Regarding "Minimalist Approaches to Locality of Movement" and the tucking in phenomenon

Tyler Peckenpaugh
CUNY Graduate Center, Department of Linguistics

0. Introduction

In his paper *Minimalist Approaches to the Locality of Movement*, Justin M. Fitzpatrick explores various methods of defining a constraint on locality of movement which are compatible with the Minimalist approach to syntax. He outlines five of these possible methods, the minimal requirement of which is to account for crossover effects, as in (1):

(1) a. Who bought what?
b. *What did who buy (what)?

The methods he illustrates are: C-Command Attract Closest (Com), Dominance Attract Closest (Dom), Relation Quantification (MutCom), Path Quantification (PathQuan) and Interpath Relations (SubPath). He then illustrates five phenomena which he thinks should be accounted for by a constraint on the locality of movement, and details if and how each locality constraint presented can or cannot account for such phenomena. The phenomena are A-over-A effects, the tucking-in phenomenon, head movement, movement of a more distant candidate, and freezing effects. He also makes certain theoretical considerations regarding each definition of locality, such as whether they make use of otherwise non-motivated theoretical machinery (like quantification or look ahead).

I will outline the material Fitzpatrick covers, beginning with the relevant mechanisms for movement in Minimalist syntax, then the locality phenomena he discusses, and finally I will describe each locality constraint, and which of the phenomena each constraint can account for. I will then discuss some of the concerns I have regarding the conclusions made by Fitzpatrick stemming from his treatment of the tucking in phenomenon. Specifically, I feel he did not adequately support his suggestion that tucking in be divorced from the other phenomena accounted for by a locality of movement constraint. In discussing these concerns, I pose two questions that I feel should be more fully explored before Fitzpatrick's suggestion be carried forward.

1. Movement in Minimalist Syntax

Fitzpatrick provides an overview of the mechanics of movement in the theory based on the Minimalist Program (Chomsky 1995). I will briefly outline those same points.

Movement is motivated by feature attraction[1]. This is formalized as follows:

(2) *Attract*
A head H bearing an uninterpretable feature F attracts the closest category G in the sister of H that bears a feature matching F, where (a) matching is feature identity, and (b) G moves to the checking domain of H.

The job of a locality constraint on movement is to define *closest* in the formulation of Attract.

A syntactic tree is built using iterative application of Merge:

(3) *Merge*
Take two syntactic objects A and B and replace them with the set K whose sole members are A and B (i.e. Merge(A,B) = K = {A,B}).

The process of movement is done using Move[2], defined below:

(4) *Move*
A and B are terms of a category E. Form E' by replacing A in E with the set L whose sole members are B and A (i.e. Move(A, B) = L = {B,A}).

The use of these methods has also lead to a modification of the notion of c-command. Following Epstein (1999), Fitzpatrick rejects the notion of c-command as a stipulated structural relationship, and instead derives it from the applications of Merge and Move in building a given structure.

(5) *Derived c-command* (Epstein 1999)
X c-commands all and only the terms of a category Y with which X was paired by Merge or Move in the course of the derivation.

In this method, c-command is a relationship between syntactic objects and not between particular positions.

It is also important to note the way in which moved material is treated in this

[1]The Extended Projection Principle (EPP) may be an exception.
[2]Fitzpatrick uses Move distinct from Merge for expository purposes, and I will do the same, although ultimately Move is not needed.

theory. Fitzpatrick says that movement is treated “as remerger […] leaving the moved category itself unduplicated.” A given category might occur twice in a structure, but it exists as just one syntactic object which occupies both positions.

An interesting result of this non-duplicative movement is that when material is moved, it enters into mutual c-command or dominance relationships from its higher position with material that c-commands or dominates its lower position. Fitzpatrick makes use of this fact, and I will discuss it in more detail later.

2. Empirical evidence for selecting a locality definition

In this section I will outline the five phenomena that Fitzpatrick uses in his paper to test the various proposals for defining locality of movement.

2.1 A-over-A effects

Fitzpatrick makes reference to a phenomena noted by Kitahara (1997) among others, which he calls a “generalized A-over-A effect.” I will codify this as a constraint on locality, using Fitzgerald's formulation, with a slight modification:

(6) *A-over-A effect*

A category K that dominates a category G is closer to an attracting head not dominated by K[3].

