On Minimalist Approaches to the Locality of Movement

Justin M. Fitzpatrick

Several distinct approaches to the locality of movement have emerged within the Minimalist Program, but little attention has been paid to their formal and empirical differences. I examine a range of possible locality constraints on movement that are representative of current approaches. I expose some important formal differences among these alternatives and examine five phenomena, some previously unnoted, that distinguish them empirically. No single approach succeeds in capturing all of the facts that should arguably follow from a theory of locality, but the bulk of the evidence seems to support a theory that defines locality using only simple and independently motivated syntactic objects and relations.

Keywords: movement, locality, minimalism, path theory, copy theory

Several distinct approaches to the locality of movement have emerged within the Minimalist Program, but little attention has been paid to their formal and empirical differences. Chomsky (2000:111ff.) summarizes the intuition that locality constraints are necessary in order to minimize computational complexity. However, this sort of a priori observation cannot adjudicate between particular versions of movement locality. In the end, assuming such intuitions are well motivated, the correct version of locality is an empirical issue. In this article I examine a range of possible approaches to movement locality, giving examples of each from contemporary literature and exposing some important formal and empirical differences among them.[1] Since considerations of parsimony lead me to believe that at most one of these hypotheses is correct, it is important to examine and reveal their differences in the interest of determining which, if any, is part of the

For insightful discussion of many of the issues presented here, I wish to thank Noam Chomsky, Avi Derrow, Mark Hale, Hisa Kitahara, Christina Tortora, and Gabe Williams. I am especially indebted to Sam Epstein for extensive discussion and written comments on earlier drafts of this article. I am also grateful to Calixto Agüero-Bautista, Daniel Arroyo, José Benkí, Cristiano Chesi, Andrés Salanova, John Sánchez, and Esther Torrego for their help with the Spanish and Italian data and judgments. Versions of this article were presented at the University of Michigan in fall 2000 and winter 2001, as well as at the Michigan Linguistic Society (Eastern Michigan University) 2001, and I thank the audiences for their comments. I also benefited greatly from the comments of two anonymous *LI* reviewers. Any mistakes are of course my own. Portions of this work were supported under a National Science Foundation Graduate Research Fellowship.

[1] Unfortunately, the scope of this article is limited to locality constraints within a fairly narrow range of syntactic theories. Therefore, I do not discuss alternative constraints that require more radical departures from derivational theories similar to the one in Chomsky 1995 (e.g., Williams's (in press) shape conservation theory). However, the derivational nature of the theory assumed here is in many cases not crucial, and the findings reported here may apply equally to representational approaches with similar properties. I have also failed to discuss possible approaches to locality that would require extensive discussion of alternative grammatical architectures (e.g., closest = most recently merged, or Nissenbaum's (2000) nonextension condition).

Linguistic Inquiry, Volume 33, Number 3, Summer 2002
443–463

correct theory of Universal Grammar and to inform future research concerning the locality of movement. I conclude that none of the approaches discussed here account for all of the phenomena in question. However, I believe the evidence suggests that the correct theory of locality will ultimately consist of simple and independently needed relations and mechanisms.

The article is organized as follows. Section 1 introduces some basic assumptions. Section 2 discusses various types of movement locality constraints that are representative of current approaches to the problem and exposes a number of important formal differences among them. Section 3 examines five phenomena, some of them previously unnoted, that distinguish these approaches empirically. Section 4 contains a summary.

1 Syntactic Movement and Syntactic Relations

Theories of movement locality depend crucially on an explicit theory of movement. Therefore, I will first outline some assumptions regarding movement, attraction, copy theory, and syntactic relations.

All of the locality constraints discussed here either depend on or are compatible with an Attract-based theory of movement (Chomsky 1995). Under Attract (1), a movement operation is motivated by some requirement of a head. Movement of a category G applies only when G is ''attracted'' by a head H for feature checking and/or satisfaction of the Extended Projection Principle (EPP). We are concerned here with the proper formulation of *closest* in (1).[2]

(1) *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.

