Why Movement?

Kyle Johnson
University of Massachusetts at Amherst

April 5, 2009

There is a certain set of locality conditions that seem to hold just of movement operations. Some of the islands described in Ross (1967) appear to be of this kind.

(1) Complex NP Constraint
   a. Every linguist$_1$ reviewed [the claim that this interested her$_1$].
   b. * Which linguist$_1$ did you reviewed [the claim that this interested t_1]?

(2) Sentential Subject Condition
   a. No linguist$_1$ realizes that [to understand her$_1$] isn't easy.
   b. * Which linguist$_1$ did you realize [to understand t_1] isn't easy?

(3) Adjunct Condition
   a. Some linguist$_1$ slept [while another linguist spoke to her$_1$].
   b. * Which linguist$_1$ did you sleep [while another linguist spoke to t_1]?

(4) Coördinate Structure Constraint
   a. Some linguist$_1$ [talked about her$_1$ work] and [ate lunch] (at the same time).
   b. * Whose work$_1$ did she [talk about t_1] and [eat lunch] (at the same time).

What property of movement is responsible for making it susceptible to these islands? It's clear that it's not the binder-variable relationship that movement establishes, because that is not subject to island constraints when it is expressed with a pronoun, as the a-examples illustrate. Instead, it is something about how movement establishes this binder-variable relationship that islands seem to interrupt.

My goal in this paper is to identify what properties of movement are the ones responsible for making it uniquely sensitive to islands. The bulk of the paper, then, will be in coming up with a good model movement. I will be guided by the following two idiosyncratic properties of movement.

- Typically, only one part of the binder-variable pair is spoken. Let's call this the "silence effect."
- Part of the moved phrase can be semantically interpreted as if it hadn't moved. Let's call this the "reconstruction effect."

I will sketch a way in which these two properties of movement can be traced back to the same source. I will then suggest that this source is also responsible for making movement susceptible to islands. The source is multidominance.

1 Deriving the Silence Effect from Multidominance

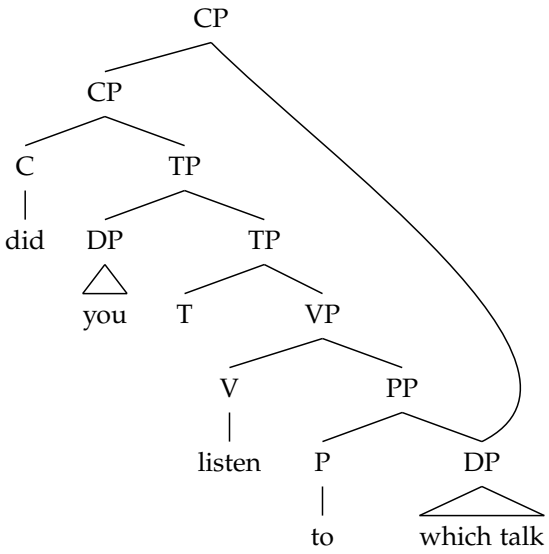
Nunes (1995, 1996) has suggested that the silence effect should be derived from the combined effect of the two assumptions in (5).

(5) a. When a phrase is moved from position α to position β, it occupies both of those positions simultaneously.

b. The linearization algorithm that converts a syntactic representation into a string cannot tolerate putting something both before and after something else. (The linearization algorithm creates strings that are both transitive and irreflexive.)

(5b) is a commonplace assumption about how linearization algorithms behave. (5a) is a stipulation. It could be derived if we take movement to assign to one term two positions. This means we should allow syntactic representations to give one phrase more than one mother – that is, they should allow for multidominance – an idea that has been entertained by a variety of people.[1] This would give to a simple instance of wh-movement, say, a representation something along the lines of (6).

(6) Which talk did you listen to?



1 See McCawley (1982), Blevins (1990), Nunes (2001), Epstein (1995), Starke (2001), Frampton (2004), Fitzpatrick and Groat (2005), Gärtner (1997, forthcoming), de Vries (2007) and Citko (2005)

If the linearization algorithm runs in its normal way, it will calculate a linear position for every terminal that reflects the position(s) that terminal is in. Because *which* and *talk* are in two positions, it will cause *which* and *talk* to both precede and follow *listen* (as well as every other terminal in (6)). Multidominant representations like these, then, are incompatible with a linearization algorithm that assigns a terminal a linear position according to every structural position it occupies. If we relax the linearization algorithm so that it assigns a linear position to a terminal by looking at any position it occupies, then the consistency requirement will produce the silence effect when these structures are linearized. A term that occupies two positions will be forced, in other words, to be linearized in only one of the two positions it occupies. Which position will be determined by other factors, and could well vary from language to language.

