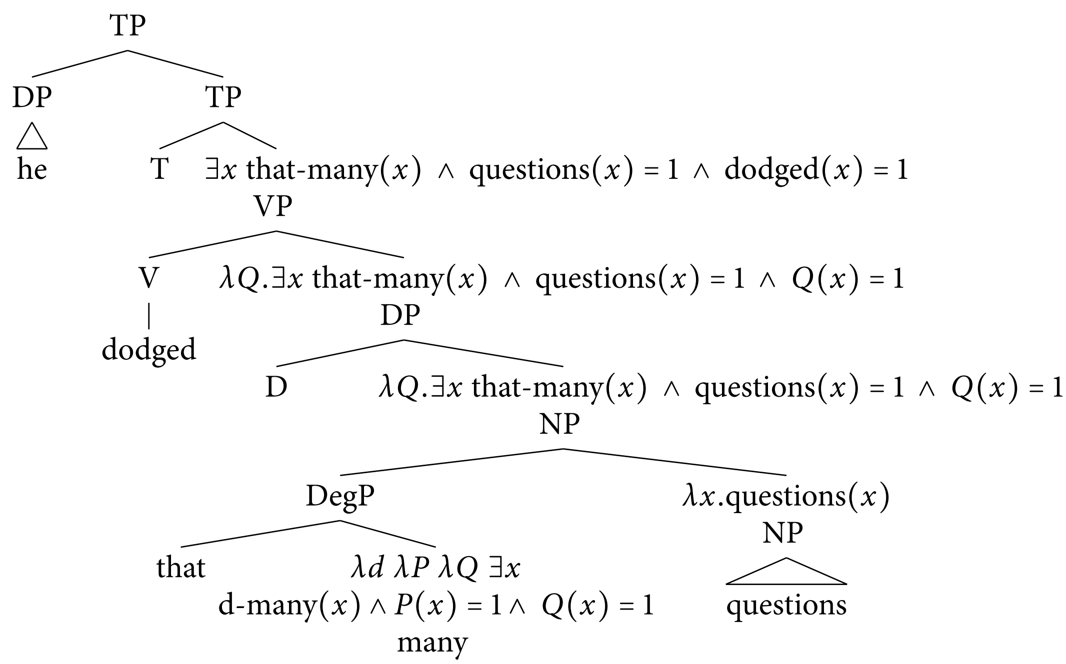
(22) a. He dodged (that) many questions.
b. He dodged more questions (than you).
c. He likes soup more than he likes kumquats.

more is the comparative form of *many*: it means what *manier* would. We can see from (22b) that *many* can be part of a comparative construction then, and from (22c), we see that these comparatives can involve comparisons of amounts that aren't numerically expressible. A standard way of expressing the amounts that are involved in these expressions is with "degrees" on a scale that is invoked by the predicates involved. So our first step is, following Hackl, to take *many* to express an amount of a degree. In the case of simple expressions like (22a), those degrees can be thought of as simple numerical amounts. So the meaning we'll assign to (22a) will be something tantamount to "He ate an amount of kumquats that is many." Here's a stab at that:

(23) He dodged that many questions.

(24) $[\![\text{that}]\!]$ = a contextually fixed degree

(25) $[\![\text{many}]\!] = \lambda d_{<\text{n}>}\ \lambda P_{<\text{e,t}>}\ \lambda Q_{<\text{e,t}>} \exists x.\text{d-many}(x) \wedge P(x) = 1 \wedge\ Q(x) = 1$



I've built into the meaning of *many* that the degrees it combines with are numerical — that's encoded in restricting the semantic type of the degree argument to numbers, i.e. <n>. I've put the existential quantification into the denotation of *many*, and that could well be wrong. More likely is that the DP containing *many* gets interpreted as existentially closed. I'm going to adopt this expedient, though, in order to shorten my narrative.

In *how many* questions, the *how* quantifies over the degree variable: $d_{<n>}$ part. That is, the part that is occupied by *that* in (23). There is evidence that we should see *how* as being equivalent to *which degree*. That evidence is that *how many* questions can be understood to seek functions as answers, just as we've seen *which* questions do.