I assume here that syntactic objects are generated by iterative application of the fundamental syntactic structure-building operations Merge (2) and Move (3).[3]

(2) *Merge* (see Chomsky 1995:243)
Take two syntactic objects α and β and replace them with the set K whose sole members are α and β. (i.e., Merge(α,β) = K = {α,β})

[2] I assume the approach to the cycle that Richards (1999) calls Feature Cyclicity (see Chomsky 1995, Bošković and Lasnik 1999). For present purposes we may assume that the checking domain of a head H includes Spec,H and categories adjoined to H. See Chomsky 1993 for one definition that includes these relations. Furthermore, assume that heads, when attracted, are paired with the attracting head and that maximal projections, when attracted, are paired with a nonminimal projection of the attracting head. (See Chomsky 2001, Fitzpatrick 2001a, and Nunes 1999 for discussion of this asymmetry.)

[3] Though I believe a unification of these operations is ultimately correct (see, e.g., Kitahara 1997, Starke 2001), I keep them separate for expository purposes.

Note that it is generally (though tacitly) assumed that $\alpha \neq \beta$ (2) and $K \neq \alpha$ (3).

The terms of a syntactic object Z are Z and the terms of the members of Z. This definition is possible in the absence of labels, which I omit here, keeping structures to the form {α,β} rather than {γ,{α,β}}. See Chomsky 1995:244, Collins 2001, Fitzpatrick 2001a, and Seely 2000 for discussion of labels.

(3) *Move* (see Chomsky 1995:250)
K and α are terms of a category Σ. Form Σ' by replacing K in Σ with the set L whose sole members are α and K. (i.e., Move(α,K) = L = {α,K})

We might wonder how (3) affects the moved category α. Is a trace left in the ''departure site'' of α? Chomsky (1995:251) notes that ''[t]he operation that raises α introduces α a second time into the syntactic object that is formed by the operation.'' That is, he treats movement of a category α as *remerger* of α, leaving the moved category itself unduplicated. Thus, although the category α occurs twice in the tree, there is only one syntactic object α; one category α is in two positions in the tree.[4] Thus, a category α may have more than one occurrence, where an *occurrence* is a relation between a syntactic object and a position in a tree (e.g., ''the category DP *is in* Spec,T''). Thus, in (4) *Mary* has occurrences ⟨Mary,arrested⟩ and ⟨Mary,[$_{T'}$ was [$_{VP}$ arrested (Mary)]]⟩.[5]

(4) [$_{TP}$ Mary [$_{T'}$ was [$_{VP}$ arrested (Mary)]]]

I believe this is the minimal assumption we can make regarding movement. Conforming to the Inclusiveness Condition (Chomsky 1995:228), this formulation of movement adds no features or markings that were not already present in items selected from the numeration (e.g., *t*s and indices); nor does it remove or in any unmotivated manner alter the featural content of the moved category.

Furthermore, unlike the proposals of Collins (1997), Nunes (1995), and many others, this approach does not require the addition of a duplication operation to the computational system, and so it avoids the complications that arise from copying (see Gärtner 1999, Kracht 2001, and references therein). For example, this approach allows us to account, without stipulation, for the fact that feature checking and Case assignment affect all ''copies'' of a category, a fact that would be mysterious or require extra, and somewhat ad hoc, theoretical devices if movement involved duplicative copying.[6]

The familiar relation of c-command plays a key role in many approaches to movement locality. I follow Epstein (1999) (and subsequent adaptations in Chomsky 2000, 2001) and note that, rather than being construed as an unexplained relation defined on trees, c-command can be derived from the independently motivated syntactic structure building operations Merge and Move and can be stated as in (5).

[4] As with any version of copy theory, this raises questions regarding where a category with multiple occurrences is pronounced. I will put aside questions of phonological positioning here. See Epstein et al. 1998:chap. 5, Nunes 1995, and Groat and O'Neil 1996 for several interesting approaches to this problem. Note that the bulk of the evidence suggesting that trace positions are independent of the head of their chain is of a semantic and phonological nature. This evidence is therefore not inherently problematic for the present remerger approach, where separate operations or relations could be mapped to separate semantic and phonological interpretations.

