Kyle Johnson
University of Massachusetts at Amherst

University College London
20 June 2016

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

(1) a. ...að Mary kaupir ekki skó?   (*Icelandic*)
       ...that 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 instances 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.

A. Difference in Semantic Displacement

   a. *Total Reconstruction:*

      Mary kaupir ekki 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

An argument that the semantic relationship between a moved verb and its underlying position is different than a moved phrase and its underlying position can be made from their different behavior in ellipsis contexts. VP Ellipsis can leave behind an object DP or a verb in contexts where that DP or verb has moved out of the VP. In English this is possible with a DP, in what goes by the name "pseudogapping."

(5) I know what kind of rice she will eat but not what kind of starch she should.

The position vacated by the moved object must be semantically equivalent to the matching position in the antecedent VP for the ellipsis to be licensed. That position must be occupied by a variable.

(6) What kind of rice$_1$ she will [$_{VP}$ eat x_1] but not

    what kind of starch$_2$ she should [$_{VP}$ eat x_2]

Because [$_{VP}$ eat x_1] and [$_{VP}$ eat x_2] are semantically equivalent for the purposes of ellipsis resolution, VPE is possible. The same is not the case, however, for (some) situations in which a verb has moved out of a VP that elides. See Goldberg (2005). The following are her examples, inspired by Doron (1999).

(7)   Q: Ha'im Miryam hevi'a et   Dvora la-xanut?
             Q   Miryam bring ACC Dvora to.the-store
             'Did Miryam bring-past Gvora to the store?'

       Ai: Ken, hi   hevi'a.
                yes,   she bring-past
                'Yes, she brought.'

       Aii: * Ken, hi   lakxa.
                 yes,   she take-past
                 'Yes, she took.'

This is derived if the movement in these cases doesn't leave a variable, but is instead, semantically vacuous.

(8)   Q: Ha'im Miryam hevi'a [$_{VP}$ hevi'a Dvora la-xanut]
                                    └──────┘

       A: Ken, hi lakxa [$_{VP}$ lakxa Dvora la-xanut]
                             └────┘

Because the VPs in these examples aren't semantically equivalent, VPE isn't allowed. (But see Hartman (2011) for evidence that suggests that verb movement does leave variables.)

B.   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 [$_{CP}$ that   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].
           └──┘

C.   Differences in how Terseness is violated

(9)   ngɔ̄nū̄ n̄   wà   nā̄ n̄   ká̀   ngɔ́nʊ̀ á̀   (*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.

(10)   * tākā̄ n̄   wà   fòtó̀   mūmʊ̀̀ n̄   tákà   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. (9) is ungrammatical if either verb isn't pronounced. When nominal material clefts, by contrast, only the higher copy may be pronounced. (11) illustrates.

(11)   ngɔ́nʊ́lì   mí́ n̄   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 Sichel forthcoming and references therein.)

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

(13)   * təlmiiz-a$_1$ lkəsleen ma   badkun tχabbro [wala mʕallme] ʕan
        student-her$_1$ the-bad NEG want.2P tell.2P [no   teacher]   about
        l-bənt   yalli *huwwe* zaʕbar   maʕ-a   b-l-faḥṣ
        the-girl that he   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, Choueiri, and Hornstein 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.

(14)   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 (15), the question seems to assume that the answer will provide individuals that meet the descriptions provided by both of the clauses.

(15)   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 – (16) is an example – seems to assume that the answer will provide the identity of individuals that meet the descriptions provided in each of the clauses separately.

(16)   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. (See Dayal 1994.)
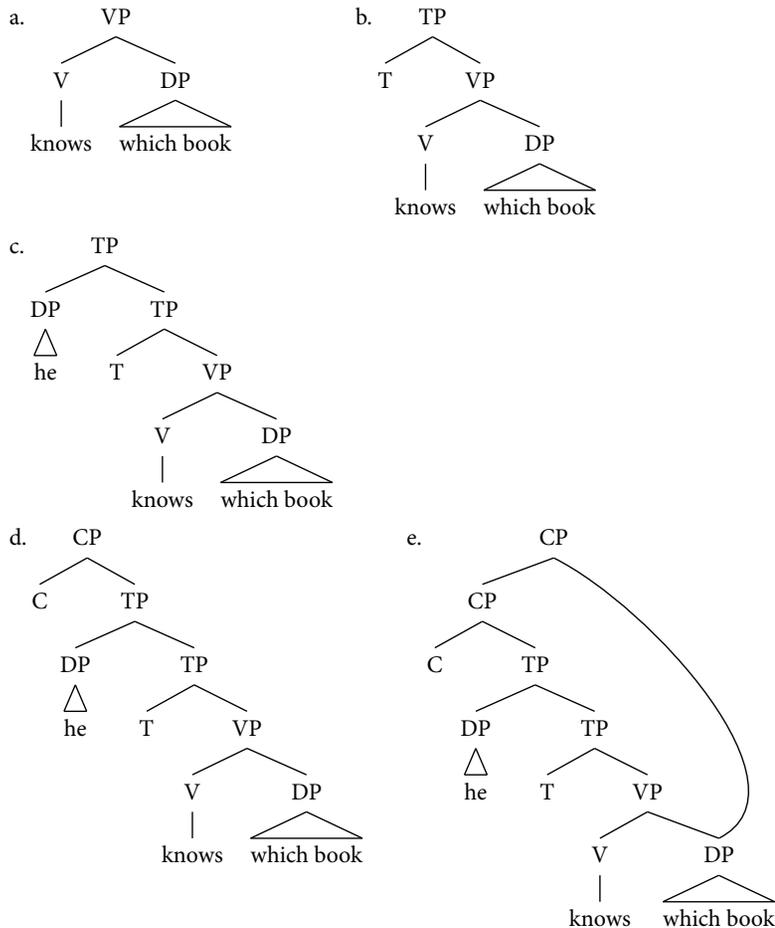
I'll assume that there is a difference between moved DPs and moved other things that is responsible for how Terseness is relaxed in them.

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.

(17)   MERGE(α,β) =$_{\text{def.}}$ γ , where the linear order of α and β is
                                           not determined.
                              α β

A derivation that involves remerge:

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

a.
```
        VP
       /  \
      V    DP
      |   /  \
   knows  which book
```

b.
```
        TP
       /  \
      T    VP
          /  \
         V    DP
         |   /  \
      knows  which book
```

c.
```
          TP
        /    \
      DP      TP
      /\     /  \
     he    T    VP
              /    \
             V      DP
             |     /  \
          knows  which book
```

d.
```
        CP
       /  \
      C    TP
          /    \
        DP      TP
        /\     /  \
       he     T    VP
                  /   \
                 V     DP
                 |    /  \
              knows  which book
```

e.
```
            CP
          /    \
        CP      \
       /  \      \
      C    TP     \
          /  \     \
        DP    TP    \
        /\   /  \    \
       he   T    VP   \
                /  \   \
               V    DP  \
               |   /  \
            knows  which book
```
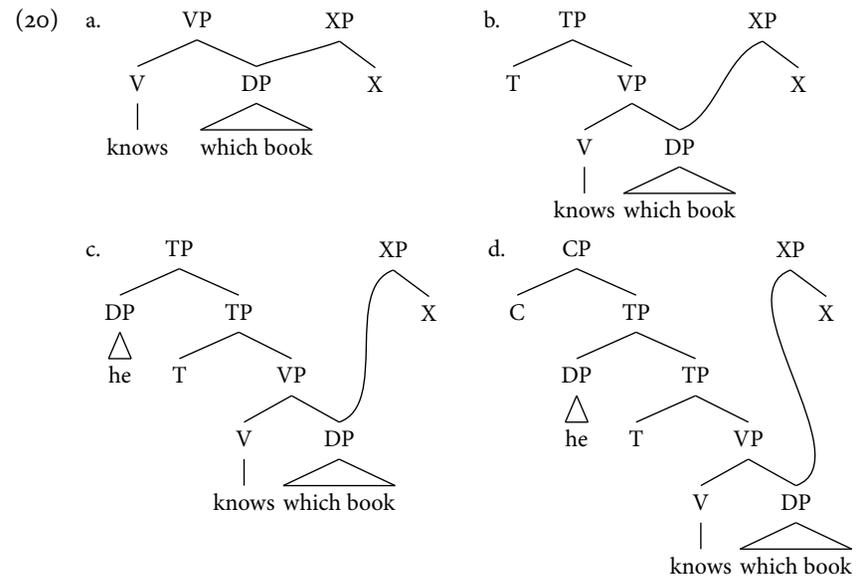
You can see how (18e) provides a way of capturing Semantic Displacement. The moved term — here *which book* — is syntactically in two positions and so its denotation has two positions where it can be applied. The differences in how Semantic Displacement arise are going to come about, I will claim, from the particular ways in which the expressions that are "moved" get broken up into two different posi-
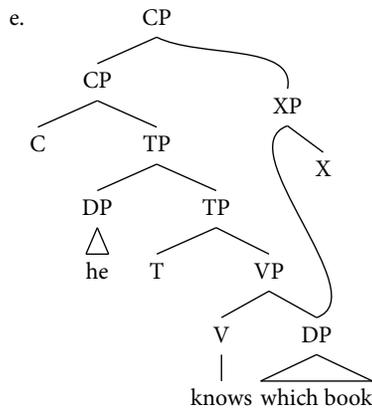
tions. That is going to be the focus of most of the seminar, I suspect, but we won't start that process until Wednesday.

It's not obvious that these representations provide any new handle on the question 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. The idea in that paper uses a constraint on derivations that Chomsky proposed which, I think, he called the "extension condition."

(19)   Extension Condition
       MERGE's arguments must be root nodes.

The Extension Condition is one attempt at deriving the fact that moved items always go "up" in a phrase marker. It would block the derivation in (18), and there is no way of getting to the representation at the end of that derivation without violating the Extension Condition. So if the Extension Condition is correct, then the closest we can come to (18) is with something like (20).

(20)   a.
```
        VP         XP
       /  \       /
      V    DP    X
      |   /  \
   knows  which book
```

b.
```
        TP         XP
       /  \        /
      T    VP     X
          /  \   /
         V    DP
         |   /  \
      knows which book
```

c.
```
          TP           XP
        /    \         /
      DP      TP      X
      /\     /  \    /
     he    T    VP  /
              /  \ /
             V    DP
             |   /  \
          knows which book
```

d.
```
          CP            XP
        /    \          /
      C      TP        X
            /   \      /
          DP     TP   /
          /\    /  \ /
         he    T    VP
                   /  \
                  V    DP
                  |   /  \
               knows which book
```

e.

```
            CP
          /    \
        CP      XP
       /  \       \
      C    TP      X
          /  \
        DP    TP
        /\   /  \
       he   T   VP
               /  \
              V    DP
              |    /\
          knows  which book
```

We might imagine that (20e) is close enough to (18e) that we can make basically the same theories around them. Of course, we should identify "X."

The first step in (20) is what Citko (2005) calls "parallel merge." For it to arise without a violation of the Extension Condition, we'd have to understand Merge to be able to apply at the same time to two pairs of root nodes. So derivations would have to be understood not as sequences of single applications of Merge, but as sequences of any number of applications of Merge. (Make me do this on the board.)
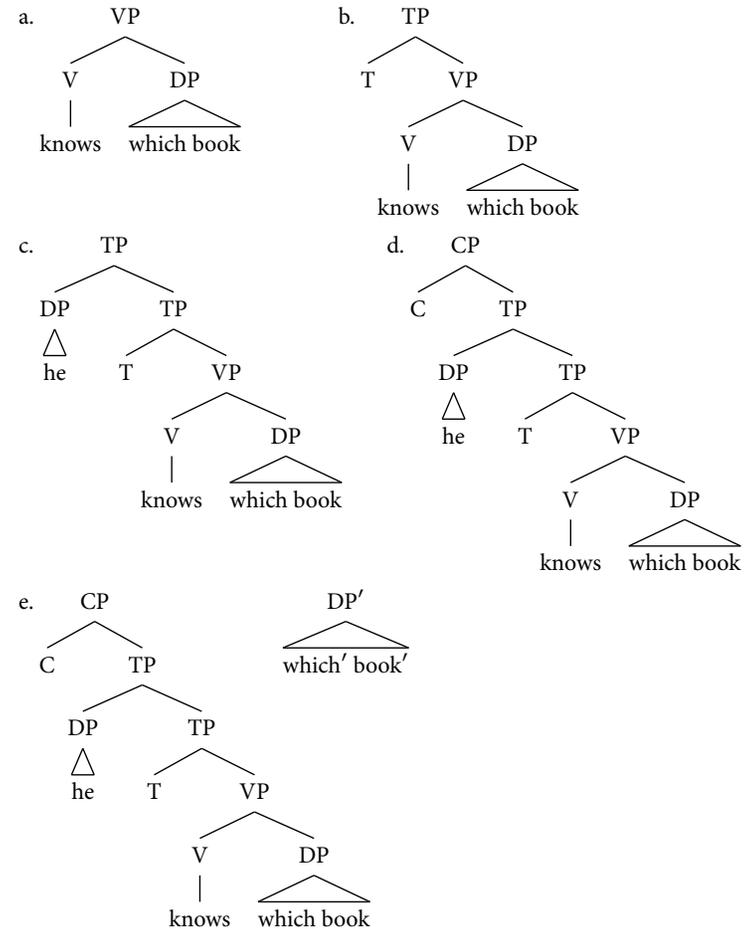
Suppose islands are those phrases at which a phonological or semantic evaluation must take place. This might be what "phases" are. If islands are phases that arise at the point in a derivation where there remain two roots, as in any but the last of the steps in (20), then arguably a semantic and phonological evaluation will not be possible, as our semantics and phonology are (perhaps) defined only for representations with a single root. This, at any rate, is one way of thinking about how to seek an answer to the question of why islands emerge uniquely with movement on an account that uses multidominant representations. I'll have nothing more to say about islands in these classes. I'll focus on the remaining two properties: terseness and semantic displacement.
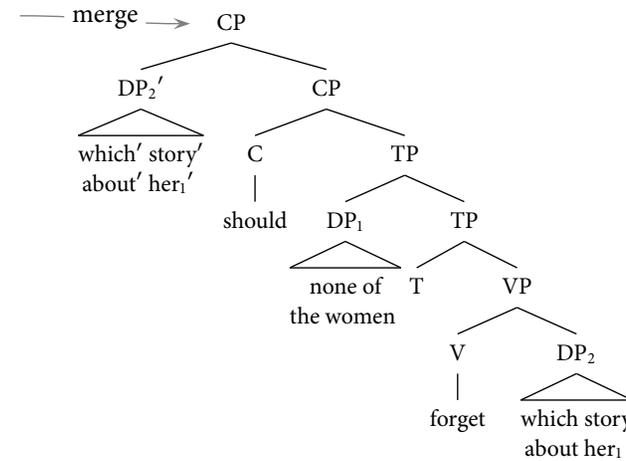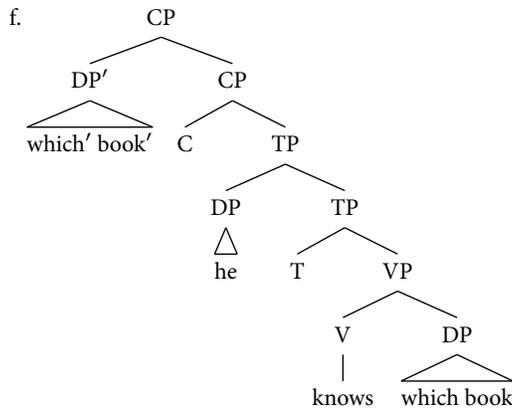
We'll start by tackling how multidominant representations derive Terseness. I'll adapt a popular explanation for Terseness due to Jairo Nunes. Nunes worked with a (possibly) different account of movement: the "copy" theory of movement. This theory does not countenance multidominant representations, like that in (18e), but instead involves a "copy" operation.

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

This gives us derivations like (22).

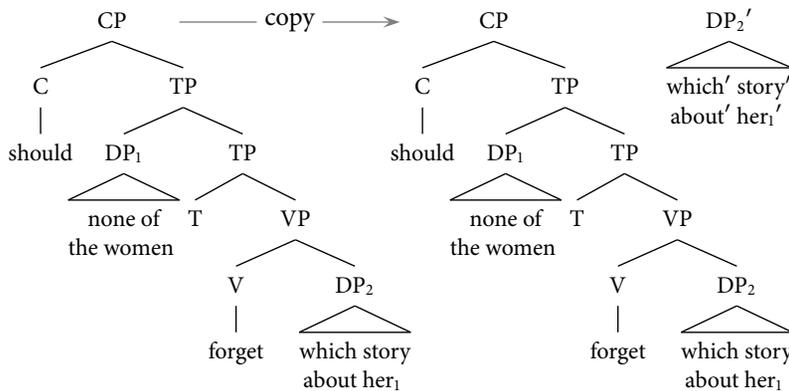(22)   (She asked) which book he knows.

a.
```
        VP
       /  \
      V    DP
      |    /\
   knows  which book
```

b.
```
        TP
       /  \
      T    VP
          /  \
         V    DP
         |    /\
      knows  which book
```

c.
```
          TP
         /  \
       DP    TP
       /\   /  \
      he   T   VP
               /  \
              V    DP
              |    /\
          knows  which book
```

d.
```
          CP
         /  \
        C    TP
            /  \
          DP    TP
          /\   /  \
         he   T   VP
                  /  \
                 V    DP
                 |    /\
             knows  which book
```

e.
```
          CP                    DP'
         /  \                  /  \
        C    TP            which' book'
            /  \
          DP    TP
          /\   /  \
         he   T   VP
                  /  \
                 V    DP
                 |    /\
             knows  which book
```

f.

CP
DP′　CP
which′ book′　C　TP
DP　TP
he　T　VP
V　DP
knows　which book

merge → CP
DP₂′　CP
which′ story′ about′ her₁′　C　TP
should　DP₁　TP
none of the women　T　VP
V　DP₂
forget　which story about her₁

This theory too is able to account for Semantic Displacement, and it does so in a way rather like that of the remerge account. So, for instance, it gives an account of "reconstruction," a special case of Semantic Displacement that (23) illustrates.

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

CP ── copy ──→ CP　DP₂′
C　TP　C　TP　which′ story′ about′ her₁′
should　DP₁　TP　should　DP₁　TP
none of the women　T　VP　none of the women　T　VP
V　DP₂　V　DP₂
forget　which story about her₁　forget　which story about her₁

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

These representations allow Terseness to arise from an operation that "deletes" one of the two phrases in the copy relation. What Terseness amounts to, then, is the observation that (with the exceptions noted) this deletion process is obligatory. 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 multidominant phrase markers. The leading idea is that a representation with undeleted copies will not be able to be linearized.

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 (24) would map onto the ordered pairs in (25).

(24)
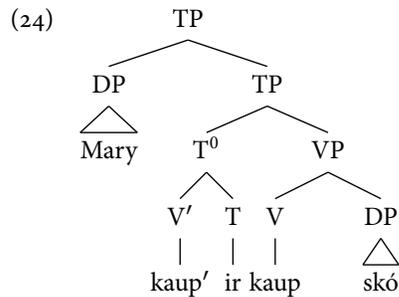
(25)
$$
\left\{
\begin{array}{lll}
\text{Mary} < \text{T}^0 & \text{T}^0 < \text{V} & \text{V} < \text{skó} \\
\text{Mary} < \text{V} & \text{T}^0 < \text{skó} & \\
\text{Mary} < \text{skó} & &
\end{array}
\right\}
$$

Note that I've represented the stem, or root, of the verb *kaupir* here with *kaup*. The set of ordered pairs in (25) involves the ordering relation "<," precedence. (25) corresponds to the string in (26).