Essentially, a given element cannot be extracted out of an element which is eligible for that operation. To illustrate the effect, Fitzpatrick provides German data from Kitahara (1997) in the form of two derivations, (7-8) below. (8) is illicit and should be ruled out by whatever definition of locality is chosen; (8) is allowed, and Fitzpatrick claims it is because the two movements are of different nature.

Scrambling is assumed to be an instance of Attract and Move; i.e. there is some phonologically null attracting head motivating the movement. He also assumes that this type of movement is distinct from topicalization, so that whatever feature attracts

[3] I added 'not dominated by K' because a category dominating the attracting head should not be considered closer. This is perhaps redundant, because in general lowering is not allowed, but I felt it better to be thorough.

topicalized elements is not the same one which attracts scrambled elements.

(7) a. daβ keiner [G das Buch] [K (das Buch) zu lesen] versucht hat.
that no-one the book to read tried has
b. *daβ [$_G$ (das Buch) zu lesen] keiner[4] [$_K$ das Buch] (das Buch zu lesen) versucht hat.

(8) a. hat keiner [$_G$ das Buch] [$_K$ (das Buch) zu lesen] versucht.
has no-one the book to read tried
b. [$_K$ (das Buch) zu lesen] hat keiner [$_G$ das Buch] (das Buch zu lesen) versucht.

In (1a) *das Buch* is scrambled out of a constituent K. In (6b) the constituent K is scrambled to a higher position. According to Fitzpatrick, the structure in (7b) is illicit (perhaps) because of the movement in (7a): the constituent K, a candidate for scrambling, was closer to the attracting head. The derivation in (8) is permitted because the movement of K is not scrambling, but rather topicalization. In other words, in (7) G and K are both categories that were selected bearing the feature which motivates scrambling; in (8), only G bears that feature, while K bears the feature which motivates topicalization; thus K is not a candidate for scrambling.

2.2 Tucking in

The next phenomenon Fitzpatrick notes is that of "tucking in," referencing Richards (1999). This essentially appears to be a constraint on landing zone for an attracted category. Here's a slight modification of Fitzpatrick's formulation:

(9) *Landing Site Constraint*
When a head H attracts multiple categories, each movement operation inserts the category in the lowest specifier of H.

He illustrates this effect with an example of multiple wh-movement from Bulgarian, shown in (10):

(10) a. Koj kogo vizda (koj) (kogo)
who whom sees

[4]In Fitzpatrick's article, *keiner* (no-one) is omitted in the second step of the derivation. I am assuming that this was typographical error.

b. *Kogo koj vizda (koj) (kogo)
Whom who sees

Both (10a) and (10b) are derived by first moving *koj* and then moving *kogo*. A definition of locality should certainly only permit that order of movement, as *koj* is clearly the closer candidate for wh-movement. Fitzpatrick later states that perhaps this phenomenon should not be accounted for by a constraint on locality of movement after all; ultimately I agree, but I feel he does not motivate this claim and that it warrants further discussion. I will return to this idea later.

2.3 Head movement

Fitzpatrick argues that a locality constraint should account for head movement of the kind one sees in English question formation. His example follows in (11):

(11) a. Who would John (would) think that Bill could see (who)?
b. *Who could John would think that Bill (could) see (who)?

This phenomenon is fairly straightforward, and won't become interesting until I begin discussing the various candidates for defining locality. One should note that Fitzpatrick is assuming head movement results in head adjunction, so that what was originally a position X, occupied by the head of that projection, becomes a branching position, bearing both the head of that projection and the moved element.

2.4 Movement of a more distant candidate

Fitzpatrick also notes an example, attributed to Fiengo et al. (1988) where it appears that only the “more distant” candidate can move. The relevant data are in (12):

(12) a. What did people from where try to buy (what)?
b. *Where did people from (where) try to buy what?

It appears to be the case that while *where* is closer than *what* in terms of string distance, something about its position makes it farther with regard to how locality should be defined. I will return to these data in a later section.

2.5 Freezing effects

The final phenomena that Fitzpatrick notes are so-called freezing effects. The observations regarding this fact stem from a prediction made by one class of candidates for the definition of locality, known as Com and Dom, which Fitzpatrick is surprised to find actually bear out. It essentially boils down to the following:

> *Freezing Constraint*
> If an element G is in a mutual c-command[5] relationship with an element K and both are candidates for attraction by a head H, neither can be attracted.