[5] Throughout this article I have used standard X-bar notation for ease of exposition. Lower occurrences will appear in parentheses for clarity. In some cases I have omitted structure irrelevant to the present discussion. I will often use the sister of an element to name an occurrence of that element (e.g., *arrested* and T′ are occurrences of *Mary* in (4)). Note that ''mother,'' also a derivationally created relation, might suffice for the identification of occurrences, as suggested in Chomsky 2001. Such a modification would be particularly plausible given the data discussed in section 3.1.

[6] For further investigation of the so-called multidominance structures that result from nonduplicative movement, see Erteschik-Shir 1987, Gärtner 1999, Kracht 2001, Moltmann 1992, Phillips 1998, Starke 2001, Wilder 1999, and other work.

(5) *Derivational 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.

Strictly speaking, only syntactic objects undergo Merge and Move, each of which is simply a pairing operation creating a two-membered set. It follows that c-command is a relation between syntactic objects, not between occurrences or between positions in a tree.[7] This distinction will become important below. However, though occurrences are not part of the structural description of Merge or Move, Attract must have access to information regarding the structural locations of a syntactic object K in order to determine the relative locality of K and other syntactic objects.

2 Locality Constraints on Movement

I will now discuss some common contemporary approaches to movement locality that fit within the general framework outlined above, focusing on their general properties and giving specific examples from the literature that exploit these approaches.

2.1 Syntactic Relations and Locality

Among other effects, movement locality has been proposed to account for (6), where we observe that only the *wh*-phrase that is ''closest'' to C_{+wh} may undergo overt movement.

(6) a. Who C_{+wh} (who) bought what?
b. *What did–C_{+wh} who buy (what)?

Many approaches to this asymmetry make use of independently motivated syntactic relations like c-command, dominance, and sisterhood.

2.1.1 Attract Closest (AC) The simplest constraint of this sort seeks a syntactic object K such that no other syntactic object intervenes between the attractor and K, where intervention is defined using a relation like c-command (7) or dominance (8).[8]

(7) *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.

(8) *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.

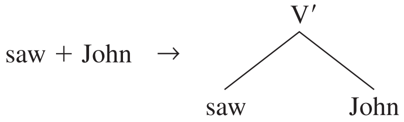
[7] Though we often speak metaphorically of a constituent moving ''from'' a particular position in the tree, if movement involves only the addition of a sisterhood relation to that constituent, then no particular occurrence of the constituent is moved, and it is technically wrong to speak of moving *from* a position.

[8] I thank an anonymous *LI* reviewer for pointing out the use of dominance here. Com is very similar to the Superiority Condition (Chomsky 1973), the Minimal Link Condition (Chomsky 1995), and the locality clause of Chomsky's (2000) probe-goal approach to Agree. Although, unlike Attract, Agree does not itself displace categories, many of the results discussed here regarding Com are also relevant for Agree.

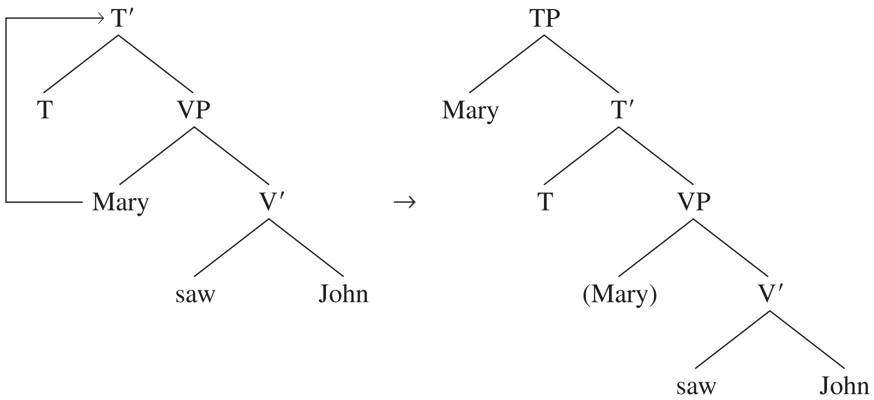
Both of these approaches predict that *what* may not be attracted by C_{+wh} as in (6b) since there is an element *who* in the sister of C_{+wh} that bears the requisite *wh*-feature, that c-commands *what,* and whose mother dominates *what.*[9]