(26)   Mary kaupir kaup skó

This isn't quite right for Icelandic, we'll return to this in a moment.

One thing to highlight about (25) is that the elements in the ordered pairs are the nodes that correspond to words – the X^0s. They are not the morphemes that make up those words. This differs from Kayne (1994), who proposes that the linearization algorithm orders the syntactic parts of a word as well. Nunes introduces the idea that linearizations don't do that to account for certain violations of Terseness, as we shall see.

In Kayne's work, the linearization procedure produces linearizations which are 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 (27).

(27)   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.

Simplifying somewhat, Kayne's procedure forms the ordered pairs in a linearization with (28)

(28) $a < b$ is in the linearization iff *a*, or a phrase containing *a*, c-commands *b*.

This has a number of interesting consequences for phrase markers, which is the focus of Kayne's work. It also has two other consequences that are of interest to us:

(29)   a.   It forces every vocabulary item in a c-command relation to be in the linearization.
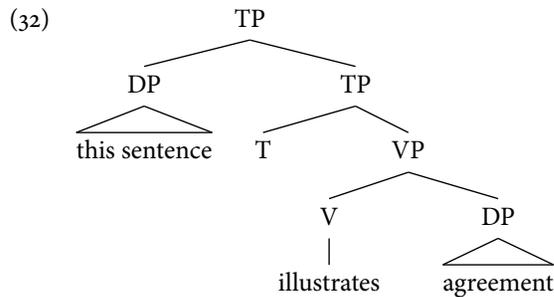
         b.   It enforces Contiguity.

(30)   Contiguity

         Let α be a word dominated by A and β be a word dominated by A's sister.

         a. $\forall \alpha, \beta, \alpha < \beta$, or

         b. $\forall \alpha, \beta, \beta < \alpha$.

Contiguity is incompatible with multidominant representations, but we don't want to completely lose its effects. We will need to modify how Contiguity is derived, and for this reason I want to jettison Kayne's procedure. Instead, let's explore how Nunes's system works if we extract from Kayne's procedure the two consequences in (29). That is let's adopt (31), and add Contiguity to our set of well-formedness conditions.

(31)   If *a* and *b* are vocabulary items in a sentence, then either $a < b$ or $b < a$ is in that sentence's linearization.

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 (32) a collection of sets that, once filtered through the constraints, would result in those listed in (33). (I will indicate the linearizations with the (more compact) strings they correspond to, rather than with the full sets of ordered pairs.)

(32)

```
                    TP
              ┌─────┴─────┐
             DP           TP
            ╱  ╲      ┌────┴────┐
    this sentence    T         VP
                           ┌────┴────┐
                           V         DP
                           │        ╱  ╲
                      illustrates  agreement
```

(33) 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

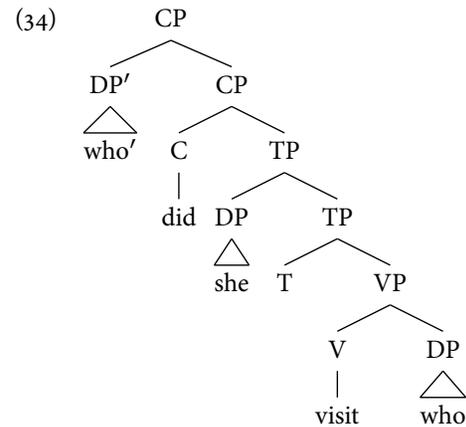p. agreement illustrates T sentence this

The ill-formed linearizations in (33) are, many of them, well formed in other languages. For instance, (33f) corresponds roughly to how German would linearize this structure, and (33i) 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

---

[1] (33n,o) are vanishingly rare according to Dryer (1996).

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."

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 on Kayne's requirement that the linearization algorithm necessarily applies to every vocabulary item in the structure being linearized. When these are coupled with the copy theory of movement, they will produce unlinearizable results.

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

(34)

```
                    CP
              ┌─────┴─────┐
            DP′           CP
            ╱ ╲      ┌─────┴─────┐
          who′       C          TP
                     │      ┌────┴────┐
                    did    DP         TP
                          ╱ ╲     ┌────┴────┐
                         she     T         VP
                                       ┌────┴────┐
                                       V         DP
                                       │        ╱  ╲
                                     visit      who
```

The linearization of (34) that satisfies the language particular component as well as Totality, Contiguity, and Transitivity is (35).

(35)   = who′ did she visit who

$$\left\{\begin{array}{lll}
who′ < did & did < she & she < T \quad\quad T < visit \quad visit < who \\
who′ < she & did < T & she < visit \quad T < who \\
who′ < T & did < visit & she < who \\
who′ < visit & did < who \\
who′ < who
\end{array}\right\}$$

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's deletion operation, which he calls "chain reduction," applies.

(36)   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 (34), it could remove DP or DP′ and thereby produce representations from which the linearization algorithm produces the strings in (37).

(37)   a.   = who did she visit

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

    b.   = did she visit who

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

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 (37a) is the correct one, which we can credit this to the language particular component.

This is how Terseness is derived.

    This method provides an interesting account for those examples of Head Movement where Terseness seems to be relaxed. For instance, Vata predicate cleft constructions involve a structure like (38), in which the verbal root has moved and joined with a functional head that encodes focus.

(38)   lī ɔ̀   dā sàká lī
       eat she/he perf rice eat
       'she/he has eaten rice.'

(39)
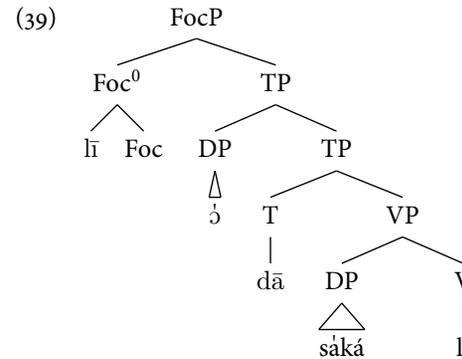
Because the result of clefting a verb in Vata makes it merely a part of the vocabulary item that is Foc[0], Antisymmetry is able to be satisfied without invoking Chain Reduction. If constraints like Antisymmetry make reference to vocabulary items, and not the terminals from which those vocabulary items are composed, then putting a copy into a vocabulary item will effectively "hide" it from Antisymmetry. The representation in (39), for instance, can be assigned the linearization in (40).

(40)
$$\left\{\begin{array}{llll}
Foc^0 < ɔ̀ & ɔ̀ < dā & dā < sàká & sàká < lī \\
Foc^0 < dā & ɔ̀ < sàká & dā < lī \\
Foc^0 < sàká & ɔ̀ < lī \\
Foc^0 < lī
\end{array}\right\}$$

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 (39) is produced.

    This is precisely what we saw for Icelandic Verb Movement as well, but in that context the result is ungrammatical.

(24)

TP
├─ DP
│   △
│   Mary
└─ TP
    ├─ T⁰
    │   ├─ V′
    │   │ |
    │   │   kaup′
    │   └─ T
    │ |
    │   ir
    └─ VP
        ├─ V
        │ |
        │   kaup
        └─ DP
            △
            skó
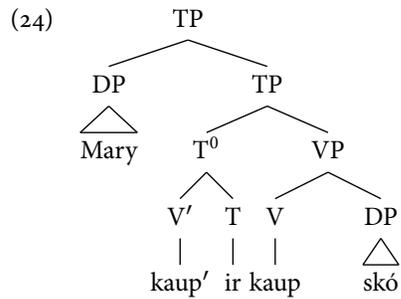
(25)
$$\left\{ \begin{array}{lll} \text{Mary} < T^0 & T^0 < V & V < \text{skó} \\ \text{Mary} < V & T^0 < \text{skó} & \\ \text{Mary} < \text{skó} & & \end{array} \right\}$$

(26)   Mary kaupir kaup skó

I suggest that the difference between Icelandic and Vata has to do with their lexicon. Vata has lexical items to match both the higher and lower copies, but Icelandic has no lexical item to match the bare root of a verb.

A virtue, then, of a Multidominant theory of movement is that it derives Nunes' stipulation that Antisymmetry treats something and its copy as the same thing. On a multidominant 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 a Multidominant 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 multidominant 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.

We do that tomorrow.

References

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.

Dayal, Veneeta Srivastav. 1994. Scope marking as indirect wh-dependency. *Natural Language Semantics* 2:137–170.

Doron, Edit. 1999. V-movement and VP ellipsis. In *Studies in ellipsis and gapping*, ed. Shalom Lappin and Elabbas Benmamoun, 124–140. New York, New York: Oxford University Press.

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.

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.

Goldberg, Lotus Madelyn. 2005. Verb-stranding VP ellipsis: A cross-linguistic study. Doctoral Dissertation, McGill University, Montreal.

Hartman, Jeremy. 2011. The semantic uniformity of traces: Evidence from ellipsis parallelism. *Linguistic Inquiry* 42:367–388.

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.

Sichel, Ivy. forthcoming. Resumptive pronouns and competition. *Linguistic Inquiry* .

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.

We left last time looking at how Nunes derives Terseness. His method uses the linearization scheme in (1).

(1)   a. If a and b are vocabulary items in P, then put either $a < b$ or $b < a$ in the linearization for P.

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

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

      d. *Contiguity*
          Let α be a word dominated by A and β be a word dominated by A's sister.
          i. $\forall \alpha, \beta, \alpha < \beta$, or
          ii. $\forall \alpha, \beta, \beta < \alpha$.

This, recall, is that portion of a linearization algorithm that is common to all languages. This won't produce a linearization for a particular sentence in a particular language. Instead, it will produce a set of linearizations for any particular sentence. The language will pick from this set the one(s) that fit the language particular settings for a linearization.

     To (1), Nunes adds (2).

(2)   a. If a and a' are copies, then Antisymmetry and Totality cannot distinguish them.

      b. *Chain Reduction*
          If a and a' are copies, then one of them can left out of the linearization.

(2a) follows from a multidominant theory of movement. So to the degree that Nunes's program is successful, it can be read as evidence for multidominance. But this requires that we do away with Chain Reduction and change the linearization procedure.

     As a first step, I suggest that we jettison Chain Reduction and (1a). (1a) is just Totality. Let's let the linearization for a sentence be any set of ordered pairs $a < b$

of vocabulary items that satisfies Totality, Antisymmetry, Contiguity and the language particular setting. Consider how that will work in a case of wh-movement, whose parse, on our theory, will be (3).

(3)
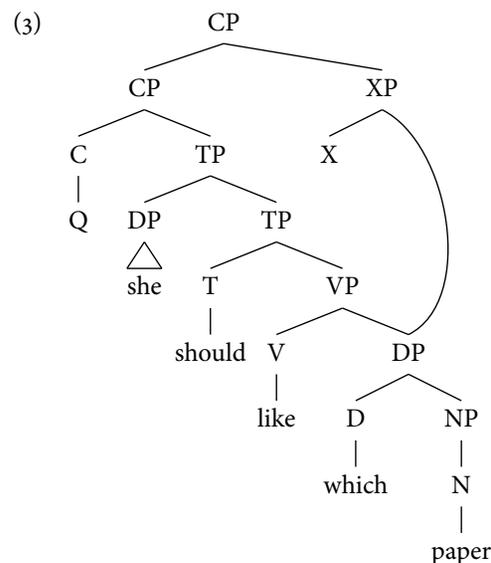
Because the wh-phrase has two positions, we can imagine that the way the ordered pairs are collected up into a linearization permits two positions for the wh-phrase: the position in the string assigned by English to Specifier of CP and the position in the string assigned by English to complement of V. Let's consider, then, the three linearizations that satisfy Totality in (4).

(4)   a. which paper Q she should like

$$\left\{\begin{array}{lllll}
\text{which} < \text{paper} & \text{paper} < \text{Q} & \text{Q} < \text{she} & \text{she} < \text{should} & \text{should} < \text{like} \\
\text{which} < \text{Q} & \text{paper} < \text{she} & \text{Q} < \text{should} & \text{she} < \text{like} & \\
\text{which} < \text{she} & \text{paper} < \text{should} & \text{Q} < \text{like} & & \\
\text{which} < \text{should} & \text{paper} < \text{like} & & & \\
\text{which} < \text{like} & & & &
\end{array}\right\}$$

   b. Q she should like which paper

$$\left\{\begin{array}{llll}
\text{Q} < \text{she} & \text{she} < \text{should} & \text{should} < \text{like} & \text{like} < \text{which} & \text{which} < \text{paper} \\
\text{Q} < \text{should} & \text{she} < \text{like} & \text{should} < \text{which} & \text{like} < \text{paper} & \\
\text{Q} < \text{like} & \text{she} < \text{which} & \text{should} < \text{paper} & & \\
\text{Q} < \text{which} & \text{she} < \text{paper} & & & \\
\text{Q} < \text{paper} & & & &
\end{array}\right\}$$

   c. which paper Q she should like which paper

$$\left\{\begin{array}{lllllll}
\text{which} < \text{paper} & \text{paper} < \text{Q} & \text{Q} < \text{she} & \text{she} < \text{should} & \text{should} < \text{like} & \text{like} < \text{which} & \text{which} < \text{paper} \\
\text{which} < \text{Q} & \text{paper} < \text{she} & \text{Q} < \text{should} & \text{she} < \text{like} & \text{should} < \text{which} & \text{like} < \text{paper} & \\
\text{which} < \text{she} & \text{paper} < \text{should} & \text{Q} < \text{like} & \text{she} < \text{which} & \text{should} < \text{paper} & & \\
\text{which} < \text{should} & \text{paper} < \text{like} & \text{Q} < \text{which} & \text{she} < \text{paper} & & & \\
\text{which} < \text{like} & \text{paper} < \text{which} & \text{Q} < \text{paper} & & & & \\
\text{which} < \text{which} & \text{paper} < \text{paper} & & & & &
\end{array}\right\}$$

Of these, only (4b) satisfies all the constraints. (4c) is a violation of Terseness, and it's clear how Antisymmetry blocks it. The linearizations that obey Totality and Antisymmetry are (4a) and (4b), and these are the two that we want to make cross-linguistically available. Terseness emerges from a linearization of multidominant structures if we let the linearizations be any collection of ordered pairs that satisfy Totality and Antisymmetry. We don't need a deletion rule like Chain Reduction.

The correct outcome in English, of course, is (4a) – so we must find a way of making Contiguity permit (4a) as well as (4b). But we don't want to give Contiguity up. Not only is it obeyed in parses that don't have movement, it's also mostly obeyed in parses that do have movement. For instance, we should block giving (3) the linearization in (5).

(5)   paper Q she should like which

$$\left\{\begin{array}{lllll}
\text{paper} < \text{Q} & \text{Q} < \text{she} & \text{she} < \text{should} & \text{should} < \text{like} & \text{like} < \text{which} \\
\text{paper} < \text{she} & \text{Q} < \text{should} & \text{she} < \text{like} & \text{should} < \text{which} & \\
\text{paper} < \text{should} & \text{Q} < \text{like} & \text{she} < \text{which} & & \\
\text{paper} < \text{like} & \text{Q} < \text{which} & & & \\
\text{paper} < \text{which} & & & &
\end{array}\right\}$$

This linearization satisfies Totality and Antisymmetry as well, but I don't think we should admit it into the class of possible linearizations. It violates Contiguity in a way that (4a) doesn't. It makes the phrase that is in two positions non-contiguous. In general, we want Contiguity to weed out random linearizations. We want it to preserve an image of the hierarchical structure of a sentence.

Contiguity presently enforces a relationship between substrings and phrases. It requires of a phrase A that dominates a phrase B that the words in B will be a substring of the words in A. It says that you can recover the dominance relations among the phrases of a sentence by looking at how strings are nested. I want to replace that idea with a very similar one. I want Contiguity to enforce a relationship between "paths" and substrings. A path is not just one phrase that dominates some set of words, but a set of phrases that dominate that set of words. What our new Contiguity will do is require that every word in a sentence be part of a path which is contiguous. A path is contiguous if all of the its members – that is all of the phrases in the set making up that path – are contiguous. The way the strings of a sentence are nested will now be an image of the paths in a sentence.

Multidominant phrase markers will give to some item more than one path. In those scenarios, Antisymmetry is violated if both paths are used in forming the linearization. Contiguity and Antisymmetry must together work to force all of the material in a phrase that has two paths to be linearized according to the same path. So the linearization procedure I will invoke is one that allows a sentence's structure to be factored into a set of paths. Those paths will be subject to Totality, Antisymmetry and Contiguity, and this will result in a set of possible linearizations. In multidominant phrase markers, there will be more than one choice of paths that will satisfy Totality, and this too will be a vector along which languages can vary.

So, here is the new definition of Contiguity, along with its attendant definitions. I've also changed Totality, so it makes reference to paths. The complexity of these definitions is a result of my attempt to remove "dominance" from them. I want the relations that a linearization looks at be entirely defined in terms that refer to paths.

(6) Let P and P' be phrases and X be a node that immediately dominates a word. $\Pi(P)$ is a set of <u>paths</u>, $p(X)$, formed from Xs in P.

  a. Let $i(\alpha)$ be a node that immediately dominates α. A path from α $(=p(\alpha))$ is the set of nodes, β, such that $\beta=i(\alpha)$ or $\beta=i(\beta)$. A path must include the root node.

  b. $d(P)$ is the set of Xs whose $p(X)$ includes P. $d(X)$ is X.

  c. $C(p)$ ("contiguous path") is a path in which, for every $P \in p$ that has a daughter $\alpha \notin p$, the linearization contains orderings such that:

    i. $\forall x, y : x < y$ for $x \in d(P) - d(\alpha)$ and $y \in d(\alpha)$, or

    ii. $\forall x, y : y < x$ for $x \in d(P) - d(\alpha)$ and $y \in d(\alpha)$.

  d. *Totality*
     For every X in P, $\Pi(P)$ must contain a $p(X)$.

  e. *Antisymmetry*
     A linearization cannot have both $X < Y$ and $Y < X$.

  f. *Contiguity*
     Every $p(X)$ in Π must be $C(p(X))$.

The language particular component – the information that languages contribute to the linearization – is in the choice between (6c-i) and (6c-ii), and how the $p(X)$ is chosen for words that have more than one path. This system is defined for multidominant trees that use the parallel merge derivation – not remerge.

Let's run through how (6) works. We'll start with a vanilla example with no multidominance.

(7)
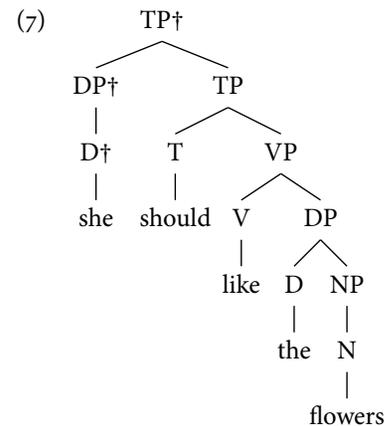
The largest $\Pi(7)$ is (8).

(8) a. $p(D\dagger) = \{DP\dagger, TP\dagger\}$

   b. $p(T) = \{TP, TP\dagger\}$

   c. $p(V) = \{VP, TP, TP\dagger\}$

   d. $p(D) = \{DP, VP, TP, TP\dagger\}$

   e. $p(N) = \{NP, DP, VP, TP, TP\dagger\}$

From (8) we can calculate the d relations. (I only look at the d of phrases, since the d of heads is trivial.)

