On the Copy Theory of Movement

EALing
Lecture One
September 2007

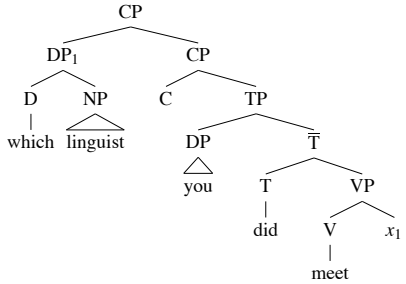
Kyle Johnson
University of Massachusetts at Amherst
kbj@linguist.umass.edu

"Movement" is the name of the operation that is central to forming a variety of sentences, including English questions of the sort that (1) illustrates.

(1) Which linguist did you meet?

This is thought to have a syntactic representation like (2), at least if a number of details are ignored.

(2)



Movement is responsible for relating *which linguist* to the x in the object position of *meet*. It creates the syntactic representation whose semantic interpretation captures the fact that we understand a phrase spoken in one position to fulfill the grammatical function defined in another position. In (1), it's responsible for letting *which linguist* be understood as the object.

Some details about (1). "T" stands for tense; it marks where the tense morpheme resides, in this case as part of the auxiliary verb *did*. You'll notice that the position of *did* in (2) is therefore not where its found in (1). In independent questions, the tense bearing item is spoken in C: this is one of the details we ignore. "C" stands for "complementizer," and it is the position in which information about the type of clause resides. In this case, C contains semantic information that indicates that (1) is a question. Throughout these lectures I will ignore the roles that C and T make in our representations. If I were to include them, several things would need adjustment, and those adjustments would become our subject matter.

What I want our subject matter to be, however, is the syntactic and semantic, content of the x part of this representation. As this is a silent piece of the representation, most of the evidence we will have about x comes by way of what it means. Today I will spell out a popular approach and review the facts that support it. We'll end with a set of problems that I want to try to solve in the lectures that follow.

We'll start with a sketch of what the non-question part of the sentence in (1) means. Determiners, as we'll see, denote relations between the denotation of the NP they combine with and the denotation of the rest of the sentence.[1] Imagine, for the sake of concreteness, that *which* is semantically equivalent to *some*. That is, imagine that (1), stripped of the contribution the question C makes, is equivalent to something like (3).

[1] See Barwise and Cooper (1981).

(3) There's some linguist you met.

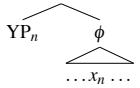
Think of a question as being the combination of (3) and something that invites one to identify the individual that makes (3) true.

So let's pretend that the question in (1) has the same meaning that (3) does and concentrate on deriving this meaning. Here's a standard way of doing that.

(4) a. $[\![\text{linguist}]\!] = \lambda x.\ x$ is a linguist
b. $[\![\text{which}]\!] = \lambda P.\lambda Q.\exists x[P(x) = 1 \wedge Q(x) = 1]$
c. $[\![x_n]\!] = \text{n}$

where x, y, z range over individuals, P and Q range over predicates, and $\lambda n.\phi(n)$ is the function f such that $f(a)$=$\phi(a)$. I'll use"1" to denote True and "0," False. I'll also use whole numbers as indices.

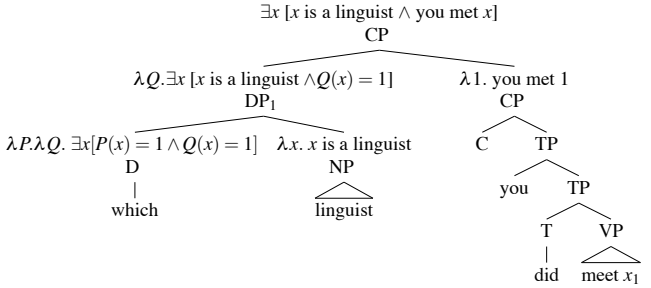
I won't work out how the meaning assigned to "you did meet x_1" is composed from its parts. To work out the denotation of the whole sentence, we need a rule that combines *which linguist* with the rest of the sentence. That rule could look like (5).

(5) The Heim/Kratzer rule. (see Heim and Kratzer (1998))
In 

change the denotation of ϕ to $\lambda n.\phi$.

These parts go together as indicated in (6).

(6)



The Heim/Kratzer rule is designed to capture the fact that the spoken phrase in a movement relation is a sister to the constituent that contains the other position. Movement cannot relate phrases to positions that are not within the phrase they are sisters to:

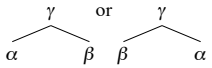
(7) * Which linguist you ran bothered.

On the Copy Theory of Movement

EALing
Lecture Two
September 2007

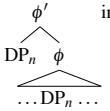
Kyle Johnson
University of Massachusetts at Amherst
kbj@linguist.umass.edu

Here are some of the conclusions we reached yesterday.

(1) a. Merge(α)(β) =$_{\text{def.}}$ 

b. Copy(α) produces α', an exact copy of α.
   i. α' may be an argument of Merge.
   ii. α must be produced by Merge.

c. A Movement relation is established by giving the output of Copy to the input of Merge.

d. TRACE CONVERSION
In 

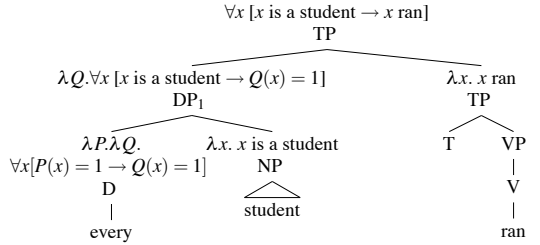
interpret ϕ as a function that maps an individual, x, to the meaning of ϕ[x/n].

ϕ[x/n] is the result of replacing the head of every constituent bearing the index n in ϕ with the head *the*$_x$, whose interpretation, $[\![\text{the}_x]\!]$, is: $\lambda P.\ [\![\text{the}]\!][P \cap \lambda y.y = x]$.
(slightly modified from Fox 2003, (52): 111)

e. Principle C
A name-like expression, α, is taken to refer to a different individual than any expression that c-commands α.
α c-commands β iff every phrase that contains α also contains β, and β is not in α.

In the first lecture we arrived at this picture from an examination of how Movement created constituent questions. Another context where it applies is in cases where the quantificational DP that is being moved isn't a DP headed by *which*, but is a DP headed by a quantifier like *every*. Like *which*, we can think of the meaning for *every* as being one that relates the denotation of the NP it combines with to a predicate that follows. Here's a sketch of how this works:

(2) Every student ran.