2.1.2 Relation Quantification (RQ) A second type of approach involves the quantification of the relations that a given movement operation will produce. Epstein et al. (1998; henceforth *DASR*) propose one such approach. *DASR* argues that mutual c-command relations carry with them a certain level of computational complexity, and so they should be minimized.[10] Under the copy theory of movement outlined above, mutual c-command is established in two ways. When two syntactic objects K and L are paired, they c-command each other. Thus, upon application of (9a), mutual c-command holds between *saw* and *John.*

(9) a. *Merge(saw,John)*



b. *Move(Mary,T′)*



In (9b) a somewhat surprising mutual c-command relation is created. Prior to (9b) T was merged with VP to form T′, and so T c-commands *Mary,* which is a term of VP. In (9b) *Mary* is paired with T′, and so *Mary* c-commands T, which is a term of T′. Thus, following application of Move, *Mary* c-commands T and T c-commands *Mary.* Note that these c-command relations also exist with a representational definition of c-command. With this in mind we can define locality in

[9] If *occurrence* is defined in terms of *mother* rather than *sister,* as discussed in footnote 5, Dom is particularly natural.

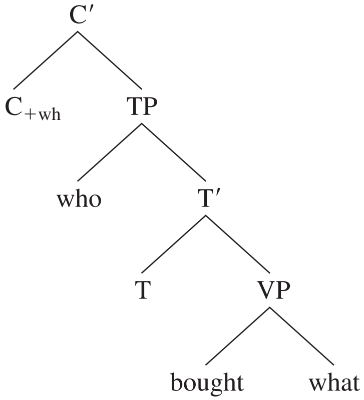
[10] The intuition behind this approach is that since mutual c-command is created by any application of Merge or Move, this relation (called *Derivational Sisterhood* in *DASR*) is the fundamental syntactic relation, and the relation under which feature checking applies. See *DASR* (pp. 114ff.) for more complete argumentation. None of the arguments below depend on *DASR*'s approach to feature checking.

terms of an economy metric that quantifies the number of new mutual c-command relations created by a given operation (10).[11]

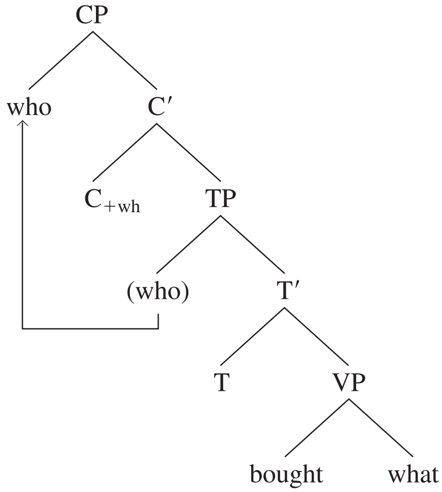
(10) *Relation Quantification: Mutual C-Command (MutCom)*
G is the closest category in the sister of H iff there is no category K that bears a feature matching F such that movement of K would create fewer mutual c-command relations than would movement of G.

The Superiority effect seen in (11) is predicted by (10), as it is by AC approaches.

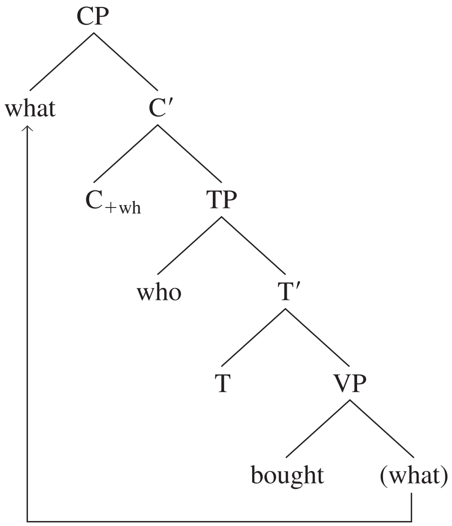
(11) a.



b.



c.



[11] *DASR* (p. 115) states this condition as ‘‘Minimize the number of derivational sisters created by [an application of] Move.’’