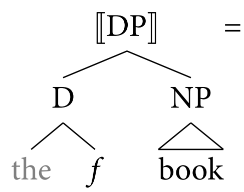
(26) How many books should everyone read?
A: 6
A: more than her professor

The second of these two answers expresses a function, one that gives amounts of books that depend on the value given to *everyone*. What we're seeing here, then, is that the degree variable which the question determiner binds is parallel to the individual variable that *which* binds in simpler questions like (27).

(27) Which book should everyone forget?
A: *Movement in Language*
A: Her first

In the case of *which*-phrases, we built the variable up from a definite determiner and an "index" that contains the function.

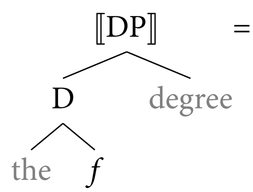
(28) ⟦which book⟧ in trace position ≈



= the value f gives, defined when that value is a book

The meaning we require for (26) is something equivalent to "which degree many books should no one read?" So we're going to want to equate *how* with the "the degree" part.

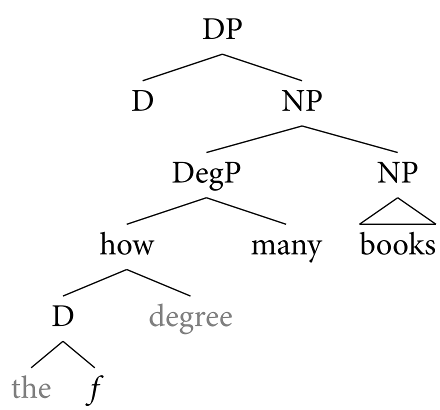
(29) ⟦how⟧ in trace position ≈



= the value f gives, defined when that value is a degree

So, this means that the DP which moves in *how many* questions looks like (30).

(30)



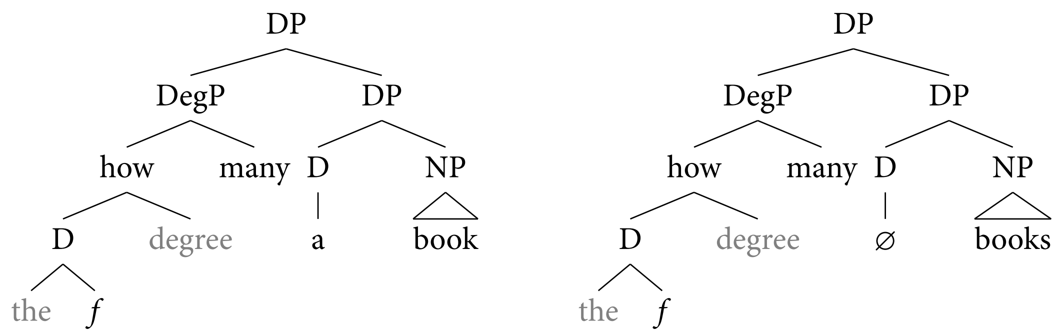
Consider next what the highest "D" position in this phrase is. I will assume that there are in fact two different *how many* expressions, and they will interface differently in movement contexts. These two *how many* expressions can be made visibly distinct by expressing the "D" in (30).

(31) a. How many a book (have you read?)

b. How many of the books (have you read?)

The D associated with these expressions can be indefinite, as in (31a), and when it is it shows up after the DegP. I'll assume that English has a silent plural definite article as well. We might parse the indefinites as (32) indicates.

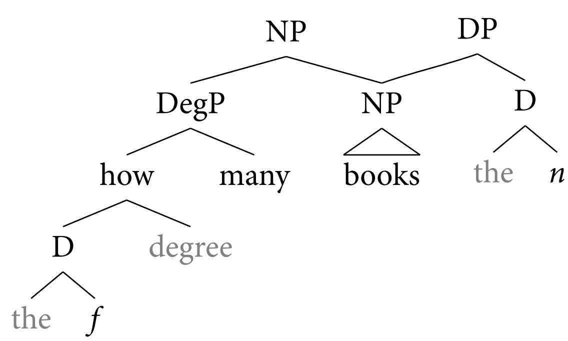
(32)



We can leave the denotation I've given for *many* the same if we understand the indefinite determiner to be semantically vacuous.