2 Deriving the Reconstruction effect from Multidominance

Some illustrations of reconstruction are in (7).[2]

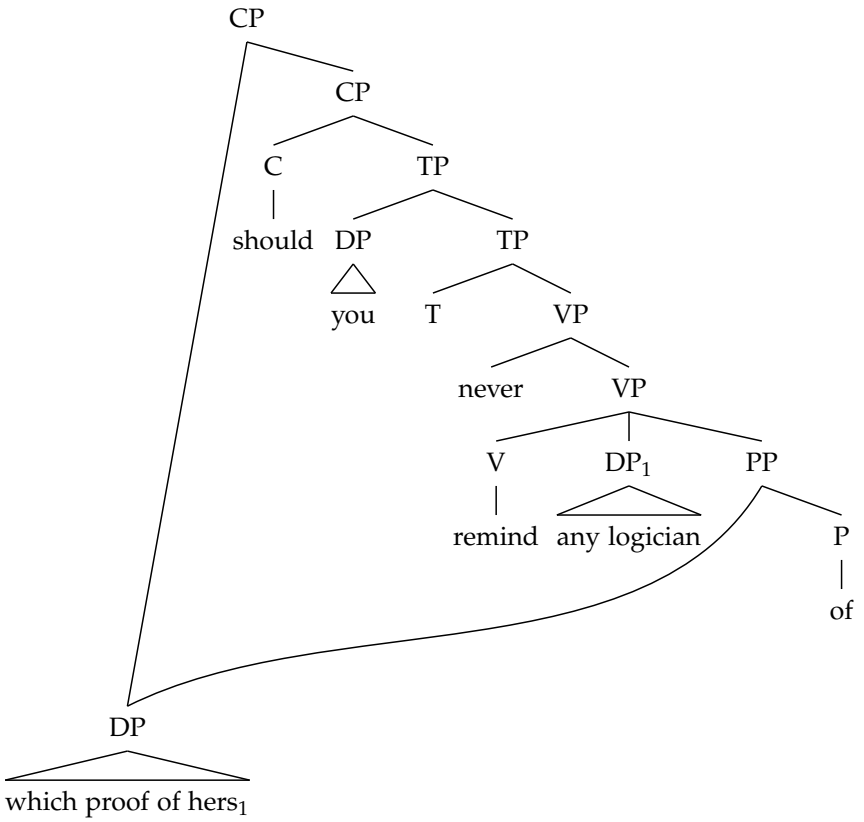
(7) a. Which proof of hers$_1$ should you never remind any logician$_1$ of *t*?
b. * Which proof of hers$_1$ did your description of every logician$_1$ portray *t* best?
c. Which picture of herself$_1$ did Sally$_1$ regret *t* most?
d. * Which picture of herself$_1$ did Sally$_1$'s brother regret *t* most?

In (7a), *hers* can be understood as a variable bound by *any logician*. This isn't made possible by *her* falling within the scope of *any logician* in its spoken position. The contrast with (7b) indicates, though, that it is made possible by the fact that the *which proof of hers* phrase has moved from a position that would put *her* in the scope of *any logician*. A similar conclusion can be reached by considering how Principle A of Chomsky's (1981) Binding Theory is satisfied in (7c). The anaphor *herself* falls within the scope of *Sally* in (7c), but not (7d), if its position before being moved is considered.

What is needed to capture these effects is a definition of movement that allows the *proof of hers* and *picture of herself* phrases to be semantically interpreted in the positions they have moved from. Chomsky (1992) suggests that movement generates "copies" of phrases, gives those copies different positions in a phrase marker, and then other principles determine which of those copies are pronounced and which are semantically interpreted. Multidominant representations are one way of conceiving of copies, and, as can be seen in (8), they take us a step towards capturing the reconstruction effects indicated in (7).

2 See Engdahl (1980, 1982, 1986), Barss (1986), Heycock (1995), Fox (1999) and many others.

(8)



As we've seen, the constraints on linearization will require that *which proof of hers*$_1$ be overtly expressed in one or the other of its structural positions. In this case, English chooses the Specifier of CP position. For the purposes of semantic interpretation, however, the *proof of hers*$_1$ phrase must be interpreted in the lower of its two positions if *hers*$_1$ is to be interpreted as bound to *any logician*. A multidominant representation like this one allows *hers*$_1$ to simultaneously be in one position (where it is pronounced) and another (where it is semantically interpreted). If movement necessarily involves multidominant representations like these, then both the Reconstruction Effect and the Silence Effect are derived.

3 Trace Conversion

While (8) provides a representation for (7) that correctly derives the Reconstruction effect from a standard semantics of variable binding, it does not offer a transparent way of giving the right denotation for the question. We should expect *which proof of hers* to provide a single denotation that is computed at its higher, Specifier of CP, position or at its lower, complement of *of*, position or in both of these positions. There is evidence, however, that we want the material in the lower of these two positions to act semantically like a variable, and the material in the higher of

these two positions to act semantically like the binder of that variable. We do not want the phrase that is assigned to these two positions to get the same meaning. In Engdahl (1980, 1986) and Fox (1999), this is achieved by a special semantic rule that gives the phrase in the lower position a special meaning. Fox calls his version of this rule "trace conversion." Trace Conversion converts the lower phrase into the kind of variable that traces are traditionally associated with. In both Fox and Engdahl's work, the result has been to form in the lower position something that has the attributes of an "anaphoric" definite description, like that which is found in sentences such as (9).

(9) Every farmer that owned a donkey fed the/that donkey.