(9) a. $d(\text{DP}\dagger) = \{\text{D}\dagger\}$

  b. $d(\text{TP}\dagger) = \{\text{D}\dagger, \text{T}, \text{V}, \text{D}, \text{N}\}$

  c. $d(\text{TP}) = \{\text{T}, \text{V}, \text{D}, \text{N}\}$

  d. $d(\text{VP}) = \{\text{V}, \text{D}, \text{N}\}$

  e. $d(\text{DP}) = \{\text{D}, \text{N}\}$

  f. $d(\text{NP}) = \{\text{N}\}$

And this allows us to evaluate Contiguity. Contiguity will be satisfied if the d of every element in every head's path in Π provides a contiguous mapping onto the linearization. For instance, suppose that the linearization of (7) is (the correct one in) (10).

(10) $\left\{ \begin{array}{llll} \text{D}\dagger < \text{T} & \text{T}<\text{V} & \text{V}<\text{D} & \text{D}<\text{N} \\ \text{D}\dagger < \text{V} & \text{T}<\text{D} & \text{V}<\text{N} & \\ \text{D}\dagger < \text{D} & \text{T}<\text{N} & & \\ \text{D}\dagger < \text{N} & & & \end{array} \right\}$

The linearization in (10) satisfies Contiguity (and Totality) with respect to the set of paths, Π, in (8). For instance, consider $p(\text{D}\dagger)$. The only member of $p(\text{D}\dagger)$ that has a sister is DP†: that sister is TP. $d(\text{DP}\dagger)$ is just D† (that is the only word containing node whose path contains DP†), and $d(\text{TP})$ is T, V and D (because these are all the word containing nodes whose paths contain TP). $C(p(\text{D}\dagger))$ is met by (10) because D† has the same ordering with respect to T, V and D (namely "<") in (10). Or consider V, whose path contains two terms with sisters: VP and TP. $d(\text{VP})$ (=V and D) are ordered the same way with everything in $d(\text{T})$ (=T) in (10) as well. And $d(\text{TP})$ (=T, V and D) are all ordered the same way with D† (the $d(\text{DP}\dagger)$) in (10) too. Apply the same calculations for $p(\text{T})$, $p(\text{D})$ and $p(\text{N})$, and you'll see that their contiguous paths are also in (10). Because (10) satisfies Antisymmetry, this is a viable linearization. It is the only one (given the English particular choices for (6c-i) and (6c-ii)), and this is how we get the right string.

Notice that $C(p(\text{V}))$ does not provide information about how V is ordered relative to its sister. That will come from the paths from heads inside that sister. Both $p(\text{N})$ and $p(\text{D})$ will contain DP, and so both $C(p(\text{D}))$ and $C(p(\text{N}))$ requires that V be ordered in the same way to all the heads in DP.

Let's now consider a tree that has multidominance in it. The first case we'll look at is an instance of Head Movement. Yesterday we saw two cases of Head Movement: one in Icelandic, which resulted in the pronunciation of the head in

only one position, and another in Vata, which resulted in the pronunciation of the head in two positions. Today we are going to look at Hebrew, which has in one language both kinds of Head Movement. As we saw yesterday, Hebrew has verb movement to T. It also has verb topicalization, rather like Vata. See Landau (2006). They combine to form a string in which the verb is pronounced in its topicalized position and in the position it gets inflected.

(11) liknot, hi kanta et ha-praxim.
  to-buy, she bought ACC the-flowers
  'As for buying, she bought the flowers.'

(Landau 2004, (8b): 37)

Note that the verb in topicalized position is pronounced as an infinitive, and the verb in T position is pronounced with tense morphology. The V position which holds the root is not pronounced. We can assume, as we did for Icelandic, that there is no vocabulary item in Hebrew that corresponds to a bare verb root.

The sentence in (11) will have the structure in (12).

(12)
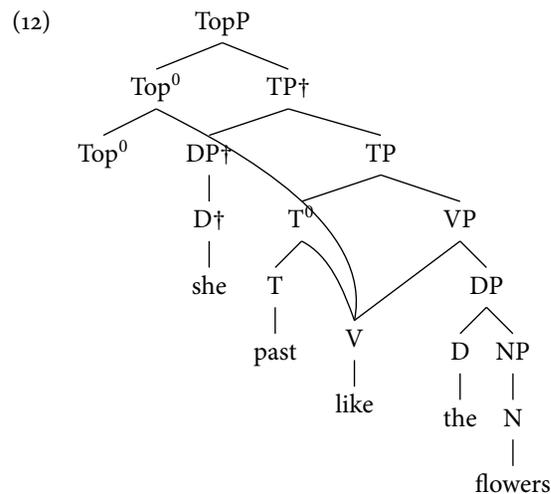
V doesn't dominate a word in Hebrew, but Top[0] and T[0] do. Top[0] is pronounced as the infinitival form of the verb and T[0] is pronounced as the inflected verb.

The largest $\Pi(12)$ is (13).

(13) a. $p(\text{Top}^0) = \{\text{TopP}\}$

b. $p(\text{D}\dagger) = \{\text{DP}\dagger, \text{TP}\dagger, \text{TopP}\}$

c. $p(\text{T}^0) = \{\text{TP}, \text{TP}\dagger, \text{TopP}\}$

d. $p(\text{D}) = \{\text{DP}, \text{VP}, \text{TP}, \text{TP}\dagger, \text{TopP}\}$

e. $p(\text{N}) = \{\text{NP}, \text{DP}, \text{VP}, \text{TP}, \text{TP}\dagger, \text{TopP}\}$

Note that there is no $p(\text{V})$, because V doesn't immediately dominate a word. From (13) we can calculate the d mappings:

(14) a. $d(\text{TP}\dagger) = \{\text{D}\dagger, \text{T}, \text{D}, \text{N}\}$

b. $d(\text{TP}) = \{\text{T}, \text{D}, \text{N}\}$

c. $d(\text{VP}) = \{\text{D}, \text{N}\}$

d. $d(\text{DP}) = \{\text{D}, \text{N}\}$

e. $d(\text{NP}) = \{\text{N}\}$

Notice, again, that V is not included.

Because V is nowhere included, you can see that this is a tree that will get linearized in the same way it would if V were absent. Because it's only V that is in a multidominant relationship, this means that the linearization process won't see the multidominance and things should run as if it weren't there. That's what we want, of course: that is Nunes's idea about why these cases end up speaking the multidominated term in more than one position. This case, then, collapses to (basically) the one in (7). The only linearizations that satisfy Contiguity are also ones that satisfy Antisymmetry, and of those the Hebrew particular choice is (11)

Now let's consider an example of wh-movement.

(15)
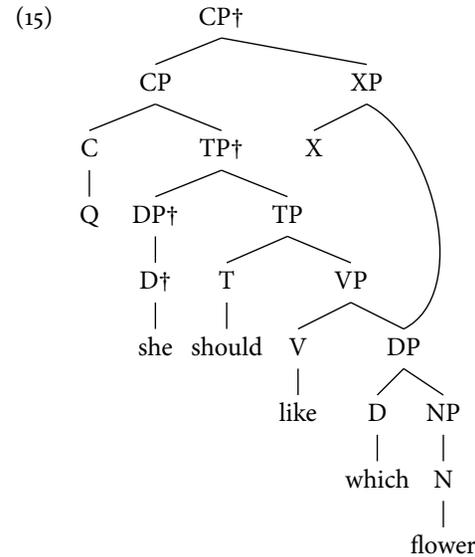
All the nodes dominating words in this sentence have just one path except for those in the DP *which flower*. The largest $\Pi(15)$ is (16).

(16) a. $p(\text{C}) = \{\text{CP}, \text{CP}\dagger\}$

b. $p(\text{D}\dagger) = \{\text{DP}\dagger, \text{TP}\dagger, \text{CP}, \text{CP}\dagger\}$

c. $p(\text{T}) = \{\text{TP}, \text{TP}\dagger \ \text{CP}, \text{CP}\dagger\}$

d. $p(\text{V}) = \{\text{VP}, \text{TP}, \text{TP}\dagger \ \text{CP}, \text{CP}\dagger\}$

e. $p(\text{D}) = \{\text{DP}, \text{VP}, \text{TP}, \text{TP}\dagger \ \text{CP}, \text{CP}\dagger\}$

f. $p(\text{D}) = \{\text{DP}, \text{XP}, \text{CP}\dagger\}$

g. $p(\text{N}) = \{\text{NP}, \text{DP}, \text{VP}, \text{TP}, \text{TP}\dagger \ \text{CP}, \text{CP}\dagger\}$

h. $p(\text{N}) = \{\text{NP}, \text{DP}, \text{XP}, \text{CP}\dagger\}$

i. $p(\text{X}) = \{\text{XP}, \text{CP}\dagger\}$

Let's consider first a language, like English, which chooses XP < CP.

For a language that requires XP < CP, (16) will not satisfy Contiguity. That's because neither the two paths for D, nor the two paths for N, can be simultaneously be made contiguous, if the other English specific choices are to be honored. Consider, for instance, the first path for D. It contains VP, whose sister is T. If this path is to be contiguous, T will have to precede every head whose path contains VP. That is, the linearization will have to contain:

(17)   { T<V, T<D, T<N }

But consider the second path for D. This includes XP; if it is contiguous, this will make the linearization contain:

(18)   { D<C, D<D†, D<T, D<V, D<D, D<N, N<C, N<D†, N<T, N<V, N<D, N<N, X<D, X<N }

(17) and (18) can't coexist in a linearization that satisfies Antisymmetry. The same problem will emerge if we consider the two paths for N. Contiguity and Antisymmetry can be jointly satisfied by a linearization from Π only if Π includes just one of the paths for D and one of the paths for N. Let's consider the possible Πs.

Both paths for D and N go through VP

Suppose we choose the first path for D and the first path for N in (16). That'll give us:

(19)   a. $p(C) = \{CP, CP†\}$
        b. $p(D†) = \{DP†, TP†, CP, CP†\}$
        c. $p(T) = \{TP, TP† CP, CP†\}$
        d. $p(V) = \{VP, TP, TP† CP, CP†\}$
        e. $p(D) = \{DP, VP, TP, TP† CP, CP†\}$
        f. $p(N) = \{NP, DP, VP, TP, TP† CP, CP†\}$
        g. $p(X) = \{XP, CP†\}$

This is the same Π we would generate from a tree in which *which flower* is not in XP. It can be contiguous, and when it is it will be Total and the corresponding linearization will meet Antisymmetry. English choices will give us an ordering equivalent to:

(20)   X Q she should like which flower.

It will be instructive to consider a couple of details. (19) produces the following *d*s:

(21)   a. $d(XP) = \{X\}$
        b. $d(CP) = \{C, D†, T, V, D, N\}$

For $C(p(X))$ to hold, {X<C, X<D†, X<T, X<V, X<D, X<N} must be in the linearization. Note that since $d(XP)$ doesn't include D or N, these don't get ordered so that they precede $d(CP)$.

(19) also gives us:

(22) $d(DP) = \{D, N\}$

Now consider how $d(D)$ will satisfies Contiguity. $C(p(D))$ requires that everything in $d(DP)$ be linearized in the same way with respect to $d(V)$. That's because contiguity says of VP, which is in $p(D)$, that everything in the daughter of VP that isn't in $p(D)$ be ordered in the same way with everything that is in the daughter of VP that is in $p(D)$. DP is in $p(D)$ and V isn't. So contiguity requires {V<D, V<N} to be in the linearization. (That particular order is an English specific choice.)

Both paths for D and N go through XP

Consider next:

(23)   a. $p(C) = \{CP, CP†\}$
        b. $p(D†) = \{DP†, TP†, CP, CP†\}$
        c. $p(T) = \{TP, TP† CP, CP†\}$
        d. $p(V) = \{VP, TP, TP† CP, CP†\}$
        e. $p(D) = \{DP, XP, CP†\}$
        f. $p(N) = \{NP, DP, XP, CP†\}$
        g. $p(X) = \{XP, CP†\}$

This Π is equivalent to the one that would be generated by a tree in which *which flower* isn't in VP. The $d(XP)$ and $d(CP)$ are:

(24)   a. $d(XP) = \{X, D, N\}$
        b. $d(CP) = \{C, D†, T, V\}$

Each of the ps in (23) are contiguous, together they satisfy Totality and Contiguity, and the resulting linearization will be Antisymmetric. With the English specific choices imposed, we'll get:

(25)   X which flower Q she should like

Again, we should consider how Contiguity applies to the phrase that has more than one sister: DP. Because both $C(p(D))$ and $C(p(N))$ have XP in them but not VP, C will require that everthing in $d(DP)$ be ordered in the same way with X, not V.

The path for D goes through XP and the path for N goes through VP

Consider finally:

(26) a. $p(C) = \{CP, CP\dagger\}$

b. $p(D\dagger) = \{DP\dagger, TP\dagger, CP, CP\dagger\}$

c. $p(T) = \{TP, TP\dagger\ CP, CP\dagger\}$

d. $p(V) = \{VP, TP, TP\dagger\ CP, CP\dagger\}$

e. $p(D) = \{DP, XP, CP\dagger\}$

f. $p(N) = \{NP, DP, VP, TP, TP\dagger\ CP, CP\dagger\}$

g. $p(X) = \{XP, CP\dagger\}$

This is the choice of paths necessary to get the kind of linearization in which the pieces of the moved phrase are linearized in different positions; for instance:

(27) X which Q she should like flower

We want to block this possibility.

The choice of paths in (26) won't be able to jointly satisfy Antisymmetry and Totality. To see this, consider the following ds.

(28) a. $d(XP) = \{X, D\}$

b. $d(CP) = \{C, D\dagger, T, V, N\}$

c. $d(DP) = \{D, N\}$

Consider how C can be satisfied by $p(N)$. C will require VP (which is in $p(N)$) to order V the same way with everything in $d(DP)$. This puts in the linearization:

(29) $\{ V{<}D, V{<}N \}$

Because CP† is in $p(D)$, $C(p(D))$ will be satisfied only if everything in $d(XP)$ precedes everything in $D(CP)$. So that will require the linearization to contain:

(30) $\{X{<}C, X{<}D\dagger, X{<}T, X{<}V, X{<}N, D{<}C, D{<}D\dagger, D{<}T, D{<}V, D{<}N\}$

A linearization that has both (29) and (30) violates Antisymmetry. Contiguity conflicts with Antisymmetry in this case. The same will be true of a Π in which $p(D)$ includes VP and $p(N)$ includes XP.

This is a linearization scheme, then, that captures Nunes's method of deriving Terseness – complete with its counterexamples in the domain of Head Movement – in a Multidominant theory of Movement.

We saw yesterday another way in which Terseness is violated. In these examples, the moved phrase is pronounced in one position and a matching pronoun is pronounced in the other.

(31) təlmiiz-a₁ lkəsleen ma baddna nχabbi [wala mʕallme]₁ ʔənno
student-her₁ the-bad NEG want.1P tell.1P [no teacher]₁ that
huwwe zaʕbar   b-l-faḥṣ
he   cheated.3SM in-the-exam
'her bad student, we don't want to tell any teacher that he cheated on the exam.'

(Aoun, Choueiri, and Hornstein 2001, (25b): 381)

These kinds of examples show up in clefts and relative clauses. I will sketch a way of capturing them that assumes that they only happen in cases that have the syntax of relative clauses.

The first thing we need to do is understand what a pronoun is. Pronouns share certain characteristics with definite descriptions. For instance, their context of use is, many times, similar. The sentences in (32) are odd in the context of this class for the same reason.

(32) a. The woman is bored.

b. She is bored.

I'm going to assume that there are bored women in my class by the time we get to this point in the handout, so these sentences should both be true. But they're not because the subjects do not refer properly, and this is because there is more than one woman in my class (unless all but one of them have left by now). This is modeled by making both *the woman* and *it* trigger the presupposition that there is one unique female in the situation being described by the sentence. This presupposition comes with *the*, so we'll build it into the denotation of *the*.

(33) ⟦the⟧ = $\lambda P\ \iota x\ P(x) = 1$, defined only when there is a unique
x such that $P(x) = 1$

The uniqueness presupposition is going to have to be restricted by context or, more likely, by the situation being described. This gives us, then, something like (34) for (32a).

(34)

$$\text{bored}(\iota x \text{ woman}(x) = 1)=1$$

```
                    bored(ιx woman(x) = 1)=1
                              IP
                 ┌────────────┴────────────┐
          ιx woman(x) = 1            λx bored(x)=1
               DP                         TP
            ┌──┴──┐                    ┌───┴───┐
         the woman                     T      AP
                                       │
                                       is   bored
```
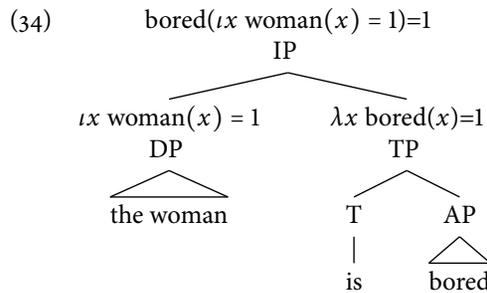
⟦DP⟧ and ⟦IP⟧ are defined just in case there is a unique x in the relevant situation which is a woman.

To capture the fact that pronouns are similar to definite descriptions in these ways, we can give them a similar composition. Imagine, for instance, that a pronoun is just the way English speaks definite description whose NP part are the relevant person, gender and number features. I'll assume that "third person" is the absence of first or second person. So the pronoun *her* on this view might look like (35). (See Elbourne 2005.)

(35)

```
           DP
        ┌──┴──┐
        D     NP
        │    ┌─┴──┐
       the  fem  sing
```
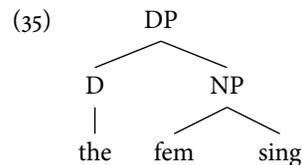
Suppose that the denotations for *fem* and *sing* are sketched by (36). (The actual denotation for number is more complex. See Sauerland (2008) and Kratzer (2009).)

(36) a. ⟦fem⟧ = $\lambda x.x$ is female
     b. ⟦sing⟧ = $\lambda x.x$ is an atom

This gives us (37) for (32b).

(32b) She is bored.

(37)

$$\text{bored}(\iota x \text{ female}(x) = 1 \wedge \text{sing}(x) = 1)=1$$

```
               bored(ιx female(x) = 1 ∧ sing(x) = 1)=1
                               IP
              ┌────────────────┴────────────────┐
     ιx female(x) = 1 ∧ sing(x) = 1        λx bored(x)=1
               DP                               TP
            ┌──┴──┐                          ┌───┴───┐
         the fem sing                         T      AP
                                              │
                                             is    bored
```
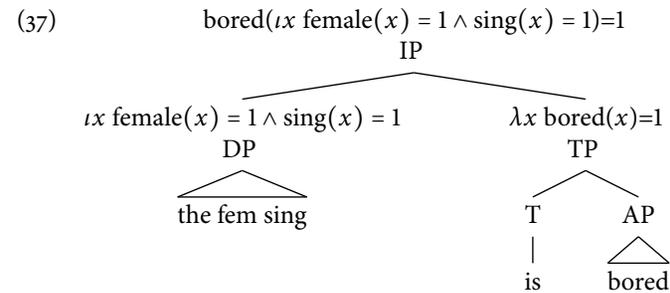
⟦DP⟧ and ⟦IP⟧ are defined just in case there is a unique x in the relevant situation which is female and singular.

Bottom line: a pronoun can be the way a definite description is pronounced when *the* combines with ϕ features.

Now let's consider the structure of a relative clause.

(38) a picture of herself that Mary is selling

Notice that we get semantic displacement here, just as in the Arabic example. The reflexive pronoun *herself* is being semantically interpreted in the position it moved from, not the position it is spoken. This shows us that the NP *picture of herself* has moved from the object position of *selling*. Suppose that is all that has moved, the NP that makes within the object of *selling* not the DP that makes the object of *selling*. We posit a structure like (39).[1]

---

[1] See Schachter (1973), Vergnaud (1974), Kayne (1994) and Bianchi (2000).