The Spanish data illustrating this phenomenon is in (8):

(13) a. qué compró quién (compró) (qué)
what bought who
"What did who buy?"
b. *qué crees (qué) que compró quién (compró) (qué)
what think that bought who
"What do you think who bought?"

In (13a) *qué* has raised to the root CP[6], resulting in a mutual c-command relationship between *qué* and *quién*. Thus, when the clause is embedded, as in (12b), further raising is illicit.

3. Proposed definitions of locality

Fitzpatrick explores five potential formulations of a locality constraint on movement. He categorizes these into three groups based on what the constraint is looking at: intervention syntactic relations, quantified syntactic relation and movement paths. The intervention syntactic relations come in two forms, which are fairly similar to each other: C-Command Attract Closest (Com) and Dominance Attract Closest (Dom). Both are essentially based on the traditional c-command approach. The quantification syntactic relation approach is concerned with comparing the number of new mutual c-command relationships made in two possible derivations. Finally, movement path approaches are based on comparing the sets of terms between a departure site and a landing site for two

[5]The same phenomenon can be defined in terms of dominance, rather than c-command.
[6]Fitzpatrick notes that this looks like it should be a Superiority violation; however, he references Bošković (1999) and follows the analysis that the raising of *qué* over *quién* is due to a focus feature borne by *qué* and not *quién*.

possible derivations.

Besides the empirical evidence provided by the phenomena outlined above, Fitzpatrick is concerned with certain theoretical assumptions that should be considered when adding new mechanisms to the theory. One is that a grammar should not rely on quantification; there is little evidence to support counting mechanisms within a grammar. The other is about whether or not "look-ahead" should be allowed; if sentence derivation is a serial process, whether it crashes or converges should probably be based on a particular step of a single derivation, and not by comparing separate derivations.

3.1 Attract Closest: Com and Dom

Both Com and Dom formulate locality in terms of an intervening closer candidate, with the only difference being the relationship between the two candidates. I provide Fitzpatrick's definitions below:

(14) ***C-Command Attract Closest (Com)***
G is the closest category in the sister of H iff there is no distinct category K such that K c-commands G and K bears a feature matching F.

(15) *Dominance Attract Closest (Dom)*
G is the closest category in the sister of H iff there is no distinct category K such that the mother of K dominates G and K bears a feature matching F.

Recall that the "feature matching F" refers to the formulation of Attract in (2), where F is an "uninterpretable feature" borne by H.

The traditional approach is Com, and the reason for formulating Dom seems to stem from phenomena such as the A-over-A effect in (6) (which states that a category K which dominates G is closer to an attracting head not dominated by K). In such instances there is no c-command relationship between the higher category (K in the formulation), and thus Com would not block movement of the lower category (G). Dom is able to predict the A-over-A effect.

An attractive feature of these methods is that they use derived, non-stipulatory relations (derived c-command, motherhood, dominance), do not require introduction of otherwise unused or unmotivated mechanisms (such as quantification or movement paths) and do not require look-ahead of any kind.

Neither Com nor Dom can predict so-called tucking-in. While both would constrain the order in which the elements can move when there are two instances of attraction by the same head, neither can constrain the landing site of the moved elements. Both can correctly predict head-movement patterns; the closer head c-commands and dominates the more distant one, ruling its movement out. Com and Dom correctly do not rule out movement of the more distant candidate *what* in (12), because there is no dominance or c-command relationship between *where* and *what*. They cannot prevent the movement of *where* and Fitzpatrick thus suggests that the illicitness of *where*'s movement is not to be accounted for by locality, but rather by some other constraint such as Nunes and Uriagereka's (2000) Condition on Extraction Domain. Finally, Com and Dom are precisely the constraints that motivated the exploration of freezing effects, and thus can and do predict them.

3.2 Relation quantification: MutCom

MutCom is also based on c-command relationships, but rather than relying on intervention, it quantifies the number of mutual c-command relationships made for a given movement, and compares it to the number made by another potential movement. Fitzpatrick, citing Epstein et al. (1998), states that the operation which creates fewer of these relationships should be preferred, because of the level of computational complexity they represent.