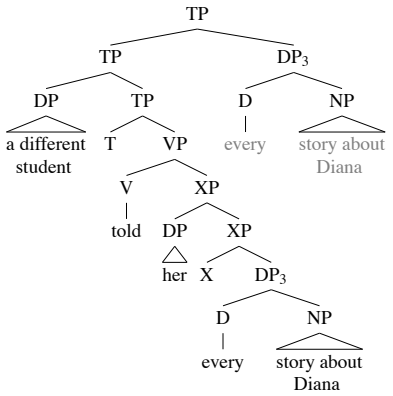
The rule that moves quantificational DPs, like those headed by *every*, plays a role in giving sentences like (3) their interpretation. This rule is known as QR (for: Quantifier Rule).[1]

(3) A different student told her [$_{\text{DP}}$ every story about Diana's parents].

The *every story* DP can have the subject in its c-command domain. We know that because the subject in (3) can get an interpretation that's only available when *different* is c-commanded by the universally quantified DP, as (4) shows.

(4) a. Every woman talked to a different student.
b. * [Her visit to every woman] disturbed a different student.
c. * A different student cried after every woman left yesterday.

So we need to let QR form a representation like (5).

(5) 

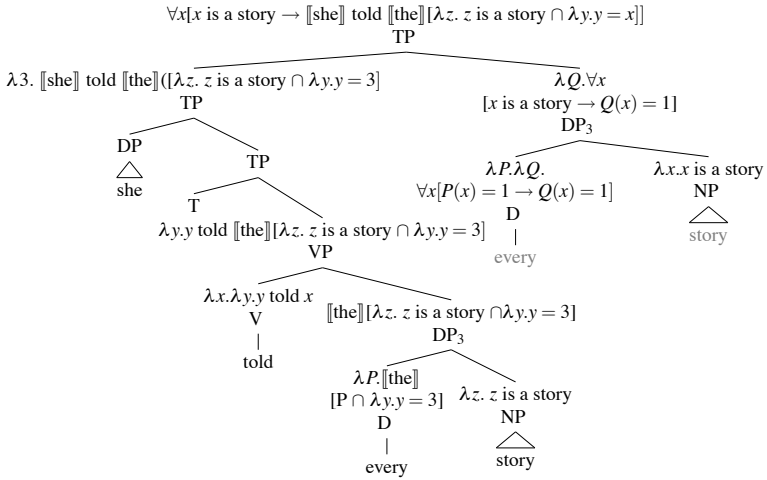
This rule does just what the rule moving Wh-phrases does. It manufactures a copy and merges it in a higher position. That copy gets interpreted as a quantificational expression, and in this case the phrase that it combines with contains the *different* expression. The lower copy gets interpreted by way of Fox's Trace Conversion rule: it is understood as a restricted variable. This rule will make from *every story about Diana's parents* something that I will paraphrase with:

(6) that x which is a story about Diana's parents

Understand the "that x" part of this paraphrase to refer to the referents picked out by the index borne by the DP. To see how this works in detail for (5), and so to see what (6) is a paraphrase of, would require wading through detail that we should avoid. But we can see it in the simpler (7).

(7) She told every story.

[1] See May (1977, 1985).



The clumsy paraphrase for the denotation of this sentence is:

(8) For all x, if x is a story, then she told that x which is a story.

Notice that, the representation in (5) places one copy of *Diana* within the c-command domain of *her*, and therefore correctly produces the environment for Principle C's effects. The referents of *Diana* and *her* in (3) do, indeed, seem to be taken distinct.

The only difference between (5) and the cases of movement we looked at previously is which copy gets pronounced. In (5) it is the lower copy, while in all of our previous examples it is the higher one. One of the questions we ended with yesterday was: what determines which copy is pronounced. We see in these examples, then, that it is not always the highest copy that gets pronounced. We'll need an answer that is more flexible than that. In this lecture, I focus on cases like (5).

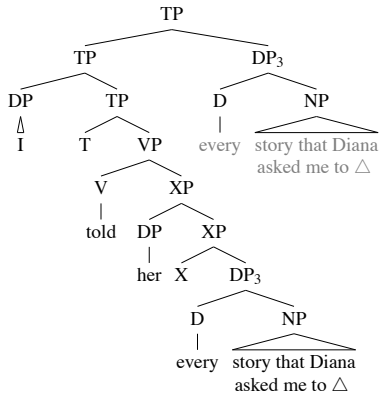
A feature of the representation in (5) that will be important for what follows is that the higher, unspoken, copy contains the NP that we see in the lower, spoken, copy. On some accounts of these cases, the higher copy contains only the quantifier and not the NP. I will therefore report the argument from Fox (2002) for this feature of the proposal.

Fox's argument comes from a phenomenon discovered by Fiengo and May (1994). There are situations where the disjoint reference effect that (3) illustrates are overcome. If the name is within a relative clause that is forced by ellipsis to be interpreted outside the phrase that contains the coreferent pronoun, as in (9), then the disjoint reference effect is modulated.

(9) ? I told her$_1$ every story that Diana$_1$ asked me to △.
△ = tell her *x*

The ellipsis in (9) appears to lie within the VP that serves as its antecedent, and this produces well-known problems. Thus, both the disjoint reference effects and the presence of ellipsis indicate that the copy theory of movement cannot have its normal outcome in this case:

(10)

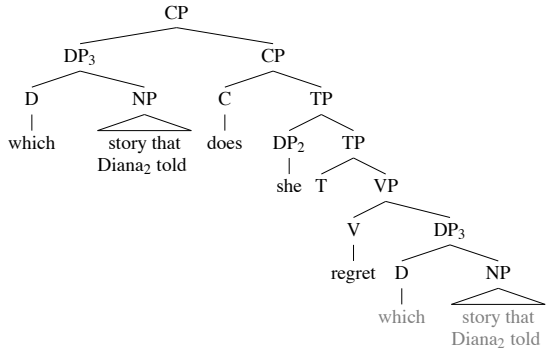


In examples where a name is within a relative clause in a moved *wh*-phrase, Principle C effects are also lost. (11) is such a case.

(11) Which story that Diana$_1$ told does she$_1$ now regret?

Without modification, the copy theory of movement would also wrongly give this sentence a representation in which a *Diana* falls within the c-command domain of *she* with an accompanying disjoint reference effect.

(12)



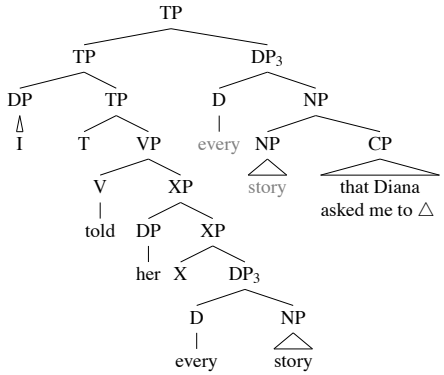
The solution to both cases is to allow for a derivation that involves the following steps.