Such an approach would allow the *which proof of hers* phrase in (8) to have the denotation of a quantifier in its higher position, and the denotation of a bound definite description in its lower position. A paraphrase that makes explicit how this method would give the desired question-meaning to such cases is (10).

(10) Which proof of hers is such that you should never show any logician that proof of hers?

We would have to find a way of ensuring that the interpretation of *that* in (10) is as a bound definite. Its values will have to co-vary with those given to the *which*-phrase.

In this paraphrase, *proof of hers* makes a semantic contribution in both its higher and lower positions. This is arguably mistaken. There is no straightforward way to ensure that *her* in the higher position is interpreted as the bound variable it is in the lower position. That problem can be sharpened in cases, like (7c), in which the bound variable is an anaphor. An account which merely converts the denotation of the phrase in its lower position to something variable-like would deliver a meaning whose paraphrase would cause the anaphor to be interpreted in the higher position. (7c), for instance, would get the paraphrased meaning in (11).

(11) Which picture of herself is such that Sally regretted that picture of herself most?

Interpreting the anaphor in the higher position should cause a violation of Principle A.[3] One of our goals, then, should be to let this part of the phrase not be semantically interpreted in the higher position.

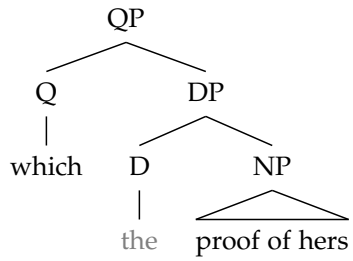
I feel we should also do away with a semantic rule that stipulates how a moved phrase is interpreted in each of the positions it resides in. This part of the account belongs to a framework that does not draw a line between syntactic and semantic processes. Rules like Trace Conversion simultaneously change syntactic representations and the denotations of the terminals in those representations. I wish to

3 That Principle A is satisfied with respect to where an anaphor is semantically evaluated is argued for in Fox and Nissenbaum (2004).

avoid rules with this much power. I want to do this for the simple – I hope not simplistic – reason that it strikes me as more constrained. On the frameworks I favor: terminals have fixed denotations, syntactic rules determine how those terminals combine, and semantic rules determine how those combinations yield phrasal denotations. I do not know how to constrain departures from these roles and, in any case, regard this as a natural starting hypothesis. For these reasons, I will design an account that mimics the effects of a Trace Conversion rule, but does so without the rule.

We have two desiderata. We must allow a moved DP to get a variable-like interpretation in its lower position. And we must allow a moved DP to be partially ignored by the semantics in its higher position. To meet the first of these goals, I suggest that we follow Matthewson (2001) and allow (some) quantificational DPs to have hidden within them definite descriptions. A phrase like *which proof of hers*, for instance, can have the representation in (12).

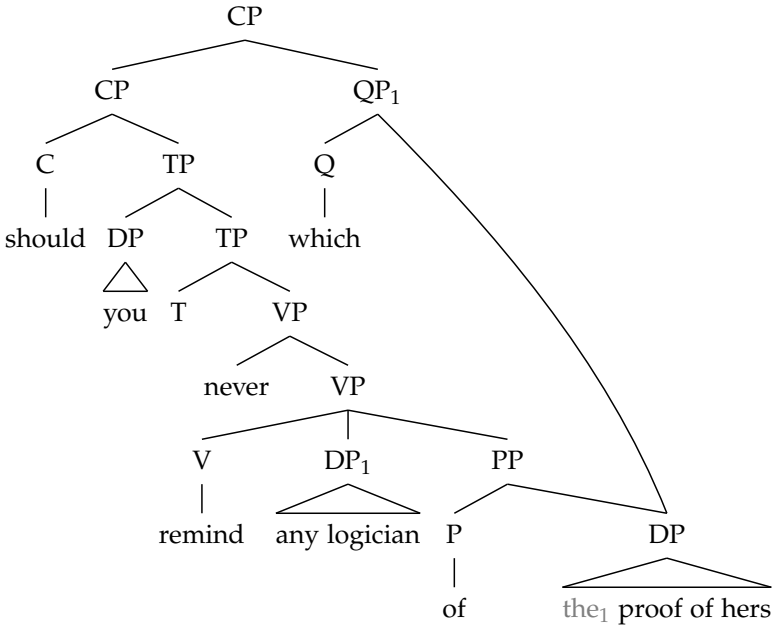
(12)



For Matthewson, who focused on certain non-interrogative quantifiers, the hidden definite determiner had the function of accounting for how quantification is contextually constrained. That role could also be taken up by the hidden *the* in our question phrases. In addition, however, it produces a phrase that can serve the role of variable in movement contexts. Rather than (8), then, let movement put the phrase headed by *the* in two positions, and put *which* only in the higher position. The representation for (7a) would now be (13) on the facing page. If *the proof of hers* gets the anaphoric meaning that definite descriptions can get, then it will be capable of serving the role of variable here, and it will do so in a way that captures the reconstruction effect. We also still derive the silence effect; the lower position does not get mapped onto anything phonological because all of the material in that position gets mapped onto a different position in the string. All that remains is to ensure that the anaphoric *the* which heads *the proof of hers* is bound by *which*, and nothing else. I will assume that this is achieved in the conventional way: by giving *which* and *the* the same referential index, as indicated in (13). Let us assume that the meaning assigned to *the*$_1$ *proof of hers* is something that can be paraphrased as "that 1 which is a proof of hers."[4] The index "1" will end up