(39)

(40)
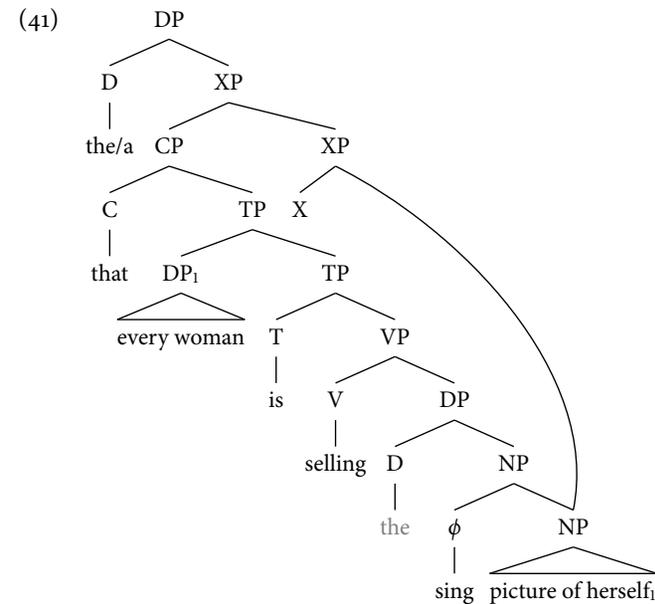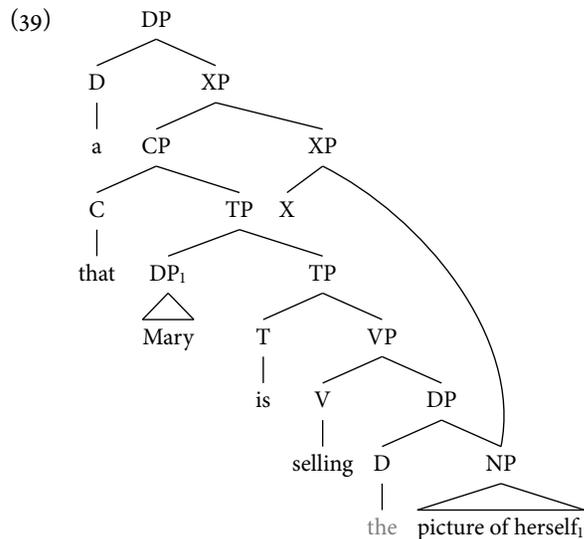
Our linearization scheme will force the NP in (39) to be linearized in just one of its positions, and this is what happens in English relative clauses. If this is the right structure for relatives, it means that English must have a silent definite determiner which is used in the lower position, because nothing is spoken in that position in English.

Consider now a language in which ϕ-features can be part of a DP that has an overt NP. This would be a language, then, which allows DPs like (40).
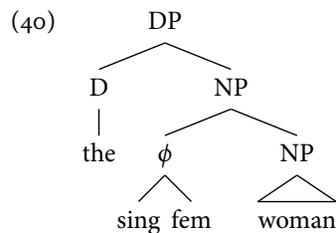
We can use predicate conjunction to put the parts of the determiner's complement together. If DPs like this are available – and if DPs can Agree with ϕ features, then this seems plausible – we should expect languages that build relative clauses as in (41).

(41)

Recall that pronouns are just the way that definite descriptions are pronounced when their NP part just has ϕ features. The NP that is the complement to the lower *the* in (41) consists only of ϕ features – at least in terms of its string information that is all that is inside it – and perhaps this allows it to be pronounced as a pronoun.

9

References

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

Bianchi, Valentina. 2000. The raising analysis of relative clauses: a reply to Borsley. *Linguistic Inquiry* 31:123–140.

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

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

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.

Landau, Idan. 2004. Chain resolution in Hebrew V(P)-fronting. In *Proceedings of the North East Linguistic Society 34*, ed. Keir Moulton and Matthew Wolf, volume 2, 357–372. University of Massachusetts at Amherst: Graduate Linguistic Student Association.

Landau, Idan. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9:32–66.

Sauerland, Uli. 2008. On the semantic markedness of Phi-features. In *Phi theory*, ed. Daniel Harbour, David Adger, and Susana Béjar, 57–82. Oxford: Oxford University Press.

Schachter, Paul. 1973. Focus and relativization. *Language* 49:19–46.

Vergnaud, Jean-Roger. 1974. French relative clauses. Doctoral Dissertation, Massachusetts Institute of Technology.

Kyle Johnson

University College London

University of Massachusetts at Amherst

22 June 2016

Today we'll start to take a closer look at how movement structures are semantically interpreted. Our focus will be Wh Movement.

Simple wh-questions have the shape in (1), on the multidominant model of movement that we're exploring.

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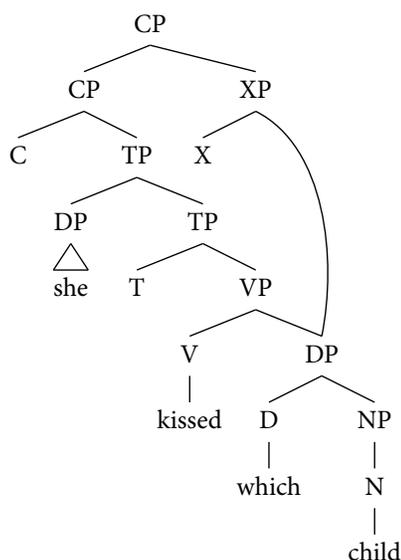
A standard, simple, view of the meaning of questions is that they denote a set of propositions, each proposition offering a kind of answer in those cases where the question is answer-seeking. This is the view introduced by Hamblin and modified by Kartunnen. One way of representing a set is with the λ-operator,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

(2) $\lambda x P(x)$ = that function which, when applied to a, gives $P(a)$ is true (or, $P(a)=1$).

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 & p = she kissed x

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

The central problem a multidominant definition of movement 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 works fine for moved heads, but it 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$_1$ does the thing no one$_1$ heard indicate I should bring?

c.   * Which picture of himself$_1$ indicates that no one$_1$ should bring it?

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)

```
                    CP
          ┌─────────┴──────────┐
        DP₁                     CP
     ┌───┴───┐             ┌────┴────┐
     D       NP            C         TP
     │       △             │      ┌──┴──┐
   which  picture of     shouldn't DP₂   TP
          himself              △      ┌──┴──┐
                             anyone   T     VP
                                         ┌──┴──┐
                                        VP     PP
                                     ┌──┴──┐    △
                                     V    t₁  on his website
                                     │
                                    put
```

$\llbracket t_1 \rrbracket \approx$ picture-of-himself$_2(x_1)$
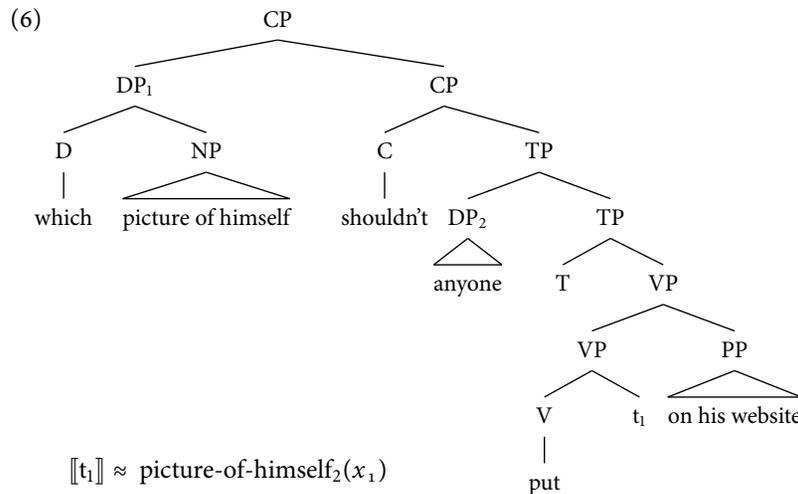
If we adopt a Hamblin/Kartunnen style analysis of questions, and make the NP part of the moved phrase interpreted in its unmoved position, then for (6) we might get an interpretation along the lines of (7).

(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 the *x*, such that $f(x)$, & 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 in (8) the donkeys are made to 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$_1$ the $x [f(y_1)](x)$
          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

Some questions describe functions that can only depend on the values given to nearby quantifiers.

(11) Which whole number does every whole number precede?

The only function in the actual worlds that fit this question are ones like *its successor*, which vary as according to the value give *every whole number*.

We get from this model an account of why the trace can only get a value that varies with the values given to expressions which c-command it.

(12) Which picture did you show the guy every girl$_1$ knows?
      *$\lambda p \exists f$ p = you showed the guy every girl$_1$ knows the $x [f(y_1)](x)$
      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 DP. 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:

(13)   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 functional 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 the view we saw yesterday that pronouns are the way definite descriptions are pronounced when they contain just ϕ features. Let's look at such a way that builds on Heim (2013), Fox (1999), Elbourne (2005), Chierchia (1995) and Rullmann and Beck (1998).

The first innovation will be to let a definite description come with an index. I will adopt the view in Fox (2003) where indices are treated as kinds of adjectives. (We will see evidence for this view in the lectures on QR.) An index is a variable that gets interpreted either by being bound or by way of a function that assigns a free variable a referent in the discourse. The function that does this is sometimes called "an assignment function," and it is standardly represented with "g." Fox's idea, then, is that an index is a predicate that holds of those things that the assignment function or binder make the index refer to. So just as the adjective "red," for instance, is a predicate that holds of all those things that are red, the index "n" is a predicate that holds of all those things that g or a binder assigns "n" to.

(14)   a. $[\![red]\!] = \lambda x\ red(x){=}1$

       b. $[\![n]\!]^g = \lambda x\ x = g(n)$

A definite description that has an index in it will look like (15).

---

(15) $\iota x\ x = g(n) \wedge flower(x){=}1$
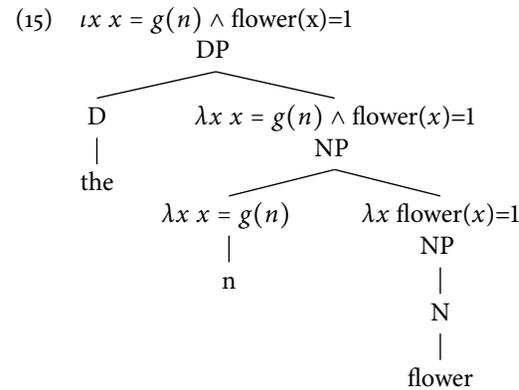
(16) $[\![the]\!] = \lambda P\ \iota x\ P(x) = 1$, defined only if there is a unique x such that $P(x) = 1$

(17) $[\![flower]\!] = \lambda x\ flower(x){=}1$

The denotation for (15) is only defined if there is a unique x which is both g(n) and a flower. This expresses the presupposition that we saw definite descriptions invoke. I will ignore that part of the meaning for the rest of this lecture.

An index can be a function, along with its arguments; when the index is a function, I will represent it with "f." This is what we have in (13b).[1]

(13b)   Every man who owns a donkey kisses the donkey.

$\iota x\ x = g(f)(g(z)) \wedge donkey(x){=}1$
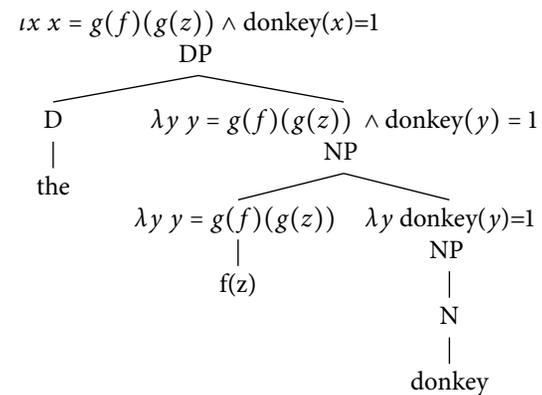
Notice that there are two variables here: f and z. f is a free variable, and will get its value from g. In this context, g assigns to f a value paraphrased by "owned by"

z, by contrast, will be bound by *every man*, and that's how we get the donkeys this definite description refers to to vary with the men.

The case in (13a) is parallel. The sole difference being that the NP contains just ϕ features in (13a).

(13a) Every man who owns a donkey kisses it.

$$\iota x\; x = g(f)(g(z)) \wedge \text{sing}(x)=1$$
DP

D $\lambda y\; y = g(f)(g(z)) \wedge \text{sing}(y) = 1$
| NP
the

$\lambda y\; y = g(f)(g(z))$ $\lambda y\; \text{sing}(y)=1$
| NP
f(z) |
       N
       |
       it
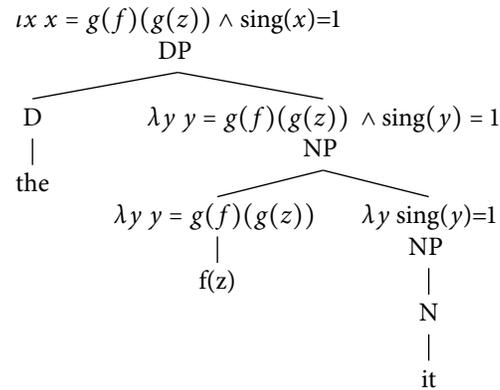
(Note that I assume here that neuter 3rd person are unmarked features, so *it* is merely singular.)

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 bind f. 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 (18).

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

$$\exists f\; \neg\exists y\; \text{linguist}(y)=1 \wedge y \text{ should forget}$$
$$\iota x\; x = f(y) \wedge \text{book\_about\_}y(x)=1$$
CP

$\lambda P \exists f\, P(f) = 1$ $\lambda 1\; \neg\exists y\; \text{linguist}(y)=1 \wedge y \text{ should forget}$
DP$_1$ $\iota x\; x = g(1)(y) \wedge \text{book\_about\_}y(x)=1$
        CP

$\lambda P \exists f\, P(f) = 1$ NP
D
| book
which   about her

C $\neg\exists y\; \text{linguist}(y)=1 \wedge y \text{ should forget}$
   $\iota x\; x = g(1)(y) \wedge \text{book\_about\_}y(x)=1$
     TP

DP$_2$ TP

no linguist   T   VP
      |
    should   V $\iota x\; x = g(1)(g(2)) \wedge$
         $\text{book\_about\_her}_2(x)=1$
      | DP
    forget

D $\lambda x\; x = g(1)(g(2)) \wedge$
| $\text{book\_about\_her}_2(x)=1$
which   NP

1(2)   NP
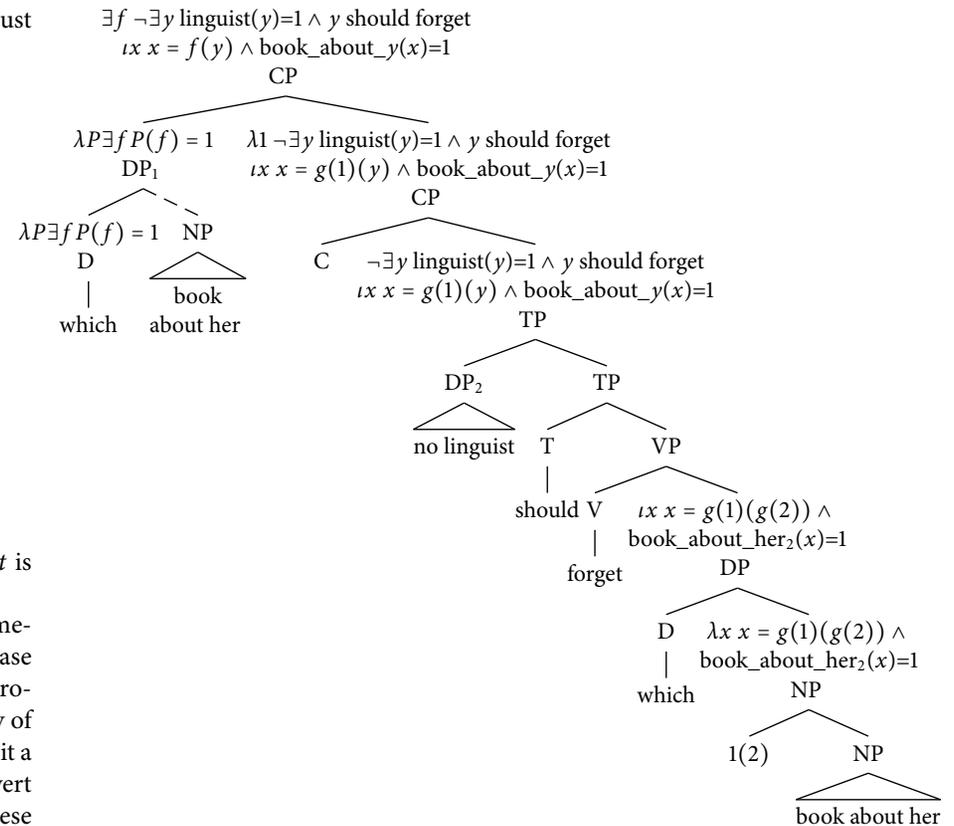
book about her

Note that the NP *book about her* in the higher copy is not semantically interpreted there. It can't be, since the pronoun it contains must be interpreted within the scope of *no linguist* if it is to be interpreted as a variable. I've indicated that this NP isn't semantically interpreted in this position with the dashed line.

What's missing from (18) 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. Let's hold off on completing this picture for a little while.

---

[2] 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).

The standard account of the semantics of movement does precisely what we see in (18). That account is due to Danny Fox,[2] 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 (18). He calls that rule "Trace Conversion," and the way he formulates it is in (19), which is slightly different from how I've built the meaning of the lower copy, but close enough to be roughly equivalent.

(19)   TRACE CONVERSION

In ϕ' interpret ϕ as a function that maps,
x, to the meaning of $\phi[x/n]$.

$$\begin{array}{c} \phi' \\ \diagup\diagdown \\ DP_n \quad \phi \\ \diagup\diagdown \\ \ldots DP_n \ldots \end{array}$$
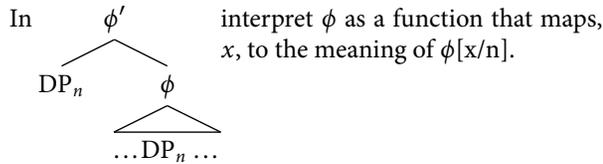
$\phi[x/n]$ is the result of replacing the head of every constituent bearing the index n in ϕ with the head the_x, whose interpretation, $[\![the_x]\!]$, is: $\lambda P : P(x).x$.

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

What this rule says is that the two copies of *which* can get different interpretations in the different places they reside in. We can't do this on a multidominant picture of movement because there aren't two *which*'s: there's just the one in to different positions. Even if we were forced to a different solution, however, I think we should be worried about the power of a rule like Trace Conversion.

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. That Q morpheme is going to be the X in our parses for questions. As a first approximation, I suggest something like (20).