(13) Late Merger
a. Build the D+NP phrase that will move
b. Move that D+NP

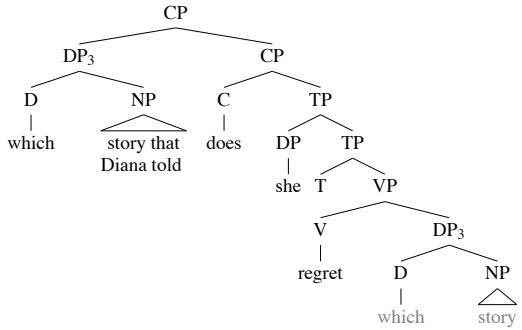
c. Build and merge the relative clause into the higher copy only.

This is David Lebeaux's solution,[2] and it is adopted by Fox. These derivations would give to (9) and (11) the representations in (14).

(14) a.



b.



As (14) indicates, this account claims that in (9) the relative clause is not in the spoken copy of the object, but is instead part of the higher, unspoken, copy. It's this part of the account that serves as evidence that the higher, unspoken, copy in such cases contains the NP part of the object. That is necessary because under standard assumptions, a restrictive relative clause of this kind must attach to an NP to produce the right meaning. Fox (2002) produces a variety of arguments that the relative clause in such examples is indeed not in the spoken copy of the DP. (See also Baltin (1987), and Fox and Nissenbaum (1999) for arguments that "extraposition from NP" has the syntax in (14a).) One of these is based on the contrast in (15), from Tiedeman (1995).

[2] See Lebeaux (1988), and also Freidin (1986).

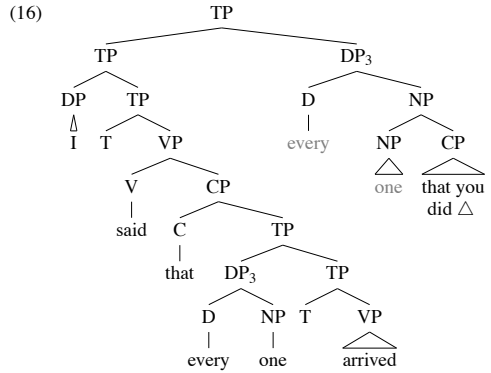
(15) a. * I said that everyone you did △ arrived.
△ = say that *x* arrived

b. I said that everyone arrived that you did △.
△ = say that *x* arrived

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

Just as this proposal predicts, the relative clause containing an elided VP cannot be spoken within the antecedent VP. That is what makes (15a) ungrammatical. Instead, that relative clause must be spoken in a position outside the antecedent VP and, more particularly, as part of the material that determines the scope of the quantificational DP the relative clause modifies. That's what's happened in (15b), which has the representation in (16).

(16)



For this account to be complete, it requires an explanation for when late merger derivations are possible and when they are not, for otherwise all of the effects gained by the copy theory of movement will be lost. See Takahashi (2006) for many steps in this direction.

1 Multidominance

Fox's Trace Conversion has the undesirable property of letting a whole class of lexical items be ambiguous: in the cases we're examining, those lexical items are determiners. It resolves that ambiguity by syntactic rule. It claims that the meanings of determiners are not fixed, but change according to their position. It gives syntax the power to change lexical content. That's more than syntax should be allowed. So let me offer a variant of Fox's proposal that avoids these consequences.

This variant builds on ideas many have had about the syntax of quantification.[3] Perhaps it's closest to ideas in Beghelli (1993, 1995), Sauerland (1998), Sportiche (2003), Butler (2004), Kratzer (2005) and Adger and Ramchand (2005). Imagine, as in Matthewson (2001), that quantificational expressions make use of two functional heads. One has the denotation of quantifiers, and the other is a choice function that provides the domain for the quantification. For concreteness, we can assume this choice function to have

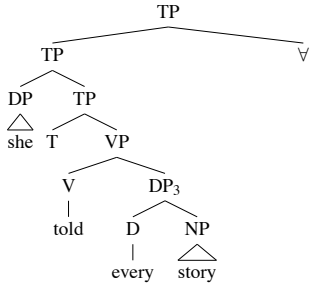
[3] See, e.g., Williams (1986, 1988), the papers in Szabolcsi (1997), Giannakidou and Merchant (2002), and Hallman (2000).

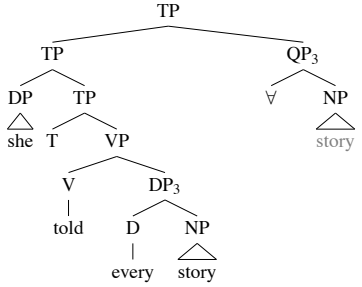
just the meaning given to determiners by Fox's Trace Conversion rule. Unlike Matthewson, but like those cited above, let's separate these two functional heads, putting the term that expresses the quantification in the position where its scope is computed, while the choice function is in construction with the NP. The morphological form of the choice function varies depending on the quantificational term. Let's follow Kratzer (2005) and Adger and Ramchand (2005) and let this dependency be mediated by AGREE. AGREE will determine the morphological form of the choice function part, and make both heads share an index. On this view, then, there is only one determiner — the one that Fox's Trace Conversion creates — with a morphological form that is fixed by AGREEing with a silent quantifier.

(17) The only (quantificational) determiner is $[\![\text{the}_x]\!]$. Its morphological form is determined by the silent Q it agrees with.

This proposal would give to (18) the derivation indicated. (I will use "∀" to represent the silent universal quantifier that AGREEs with *every*.)

(18) She told every story.

a.


b.


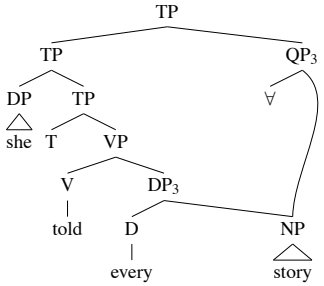
Let's now consider what the copies are in the copy theory of movement. There are two criteria that any successful characterization of copies must meet.

(19) a. Only one copy should be able to be pronounced.
b. Every copy must be absolutely identical.

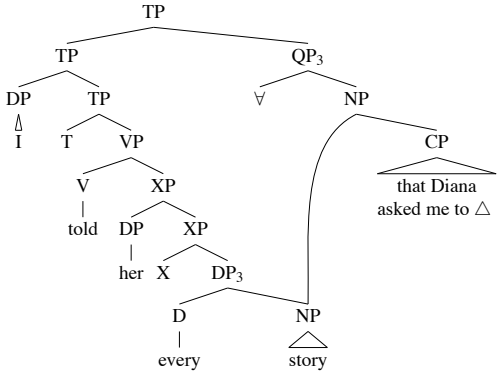
One proposal that achieves both these goals is (20).

(20) Copies of α are one and the same α in different syntactic positions.

A simple implementation of this idea is to let phrase-markers allow for multidominance, i.e., relax the requirement that a term have no more than one mother. (See Nunes (2001), Starke (2001), Frampton (2004) and Fitzpatrick and Groat (2005) for recent proposals along these lines, Citko (2005) for an application of the idea to across-the-board movement, and McCawley (1982) and Blevins (1990) for some early versions of the idea. Much of the same ground I cover here is covered in de Vries (2007).) This would give to (18) the representation in (21).