(16) *Relation Quantification: Mutual C-Command*
G is the closest category in the sister of H iff there is no category K that bears a feature matching G such that movement of K would create fewer mutual c-command relations than would movement of G.

While this is certainly a well-motivated concept, it makes use of the otherwise unused mechanism of quantification, and also uses look-ahead, which might increase computational load, which it was formulated to avoid. It also does not fare as well empirically.

While MutCom easily accounts for A-over-A effects and tucking-in, it cannot account for head-movement. Fitzpatrick states “the only new mutual c-command relation formed by pairing two heads H and K is H-K,” thus it is the case that any two head movements will be indistinguishable by MutCom. MutCom also cannot predict movement of a more distant candidate as in (12): movement of *where* would create fewer mutual c-command relationships than would *what*, and so movement of *what* would be blocked. Finally, it cannot account for freezing effects: an element which creates fewer new mutual c-command relationships than alternative candidates, even if it is already in such a relationship, is allowed to move on under MutCom.

3.3 Movement path based approaches

Fitzpatrick lists two main approaches to locality which use movement paths, path quantification (PathQuan) and interpath relations (SubPath). Both rely on the concept of movement path defined here, adapted from Richards (1999) by Fitzpatrick:

> (17) *Movement Path*
> The set of terms of the position (i.e. sister) A of a category G that dominate the position B of G.

Assume that A is the landing site and B is the point of origin (actually, the lower occurrence, since "movement" is not actually movement and the syntactic object occupies both positions). The movement path is the set of projections which intervene between the landing site and the departure site.

PathQuan compares the movement path created from one movement to another, and only permits the one with fewer members. It is defined as follows:

> (18) *Path Quantification (PathQuan)*
> G is the closest category in the sister of H iff there is no category K bearing a feature matching F such that the path P' that would result from movement of K is smaller than the path P that would result from movement of G.

This is a fairly intuitive and straightforward way of defining locality, since it directly evaluates what we perceive as distance. To do so, however, it makes use of quantification as well as look ahead, which, as discussed previously, may be undesirable features.

The interpath relations approach makes a slight refinement, which actually has an interesting effect in terms of empirical predictions. Rather than comparing the size of paths, it compares the members of two paths: if one is a proper subset (contains only members of and fewer members than) of the other, it is the preferred path.

(19) *Subset Paths (SubPath)*
G is the closest category in the sister of H iff there is no category K bearing a feature matching F such that the movement path that would be created by movement of K is a proper subset of the movement path that would be created by movement of G.

This method of comparing paths eliminates the need for quantification, but still requires look-ahead.

Both path-based methods account for A-over-A effects, as it is a condition based solely on the kind of distances these methods evaluate. Both methods also predict the tucking-in effect, as the closest spec position will always yield the smaller path (i.e. a proper subset of the path to the other potential landing sites).

Head movement is a problem for these methods: because of the way in which Fitzpatrick assumes head movement derives (i.e. via head adjunction), and the way in which movement paths are defined, head movement does not generate movement paths with any members. The sister of A from the definition of movement path will be the head previously occupying that position; it does not dominate any terms which dominate B. Fitzpatrick notes that an alternative definition of movement paths would allow such methods to predict head movement, but because the concept of path is not empirically motivated, Fitzpatrick feels there is no sound way to choose between definitions. Ultimately, this is one of the issues that leads him to come down on the side of approaches such as Com and Dom.

With regard to the movement of more distant candidates illustrated in (12), the two path-based approaches diverge. Whereas PathQuan has the same issue as MutCom, in that the more distant candidate will clearly have a larger path, ruling out the movement of *what*, SubPath correctly does not rule out movement of *what*. While the movement path of *where* would be smaller, it is not a proper subset of *what*'s movement path.

Both path-based approaches are unable to predict freezing effects: as with MutCom, there is nothing in their mechanics which prevents a mutually c-commanded element from moving on as long as it is the closest candidate for a given operation.

3.4 Summary

Fitzpatrick concludes, based on performance in predicting the phenomena he outlined, as well as theoretical considerations, that the simpler intervention-based approaches like Com and Dom are to be preferred over the more complex approaches which require quantification and/or look ahead. Indeed, the only phenomena that Dom cannot account for is that of tucking in.