(20)

$\exists f \neg \exists y \text{linguist}(y) = 1 \wedge y\text{should forget}$
$\iota x \, x = f(y) \wedge \text{book\_about\_}y(x)=1$
CP

$\lambda P \exists f.P(f) = 1$
QP$_2$

$\lambda 2.\neg \exists y \text{ linguist}(y)=1 \wedge y \text{ should forget}$
$\iota x \, x = g(2)(y) \wedge \text{book\_about\_}y(x)=1$
CP

$\lambda P \exists f.P(f)$
Q

C   TP

should   DP$_1$ TP

no linguist   T   VP

V

$\iota x \, x = g(2)(g(1)) \wedge$
$\text{book\_about\_her}_1(x)=1$
DP

forget

D   NP

2(1)   NP

book
about her$_1$
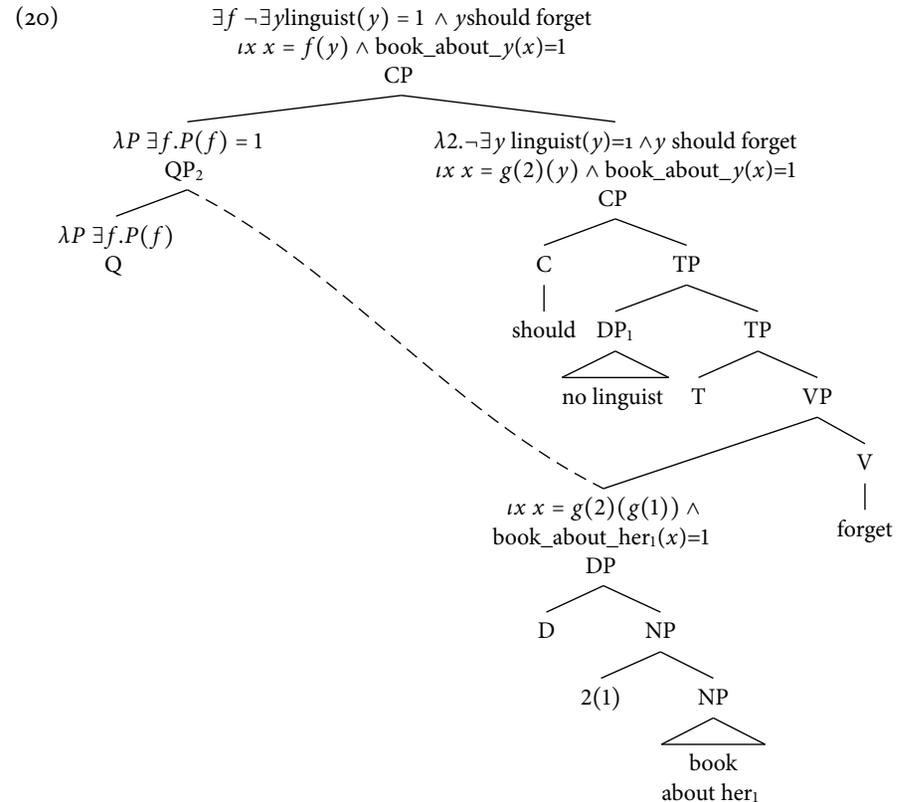
On this view, technically what has moved is just the DP portion that is interpreted as a variable. This denotation it supplies to the object position of *forget*. That DP has merged with the higher Q, which is the binder of the index within the DP in object position. It has merged with that Q, but its 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 (20) 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. We need this because it is possible for *which* to show up on unmoved DPs, as in (21).

(21)   Which student asked which question?

But I want the form this D has to reflect the fact that there is the question morpheme, Q, in the sentence. We need to ensure that when Q is present, there is a *which*. So I suggest that we let *which* be the exponent of the Q morpheme. When the exponent of some morpheme is expressed at the position that another morpheme has, we say that "Agreement" holds between them. So I will assume that Agree holds between Q and the D containing *which*.

(22)
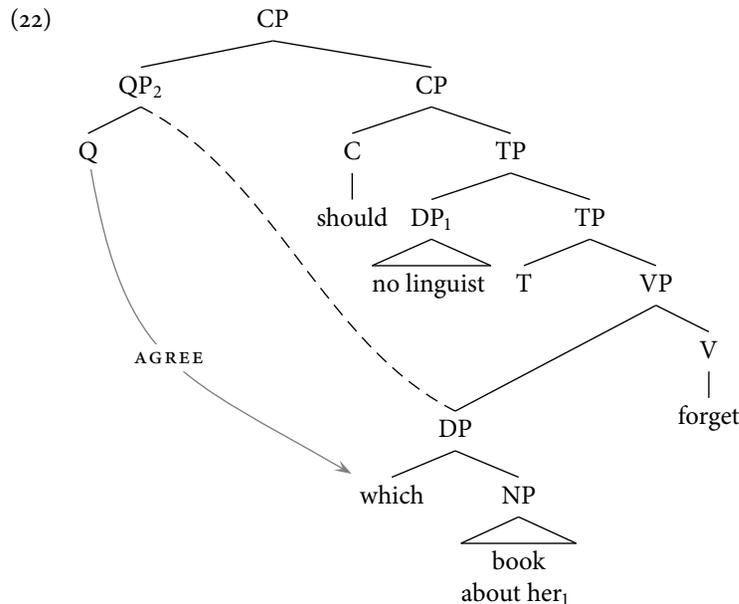
The *which*-phrase in (22) has paths. The grammar of English requires that the linearization choose the path that includes QP, and this is why the wh-phrase is linearized according to its status as part of the Specifier of CP.

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 (22)?

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.

(23) (Kimi-wa) dono-gakusei-ga   nattoo-o tabe-tagatte-iru-to
       (you-top) which-student-nom natto-acc eat-desirous-be-C
       omoimasu-ka?
       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.[3] 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 (24) is responsible for destroying the relationship between *nuku* and *ni*, thereby causing this sentence to be ill-formed.

(24) * 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 (24) is that the focus sensitive operator *man*

---

[3] See Adger and Ramchand (2005), Kratzer and Shimoyama (2002), Hagstrom (1998, 2000), Kishimoto (2005), and Cable (2010).

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. In something like (25), then, *Sally* has a normal semantic value that allows it to refer to Sally, and the focus semantic value in (25b).

(25)   She only visited Sally$_F$.

    a. $[\![$Sally$]\!]$ = the individual named Sally

    b. $[\![$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.

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

    $[\![$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 function in the wh-word is an alternative generator, rather than a variable. We'll get representations like that indicated in (27).

(27)
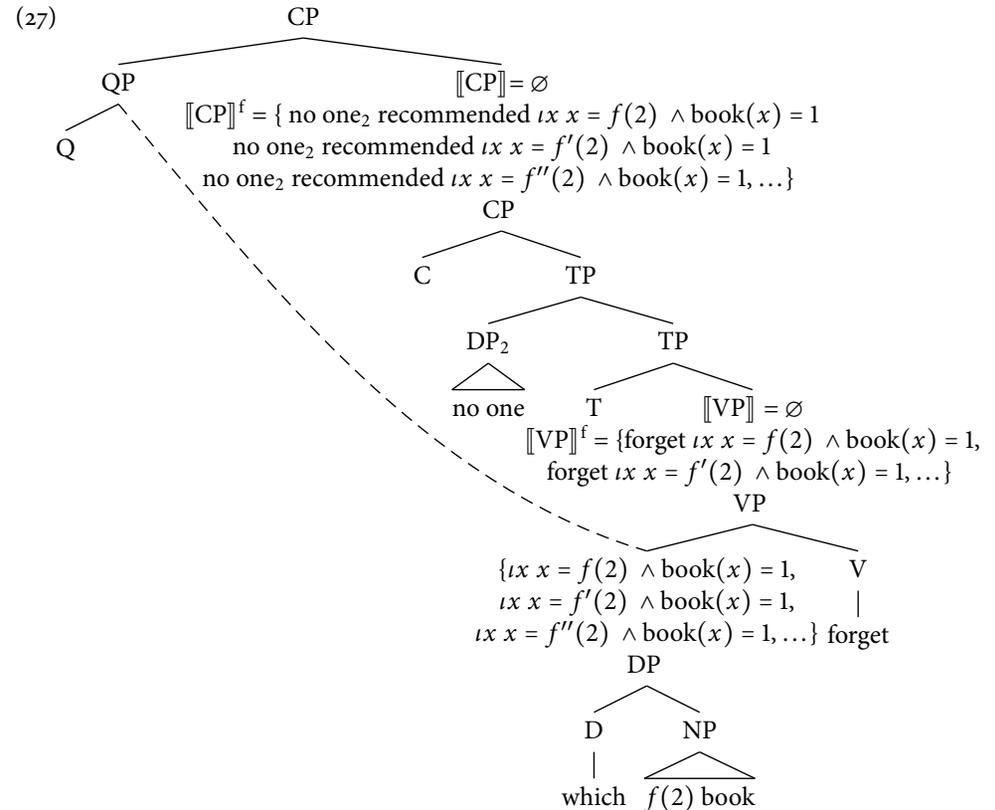
Now, Beck's proposal is that the Q morpheme, among perhaps other things, converts the focus semantic value of its sister into a normal semantic value. This gives us (28).

(28) $[\![CP]\!] = \{$ no one$_2$ recommended $\iota x\ x = f(2)\ \wedge \text{book}(x) = 1$
      no one$_2$ recommended $\iota x\ x = f'(2)\ \wedge \text{book}(x) = 1$
      no one$_2$ recommended $\iota x\ x = f''(2)\ \wedge \text{book}(x) = 1, \ldots\}$
                    $[\![CP]\!]^f = \varnothing$
                          CP

          QP $[\![CP]\!] = \varnothing$
                    $[\![CP]\!]^f = \{$ no one$_2$ recommended $\iota x\ x = f(2)\ \wedge \text{book}(x) = 1$
        Q   no one$_2$ recommended $\iota x\ x = f'(2)\ \wedge \text{book}(x) = 1$
                    no one$_2$ recommended $\iota x\ x = f''(2)\ \wedge \text{book}(x) = 1, \ldots\}$
                                  CP

                          C   TP

                            DP$_2$ TP

                          no one   T $[\![VP]\!] = \varnothing$
                                              $[\![VP]\!]^f = \{$forget $\iota x\ x = f(2)\ \wedge \text{book}(x) = 1,$
                                                forget $\iota x\ x = f'(2)\ \wedge \text{book}(x) = 1, \ldots\}$
                                                        VP

                          $\{\iota x\ x = f(2)\ \wedge \text{book}(x) = 1,$ V
                          $\iota x\ x = f'(2)\ \wedge \text{book}(x) = 1,$ |
                          $\iota x\ x = f''(2)\ \wedge \text{book}(x) = 1, \ldots\}$ forget
                                        DP

                                  D   NP
                                  |
                                which $f(2)$ book

(29)   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. We have to learn about QR before I can address that problem, so we have to leave this question unanswered for a while.
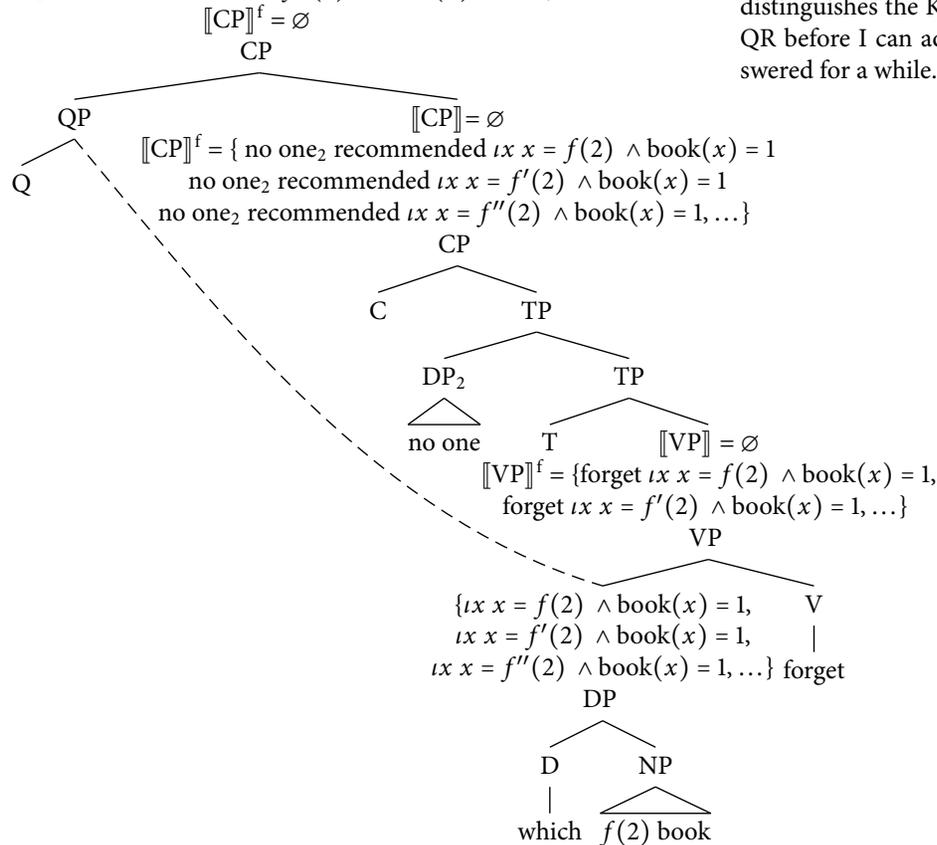
A set of alternatives should be thought of as a long disjunction: each member in the set is a separate disjunct. This means that the set of alternatives Beck assigns to a question is semantically equivalent to an existential quantification over functions. We have achieved the denotation for questions that we wanted.

What goes wrong in (24), 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 (24) is perfectly grammatical.

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 and Hadas Kotek. 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. (30) illustrates.

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

                              (Cable 2007, (212)–(213): 155-6)

In (30), 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.

(31)

aadóo yaagú   sá   ysiteen
whose boat   you saw t_1

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 (32).

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

He argues 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. (33) illustrates.

(33) 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.

(34) 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.

This how a multidominant model of wh-movement can be formulated. It harbors two problems, though. One is the one already acknowledged that we should expect English wh-questions to show Beck effects. Because the wh-phrase is semantically interpreted in just its lower position, even when it has overtly moved,

just like Korean, its lower position should be separated from its higher position by a focus marked thing. That is not the case:

(35)		Who did only Minsu visit?

A second, related, problem is that we should expect everything within a moved wh-phrase to obligatorily be interpreted in its lower. This means we should get the same disjoint reference effect in both the examples of (36).

(36)		a.		* She₁ hates a book that Mary₁ brought.
		b.		Which book that Mary₁ brought does she₁ hate.

But that isn't the case. Both of these examples suggest that a moved wh-phrase is not interpreted wholly in its lower position.

The solution to these problems that I will explore is suggested in Kotek (2014). She suggests that the derivations that form wh-questions in English can involve a QR step. That is, she suggests that the wh-phrase can first QR past a focus marked term, or a disjoint reference trigger, and from that QR'd position wh-move. To see how that solution might work, we will first have to learn how to do QR in a multidominant theory.

References

Adger, David, and Gillian Ramchand. 2005. Merge and move: *Wh*-dependencies revisited. *Linguistic Inquiry* 36:161–193.

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. 2010. *The grammar of Q: Q-particles, Wh-Movement and Pied-Piping*. Oxford University Press.

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.

Heim, Irene. 2013. Functional readings without type-shifted noun phrases. Unpublished paper, MIT.

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

Kotek, Hadas. 2014. Composing questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *Proceedings of the Tokyo conference on psycholinguistics*, ed. Yukio Otsu, volume 3, 1–25. Tokyo: Hituzi Syobo.

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.

Kyle Johnson
University of Massachusetts at Amherst

University College London
23 June 2016

Quantifier Raising is the rule that allows the quantificational determiner *every* in (1) to be interpreted in a position higher than where it is pronounced. This is required to give us the interpretations where the indefinite's value varies as a function of the universal quantifier.

(1)   A guard stood in front of every bank.

In it original formulation – which I think goes back to Chomsky (1977), but is usually credited to May (1977) – movement is allowed to create the representation in (2) "covertly" from (1).

(2)
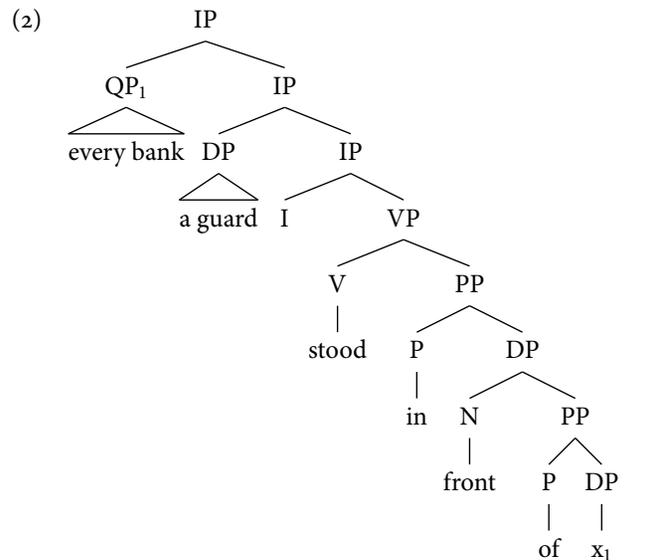
A standard semantics for *every* gives us the interpretation in (4).

(3) $[\![every]\!] = \lambda Q.\lambda P.\forall y\, Q(y) \rightarrow P(y)$

(4) $\forall y.y$ is a bank $\rightarrow$ a guard stood in front of y
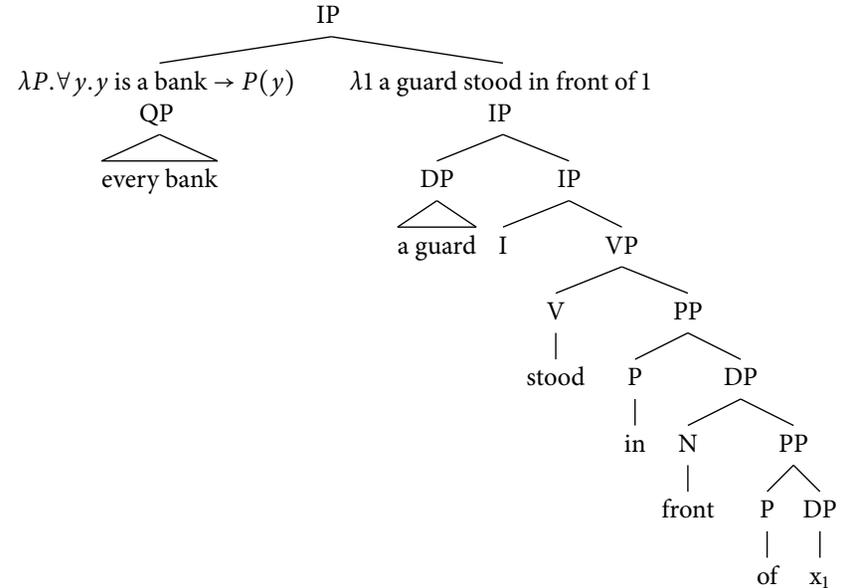
The proposal in Fox (2002) is to use the copy theory of movement and do "Trace Conversion" on the lower copy.