(21)


An example involving late merger, such as (14a), will get a representation like (22).

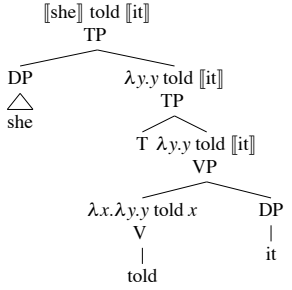
(22)


These representations ensure that every copy has exactly the same material in it, and that the semantic contribution it makes is precisely the same in each position. It also provides a way of deriving that only one copy may be pronounced, along the lines described in Nunes (1995, 1996, 1999). The procedure that maps syntactic structures onto strings will (sometimes) get conflicting information from phrases that have more than one mother. In those cases, that procedure will have no choice but to resolve the conflicting information so that it puts such a phrase in just one of the positions it occupies. We'll look at details in the next lecture.

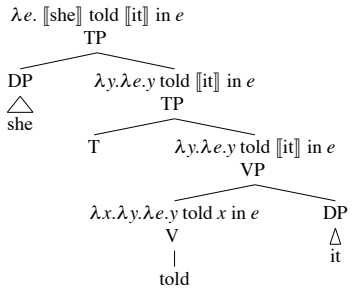
2 Events

There is now considerable evidence that predicates like *ran*, *read*, *tell*, etc. do not describe properties of individuals, as in (23), but instead relate individuals to events, as in (24).

(23) ⟦she⟧ told ⟦it⟧

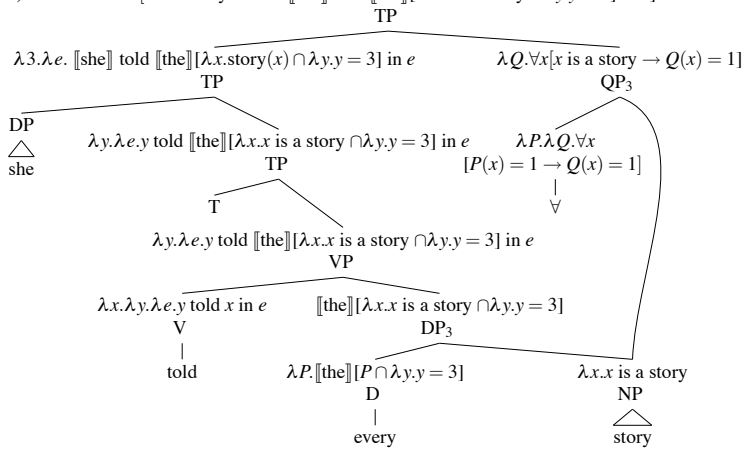


(24) λe. ⟦she⟧ told ⟦it⟧ in e



On this view, sentences refer to the events in which the relations named by their predicates hold. The account sketched in the previous two sections is built upon a non-event view of sentence meanings and won't work with an event based semantics, as (25) demonstrates.

(25) ∀x[x is a story → λe. ⟦she⟧ told ⟦the⟧[λx.x is a story ∩λy.y = x] in e]



The denotation for (25) is not what we want, and is, in any case, ill-formed. In paraphrase form, it's:

(26) For every x, if x is a story then the set of events, e, such that she told the things that are stories and x in e are true.

What we need to do is change the meaning given to ∀ so that in (26) it relates stories in some event to those things that she told in some event.[4]

To do this, I will use Elbourne (2005) as my model. Elbourne's system is designed to capture the reading that the definite description gets in examples such as (27).

(27) Every man who owns a donkey beats the donkey.

In (27), *the donkey* is anaphoric on the donkeys described in the subject: the donkey each man beats is the donkey that he owns. This can be achieved by letting the universal quantifier associated with *every* quantifier over the events that *man who owns a donkey* and *beats the donkey* describe. Simplifying somewhat, Elbourne's proposal is that ∀ has the denotation in (28).

(28) $[\![\forall]\!] = \lambda f_{<e,<s,t>>}.\lambda g_{<e,<s,t>>}.\lambda s.$ for every x and every minimal s' such that $s' \leq s$ and $f(x)(s') = 1$, there is a minimal s'' such that $s' \leq s'' \leq s$ and $g(x)(s'') = 1$.
(compare: Elbourne 2005, section 2.2.4)

The variables s, s' & s'' range over "situations," which we can equate with events, and "$\leq$" is the reflexive part-of relation. (28) says that ∀ takes two relations between individuals and situations (*qua* events), f and g, and describes those situations, s, in which, for every x, all of the smallest sub-situations of s that contain

[4] There are less sweeping ways of addressing the problem in (25) than I will now propose. For instance, we could assume that the material I've been ignoring in T, and whatever else might lie between the subject and VP, existentially closes the event variable. This would give the right meaning for these cases. I'll go a different route, then, not because it's demanded by this problem, but because it yields results that can't be derived otherwise and I think are useful.

x and make f true are part of the smallest sub-situation of s that contains x and makes g true. What this will do in the case of (27) is give it a meaning that can be paraphrased with (29).

(29) There is a situation, s, such that for every x and every minimal situation s' in s such that x is a man who owns a donkey in s', there is a larger minimal situation, s'' in s that contains s' such that x beats the donkey in s''.

A looser, but perhaps more revealing, paraphrase is (30).

(30) For every x, all of the minimal situations of a man, x, owning a donkey are part of a larger minimal situation in which x beats the donkey.

The notion of "minimal situation" does the work of getting the donkeys that are beaten to be the donkeys that are owned in (27). Think of a situation as being made up of individuals, relations and properties. A minimal situation can be informally described as one which is made up of only those individuals, relations and properties necessary to make some proposition true. (See Berman (1987), Schein (1993), and Kratzer (1989, 1990, 2002, to appear).)

(31) A situation, s, is a minimal situation in which $P(s) = 1$ iff there is no $s' < s$ such that $P(s') = 1$.

Therefore, a situation that makes "a man, x, owns a donkey" true will be a minimal situation just in case it contains only that man and one donkey and the 'own' relation between them. A situation that makes "x beats the donkey" true will be a minimal situation just in case it contains only x and the donkey and the 'beats' relation between them. If the 'owns' situation is a part of the 'beats' situation, then, because they are each allowed only one donkey, the donkey in both situations will be the same.

This treatment of (27) extends to cases like (32).[5]

(32) Every man who owns a donkey beats it.

The *it* in (32) is anaphoric to *a donkey* in the very same way that *the donkey* is in (27). Elbourne (2005) argues that this is because *it* is, in fact, *the donkey* with the NP containing *donkey* elided. His proposal, following Postal (1969), is that pronouns are how definite determiners get pronounced when the NP they are in construction with is elided. The transformations in (33) are all on a par.