4. Concerns regarding a landing site constraint

In his conclusion, Fitzpatrick states that most of the data supports Dom as the optimal candidate for defining a locality constraint on movement, in that it can account for each phenomena as well as any other candidate, with the exception of tucking in, and that Dom is also favored by considerations regarding the formal machinery required (it does not make use of quantification or look ahead, and only uses derived non-stipulatory relationships). He goes on to say that "this suggests that locality of attraction (i.e., choice of mover) should be divorced from choice of landing site." While ultimately I believe this suggestion has merit, I do not think it necessarily follows from what Fitzpatrick has presented in his paper. Indeed, he presents no argument or evidence for such a claim, other than the fact that Dom can account for all other phenomena besides tucking in. It is certainly true that it is worth exploring other ways with which to explain the phenomena, but it is dangerous to claim that something your theory does not account for should simply be overlooked by it. Before such a claim can be solidified, two questions should be answered: is there a way to account for tucking in with a locality constraint, that maintains the effectiveness and formal superiority of Dom? Is there a viable alternative way to explain the tucking in phenomena that does not require the formal apparatus that the choice of Dom allows us to avoid? I will elaborate and explore these questions in what follows.

Fitzpatrick also notes that tucking in does not always occur; in cases where a single head attracts multiple elements for checking of different features, it is sometimes the case that the landing site constraint appears to be violated. I suggest a possible analysis that accounts for this variation, without complicating efforts for defining a constraint on landing sites.

4.1 A locality account of the landing site constraint

The question is whether or not a constraint on the locality of movement can be formulated in such way that it maintains the desirable features of Dom, both in terms of its formal advantages and its empirical power, while also accounting for the landing site constraint. A logical first step then would be to reformulate Dom so that rather than being a relationship between candidates for movement, it is a relationship between a landing site and other potential candidates. It should still in essence be an intervention relationship, but it should apply after the movement has happened rather than before.

What we must consider is what the relationship would be between an intervening candidate for movement such as K from the definition of Dom in (15) and the landing site. Assuming that movement only obtains when the destination position dominates (i.e. the mother of that position dominates, I will continue using "dominate" to mean this) the point of origin, then the movement of some G will result in a mutual dominance relationship with any intervener K. That is, any intervener K will dominate the lower position of G, and the movement of G will cause it to dominate K. Perhaps, then, a constraint such as the following can maintain the desirable features of Dom while gaining the ability to account for landing site constraints:

(20) *Mutual Dominance Constraint*
The derivation cancels if the movement of G to check the uninterpretable feature F results in there being an element K bearing a feature matching F such that the mother of G dominates K and the mother of K dominates G.

In other words, if G moves into a position where it is in a mutual dominance relationship with another potential candidate for the same movement, the structure crashes. Let us check to see if this accounts for the tucking in effect. Recall the multiple wh-movement Bulgarian example in (10), repeated below. If *koj* moves first, and then *kogo* moves to the outer spec position, as in (10b), the structure will crash. Likewise, if *kogo* moves first, the structure will crash.

(10) a. Koj kogo vizda (koj) (kogo)

who whom sees
b. *Kogo koj vizda (koj) (kogo)
Whom who sees

The only derivation that does not crash is the tucking in case. Notice though that in order for this to account for the tucking in phenomenon, it must be the case that even if K has already checked its feature F, it must still prevent G from moving to a higher position (to check the same feature). We should now check to see if the constraint in (20) can account for the other phenomena Fitzpatrick discussed.

The constraint in (20) does account for the A-over-A effect data Fitzpatrick discussed: scrambling a constituent G out of another constituent K which is also a candidate for scrambling creates the relationship disallowed by (20). It also accounts for movement of a more distant candidate as in (12) at least as well as Dom: neither prevent the movement of *what* because *where* does not dominate *what*, but it is also true that neither can prevent the movement of *where*. The constraint in (20) is equally well equipped to handle freezing effects: indeed, in such cases the two candidates are already in a mutual dominance relationship, so any movement motivated by a feature borne by both will crash.

Unfortunately, this method cannot account for head movement locality, unless a more stipulatory relationship such as m-command is used. From the new head-adjoined position, G will not create any new dominance relationships, and thus there will not be mutual dominance between G and K. This method thus represents a trade off, explaining one phenomenon but losing the ability to explain another; thus, there is no reason to prefer it over Dom. Whether or not (20) can account for further data, or can be refined to better account for head movement, I leave open to further research.