(5) $\forall y. y$ is a bank $\rightarrow$ a guard stands
in front of $\iota x, x = y$ such that x is a bank
IP

$\lambda Q. \forall y. y$ is a bank $\rightarrow Q(y)$ $\lambda n.$ a guard stands in front
QP   of $\iota x, x = n$ such that x is a bank
IP*

every bank

DP   IP

a guard   I   VP

V   PP

stand   P   DP

in   N   PP

front   P $\iota x. x = g(n)$
                 $\wedge\ x$ is a bank
             of   DP

every$_n$ bank
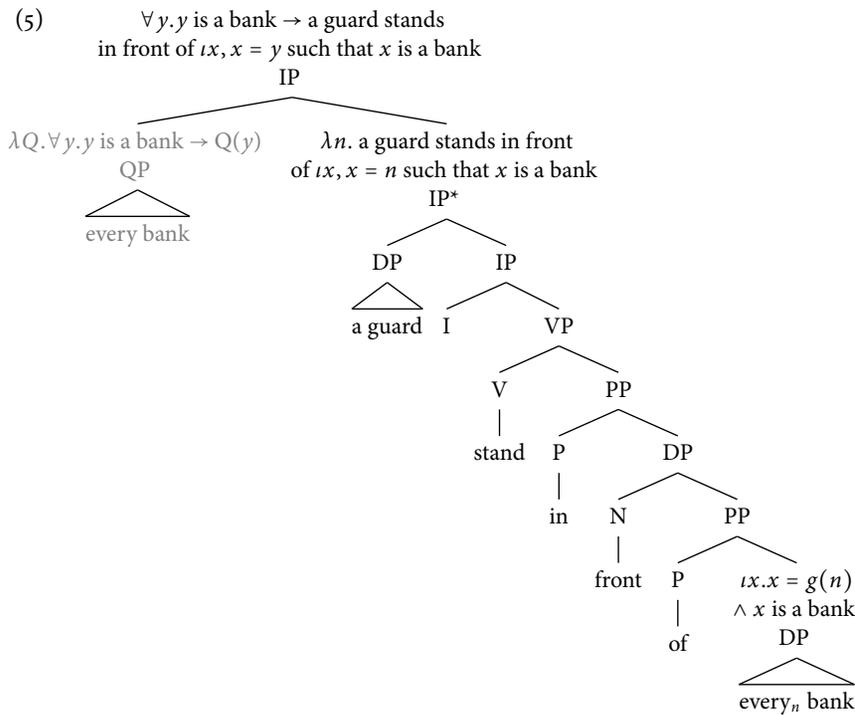
(6) Trace Conversion
Interpret $[_{DP}$ D NP$]$ as $[_{DP}$ the n NP$]$.

Trace Conversion causes the lower copy to be interpreted as a definite description with a variable.

The proposal in Johnson (2012) is to translate this into a multidominant structure without changing the semantics. That's done by sharing the NP part with two separate heads: a definite determiner downstairs and a quantifier upstairs.

(7) $\forall y. y$ is a bank $\rightarrow$ a guard stands
in front of $\iota x, x = y$ such that x is a bank
IP

$\lambda n.$ a guard stands in front $\lambda Q. \forall y. y$ is a bank $\rightarrow Q(y)$
of $\iota x, x = n$ such that x is a bank   QP
IP*

Q

DP   IP $\forall$

a guard   I   VP

V   PP

stand   P   DP

in   N   PP

front   P $\iota x. x = g(n)$
                $\wedge\ x$ is a bank
            of   DP

D   NP

the   n   NP

bank
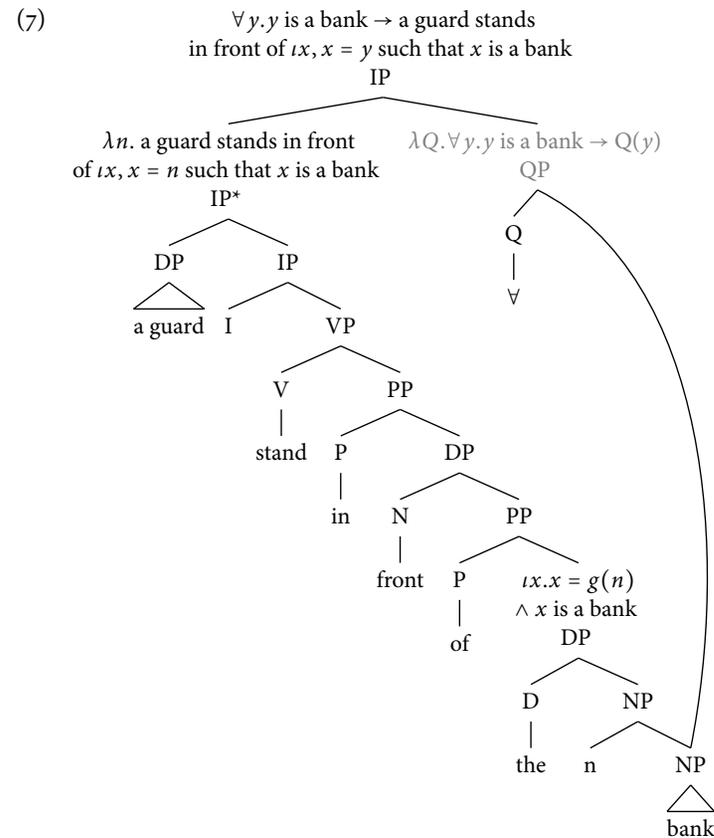
the gets pronounced as *every* when it shares an NP with $\forall$

Note that on all these accounts, the NP part of the moved phrase is interpreted in both of its positions. Unlike what we saw for questions, the NP part of the moved phrase must be interpreted in its higher position if we are to restrict the quantification. We can see that this is a desirable outcome by considering the fact that (8) does not allow an inverse scope reading.

(8) A guard stands in front of every picture of himself.

Because *picture of himself* must be interpreted together with the universal quantifier, it cannot be interpreted lower than where the universal quantifier is interpreted. And because *himself* must be interpreted within the scope of *a guard*, the result is that the universal quantifier must too. We should keep a semantics for the

universal quantifier – the standard one will do – that has this effect. Rather than Trace Conversion, this model separates the two meanings of QR: the quantificational part (=Q) from the variable part (=D). Q, then, has the denotation that the other models ascribe to *every*. *every* is the way that the definite determiner gets pronounced when its expressing the presence of a universal quantifier somewhere else.

But unlike questions, we cannot rely on Agreement to characterize when the definite determiner is in the right relation to Q to be pronounced as *every*. Because Q does not c-command *every*, we cannot use Agreement. So, that's an open problem we will have to solve:

(9)   How to get D and $\forall$ to jointly spell out as *every*

QR is thought to be responsible for resolving ACD.

(10)   Joan accused everyone that Sherlock did △.
       △ = accused x

That's because there seems to be a match between the scope of the quantifier and the size of the VP that can be elided. (see Williams (1977) and Sag (1976).)

(11)   Joan wants to eat everything Sherlock does △.
       a. *every* » *want*
           i.   Everything that Sherlock wants to eat is something that is desirable for Joan to eat.
           ii.   ? Everything that Sherlock eats is something that is desirable for Joan to eat.
       b. *want* » *every*
           i.   Eating everything that Sherlock eats is what is desirable for Joan.
           ii.   * Eating everything that Sherlock wants to eats is what is desirable for Joan.

On the copy theory and multidominant theories of movement, this isn't trivial. If the NP part of a QR'd phrase is semantically in both the lower and higher position, then we might expect a relative clause to be in both those positions too. That would not resolve ACD. Fox (2002) argues that Late Merge is responsible for resolving the ellipsis.

(12) $\forall y.\, [y$ is a human that Sherlock accused$] \rightarrow$ Joan accused $\iota x.x = y \wedge\ x$ is a human
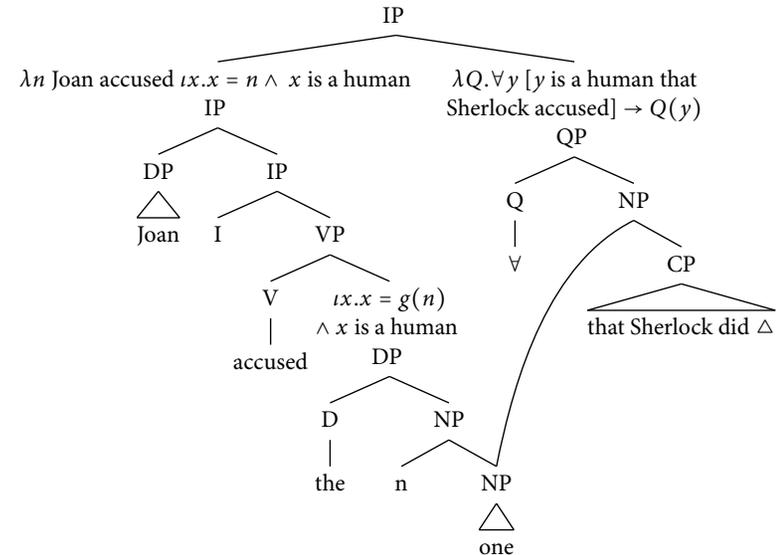
Fox and Nissenbaum (1999) argues that this is the same syntax we see for "extraposition from NP," as in (13).

(13)   John accused everyone yesterday that Sherlock did △.

(14) $\forall y.\ [y$ is a human that Sherlock accused$] \to$ Joan accused $\iota x.x = y \wedge\ x$ is a human

IP

λn Joan accused $\iota x.x = n \wedge\ x$ is a human $\lambda Q.\forall y\ [y$ is a human that

IP   Sherlock accused$] \to Q(y)$

QP

DP   IP   Q   NP

Joan   I   VP $\forall$ CP

VP   yesterday   that Sherlock did △

V $\iota x.x = g(n)$

    $\wedge\ x$ is a human

accused   DP

D   NP

the   n   NP

one

Overfelt (2015b) points out that this analysis of clausal extraposition from NP solves the mystery of why relative clauses do not seem to be movable otherwise:

(15)   * That Sherlock accused, John accused everyone.

And, as Overfelt (2015a) demonstrates, this also accounts for why an extraposed relative behaves semantically as if it is part of the restrictor for the quantifier with respect to NPI licensing. The NPI *ever* is licensed when it is in the restrictor of the quantifier *every*, but not when it is in the nuclear scope of *every*.

(16)   a.   We met every biker who has ever ridden on these trails.
       b.   * We met every biker while he has ever ridden on these trails.

This is a fact about *every*, and not a fact about relative clauses more generally, as we can see from the contrast in (17).

(17)   a.   We met every biker who has ever ridden on these trails.
       b.   * We met some biker who has ever ridden on these trails.

This is preserved under extraposition:

(18)   a.   We met every biker yesterday who has ever ridden on these trails.
       b.   * We met some biker yesterday who has ever ridden on these trails.

Moreover, we can find examples where the relative clause cannot be interpreted as if it hadn't moved out of the VP, and yet even here it behaves like it is part of the restrictor of the universal quantifier.

(19)   a.   Kim bought every book yesterday that she'd ever been told to △.
       b.   * Kim bought some books yesterday that she'd ever been told to △.

(Overfelt 2015a, (32): 36)

There are other subtle effects of Extraposition from NP that follow from our syntax. One that Fox (2015) discusses is a contrast in (20) first noticed, to my knowledge, by Rochemont and Culicover (1990).

(20)   a.   Mary saw an alleged mouse from Mars yesterday.
       b.   Mary saw an alleged mouse yesterday from Mars.

The fact reported in Rochemont and Culicover (1990) is that *from Mars* isn't alleged in (20b), but can be in (20a). This suggests that the extraposed PP must be related to a position in the DP that is higher than *alleged*. That's achieved in our system with the representation in (21).

(21) $\exists y.\ y$ is an alleged mouse from Mars $\wedge$ Mary saw $\iota x.x = y \wedge\ y$ is an alleged mouse   (22)

IP

λn Mary saw $\iota x.x = n \wedge\ n$ is an alleged mouse $\lambda Q.\exists y.y$ an alleged
IP   mouse from Mars $\wedge\ Q(y)$
                                                              QP

DP   IP   Q   NP

Mary   I   VP $\exists$ PP

          VP   yesterday   from Mars

       V $\iota x.x = g(n)$
       | $\wedge\ x$ is an alleged mouse
      saw   DP

             D   NP

             |
            the   n   NP

                 alleged   NP

                          mouse
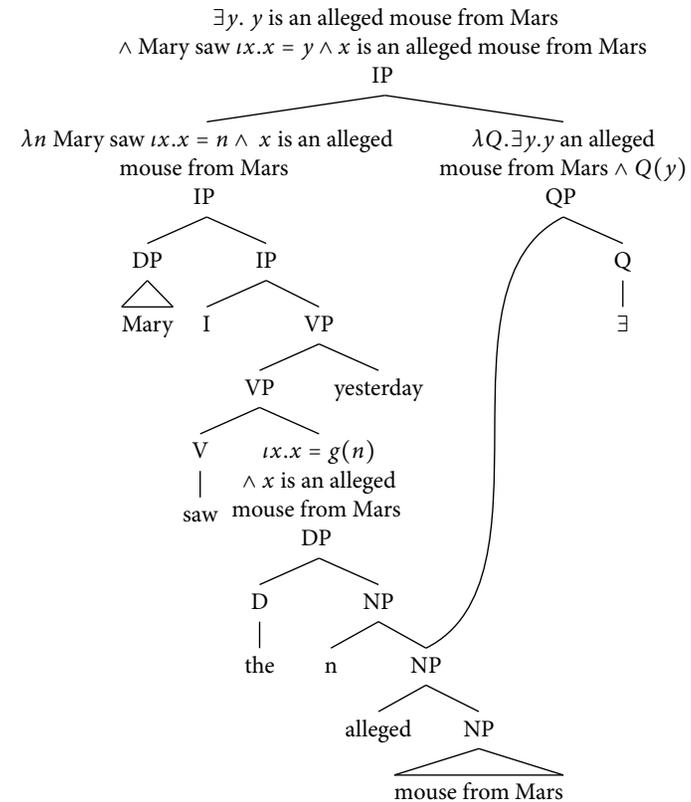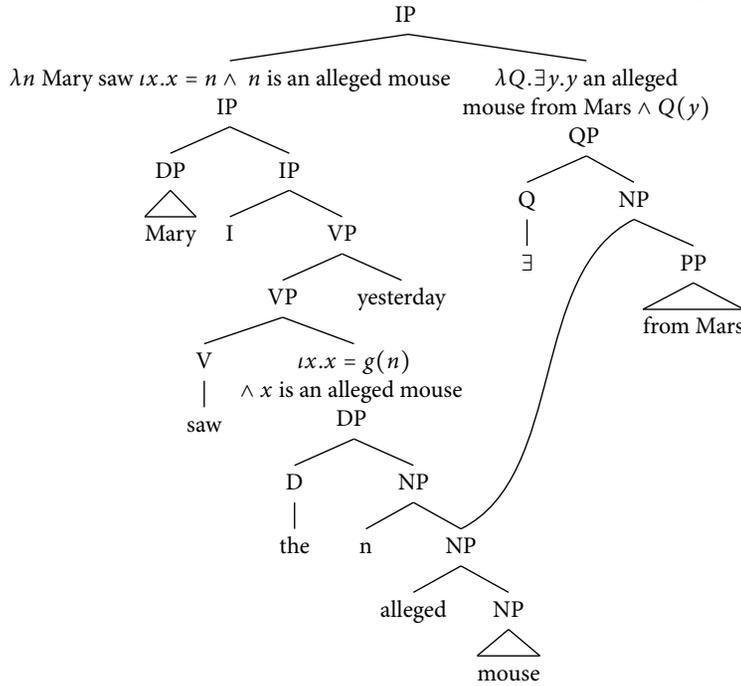
By contrast, the case in (20a) can have the structure in (22), which would allow
from Mars to be what is alleged.

(22)

$\exists y.\ y$ is an alleged mouse from Mars
$\wedge$ Mary saw $\iota x.x = y \wedge x$ is an alleged mouse from Mars
IP

λn Mary saw $\iota x.x = n \wedge\ x$ is an alleged $\lambda Q.\exists y.y$ an alleged
mouse from Mars   mouse from Mars $\wedge\ Q(y)$
IP   QP

DP   IP   Q

Mary   I   VP $\exists$

          VP   yesterday

       V $\iota x.x = g(n)$
       | $\wedge\ x$ is an alleged
      saw mouse from Mars
                DP

             D   NP

             |
            the   n   NP

                 alleged   NP

                          mouse from Mars

This is roughly the characterization of the situation that is in
Culicover and Rochemont (1990).

Fox (2002) argues, following Baltin (1987), that relative clauses which have
ACD in them are always extraposed. That is, he argues that the position a rela-
tive clause is pronounced in is the position that resolves its ellipsis. One piece of
evidence on behalf of that idea is (23), which is from Tiedeman (1995).

(23)   a.   * I said that everyone you did $\triangle$ arrived.

        b.   I said that everyone arrived that you did $\triangle$.

              $\triangle$ = said that x arrived

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

The word order in (23a) requires that the relative clause be inside the VP headed
by *said*, whereas in (23b), it's possible for the relative clause to be outside this VP.

5

If QR moves things rightwards, as these examples would suggest, it is often very difficult to tell when the relative clause is in or outside the VP. In (24), it could have moved string-vacuously out of the VP, and in (25), the PP could have itself moved farther to the right than the relative clause is.

(24)   She [$_{VP}$ read every book   ] that you did.

(25)   She [$_{VP}$ read every book   ] that you did to me.

   What this means for us is that we need the linearization scheme to assign to (26) the string in (27), and to (28) the string in (29).

(26)
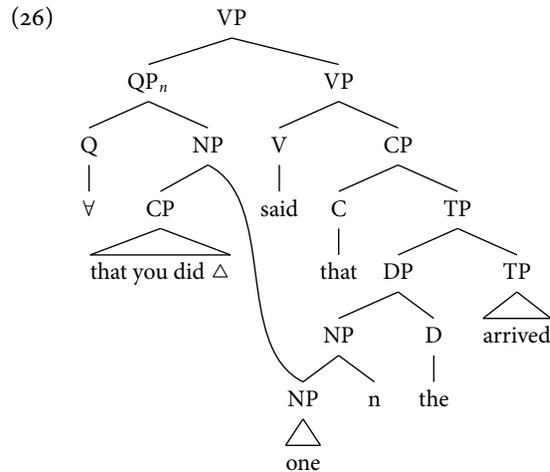
(27)   She said that everyone arrived that you did △

(28)
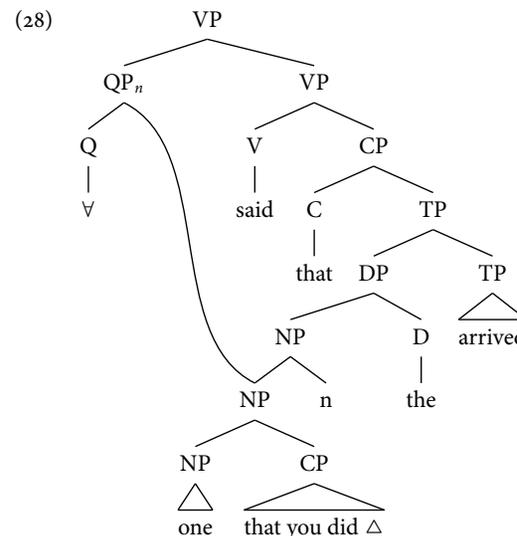
(29)   She said that everyone that you did △ arrived

If we find a way of ensuring that the NP which is a sister to the D always gets linearized according to the path that contains DP, then this contrast does, indeed, emerge. (Though the fact that the QR'd material shows up to the right of its sister, rather than to the left, does not.)

   This is the method of accounting for QR in a multidominant framework that Johnson (2012) suggests. Fox and Johnson (2016) tries to show how this method captures a puzzle that Perlmutter and Ross (1970) discover:

(30)   A man entered the room and a woman went out who were quite similar.
                                                    (Perlmutter and Ross 1970, (3), p. 350)

The relative clause here seems to be related to both subjects of the preceding coör-dinaton. We're going to look at a slightly different example because it's difficult to be sure in (30) that we have a restrictive relative clause. That's clearer, I think, in (31).

(31)   Every woman is smiling and every man is frowning who came in together.

This has the same meaning as (32).

(32)   Every woman and man who came in together are smiling and frowning respectively.

The theory of relative clause extraposition that we have is part of a theory about QR, so what we expect in (31) is that QR has (somehow) related the subjects of the conjuncts to a DP that the relative clause modifies. This, it turns out, is possible on the present theory of QR, when we add to it one additional hypothesis.

But before adding that hypothesis, I have to show you one necessary refinement. We've modeled the "trace" left by QR as a definite description equipped with an index, and we've given a semantics for the index that makes it a kind of adjective: it is the property of being identical to what the assignment function maps the index onto. But we can see from examples like (33) that the meaning of an index in a definite description is not exactly that.