(33) a. She saw some books and he read some books. → She saw some books and he read some △.
b. She wrote no books and he read no books. → She wrote no books and he read none △.
c. She wrote the book and he read the book → She wrote the book and he read it △.

Thus, (32) is actually (34), where *it* has the meaning of *the*, and the semantics has the same consequences for (34) that it does for (27).

(34) Every man who owns a donkey beats [$_{DP}$ it [$_{NP}$ ~~donkey~~]].

The work done by minimal situations in guaranteeing that *the donkey* in (27), or *it* in (32), are anaphoric on the descriptive content of the preceding quantified expression mimics the work done by indices in Fox's Trace Conversion rule. In translating Fox's system into an event based semantics, I propose that we dispense with the indices in his rule and exploit the minimal situation technique. Elbourne's system is also built on quantificational expressions binding indexed variables, so I will have to change his denotation for $\forall$ as well. I propose changing it to (35).

(35) $[\![\forall]\!] = \lambda f_{<e,<s,t>>}.\lambda g_{<s,t>}.\lambda s.$for every minimal s' such that $s' \leq s$ and $\exists x\, f(x)(s') = 1$, there is a minimal s'' such that $s' \leq s'' \leq s$ and $g(s'') = 1$.

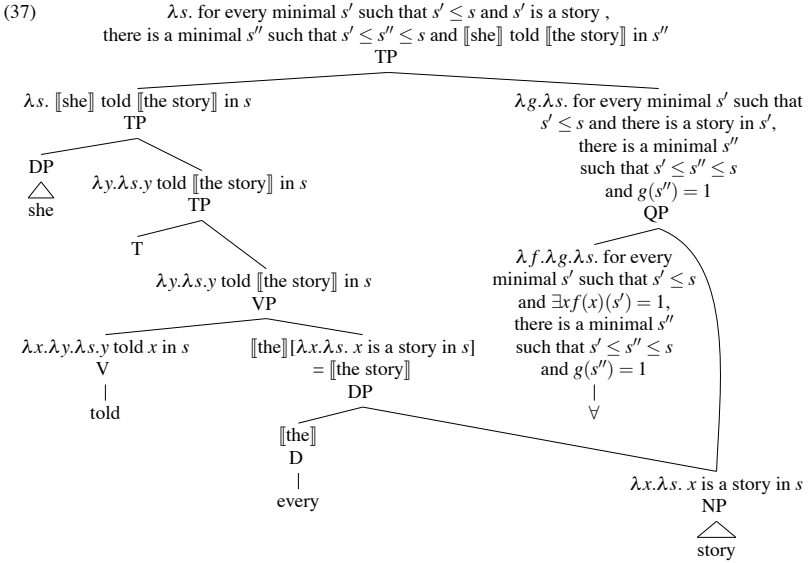
[5] See Geach (1962), Evans (1977), Cooper (1979), Heim (1990).

This assumes that NPs are functions from individuals to predicates of situations (or events).

(36) $[\![\text{story}]\!] = \lambda x.\lambda s.\ x$ is a story in s.

Not only will this proposal dispense with the indices in Fox's Trace Conversion rule, it will dispense with Fox's Trace Conversion rule entirely. Here's an illustration.[6]

(37)



My proposal, then, is that lower copies are plain definite descriptions — or the pronoun version of them — which share an NP with a higher quantificational expression. The connection between the higher quantifier and the lower definite description comes by way of quantifying over situations. The only difference between DP^a and DP^b being in a movement relation and DP^a and DP^b being independent is the shared NP that exists in the movement case. This achieves Chomsky (1993)'s goal of reducing the movement relation entirely to structure building; it does so in a way that is close in spirit to Starke (2001).

[6] My proposal has certain features in common with the "choice function" version of the Trace Conversion rule that Sauerland (2004) argues for. On Sauerland's view, the definite determiner in the lower copy is a variable over choice functions that is quantified over by the higher copy. As he shows in Sauerland (1998), this approach gives an interesting explanation of certain crossover phenomena. My proposal does not obviously extend to these cases.

References

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

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

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In *Proceedings of the North East Linguistic Society 23*, ed. Amy J. Schafer, 65–80. University of Ottawa: Graduate Linguistic Student Association.

Beghelli, Filippo. 1995. The phrase structure of quantifier scope. Doctoral Dissertation, UCLA, Los Angeles.

Berman, Stephen. 1987. Situation-based semantics for adverbs of quantification. In *University of Massachusetts occasional papers vol. 12*, ed. Jim Blevins and Ann Vainikka, 8–23. University of Massachusetts at Amherst.

Blevins, James. 1990. Syntactic complexity: Evidence for discontinuity and multidomination. Doctoral Dissertation, University of Massachusetts at Amherst.

Butler, Jonny. 2004. Phase structure, phrase structure, and quantification. Doctoral Dissertation, The University of York.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Ken Hale and Jay Keyser, 1–52. Cambridge, Massachusetts: MIT Press.

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

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.

Evans, Gareth. 1977. Pronouns, quantifiers, and relative clauses. *Canadian Journal of Philosophy* 7:467–536.

Fiengo, Robert, and Robert May. 1994. *Indices and identity*. Cambridge, Massachusetts: MIT Press.

Fitzpatrick, Justin, and Erich Groat. 2005. The timing of syntactic operations: Phases, c-command, remerger, and Lebeaux effects. Paper presented at ECO5, March 2005.

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

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

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.

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

Freidin, Robert. 1986. Fundamental issues in the theory of binding. In *Studies in the acquisition of anaphora*, ed. Barbara Lust, volume 1, 151–188. Dordrecht: Reidel.

Geach, Peter. 1962. *Reference and generality*. Ithaca, NY: Cornell University Press.

Giannakidou, Anastasia, and Jason Merchant. 2002. Modularity in the Minimalist Program. Paper presented at the Maryland Mayfest, May 2002.

Hallman, Peter. 2000. The structure of predicates: Interactions of derivation, case and quantification. Doctoral Dissertation, University of California, Los Angeles.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13:137–178.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12:607–653.

Kratzer, Angelika. 1990. Uniqueness. *Linguistics and Philosophy* 13:273–324.

Kratzer, Angelika. 2002. Facts: Particulars or information units? *Linguistics and Philosophy* 25:655–670.

Kratzer, Angelika. 2005. Indefinites and the operators they depend on: From Japanese to Salish. In *Reference and quantification: The Partee effect*, ed. Gregory N. Carlson and Francis Jeffrey Pelletier, 113–142. CSLI Publications.

Kratzer, Angelika. to appear. Situations in natural language semantics. In *The Stanford encyclopedia of philosophy*.

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

Matthewson, Lisa. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9:145–189.

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

May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge, Massachusetts: MIT Press.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13:91–106.