4.2 The mechanics of Attract

The other question to consider before claiming two phenomena be divorced in terms of how they are constrained is whether or not an alternative explanation is feasible. It is likely that Fitzpatrick had some method in mind, but because he did not discuss it, his claim was left unsupported.

One possibility is that the phenomena is simply an artifact of the mechanics of attraction. If a head H is attracting an element K, it seems fairly logical to assume that the resulting operation will pair K with the most immediate projection of H, and not with some higher projection of H. This could be formalized by defining “checking domain” in the definition of Attract (2) such that it extends to only the most immediate projection of H, thus movement to a higher projection would prevent checking of an uninterpretable feature and the derivation would crash. I leave the best formulation of this as an open question.

4.3 Non-tucking in cases

Fitzpatrick notes in his footnote 16 that tucking in might only occur “when multiple categories are attracted for the checking of the same feature (e.g., multiple-wh or QR movement; see Bruening 2001), in which case the phenomenon would motivate a more nuance treatment of the locality of landing site choice.” While this certainly has serious repercussions for an account of the landing site constraint, it seems to me that there is an analysis which does not pose any problem to either treatment of landing site choice (locality constraint, or mechanics of attraction). Consider the abstract example in (21), where the element base-generated higher in the tree, K, has moved to the inner spec of HP, while the lower element, G, has moved to the outer spec of HP.

(21) $[_{HP}$ $[_{G}$ Foc] $[_{HP}$ $[_{K}$ Wh] [*H*] $[_{TP}$ (Wh) (Foc)]]]

Because K originates from higher in the tree than G, it would be logical to assume that K is attracted first (it's closer), and (21) would thus be an instance of attract without tucking in. If the movement of K and G are distinct instances of attraction (i.e. motivated by the checking of different features) it could be argued that the order of movement would not be affected by a constraint on movement locality formulated around alternative candidates, such as those discussed, because different features are in question. As such, one could propose that any constraint on landing site can be uniform in prediction and still account for the data if variations can be attributed to the order in which the features where checked.

5. Conclusion

Fitzpatrick very clearly evaluates five apparently similar constraints on locality of movement in terms of five well described phenomena. He succeeds in differentiating them on empirical grounds, and convincingly argues in favor of the Dom formulation of locality. He then makes a suggestion which is to some degree unsupported by his paper, which is that one of the phenomena he describes (tucking in, or the landing site constraint) not be accounted for by a constraint on locality of movement. I have explored his suggestion, explaining my problem with his argumentation, and posing and exploring two questions that should be answered before going further. I have also noted that some data which Fitzpatrick felt to be problematic for an account of the landing site constraint can perhaps be analyzed in a way that does not complicate the issue. While I do not think he provided the necessary explanation, ultimately I believe his suggestion has merit.

References

Bošković, Željko. 1999. On multiple feature checking: Multiple wh-fronting and multiple head movement. In *Working minimalism*, ed. By Samuel David Epstein and Norbert Hornstein, 159-187. Cambridge, Mass.: MIT Press.

Bruening, Benjamin. 2001. QR obeys Superiority: Frozen scope and ACD. *Linguistic Inqiury* 30: 691-703.

Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*, 219-394. Cambridge, Mass.: MIT Press.

Epstein, Samuel David. 1999. Un-principled syntax: The derivation of syntactic relations. In *Working minimalism*, ed. By Samuel David Epstein and Norbert Hornstein, 27-34. Cambridge, Mass.: MIT Press.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations.* Oxford: Oxford University Press.

Fiengo, Robert, James C.-T. Huang, Howard Lasnik, and Tanya Reinhart. 1988. The syntax of *wh*-in-situ. In *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, ed. By Hagit Borer, 81-9. Stanford, Calif.: CSLI Publications.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations.* Cambridge, Mass.: MIT Press.

Nunes, Jairo, and Juan Uriagereka. 2000. Cyclicity and extraction domains. *Syntax* 3:20-43.

Richards, Norvin. 1999. Feature cyclicity and the ordering of multiple specifiers. In *Working minimalism*, ed. By Samuel David Epstein and Norbert Hornstein, 217-249. Cambridge, Mass.: MIT Press.