(33)   A friend of every couple$_1$ thinks that the woman$_1$ dances better than the man$_1$.

Instead, we need to let an index hold of things that are part of what the assignment function assigns to the index.

(34) $[\![n]\!]^g = \lambda x\,[x \leq g(n)]$, where "$\leq$" means equal to or part of.

I need to also reveal to you now another fact about the meaning of *the* that we can see when it heads a plural DP. In a sentence like (35), we need to ensure that the set that *the students* refers to includes all the students in the context of utterance.

(35)   The students are irritated.

The meaning of *the* builds in a maximality operator that ensures the phrase it heads refers to every individual that satisfies the property denoted by the NP it contains. (It invokes a similar presupposition, namely that it can refer only when there is just one set of individuals in the context that has everything that satisfies the property denoted by the NP.) So let's build this into our meaning of *the*.

(36) $[\![the]\!] = \lambda P.\iota x\,\max(P(x))$, defined when there is exactly one $x\,\max(P(x))$.

       $\max(P(x)) = [P(x) = 1 \;\wedge\; \forall y[P(y) = 1 \rightarrow y \leq x]]$

That will give (33) the meaning in (37).

(37) $\forall y\,\text{couple}(y){=}1 \rightarrow$ a friend of y
      thinks that $\iota x\,\max(\text{woman}(x){=}1 \wedge x \leq y)$
    dances better than $\iota x\,\max(\text{man}(x){=}1 \wedge x \leq y)$
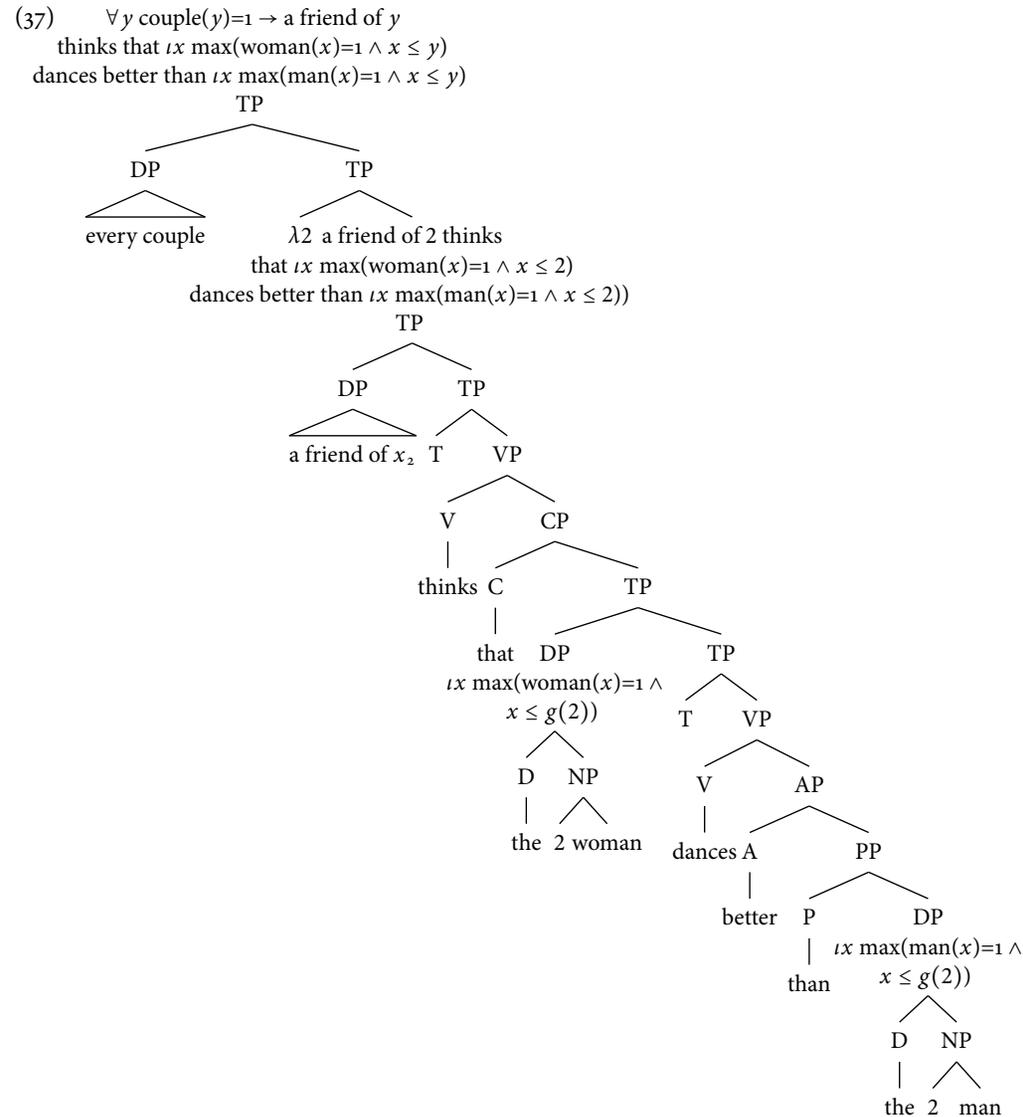
With these refinements to the meaning of a definite description that comes with an index, we now add a theory of coördination. This theory has two parts. The first is (38).

(38) If two merged phrases are of the same morpho-syntactic kind, then prepend *and* to the last part of the second.

DP → DP

DP
├ D — the
└ NP
  ├ NP — cats
  └ NP
    ├ NP — dogs
    └ NP — elephants

→

DP
├ D — the
└ NP
  ├ NP — cats
  └ NP
    ├ NP — dogs
    └ NP — and elephants

DP → DP

DP
├ D — the
└ NP
  ├ AP
  │ ├ AP — fat
  │ └ AP — hairy
  └ NP — mice

→

DP
├ D — the
└ NP
  ├ AP
  │ ├ AP — fat
  │ └ AP — and hairy
  └ NP — mice

The second part is the thesis that there is a process which derives the sums reading of conjoined NPs from Predicate Modification. (See Winter 2001 and Champollion 2015.)

(39) DP → DP

DP
├ D — the (λx woman(x)=1 ∧ man(x)=1)
└ NP
  ├ NP — woman
  └ NP — and man

→

DP
├ D — the (λx woman⊕man(x)=1)
└ NP
  ├ NP — woman
  └ NP — and man

woman⊕man holds of a plurality x if x contains a woman and a man.

QR now provides the following structure for (31).

(40)

TP
$\forall y$ [y is woman⊕man ∧ y came in together] →
[ιx max(x is a woman ∧ $x \leq y$) is smiling and ιx max(x is a man ∧ $x \leq y$) is frowning]

TP
  TP
    ιx max(x is a woman ∧ $x \leq 2$) is smiling
    and ιx max(x is a man ∧ $x \leq 2$) is frowning
    $\lambda 2$

  TP
    DP
      D — every
      NP
        2 NP — λx x is a woman
        woman
    TP
      T
      VP — is smiling

  IP
    and
    TP
      DP
        ιx max(x is a man ∧ $x \leq g(2)$)
        D — every
        NP
          2 NP — λx x is a man
          man
      TP
        T
        VP — is frowning

QP
  $\lambda Q \forall y$ [y is woman⊕man ∧ y came in together] → $Q(x)$
  NP
    λx x is woman⊕man
    NP
  Q
    CP — who came in together
    ∀

ιx max(x is a woman ∧ $x \leq g(2)$)

This is the structure that Zhang (2007) argues is correct for the Perlmutter/Ross examples, and it emerges out of our theory of QR, once we grant this theory of coördination. To the extent that this is successful, I think we can see it as evidence for a multidominant theory of movement. This isn't a structure that would be easily achievable on a, say, copy theory of movement.

8

References

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

Champollion, Lucas. 2015. Ten men and women got married today: noun coordination and the intersective theory of conjunction. *Journal of Semantics* 1–62.

Chomsky, Noam. 1977. *Essays on form and interpretation*. Amsterdam, The Netherlands: North-Holland Publishing Company.

Culicover, Peter W., and Michael S. Rochemont. 1990. Extraposition and the complement principle. *Linguistic Inquiry* 21:23–47.

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

Fox, Danny. 2015. The alleged reading of certain sentences with extraposed modifiers – evidence for embedded late merge. MIT.

Fox, Danny, and Kyle Johnson. 2016. QR is restrictor sharing. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, ed. Kyeong-min Kim, Pocholo Umbal, Trevor Block, Chan Queenie, Tanie Cheng, Finney Kelli, Mara Katz, Sophie Nickel-Thompson, and Shorten Lisa, 1–16.

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.

Johnson, Kyle. 2012. Toward deriving differences in how *Wh* movement and QR are pronounced. *Lingua* 122:529–553.

May, Robert. 1977. The grammar of quantification. Doctoral Dissertation, Massachusetts Institute of Technology.

Overfelt, Jason. 2015a. Extraposition of NPIs from NP. *Lingua* 164:25–44.

Overfelt, Jason. 2015b. Rightward movement: A study in locality. Doctoral Dissertation, University of Massachusetts at Amherst.

Perlmutter, David, and John Robert Ross. 1970. Relative Clauses with Split Antecedent. *Linguistic Inquiry* 1:350.

Rochemont, Michael, and Peter W. Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.

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.

Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry* 8:101–139.

Winter, Yoad. 2001. *Flexibility principles in Boolean semantics*. Cambridge, Massachusetts: MIT Press.

Zhang, Niina Ning. 2007. The syntactic derivations of split antecedent relative clause constructions. *Taiwan Journal of Linguistics* 5:19–48.

Kyle Johnson   University College London
University of Massachusetts at Amherst   24 June 2016

We left with a theory of QR that had these parts.

(1) $[\![\text{the}]\!] = \lambda P.\iota x\ \max(P(x))$, defined when there is exactly one x $\max(P(x))$.

  $\max(P(x)) = [P(x) = 1 \wedge \forall y[P(y) = 1 \to y \leq x]]$

(2) $[\![\text{n}]\!]^g = \lambda x\ [x \leq g(n)]$, where "$\leq$" means equal to or part of.

(3) $[\![\forall]\!] = \lambda Q.\lambda P.\forall y\ Q(y)$

(4) QR

  a. Merge QP with a λ expression,

  b. Make the NP part of that QP dominate the sister of the index the λ binds,

  c. Make the exponent of the head of QP be the D whose sister contains the index bound by the λ.

(5) A guard stands in front of every bank.
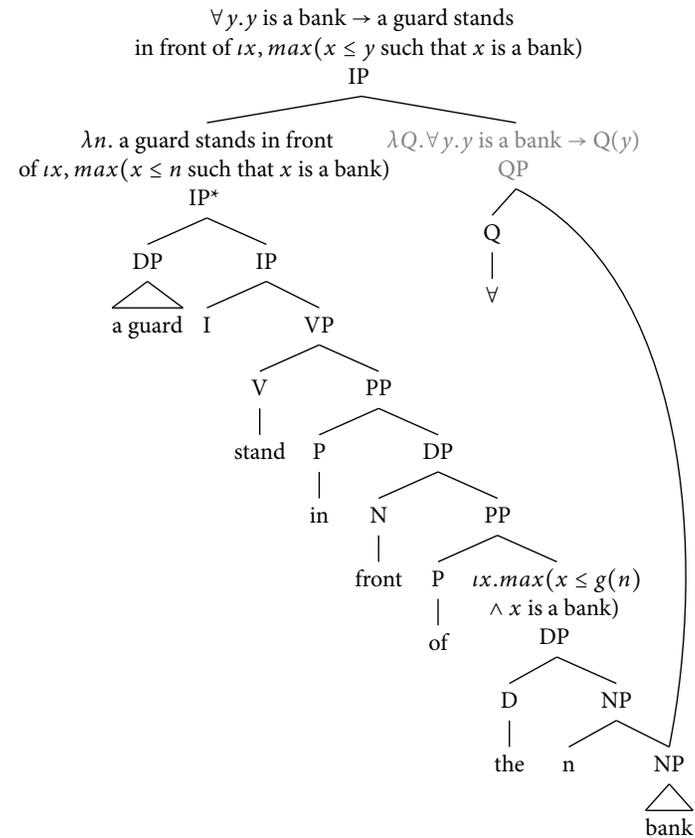
We saw that this theory gave us a way of capturing the Perlmutter/Ross examples:

(6)　Every woman is smiling and every man is frowning who came in together.

TP
$\forall y$ [y is woman$\oplus$man $\wedge$ y came in together] $\rightarrow$
[ιx max(x is a woman $\wedge$ $x \leq y$) is smiling and ιx max(x is a man $\wedge$ $x \leq y$) is frowning]

TP

QP
$\lambda Q \forall y$ [y is woman$\oplus$man $\wedge$ y came in together] $\rightarrow Q(x)$

TP　　$\lambda 2$
ιx max(x is a woman $\wedge$ $x \leq 2$) is smiling
and ιx max(x is a man $\wedge$ $x \leq 2$) is frowning

NP　　Q

λx x is woman$\oplus$man　　CP　　$\forall$
NP

who came in together

TP　　　　　　IP

DP　　TP　　and　　TP
ιx max(x is a woman
$\wedge$ $x \leq g(2)$)　　T　　VP

DP　　TP
ιx max(x is a man
$\wedge$ $x \leq g(2)$)　　T　　VP

D　　NP
is smiling

D　　NP
is frowning

every　　2　　NP
λx x is a woman

every　　2　　NP
λx x is a man

woman

man

This predicts, correctly, that the exponents of the quantifier in each DP must be the same, a fact noted by Moltmann (1992), McKinney-Bock (2013), and Zhang (2007).

(7)　a.　* A woman is smiling and every man is frowning who came in together.

　　b.　A woman is smiling and a man is frowning who came in together.

　　c.　* Every woman is smiling and some man is frowning who came in together.

　　d.　Some woman is smiling and some man is frowning who came in together.

　　e.　* Most women are smiling and every man is frowning who came in together.

　　f.　? Most women are smiling and most men are frowning who came in together.

　　g.　* Every woman is smiling and Few men are frowning who came in together.

　　h.　few women are smiling and few men are frowning who came in together.

　　i.　* No woman is smiling and every man is frowning who came in together.

　　j.　No woman is smiling and no man is frowning who came in together.

Link (1984)'s hydras teach us that (4c) needs changing.

(8)　Every woman and every man who came in together are now dancing.

This should be given the same meaning that (9) has.

(9)　Every woman and man who came in together are now dancing.

We want the expressions *every woman* and *every man* to be predicates which sum in the same way that *woman* and *man* do. This means that *every* cannot be $[\![\text{the}]\!]$. I suggest that we take *every* to be the index. Our new theory of QR:

(10)   QR

  a. Put QP in the Specifier of LP, a phrase headed by a term that intro-
   duces a λ binder,

  b. QP agrees with L which agrees with the index it binds,

  c. Make the sister of Q dominate the sister of Q's exponent: the index.

This gives us (11).

(11)
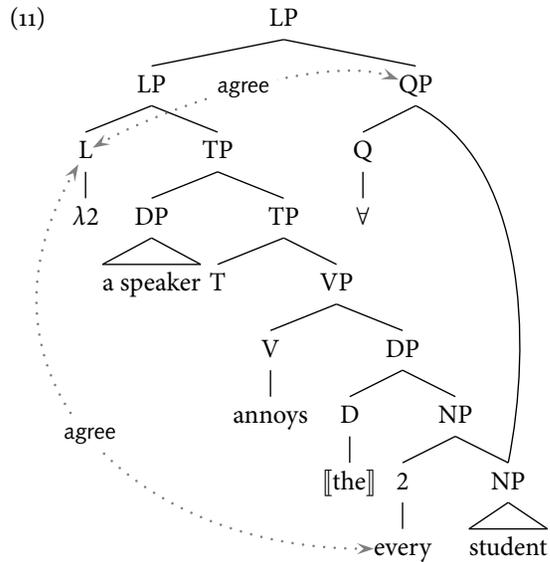
Note that the head that holds $[\![\text{the}]\!]$ is (always?) silent.

  My role model for (10a) and (10b) is subject agreement:

(12)
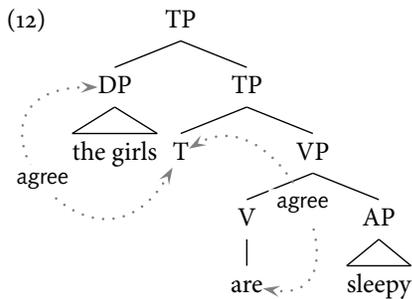
Hydras can now be given the structure in (13) on the next page.

(13) $\forall y\,[y$ is woman$\oplus$man$]\rightarrow[\iota x$ max$(x$ is [woman $\wedge\ \le y]\oplus$[man $\wedge\ \le y]\wedge x$ came in together$)]$ are now dancing
LP

LP $\lambda Q\forall y\,[y$ is woman$\oplus$man$]\rightarrow Q(y)=1$
                                                                                      QP

L   TP $\lambda x\,x$ is woman$\oplus$man Q

| NP |

$\lambda 2$ ιx max$(x$ is [woman $\wedge\ \le g(2)]\oplus$[man $\wedge\ \le g(2)]$ TP $\forall$
             $\wedge\ x$ came in together$)$
                          DP   are now dancing

D $\lambda x\,x$ is [woman $\wedge\ \le g(2)]\oplus$[man $\wedge\ \le g(2)]$
| $\wedge\ x$ came in together
$[\![\text{the}]\!]$ NP

$\lambda x\,x$ is [woman $\wedge\ \le g(2)]\oplus$[man $\wedge\ \le g(2)]$ $\lambda x\,x$ came in together
                        NP   CP

$\lambda x\,x$ is a woman   who came in together
       $\wedge\ x \le g(2)$ and $\lambda x\,x$ is a man
              NP $\wedge\ x \le g(2)$
                                                                     NP

   2   NP   2   NP
   | $\lambda x\,x$ is a woman | $\lambda x\,x$ is a man
 every   every
              woman   man

Note that "[woman $\wedge\ \le y]\oplus$[man $\wedge\ \le y]$" is a predicate that holds of a plurality x if x contains a woman who is part of y and a man who is part of y.

4

Note that Hydras are just special cases of the Perlmutter/Ross examples, ones in which the relative clause is merged with the NP building the DP and not the NP building the QP. We expect, therefore, that the exponents of the quantifier will be the same in hydras as well, and as Link (1984), Zhang (2007), and Moltmann (1992) note, this seems to be correct.

(14)   a.   * A woman and every man who'd agreed to dance with each other got out on the floor.

       b.   A woman and a man who'd agreed to dance with each other got out on the floor.

       c.   * Every woman and some man who'd agreed to dance with each other got out on the floor.

       d.   Some woman and some man who'd agreed to dance with each other got out on the floor.

       e.   * Most women and every man who'd agreed to dance with each other got out on the floor.

       f.   ? Most women and most men who'd agreed to dance with each other got out on the floor.

       g.   * Every woman and few men who'd agreed to dance with each other got out on the floor.

       h.   Few women and few men who'd agreed to dance with each other got out on the floor.

       i.   * No woman and every man who'd agreed to dance with each other got out on the floor.

       j.   No woman and no man who'd agreed to dance with each other got out on the floor.

We detoured into QR in order to explore the idea that it might be used in the way that Kotek (2014) does in forming derivations for wh movement. Kotek adopts a semantics for questions that uses alternatives and combines it with QR to explain why we don't get the expected intervention effects. Let's first update our system for questions. I've added our new idea about what indices mean to the special index that we used for questions, and I've updated our meaning for the definite determiner as well.