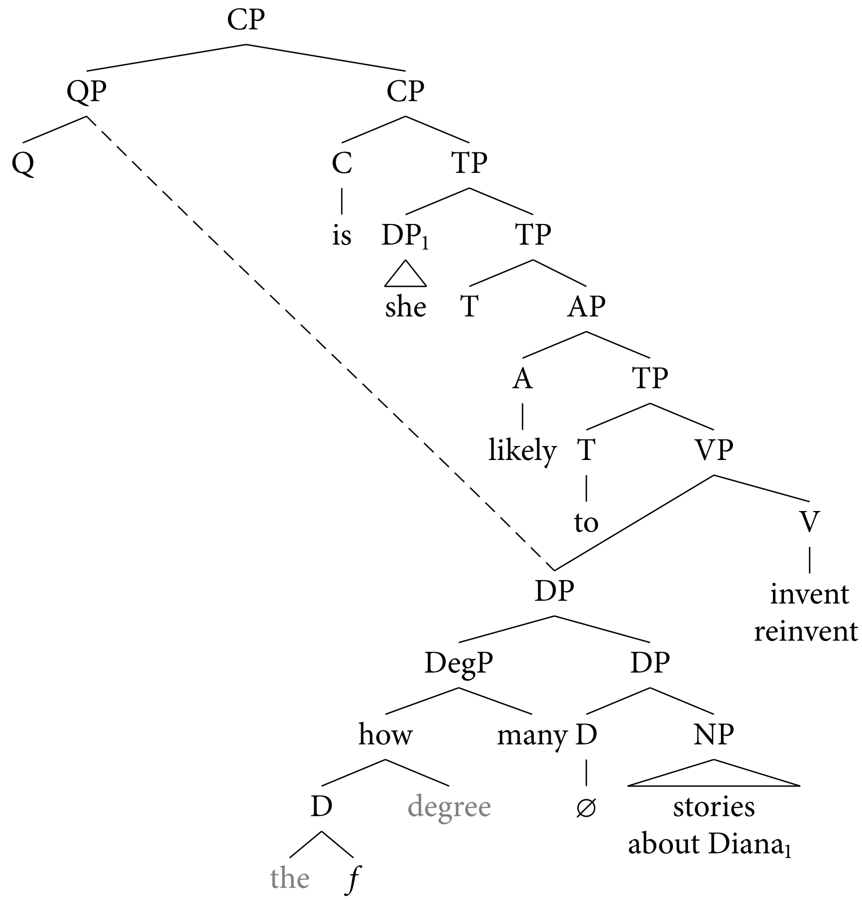
The other *how many* expression, the one with the definite determiner associated with it, will be what is involved in moving *how many* phrases so that the *many* is not interpreted in its lowest position, but instead has QR'd into a higher spot. So, we should let these expressions have the shape in (33), which is parallel to what we've seen for QRable nominals.

(33)



These DPs provide us with a way of forming representations for *how many* questions that captures the contrasts in (20) and (21). In one representation, there has been no successive cyclic movement, and *many* is interpreted in the lower position.

(34)