4 Let *the* combine with the index first, and the resulting phrase then combine with the NP. Then the denotation for *the* could be: "$\lambda n \lambda P.n$, if P holds of the value given to n." This would give to the expression *the*$_1$ *proof of hers* the value given to "1," and introduce the presupposition that this value

(13) Which proof of hers should you never remind any logician of? (=(7a))



being a variable that ranges over the things that the question is an inquiry about. We have produced the right sort of variable-like denotation in the lower position: we have satisfied the first desideratum.

Our second goal is to let the denotation of the shared DP – in (7a), *the proof of hers* – be able to be semantically uninterpreted in the higher of its positions. There are a variety of ways this could be achieved. Choosing among them involves, among other things, determining the proper semantics for questions. I will assume a Hamblin semantics for questions, and take a simple, and perhaps overly strong, direction and give *which* a denotation that prevents it from combining semantically with *the proof of hers*. This is what's done in Sauerland (1998), though my execution of the idea follows the details of a handout I've come upon by Irene Heim from a course she taught with Polly Jacoboson at the LSA Summer Institute in 2006. The idea is that *which* combines with a predicate and introduces an existential quantifier that closes that predicate. The predicate it combines with should be the clause from which the *which*-phrase has "moved." That clause will get assigned a denotation that makes it a predicate with a variable standing in the position of the index. For our exemplar *which proof of hers should you never remind any logician of*, the clause that the *which*-phrase combines with will have the denotation paraphrased by (14).

is a proof of hers. See Elbourne (2005).

(14) $\lambda 1$ you should never remind any logician of that 1 which is a proof of hers

Let the denotation of *which* be (15a), and it can combine directly with (14) to form the denotation in (15b).

(15) a. $[\![\text{which}]\!] = \lambda P.\exists x.P(x)$
b. $\exists 1$ you should never remind any logician of that 1 which is a proof of hers

The question morpheme (silent in English) we can assume converts (15b) into the set of propositions that is a Hamblin treatment of questions. The result is something like (16).

(16) $\lambda p\ \exists 1\ p =$ you should never remind any logician of that 1 which is a proof of hers.

Because *which* is designed to combine directly with the clause that follows, it cannot combine semantically with the DP that is its syntactic complement. This will have the desired effect of forcing that DP to only be interpreted in the lower of its two positions.

This is the general approach I will take. The syntax allows some phrases to occupy more than one position in the phrase marker. This is what movement is. We will need to require that those phrases are semantically interpreted in at least one of their positions. Let's assume that this is a result of a Principle of Full Interpretation, formulated for multidominant representations in (17).

(17) Principle of Full Interpretation

Every term in a phrase marker must combine semantically with at least one of its sisters.

The semantics will determine which of those positions a phrase is semantically interpreted in. Certain DPs, for example, are constructed in such a way that their parts can be distributed across several positions and a semantically coherent interpretation still result. This is what wh-movement is. One of those parts is pressed into the role of a variable, and the other plays the role of binder.

4 Late Merge

The system described in the previous section has the effect of forcing all of the material of a wh-moved phrase to be semantically interpreted in the lower position. There is evidence that this is the desired consequence. When the head nouns in moved DPs are strongly transitive, their complements trigger disjoint reference effects under reconstruction. Thus, many speakers find examples like (18) ungrammatical on a reading in which the pronoun and name corefer.

(18) a. * Whose evaluation of Jane does she agree with?
b. * Whose kissing (of) Jane does she want to forget?
c. * Whose painting (of) Jeremy blue does he remember?

But there are cases where this disjoint reference effect is lost as well. When the head noun is "concrete," or otherwise optionally intransitive, the post-nominal material tends not to trigger a disjoint reference effect. The name and pronoun can be understood to corefer in examples such as (19).

(19) a. Which evaluations in front of Jane does she dislike?
b. Whose cans near Jane did she label?
c. Which painting near Jeremy will he soon forget?