As with T and *Mary* in (9b), *who* and C_{+wh} c-command each other in (11b). At one point in the derivation C_{+wh} was paired with TP, of which *who* is a term, and so C_{+wh} c-commands *who*. *Who* c-commands C_{+wh} due to the operation that pairs *who* with C′ (of which C_{+wh} is a term) to form (11b). The movement operation that forms (11b) creates two mutual c-command relations: *who*–C_{+wh} and *who*–C′. The operation that forms (11c) creates four such relations: *what*–T, *what–who, what*–C_{+wh}, *what*–C′. By MutCom, movement of *what* cannot apply since movement of *who* would check the same feature but create fewer relations in doing so. Thus, MutCom predicts the attraction of *who* (11b) rather than *what* (11c).

Since they necessitate comparison of the number of relations that *would be created were movement to apply,* RQ approaches like MutCom are not entirely local in their application. Rather, they require a species of limited look-ahead: relations that *would be* created by various movement operations must be quantified and compared to determine the most local operation. It is not immediately obvious whether this characteristic, which is shared by the following approaches, is a problem. However, note that AC approaches do not employ look-ahead even to this limited degree. Additionally, note that since, by hypothesis, no other aspect of the grammar employs a counting mechanism such as that found in RQ approaches, AC is preferable to RQ, all else being equal.

2.2 Movement Paths

Since Pesetsky 1982 movement paths have been a common way to measure ‘‘distance’’ of movement. Paths can be used for this purpose in at least two ways: path quantification and interpath relations.

2.2.1 Path Quantification A path quantification constraint such as (12), adapted from Richards 1999,[12] states that the path-theoretic distance of movement—that is, the cardinality of the movement path between the ‘‘departure site’’ of an element and its ‘‘landing site’’—must be minimized. For clarity, assume that α is the ‘‘landing site’’ occurrence of the moved element G and β is the ‘‘departure site’’ occurrence of G.[13]

[12] Richards’s (1999) Shortest (i) uses a path-quantificational approach.

(i) *Shortest* (Richards 1999:154)
A pair P of elements {α,β} obeys Shortest iff there is no well-formed pair P′ that can be created by substituting γ for either α or β, and the set of nodes c-commanded by one element of P′ and dominating the other is smaller than the set of nodes c-commanded by one element of P and dominating the other.

However, the metric shown here differs from Richards’s in (at least) two ways: (a) it conforms more fully with the present approach to the copy theory of movement, where positions in a tree are defined by the sister of a category; (b) it conflates all locality of movement into one metric, rather than computing attractor-attractee (choice of mover) and departure site–landing site (choice of landing site) relations separately. Though look-ahead is not required for the former, it is required for the latter, even in Richards’s original formulation, unless empty Δ nodes are used to mark future specifier positions. Note that under the present definition of path PathQuan is not empirically distinct from MutCom (but see section 3.3 for a possible modification). However, their formal differences may be important since each is compatible with a different (though overlapping) set of theories.

[13] As discussed in footnote 7, the terms *landing site* and *departure site* are somewhat misleading metaphors. While it is trivial to obtain a path for a category that has only two occurrences, for three or more occurrences we

(12) *Path Quantification (PathQuan)*

a. 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.

b. Path = the set of terms of the position (i.e., sister) α of a category G that dominate the position β of G.

As with AC and RQ, PathQuan accounts for (11). Movement of *what* (11c) results in the set of occurrences L = {C′,bought}, whereas movement of *who* (11b) results in L′ = {C′,T′}. The path P = {C′,TP,T′,VP} (derived from L) is the set of terms of C′ (the ‘‘landing site’’ of *what*) that dominate *bought* (the ‘‘departure site’’ of *what*). $|P|$ is clearly larger than the cardinality of the path P′ = {C′,TP} (derived from L′), which is the set of terms of C′ (the landing site of *who*) that dominate T′ (the departure site of *who*). Since P′ has a lower cardinality than P, only movement of *who* (11b) may apply.