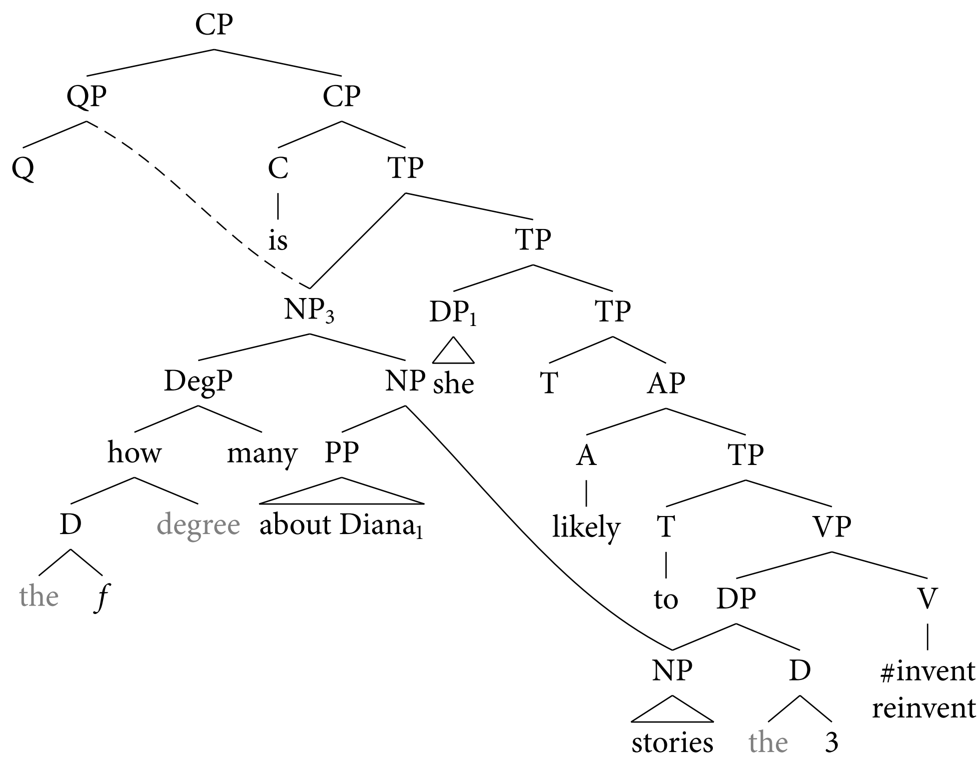


Presupposition: The output of *f* is a degree

This representation is the only one that will fit creation verbs like *invent*, as we'll see in a moment. As you can see, this representation puts *Diana* within the scope of *she*, and this is the reason we find a disjoint reference effect in the case where *invent* is used.

The other representation is one in which the *how many* phrase QRs first.

(35)



Presupposition: The output of f is a degree.

Presupposition: 3 is a story.

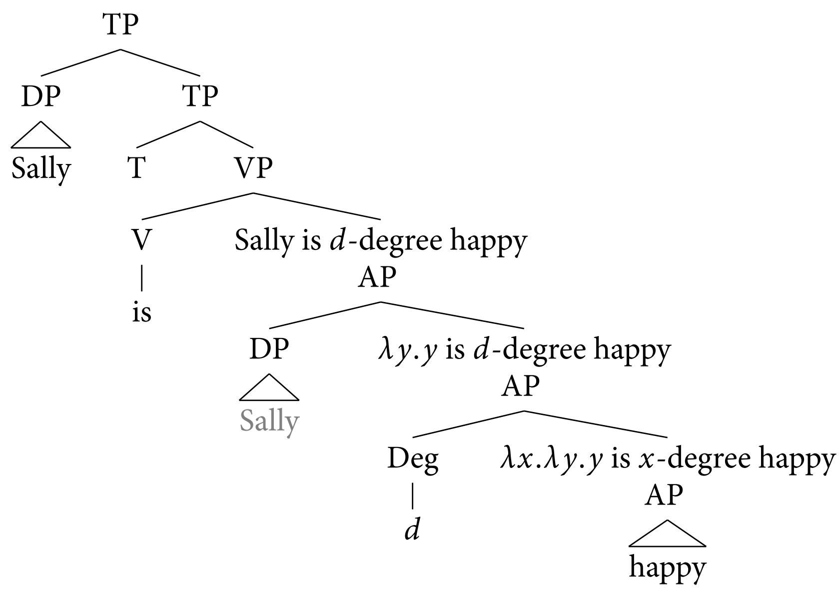
This representation doesn't fit well with creation verbs like *invent*. When we work out what (35) means, we'll see that *many* introduces an existential quantification, and the DP in object position is a variable bound by this existential quantification. So, this sentence is going to say that there exist some stories about Diana, and then adds that she is likely to invent them. But inventing stories that already exist is an anomalous thing to say. By contrast, when indefinites are interpreted in the object position of creation verbs, as in (34), they lose their existential force, and that is why (34) is not anomalous. Because this is the only representation that puts *Diana* outside the scope of *she*, it is the only representation in which *Diana* and *she* can corefer. This explains, then, the contrast in (19).