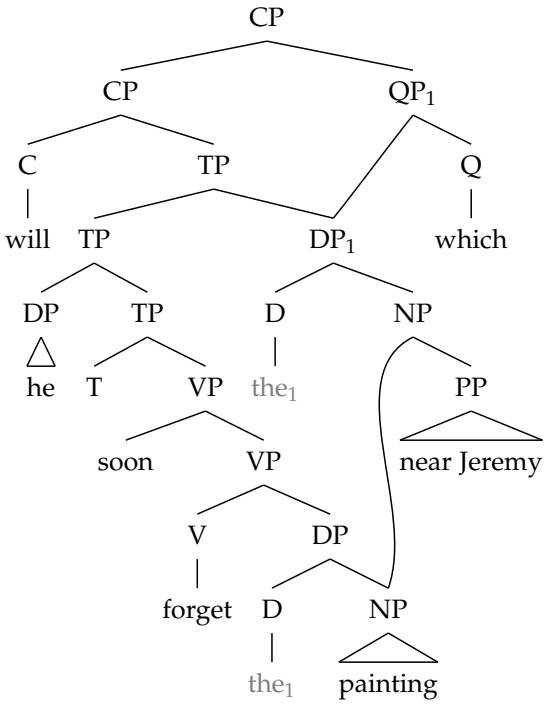
The best treatment of this contrast that I know of is based on Lebeaux (1990) and makes use of the idea that phrases can be merged into derivations after movement. This idea derives the contrast between (18) and (19) in the following way. In (19), the lower position may contain just the head noun. The post-nominal material in these examples can be merged after movement in the higher position. As a consequence, the lower position will not contain material whose denotation includes the name, and no disjoint reference effect emerges. In (18), by contrast, the lower position cannot include just the head noun. Because these nouns are transitive, they are semantically incomplete without their complements. A derivation in which these complement phrases are "late merged" in a higher position is therefore not available. The names in these moved DPs, therefore, will be part of the denotation given to the phrases in the lower positions, and this will invoke a disjoint reference effect. (See Chomsky 1995, and Fox 1999, 2003.)

Beyond allowing derivations in which phrases merge into the structure after movement has occurred, this account relies on two other premises. First, it must be possible for a phrase to be attached, and semantically interpreted, in just the higher position. Our present system does not provide for that possibility. We have adopted a denotation for *which* that does not permit it to combine semantically with its sister, thereby depriving a DP that is assigned that position from being interpreted there. And second, the Lebeaux account requires that the head noun of the moved DP be forced to be semantically interpreted in the lower position. This could be achieved if the meaning we assign to the silent *the* requires a nominal complement as its argument. Because the silent *the* must be put in the lower position, the head noun will be required in the lower position as well. Let's adopt this method of forcing the head noun to be interpreted in the lower position,[5] and concentrate on finding a way of letting the higher position contain semantically interpreted material.

5 The denotation for *the* in note 4 has the desired consequence. See Takahashi (2006) for an alternative approach to this issue.

In cases where Late Merge arises, I suggest that movement is "successive cyclic," thereby creating another position in which a silent *the* is placed.[6] There are a couple of ways to do this. Anticipating a proposal I will make below, let's let instances of successive cyclic movement use as many silent *the*s as there are variables. For an example like (19c), then, a representation that involves Late Merge, and allows for coreference between name and pronoun, is (20).

(20) Which painting near Jeremy will he soon forget? (=(19c))



The "chain" created by movement in this representation has three parts.

(21) Parts of the chain:

- The top phrase: "which the$_1$ painting near Jeremy," has the denotation assigned to *which* and is a binder.
- The intermediate phrase: "the$_1$ painting near Jeremy," has the denotation assigned to *the*$_1$ *painting near Jeremy* and both binds and is bound.
- The lowest phrase: "the$_1$ painting," has the denotation assigned to *the*$_1$ *painting* and is bound.

The index on QP, the top phrase, causes it to be a binder and, following the conclusion of the preceding section, has just the denotation assigned to *which*. It is in

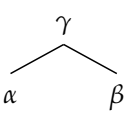
6 An alternative would be to let *which* have a flexible enough denotation for it to semantically combine with its sister when that is necessary (as in (19)) and avoid doing so with it isn't possible (as in (7)).

this position that the linearization rules of English will cause all of the shared material in this phrase to be spelled out. The intermediate phrase also has an index on it, and this will cause it to be a binder as well. In addition, the silent determiner heading this phrase will also have an index, and this index gets interpreted as a variable bound by the top phrase. Again, following the conclusions of the preceding section, this phrase will get a denotation whose paraphrase is: "that 1 which is a painting near Jeremy." Finally, the lowest phrase only has an index on the determiner, causing this phrase to be a variable bound by the intermediate phrase. The denotation the lowest phrase gets can be paraphrased as: "that 1 which is a painting." When the denotations of these three parts are knit together, the right meaning emerges. And because *Jeremy* is interpreted in a position outside the scope of *he*, coreference between the two is allowed.

A parallel representation is not available to (18), however. If the material in the lowest copy does not include the complement to the head noun, then it will be semantically incomplete. The postnominal material must, in these examples, be part of the phrase in the lowest position and this will put the name in the scope of the pronoun, causing a disjoint reference effect between them.

5 Cyclic Merge

On the definition of Movement that I have just sketched, Movement devolves entirely to Merge, an operation that builds syntactic structure by taking two syntactic objects and combining them to form another syntactic object.

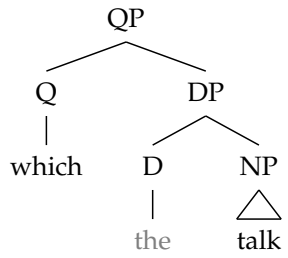
(22) Merge (α,β) $\rightarrow$ 

Movement is merely a case of Merge in which the two phrases brought together are already part of the same phrase marker. As the cases of Late Merge illustrate, these special instance of Merge (known as "internal Merge") can be interleaved with those instances of Merge that bring new material into a phrase marker (known as "external Merge"). If we take a derivational view of how Merge is deployed, then phrase markers involving internal Merge (i.e., Movement) will transit through representations in which the tree has more than one root. To see this, consider the derivation that arises from cyclic applications of Merge in forming (23).