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

Nunes, Jairo. 1996. On why traces cannot be phonetically realized. In *Proceedings of North East Linguistic Society*, ed. Kiyomi Kusumoto, 211–226. Harvard University and MIT: Graduate Linguistic Student Association.

Nunes, Jairo. 1999. Linearization of chains and phonetic realization of chain links. In *Working minimalism*, ed. Samuel Epstein and Norbert Hornstein, 217–249. Cambridge, Massachusetts: MIT Press.

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

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

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.

Schein, Barry. 1993. *Plurals and events*. Cambridge, Massachusetts: MIT Press.

Sportiche, Dominique. 2003. Reconstruction, binding and scope. Unpublished manuscript, University of California, Los Angeles.

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

Szabolcsi, Anna. 1997. *Ways of scope taking*. Dordrecht ; Boston ; London: Kluwer Academic Publishers.

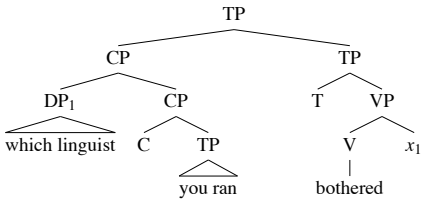
Takahashi, Shoichi. 2006. Decompositionality and identity. Doctoral Dissertation, Massachusetts Institute of Technology.

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.

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

Williams, Edwin. 1986. A reassignment of the functions of LF. *Linguistic Inquiry* 17:264–300.

Williams, Edwin. 1988. Is LF distinct from S-structure? a reply to May. *Linguistic Inquiry* 19:135–146.



Under this scheme, x is just a vehicle for an index. By virtue of carrying that index, it ends up getting bound by the lambda-operator which is appended by the Heim and Kratzer rule to a phrase containing x. This lambda operator is invoked by the phrase that binds x. For this to work, then, we must let the relation that connects a moved phrase its argument position ensure that the phrase and the variable share an index. This task is assigned to Movement.

(8) If XP and x are in a movement relation, then give both the same index.

There is evidence, however, that the x in representations such as these is not just a vehicle for an index, but has some semantic content of its own. This can be appreciated by considering examples like (9).[2]

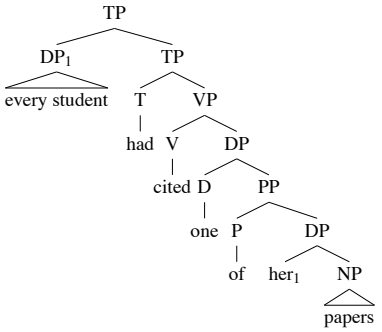
(9) a. Which of her$_1$ papers did you say every student$_1$ had cited?
compare:
b. *Which of her$_1$ papers did you say that the teacher of every student$_1$ had cited?

Personal pronouns can also carry indices, and when they are in the scope of a lambda operator they behave semantically just like the xs in our movement relations. The pronouns in these examples behave as if they are being interpreted in the argument position for the moved phrase. Compare these cases to (10).

(10) a. Every student$_1$ had cited one of her$_1$ papers.
b. * The teacher of every student$_1$ had cited one of her$_1$ papers.

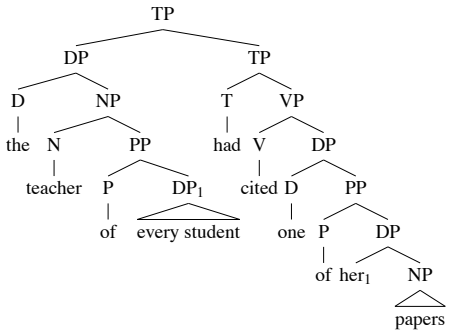
As expected from the Heim/Kratzer rule, the pronoun in (10a), but not (10b), can be interpreted like the variable, x, in our movement case.

(11)



[2] There's a large literature on effects like these. Some more recent examples are Barss (1986), Heycock (1995), Cresti (1995), Romero (1998), Fox (1999), Sharvit (1998), Sauerland (1998, 2004) and Sauerland and Elbourne (2002).

(12)

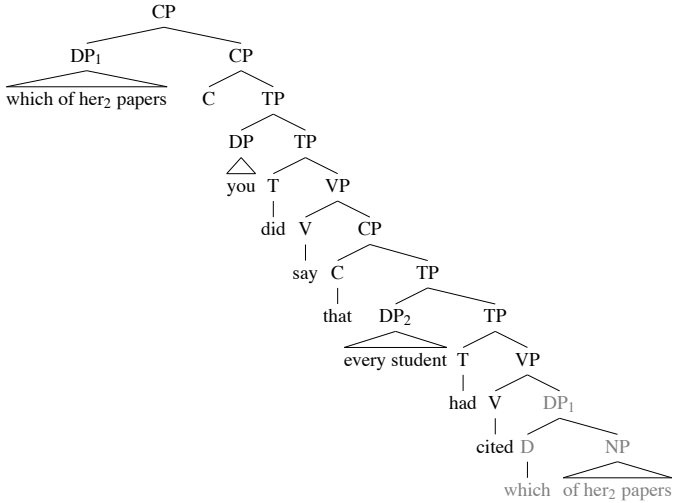


Chomsky (1993) suggests that we capture this fact by letting the movement relation generate copies of the spoken phrase. If we preserve the index sharing part of movement, we can express this thesis with (13).

(13) Movement (revised)
Put identical copies of XP in positions α and β, and give each copy the same index.

This would give to the example in (9a) the representation in (14).

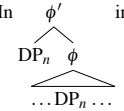
(14)



I will use the convention of putting unspoken copies into a shaded font.

This representation provides a way of interpreting the lower copy of the pronoun *her*, so that it gets bound by the lambda operator invoked by *every student*. To get a final interpretation for this, we'll need

two other things. A way of interpreting the higher copy of *her*, and a new rule to interpret the movement operation. I will ignore the first of these tasks until the next lecture. The new rule for interpreting the movement relation I will take from Fox (1999, 2002, 2003) (and see Sauerland (1998) for some close alternatives, and Elbourne (2005) for a more general version Fox's proposal). Fox suggests replacing the Heim/Kratzer rule with what he calls the "Trace Conversion" rule.

(15) TRACE CONVERSION
In 

interpret ϕ as a function that maps an individual, x, to the meaning of ϕ[x/n].

ϕ[x/n] is the result of replacing the head of every constituent bearing the index n in ϕ with the head the_x, whose interpretation, $[\![the_x]\!]$, is: $\lambda P.\ [\![the]\!][P \cap \lambda y.y = x]$.
(slightly modified from Fox 2003, (52): 111)

What Fox's rule does is change the meaning of the determiner in a lower copy into a kind of definite determiner. This definite determiner combines with the NP in that lower copy in the usual way. Suppose this means that the definite determiner refers to an individual that is among the set denoted by the NP it combines with. In (14), this will mean that it picks out an individual that is among the set denoted by *her papers*. To this, the Trace Conversion rule adds the restriction that whatever individual the definite determiner picks out must also be the individual assigned to the index borne by the copy. This will give a meaning to (14) that could be roughly paraphrased by (16).