It is useful to compare these examples with ones in which *how* is part of a predicate, not a DP, as in questions like (36).

(36) How happy is Sally?

The idea about adjectives like *happy* is that they are relations between degrees and individuals. So a sentence like (37) has the representation in (38).

(37) Sally is happy.

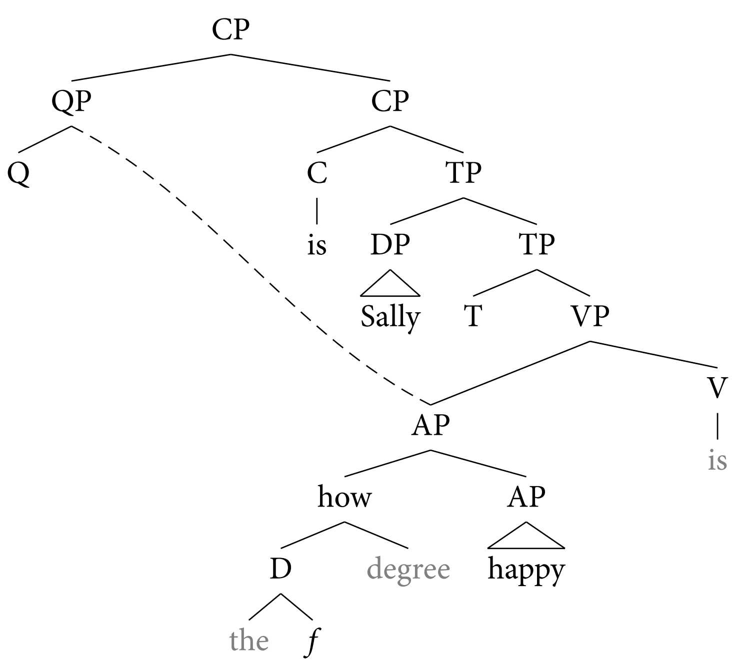
(38) 

The *how* in (36) is the same one that we saw in *how many* questions. As in *how many* questions, questions like (36) can get a functional reading.

(39) How happy is every student?

A: Way more than her professor!

So (36) gets an analysis like that indicated in (40).

(40) 

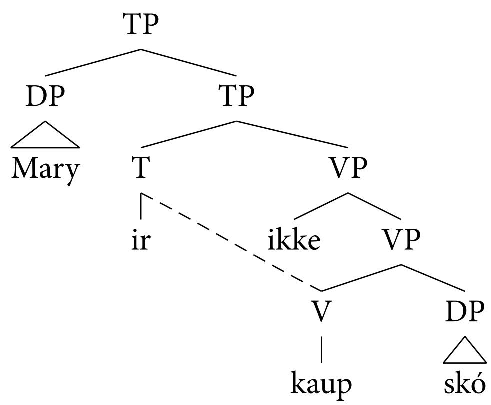
Now, an interesting difference between examples like these, where *how* is part of an AP, and those like (19), in which *how* is part of a DP, is that in these AP cases there is no possibility of QR, and therefore no possibility of successive cyclic movement. If successive cyclic movement is not possible, then everything in the moved phrase (except Q) must be interpreted in its lowest position. This predicts that Principle C effects should be unavoidable, and this is known to be the case.

(41) a. * How satisfied by Sarah$_1$'s debate was she$_1$?
b. * How satisfied because of Sarah$_1$'s performance did she$_1$ say you were?