(23) Which talk should I finish?

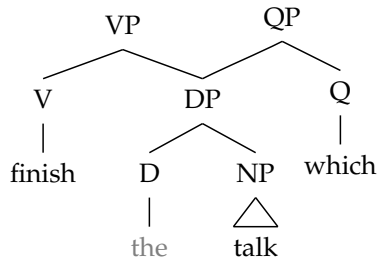
Imagine that the phrase that is to be moved is constructed first. This phrase is (24).

(24)



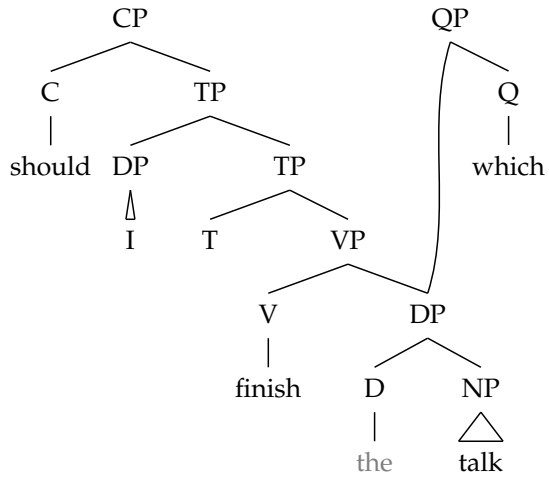
The semantics will require that this phrase be merged with CP to form the question. That CP, in turn, will have to include the DP, *the talk*, to which the verb *finish* has merged. The first step in creating that CP, then, will require Merge to combine *finish* and *the talk*. This result is (25).

(25)

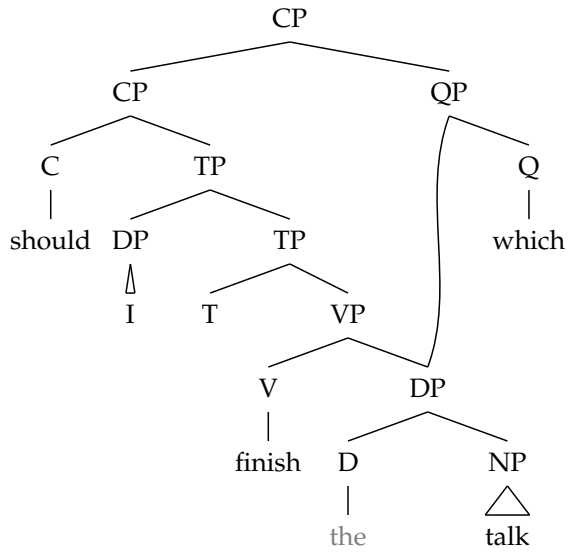


This phrase marker has two roots, as will every subsequent phrase marker until the CP that the QP needs to Merge with has been constructed. Merge will apply recursively to the VP root until (26) is formed, from which (27) can be created by Merge applying to the two roots of (26).

(26)



(27)



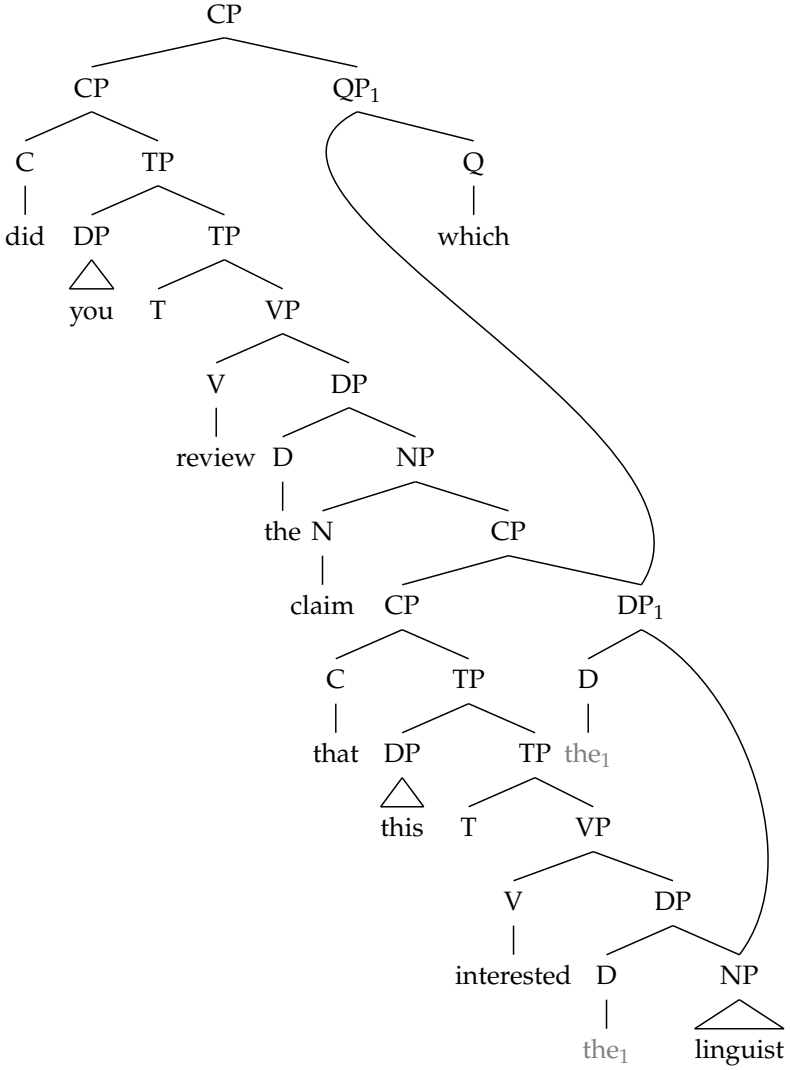
There are other ways of ordering the applications of Merge. But under every ordering, there will be at least one step in the derivation in which the phrase marker has two roots. That is made necessary by the twin goals of creating a VP that is made up of just the verb and the DP headed by *the*, and putting that DP in a QP that merges with some phrase larger than the VP. This is quite general. Any instance of wh-movement – those that involve Late Merge, those that are successive cyclic – they will all create structures in the course of the derivation that have more than one root. This, I propose, is why movement uniquely triggers island effects.