(15) $which = [\![the]\!] = \lambda P \, \iota x \, \max(P(x))$

(16)   The index in a wh-phrase is:
       $[\![f(x)]\!]^{g,f} = \{\lambda x \, x \leq f(g(x)), \lambda x \, x \leq f'(g(x)), ...\}$

(17)   …which book no one forgot.

$[\![CP]\!] = \{$ no one$_2$ forgot $\iota x \max(x \leq f(2) \wedge \text{book}(x) = 1)$
        no one$_2$ forgot $\iota x \max(x \leq f'(2) \wedge \text{book}(x) = 1)$
   no one$_2$ forgot $\iota x \max(x \leq f''(2) \wedge \text{book}(x) = 1), ...\}$
              $[\![CP]\!]^f = \varnothing$
                   CP
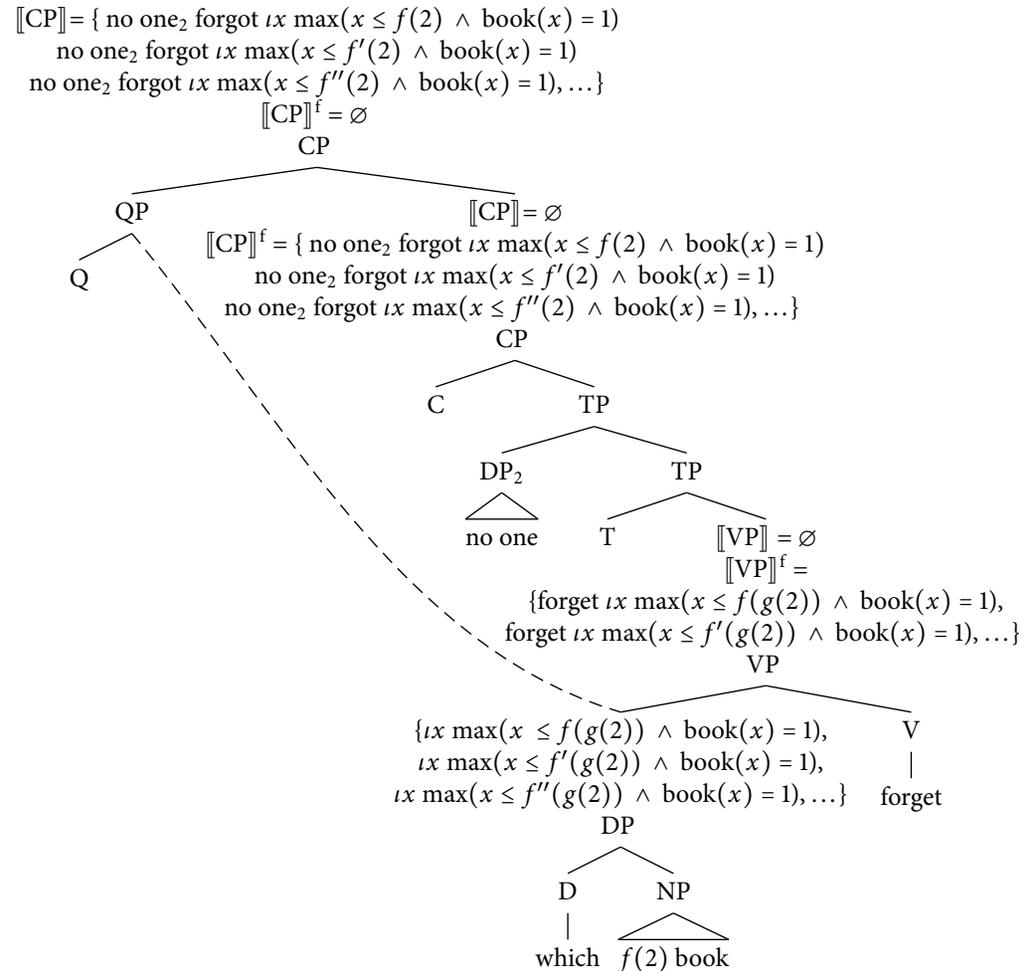
In languages in which a question can be formed with a wh-phrase in situ, this computation fails because *no one* is one of the things that breaks the set of alternatives from percolating up. (We looked at an example parallel to this one with *only Minsu* as subject, which also blocks the alternatives from percolating up.) The subject here is one of the class of things that produces "Beck effects." So there is a problem in giving English the same kind of meaning that we give questions with wh-in-situ. Kotek's idea is that we solve that with QR.

We can see that a wh-phrase is an expression that can undergo QR from cases like (18).

(18)   Which woman won't read which book that you have △.

So let's redesign wh-phrases so that they are quantificational expressions.

We'll make the *which*-phrase that we speak exactly like the definite descriptions with indices in them that we have made the quantificational phrases that we speak. The index in wh-phrase, however, will be one of those that can be a variable over functions that takes another variable as its argument.

(19) $[\![f(n)]\!]^g = \lambda x.x \leq g(f)(g(n))$

Like all indices, this a predicate that holds true of anything that is a part of the set of things which the function f picks out given the value given to n. Unlike what we had before, this function f won't be an alternative generator, it will be a garden variety variable over functions.

The alternative generator will be in the DP that QRs. I'll assign that meaning to *which*, and I'll *which* combine with a predicate, like other determiners that QR do, but rather than take that predicate as an argument it will use that predicate to introduce a presupposition. Here's a stab at that:

(20) $[\![which]\!] = \lambda P.\{f, f', f'', \ldots\}$, defined if $P(f) = 1$
       f is a function that picks an entity

So, for instance:

(21)

$$\{f, f', f'' \ldots\}$$
$$\text{DP}$$

$\lambda P\{f, f', f'', \ldots\}$ $\lambda x\, \text{book(x)=1}$
       D   NP
       | △
     which   book
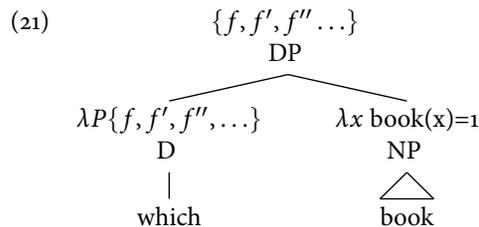
where $[\![D]\!]$ and $[\![DP]\!]$ are defined only if the thing that f picks out is a book.

I make this function defined only if the thing f picks out is something that $[\![NP]\!]$ holds of so that the content of NP will restrict what f can range over. That this can be done with a presupposition in this way is what Heim (2013) suggests.

Okay, putting this altogether, here's how Wh-movement will look. We'll QR a wh-phrase, and wh-move that QR'd phrase.

(22)

$[\![CP]\!]^o =$
{ no one$_3$ forgot ιx max($\lambda x\ x \le f(3)\ \wedge\ x$ is a book),
no one$_3$ forgot ιx max($\lambda x\ x \le f'(3)\ \wedge\ x$ is a book),
no one$_3$ forgot ιx max($\lambda x\ x \le f''(3)\ \wedge\ x$ is a book),…}
$[\![CP]\!]^f = \varnothing$
CP

CP   QP

C $[\![TP]\!]^f =$ Q
{ no one$_3$ forgot ιx max($\lambda x\ x \le f(3)\ \wedge\ x$ is a book),
no one$_3$ forgot ιx max($\lambda x\ x \le f'(3)\ \wedge\ x$ is a book),
no one$_3$ forgot ιx max($\lambda x\ x \le f''(3)\ \wedge\ x$ is a book),…}
$[\![TP]\!]^o = \varnothing$
TP

$\lambda 2$ no one$_3$ forgot $[\![DP\!\star]\!]^f = \{f, f', f'', …\}$
ιx max($\lambda x\ x \le 2(3)\ \wedge\ x$ is a book) $[\![DP\!\star]\!]^o = \varnothing$
LP   DP$\star$

L   no one$_3$ forgot   D
| ιx max($\lambda x\ x \le g(2)(3)\ \wedge\ x$ is a book)
$\lambda 2$ TP   which

DP$_3$ TP

no one   T   VP

V ιx max($\lambda x\ x \le g(2)(g(3))\ \wedge\ x$ is a book)
| DP
forgot

D $\lambda x\ x \le g(2)(g(3))\ \wedge\ x$ is a book
| NP
$[\![the]\!]$

2(3)   NP

book

DP$\star$ and every node above it has a value only if the entity f picks out is a book.

These structures allow us to do late merge too:

(23)   Which book that John brought does he recommend?

$$[\![CP]\!]^o =$$
$$\{ he_3 \text{ recommend } \iota x \max(\lambda x\; x \leq f(3) \;\wedge\; x \text{ is a book}),$$
$$he_3 \text{ recommend } \iota x \max(\lambda x\; x \leq f'(3) \;\wedge\; x \text{ is a book}),$$
$$he_3 \text{ recommend } \iota x \max(\lambda x\; x \leq f''(3) \;\wedge\; x \text{ is a book}),\ldots\}$$
$$[\![CP]\!]^f = \varnothing$$

CP

CP   QP

C $[\![TP]\!]^f =$ Q
$\{ he_3 \text{ recommend } \iota x \max(\lambda x\; x \leq f(3) \;\wedge\; x \text{ is a book}),$
$he_3 \text{ recommend } \iota x \max(\lambda x\; x \leq f'(3) \;\wedge\; x \text{ is a book}),$
$he_3 \text{ recommend } \iota x \max(\lambda x\; x \leq f''(3) \;\wedge\; x \text{ is a book}),\ldots\}$
$[\![TP]\!]^o = \varnothing$

TP

$\lambda 2\; he_3 \text{ recommend}$ $[\![DP^*]\!]^f = \{f, f', f'', \ldots\}$
$\iota x \max(\lambda x\; x \leq 2(3) \;\wedge\; x \text{ is a book})$ $[\![DP^*]\!]^o = \varnothing$

LP   DP*

L $he_3 \text{ recommend}$ D   NP
| $\iota x \max(\lambda x\; x \leq g(2)(3) \;\wedge\; x \text{ is a book})$
$\lambda 2$ TP   which   CP

DP_3   TP   that John_3 brought

he   T   VP

V $\iota x \max(\lambda x\; x \leq g(2)(g(3)) \;\wedge\; x \text{ is a book})$
| DP
recommend

D $\lambda x\; x \leq g(2)(g(3)) \;\wedge\; x \text{ is a book}$
| NP
$[\![the]\!]$

2(3)   NP

book

DP* and every node above it has a value only if f is a book that John brought.

Let's consider now how pied-piping will work.

(24)   On which table did no one stand?

The problem here will be fitting our theory – which is designed to let variables only be DPs – to a case where something other than a DP moves. One solution, indicated in the parse on the next page, is to let the preposition move semantically vacuously.

(25)

$\llbracket CP \rrbracket^o =$
{ no one$_3$ stand on ιx max($\lambda x\ x \leq f(3)\ \wedge\ x$ is a table),
no one$_3$ stand on ιx max($\lambda x\ x \leq f'(3)\ \wedge\ x$ is a table),
no one$_3$ stand on ιx max($\lambda x\ x \leq f''(3)\ \wedge\ x$ is a table),…}
$\llbracket CP \rrbracket^f = \varnothing$
CP

CP　　　　　　　　　　　　　　　　　　　QP

C　　　　　　　$\llbracket TP \rrbracket^f =$
{ no one$_3$ stand on ιx max($\lambda x\ x \leq f(3)\ \wedge\ x$ is a table),
no one$_3$ stand on ιx max($\lambda x\ x \leq f'(3)\ \wedge\ x$ is a table),
no one$_3$ stand on ιx max($\lambda x\ x \leq f''(3)\ \wedge\ x$ is a table),…}
$\llbracket TP \rrbracket^o = \varnothing$
TP
　　　　　　　　　　　　　　　　　　Q

$\lambda 2$ no one$_3$ stand on
ιx max($\lambda x\ x \leq 2(3)\ \wedge\ x$ is a table)
LP

$\llbracket PP\star \rrbracket^f = \{f, f', f'', …\}$
$\llbracket PP\star \rrbracket^o = \varnothing$
PP$\star$

L　　　　no one$_3$ stand on
|　　ιx max($\lambda x\ x \leq g(2)(3)\ \wedge\ x$ is a table)
$\lambda 2$　　　　　TP

$\{f, f', f'', …\}$
DP$\star$

DP$_3$　　　TP

D

no one　T　　VP

which

V　　　　PP

stand　　P　　ιx max($\lambda x\ x \leq g(2)(g(3))\ \wedge\ x$ is a table)
|　　　　　DP
on

D　$\lambda x\ x \leq g(2)(g(3))\ \wedge\ x$ is a table
|　　　　　NP
$\llbracket the \rrbracket$

2(3)　　NP

table

DP$\star$ and every node above it has a value only if f is a table.

10

The preposition is "late merged" into the QR'd term, but it isn't semantically interpreted in that position. To model the difference between languages like English which allow Pied-Piping as well as preposition stranding and languages, like those descended from Latin, which don't allow preposition stranding we can find a way of constraining what Q can attach to. This is what Cable (2010) suggests, though the method he proposes is incompatible with our multidominant representations.

This may look alarmingly complex, but there are few alternatives. If the variable that is bound in these cases corresponds to the argument of the preposition, then every syntax-to-semantics mapping is going to do something like what (25) does. On the copy theory of movement, for instance, the copy that is spoken in the higher position will have to be designed so that the semantics ignores the preposition, and that is precisely what (25) is expressing. The only way I can see to avoid this consequence is to adopt a theory of movement that would allow the PP to be translated into a variable in its lower position. One way we could do that is to interpret the PP in situ and use the alternative type semantics for questions that I described on Wednesday. But that would wrongly predict that we should get Principle C effects in examples like (26).

(26)   On which table that John hates will he stomp?
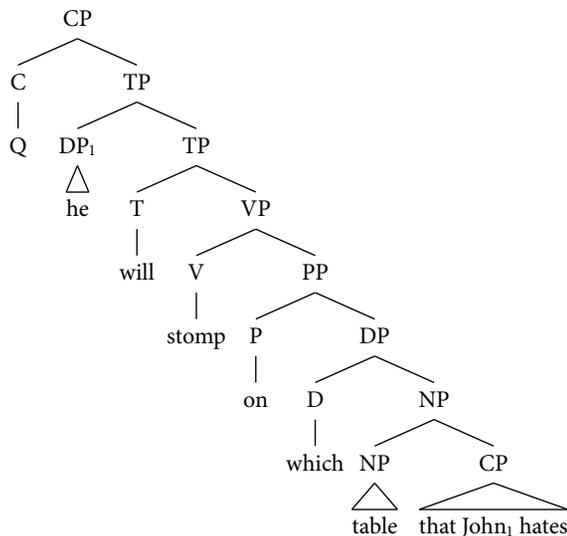
Another way we could do that is to revert to something closer to the trace theory of movement and let the PP in the lower position get translated into a variable of the same semantic type that the PP has. But that will wrongly predict that reflexives inside the PP cannot be bound from their lower position.

(27)   On which photo of herself has Mary put a frame?
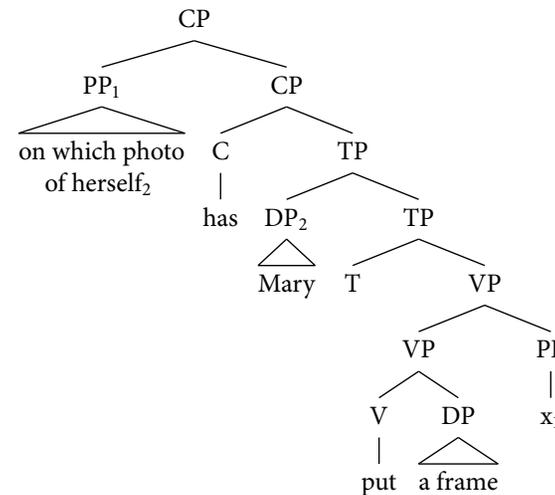
We need a syntax that will for part of the moved phrase to be interpreted below and part of it to be interpreted above, and that is difficult to achieve on a syntax that translates the moved PP into some particular semantic object. What we need is something that allows the DP within that PP to be broken apart. And that is what our syntax achieves. For the case in (26), for instance, we can do plain old late merge of the relative clause.

(28)

$\llbracket CP \rrbracket^o =$
$\{$ he$_3$ will stomp on ιx max$(\lambda x \; x \le f(3) \; \wedge \; x$ is a table$)$,
he$_3$ will stomp on ιx max$(\lambda x \; x \le f'(3) \; \wedge \; x$ is a table$)$,
he$_3$ will stomp on ιx max$(\lambda x \; x \le f''(3) \; \wedge \; x$ is a table$)$,$\ldots\}$
$\llbracket CP \rrbracket^f = \varnothing$
CP

CP   QP

C $\llbracket TP \rrbracket^f =$ Q
| $\{$ he$_3$ will stomp on ιx max$(\lambda x \; x \le f(3) \; \wedge \; x$ is a table$)$,
will   he$_3$ will stomp on ιx max$(\lambda x \; x \le f'(3) \; \wedge \; x$ is a table$)$,
he$_3$ will stomp on ιx max$(\lambda x \; x \le f''(3) \; \wedge \; x$ is a table$)$,$\ldots\}$
$\llbracket TP \rrbracket^o = \varnothing$
TP

$\lambda 2$ he$_3$ will stomp on $\llbracket PP^\star \rrbracket^f = \{f, f', f'', \ldots\}$
ιx max$(\lambda x \; x \le 2(3) \; \wedge \; x$ is a table$)$ $\llbracket PP^\star \rrbracket^o = \varnothing$
LP   PP*

L   he$_3$ will stomp on $\{f, f', f'', \ldots\}$
| ιx max$(\lambda x \; x \le g(2)(3) \; \wedge \; x$ is a table$)$ DP*
$\lambda 2$ TP

D   NP
DP$_3$ TP
|
he   T   VP   which   CP

V   PP   that John hates

|
stomp   P   max$(\lambda x \; x \le g(2)(g(3)) \; \wedge \; x$ is a table$)$
| DP
on

D $\lambda x \; x \le g(2)(g(3)) \; \wedge \; x$ is a table
| NP
$\llbracket$the$\rrbracket$
2(3)   NP

table

DP* and every node above it has a value only if f is a table and John hates f.

Sauerland (1998) discovered that it is possible for one moved wh-phrase to have a part that is just interpreted in the lower position and another part that is just interpreted in the higher position. An example of that kind is (29).

(29)   Which picture of hers that John brought should he show to no woman?

Our system will give this the representation in (30).

(30)

[Tree diagram with root CP]

CP
- CP
  - C
    - should
  - TP
    - LP
      - L
        - λ3
      - TP
        - DP₁
          - he
        - TP
          - T
          - VP
            - VP
              - V
                - give
              - DP
                - D
                  - ⟦the⟧ 3(2)
                - NP
                  - NP
                    - pictures of hers₂
            - PP
              - to no one₂
    - DP
      - D
        - which
      - NP
        - CP
          - that John₁ brought
- QP
  - Q

References

Cable, Seth. 2010. *The grammar of Q: Q-particles, Wh-Movement and Pied-Piping*. Oxford University Press.

Heim, Irene. 2013. Functional readings without type-shifted noun phrases. Unpublished paper, MIT.

Kotek, Hadas. 2014. Composing questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Link, Godehard. 1984. Hydras. on the logic of relative clause constructions with multiple heads. In *Varieties of Formal Semantics*, ed. Fred Landman and Frank Veltman, 151–180. Reidel.

McKinney-Bock, Katherine. 2013. Deriving split-antecedent relative clauses. *University of Pennsylvania Working Papers in Linguistics* 19.

Moltmann, Friederike. 1992. Coordination and comparatives. Doctoral Dissertation, MIT.

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

Zhang, Niina Ning. 2007. The syntactic derivations of split antecedent relative clause constructions. *Taiwan Journal of Linguistics* 5:19–48.