(For a discussion of these cases, see Barss 1986, Huang 1993 and Takano 1995.) The reason there is no possibility of QR is because the AP that is moved is not a definite description. The cases we have seen where a variable exists in the lower position are both cases in which a definite description is placed in this position. If the position a term moves from is not a position in which a definite description can be interpreted, then we will not have an example of movement that gets translated into a variable-binder relationship. Instead, these cases of movement will all be instances where the moved item behaves, semantically, as if it hadn't moved.

As Heycock (1995) argued, that seems very generally true. We have a handle, then, on why verb movement behaves semantically unlike DP movement — one of the differences in the expression of Semantic Displacement that I raised in the first lecture. Verb movement will invoke a syntax like that in (42).

(42) Mary kaupir ikke skó
Mary buys not shoes
'Mary doesn't buy shoes.'

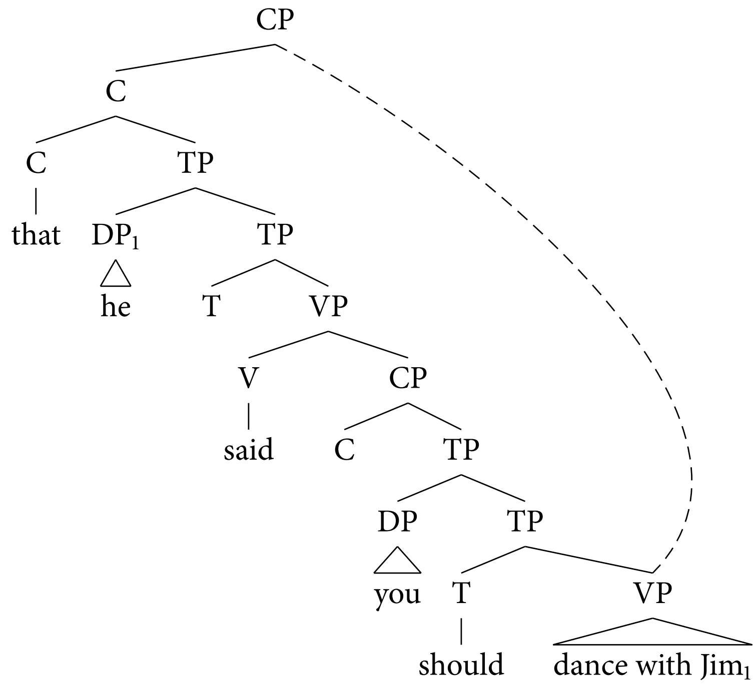


There is no way of semantically interpreting this structure except as indicated: the verb is interpreted in its lowest position only. A similar outcome will arise for cases where a VP moves, as in cleft or topicalization constructions.

(43) * It's [$_{VP}$ dance with Jim$_1$] that he$_1$ said you should.

This will get a representation like (44).

(44)



Again: the VP that has moved can only be interpreted in its lower position. There is no syntax that would allow it to combine with a definite description that could function as a variable in the lower position. As expected, there is a disjoint reference effect between *Jim* and *he*.

The proposal here explains these differences in how movement is interpreted in a simple way. Because movement is nothing more than arranging the terminals and phrases that a language has outside of movement contexts, the material that can be put into the higher and lower positions of a movement structure are just those things that exist in the language elsewhere. English has DPs that function as variables — we have concentrated here on those cases found in donkey-anaphora — and so these DPs can be an ingredient in movement relations that get a binder-variable interpretation. But English has no Vs, VPs or APs that can get interpreted as variables, and so movement relations that involve these terms cannot get a binder-variable interpretation.

References

Barss, Andrew. 1986. Chains and anaphoric dependence: On reconstruction and its implications. Doctoral Dissertation, Massachusetts Institute of Technology.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Hackl, Martin. 2000. Comparative quantifiers. Doctoral Dissertation, Massachusetts Institute of Technology.

Heycock, Caroline. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26:547–570.

Huang, C.-T. James. 1993. Reconstruction and the structure of VP: some theoretical consequences. *Linguistic Inquiry* 24:103–138.

Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral Dissertation, University of Massachusetts, Amherst.

Takano, Yuji. 1995. Predicate fronting and internal subjects. *Linguistic Inquiry* 26:327–340.