6 Islands

Let us consider what representation this account of movement would give to an island violation. A Complex NP Constraint violation, for example, might look like (28).[7]

7 This representation arises from a derivation in which the wh-phrase has moved successive cyclically through the Specifier of the embedded CP.

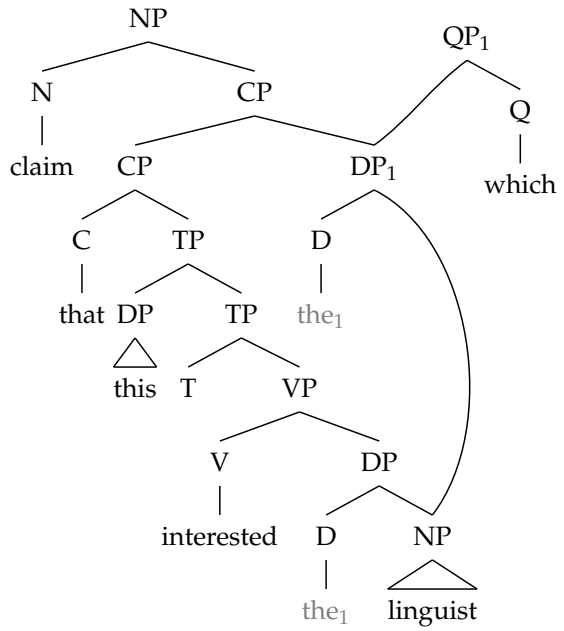
(28) * Which linguist did you review the claim that this interested? (=(1))



This representation differs significantly from one in which a pronoun, or other non-movement created variable, is bound across an island. We might now exploit this difference in seeking the cause for Movement's proprietary sensitivity to islands.

But what is that cause? As the system is presently designed, there is nothing wrong with the syntax or semantics of representations like (28). Consider, however, a stage in the derivation before the wh-phrase has merged into its higher, spoken, position. Suppose that the island in (28) is the CP complement to the noun *claim*; let's look at that point in the derivation in which that island has just been merged into the phrase-marker.

(29)



At this stage, the position that the QP will merge into has not been created, and so the phrase marker has two roots. It is plausible that phrase markers with two roots are deficient. From a semantic point of view, double-rooted phrase markers would deliver two, disjointed, denotations. And, similarly, from the point of view of Spell Out, the linearization rules plausibly produce from a double-rooted phrase marker not one string, but two disjointed ones. If islands trace back to constraints that must be met at certain stages of a derivation, this may be the reason that movement structures are uniquely susceptible to them.[8]

For instance, imagine that certain islands are the result of running cyclically the algorithm that maps structures onto strings, as in, for example, Fox and Pesetsky (2004). The proposal in Fox and Pesetsky (2004) goes as follows. At certain points in the derivation there will be created phrases that cause the linearization algorithm to run and fix the relative linear positions of every terminal within

8 In the previous section, I observed that the multidominant model of movement I am proposing here requires that at some point in the derivation a multi-rooted tree will be created. If it is the multi-rooted structures that are to blame for islandhood, then it will be necessary that these multi-rooted structures appear in the derivation at the right point. They must be present when the island environment is encountered. Nothing I have said so far will ensure that. There are a couple ways of doing this. One would be to require the QP to be built before its parts merge with other material. (This is the kind of derivation I went through in the previous section.) To the extent that Q selects DP, this could be seen as the result of Chomsky (1981)'s Projection Principle. Another possibility would be to rely on Chomsky (1995)'s Extension Condition, which only allows root nodes of phrases to be arguments of Merge. If this direction is pursued, we would have to let Merge bring together more than two terms at a time. For instance, the shared NP (*linguist*) in (29) would have to Merge simultaneously with the lower *the* and the higher *the*. I don't have the space here to examine this problem further.

that phrase. Let's call these phrases: "phases."[9] Subsequent instances of internal and external Merge will produce additional material that, when submitted to the linearization algorithm, must not conflict with these assigned positions. So, for instance, if a phrase, α, is determined to follow a term, β, at some phase, then subsequent operations may not lead to a linearization that would put α before β. This system is designed to capture the fact that the phrases which move out of phases are just the phrases that are at the left (or right) most edges of those phases. Imagine, for example, that CPs are phases. Then the system in Fox and Pesetsky (2004) will require that only that phrase which is left-most in the CP will be capable of leftward movement out of that CP. In this way, the successive cyclic movement out of CPs through their Specifier positions is derived.