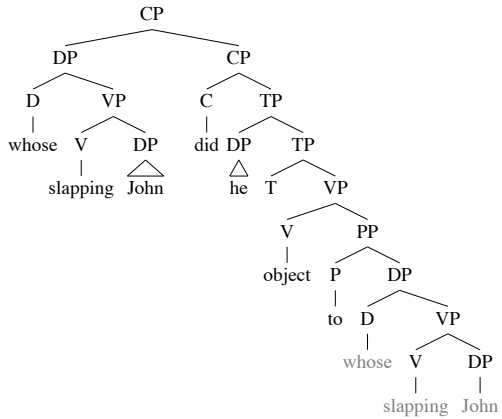
(16) (Provide the values for x that make this true:)
there is some x such that x is her paper and you said that every student had cited that x, which is her paper.

(Recall that we are focusing on that part of the meaning of questions that equates them to existentially quantified statements. I've framed the question part of the meaning in parentheses.)

This theory extends correctly to many cases. For example, it captures the fact that the referent of *John* in (17) cannot be the same as the referent of *he*.

(17) * Whose slapping $John_1$ does he_1 object to?

The representation (17) gets is (18).

(18) 

That's because names cannot corefer with terms that c-command them:

(19) Principle C
A name-like expression, α, is taken to refer to a different individual than any expression that c-commands α.

α c-commands β iff every phrase that contains α also contains β, and β is not in α.

(20) * He objects to my slapping John.
$he \neq John$

But it also misfires in certain cases. In particular, it leads to the expectation that there should be a Principle C violation in (21) as well.

(21) Which picture behind John does he object to?

But this is incorrect: *he* can refer to John in (21). In (21), we want the lower copy to have less in it than the spoken copy does. We want it to exclude *John*.

What's the difference between (19) and (21)? It looks like it should be connected to the difference between the examples in (22).

(22) a. * He objects to my slapping.
b. He objects to some picture.

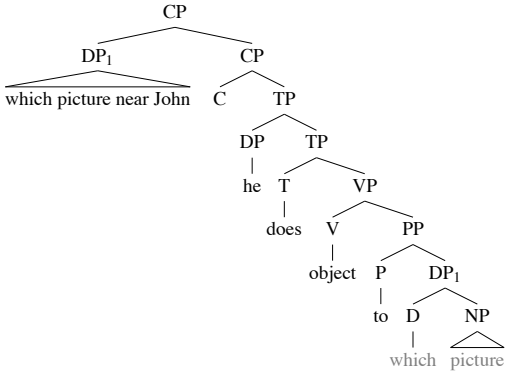
The expression "X's slapping" is semantically incomplete; the verb *slapping* requires an object. By contrast, the expression "D picture" is semantically well-formed. If we let copies differ slightly, then we could potentially let the semantics rule out situations where a copy is not complete enough. That this is what makes these two cases different is, essentially, the idea in Lebeaux (1988, 1990).

So, let's weaken our definition of Movement to:

(23) Movement (second revision)
Put copies of XP that are as identical as the semantics requires in positions α and β, and give each the same index.

This will allow (21) to get the representation in (24).

(24)



This places *John* in a position that *he* does not c-command, and Principle C is thereby not violated if *John* and *he* corefer. The Trace Conversion rule produces a meaning for (24) that could be paraphrased as (25).

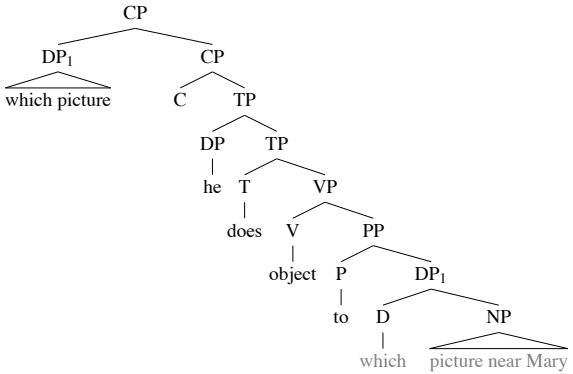
(25) (Provide the values for x that make this true:)
There is some x such that x is a picture near John and he objects to that x, which is a picture.

That's the right meaning.

There's one last fix required of our definition of Movement. As it stands, it would permit a sentence like (26) to have the representation in (27).

(26) Which picture does he object to?

(27)



Trace Conversion would produce a meaning for this sentence that could be paraphrased as (28).

(28) (Provide the values for x that make this true:)
There is some x such that x is a picture and he objects to that x, which is a picture near Mary.

That is, the representation in (27) would produce a meaning that makes (26) roughly equivalent to the meaning (29) has.

(29) Which picture near Mary does he object to?

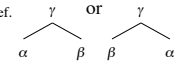
What we want to do to prevent this is find a way of ensuring (30).

(30) The material in a higher copy always contains all of the material in a lower copy.

One way of deriving (30) is from the way in which syntactic representations are constructed.

Chomsky (1995) suggests that phrase markers should be constructed from a simple, recursive, grouping operation, which he calls "Merge." This idea has enjoyed some success in modeling how syntactic operations interact, so I suggest we adopt it. I'll define Merge with (31).

(31) Merge(α)(β) =$_{\text{def.}}$



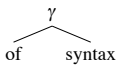
Which of the two linear orders of α and β are created by Merge will depend on what α and β are. We'll look at what's involved in making this choice in the third lecture. Until then, I will simply make the right choice when illustrating Merge. The value that γ has will depend on properties of α and β as well. I won't examine this part of the process in these lectures, and will just do the right thing. So, to illustrate how Merge produces syntactic representations, consider how it would create the DP in (32).

(32) the story of syntax

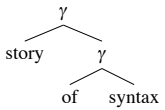
We take the words in (33) and recursively apply Merge as in (34).

(33) story, the, syntax, of

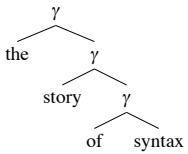
(34) a.



b.

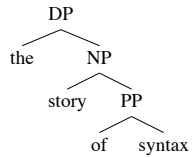


c.



If the values for the γs in (34) are filled in, we'd emerge with (35).

(35)



To derive (30), we restrict what the copying process can apply to.

(36) Copy(α) produces α', an exact copy of α.
  a. α' may be an argument of Merge.
  b. α must be produced by Merge.

We can now define Movement in terms of Copy and Merge:

(37) A Movement relation is established by giving the output of Copy to the input of Merge.