As it stands, the procedure in Fox and Pesetsky (2004) will apply to the representation in (29) without a hiccup. This is because their algorithm merely collects pairwise ordering statements for each terminal within the phrase being linearized. In (29), those statements would consist of statements that constrain how QP is arranged (namely: *which* < *the*, *which* < *linguist* and *the* < *linguist*) and how CP is arranged (namely: *linguist* < *that*, *the* < *that*, *that* < *this*, and so on). When QP is later merged in a higher position, the new ordering statements that are acquired won't conflict with these, as long as QP is linearized to the left of everything else in the relative clause. The algorithm in Fox and Pesetsky (2004) will ensure successive-cyclic movement out of phases, then, but it doesn't produce island effects.

To produce the island effect, I suggest that a stronger condition on the structure-to-string algorithm is imposed at islands. Imagine that islands are those phases at which the ordering statements gathered under the algorithm in Fox and Pesetsky (2004) are converted into actual strings. This is close to the proposal about certain islands in Uriagereka (1999) and more or less what I proposed here five years ago.[10] Let's call this process "string formation." We can picture it as a procedure that produces from a set of ordered pairs a string of concatenated terms, as indicated in (30).

(30) String Formation

$$\left\{ \begin{array}{c} \alpha<\beta \\ \gamma<\delta \\ \beta<\delta \\ \alpha<\gamma \\ \beta<\gamma \\ \vdots \end{array} \right\} \rightarrow \alpha + \beta + \gamma + \delta$$

If String Formation is a process that applies to phrasal nodes – that is, it assigns a string to a phrase based on the ordering statements gathered from that phrase – then it will produce an anomalous output from (29). It will assign one string to

9 See Chomsky (2001).
10 See Johnson (2004).

the highest CP and an independent string to QP. If islands are just those phases at which String Formation is triggered, then these double-rooted representations will produce an anomaly.

This is the merest of sketches. It requires considerable filling it before it can be fully evaluated. We'll need to know why producing two strings at a phrase is anomalous, and, of course, we should find an answer to the larger question of why the configurations at which islands are formed are the ones they are. I don't have an answer to that question, nor am able to know how to settle on the details of executing the idea. What I hope to have shown, however, is where this model of movement directs our search for answers to these questions. It provides a way of understanding why Movement is the operation that sustains island violations.

I have sketched how this approach would work for islands that devolve to the workings of the structure-to-string algorithm. But a similar sketch could be manufactured for islands that devolve to semantic constraints, should there be such. Not only will double-rooted representations like (29) fail to support a single string, they also fail to support a single, well-formed, denotation. If there are points in the derivation at which single, well-formed denotations are required, then those will be points that representations like (29) fail to deliver what's required. I can also imagine that these representations would dovetail with processing-based accounts of island effects. I am not well-enough schooled to know how to predict what the particular on-line processing algorithms that are found in the literature would produce when applied to these representations. But it is conceivable that the extra complexity posed by these multi-dominant representations would present unique processing problems for these algorithms.

The answer I've tried to give to the question of why island effects are found with movement, but not with other forms of long-distance variable binding, allows islands to have different, unrelated, causes. We'll only know if this is a virtue of the proposal once we know what the sources of islands are.

References

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

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

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.

Chomsky, Noam. 1992. *A minimalist program for linguistic theory*, volume 1. MIT, Cambridge Massachusetts: MIT Working Papers in Linguistics.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, ed. Michael Kenstowicz, 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.

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. 1982. A note on the use of lambda conversion in generalized phrase structure grammars. *Linguistics and Philosophy* 4:505–516.

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

Epstein, Samuel. 1995. A derivational approach to syntactic relations. Unpublished manuscript, Harvard University.

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. 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.

Fox, Danny, and Jon Nissenbaum. 2004. Condition A and scope reconstruction. *Linguistic Inquiry* 35:475–485.

Fox, Danny, and David Pesetsky. 2004. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–46.

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

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

Gärtner, Hans-Martin. forthcoming. *Generalized transformations and beyond*. Berlin: Akademie-Verlag.

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

Johnson, Kyle. 2004. How to be quiet. In *Proceedings from the 40th Annual Meeting of the Chicago Linguistics Society*, ed. Nikki Adams, Adam Cooper, Fey Parrill, and Thomas Wier, 1–20. Chicago, Illinois: Chicago Linguistics Society.

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.

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

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

Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

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

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

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

Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Samuel Epstein and Norbert Hornstein, 251–282. Cambridge, Massachusetts: MIT Press.

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