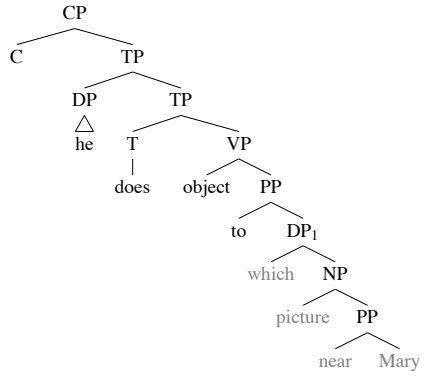
Consider how this will apply to a sentence like (26) which assembles the formatives in (38).

(26) which picture does he object to which picture near Mary?

(38) C, does, he, to, object, picture, near, Mary, which

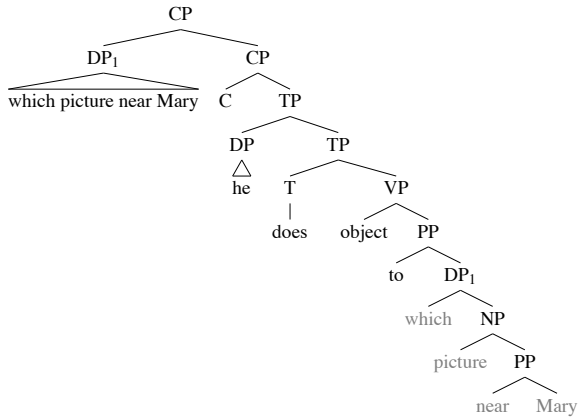
Running Merge on (38) will eventually produce (39).

(39)



If Copy applies to *which picture near Mary*, and Merge groups the output of Copy with the CP in (39), we'll end up with (40).

(40)



Note that Trace Conversion requires the higher copy to bear an index identical to the lower copy. If that doesn't happen, then Trace Conversion can't apply, and there will be not other way of giving the structure an interpretation. As a result, Copy will only be able to apply to phrases that have indices. If these are only DPs, then Copy cannot apply to some subpart of a DP. This is why sentences like (26) cannot have the representation in (27).

Let's consider, now, how this analysis of Movement will create the representation we've given to (21).

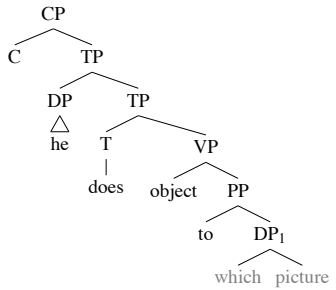
(21) Which picture behind John does he object to?

This will be constructed by Merge and Copy applying to the collection of formatives in (41).

(41) pictures, behind, which, John, C, does, he, to object

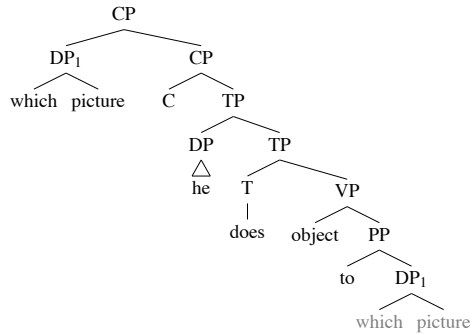
What we want to do is create a representation in which the lower copy of *which picture near John* is just *which picture*. First, Merge applies until we have created the representation in (42).

(42)



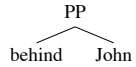
Then Copy applies to *which picture* and Merge runs on the result. From this, we get (43).

(43)



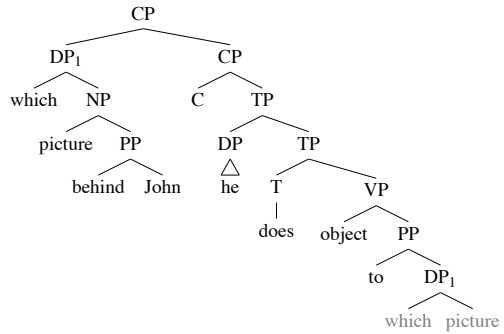
Then we will run Merge independently on *behind* and *John* to form (44).

(44)



Finally, we let Merge bring together the PP in (44) with the higher copy of *picture* in (43). The result is (45).

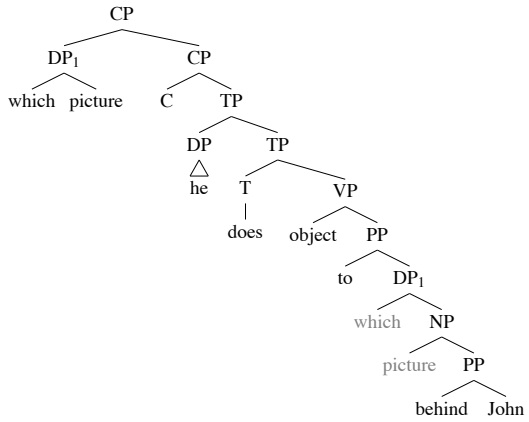
(45)



Trace Conversion applies to this and produces the desired interpretation.

This derivation requires that Merge be able to construct independent phrases along the way to producing a single sentence. It also requires that Merge be able to bring these two independent phrases together. And finally, it requires that Merge be able to apply to material within a phrase it has already constructed. This last property is dangerous. If left unconstrained, it will badly misfire. It could, for instance, produce (46) instead of (45) in the final step of this derivation.

(46) * Which picture does he object to near John?



Derivations that involve Merge bringing two independent phrases together in the way just illustrated are said to involve "Late Merge" or "Late Merger." We will require some constraint on Late Merge.

So this will be our starting point. And here are some of the questions it raises.

(47) a. What constrains Late Merge?
b. What is responsible for allowing only one of the copies to be pronounced?
c. What determines which copy is pronounced?

References

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

Barwise, John, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Ken Hale and Jay Keyser, 1–52. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995. Bare phrase structure. In *Government binding theory and the minimalist program*, ed. Gert Webelhuth, 383–439. Oxford: Oxford University Press.

Cresti, Diana. 1995. Extraction and reconstruction. *Natural Language Semantics* 3:79–122.

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

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

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

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

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

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

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

Lebeaux, David. 1990. Relative clauses, licensing, and the nature of the derivation. In *Proceedings of the North East Linguistic Society 20*, ed. Juli Carter, Dechaine Rose-Marie, Bill Philip, and Tim Sherer, 318–332. University of Massachusetts at Amherst: Graduate Linguistic Student Association.

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

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.

Sauerland, Uli, and Paul Elbourne. 2002. Total reconstruction, PF-movement, and the derivational order. *Linguistic Inquiry* 33:283–319.

Sharvit, Yael. 1998. Possessive wh-expressions and reconstruction. In *Proceedings of the North East Linguistic Society*, ed. Pius N. Tamanji and Kiyomi Kusumoto, 409–424. University of Toronto: Graduate Linguistic Student Association.