Like relation-quantifying approaches, the path-theoretic PathQuan requires quantification and look-ahead. In fact, all path-theoretic approaches require this type of look-ahead. While AC approaches require only the recognition of intervening categories, both relation-quantifying and path-theoretic approaches rely on a comparison of the distances of the candidate movement operations, *were they to apply.*

2.2.2 Interpath Relations We can also regulate the locality of movement by looking for relations between the paths that would be created by candidate movement operations. In (13), adapted from Bošković 1997,[14] we see the most straightforward implementation of this approach.

(13) *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.

---

must know which are relevant to the computation of a movement path. In (i) *who* has been attracted to Spec,C.

(i) $[_{CP}$ who $[_{C'}$ C_{+wh} $[_{TP}$ (who) $[_{T'}$ T $[_{VP}$ (who) $[_{V'}$ bought what]]]]]]

The set of occurrences for *who* is {C′,T′,V′}. But which path is relevant to attraction of *who* by C_{+wh}? It is clear that only the path between the two ‘‘highest’’ occurrences of *who* (i.e., T′ and C′) is germane to the locality constraints of Attract. The pairs {T′,V′} and {C′,V′} are irrelevant. Therefore, we must stipulate that paths are computed from the occurrence created by the movement operation in question and the highest occurrence prior to this operation, where *highest* is defined as the occurrence O of a category K such that there is no other occurrence O′ of K that dominates O.

It is unclear whether the assumptions of copy theory and the cycle will ensure that this approach is feasible even in cases of remnant movement (Kitahara 1997, Müller 1996, Sauerland 1999). The question would then be whether, in (ii), YP remains *closer* to the head F than ZP is, even after movement of XP to Spec,G.

(ii) $[_{F'}$ F $[_{GP}[_{XP}$ . . . ZP . . .] $[_{G'}$ G [. . . YP [. . . (XP) . . .]]]]]

The data needed to answer this question will likely be hard to find due to independent factors. However, if later conclusions are correct, we would expect YP to be closer to F in this case.

[14] I have modified Bošković’s (1997) formulation (i) slightly. However, as the reader can verify, (13) and (i) are equivalent under the relevant reading of *derivation* and *set of nodes crossed by movement of G.*

(i) A derivation α is to be preferred to a derivation β if and only if the set of nodes crossed by the derivation α is a proper subset of the set of nodes crossed by the derivation β. (Bošković 1997:251)

As with the previous three approaches, (13) accounts for (11). Here movement of *what* (11c) would result in path P = {C′,TP,T′,VP} and movement of *who* (11b) would result in path P′ = {C′,TP}. Since P′ ⊂ P, only movement of the closest category *who* is allowed by SubPath.

Though this approach does not require quantification, it does require look-ahead. We have already seen that relation-quantifying and path-theoretic approaches differ from AC in their look-ahead requirements. However, RQ and the path-theoretic PathQuan and SubPath also differ in that PathQuan and SubPath require the computation of movement paths, which, though definable, are unnecessary with RQ and AC approaches. Thus, the question regarding path-theoretic approaches is, why would paths, so defined, be relevant at all? I return to this below.

2.3 Summary

Though far from comprehensive, the approaches outlined so far are representative of the vast majority of contemporary treatments of the locality of movement in the Minimalist Program. These approaches are organized in (14) according to their general properties.

(14) 1. Syntactic relations
  a. Intervention: AC
    i. C-command: Com
    ii. Dominance: Dom
  b. Quantification: RQ
    i. Mutual c-command: MutCom
 2. Movement paths
  a. Quantification: PathQuan
  b. Interpath relations: SubPath

3 The Empirical Content of Locality Constraints on Movement

Though the approaches summarized in (14) differ in form, one might wonder whether they are empirically equivalent, essentially making explicit Chomsky's suggestion noted in the introduction. I argue in this section that though all account for the central effect of locality illustrated in (11), they diverge empirically in at least five cases.

3.1 A-over-A Effects

Kitahara (1997), Müller (1996), and Sauerland (1999) present evidence that a category K that dominates a category G is closer to an attracting head than G is. We observe this ‘‘generalized A-over-A’’ effect in derivation (15) (from Kitahara 1997). Here scrambling of *das Buch* is prohibited since the constituent *das Buch zu lesen,* which dominates *das Buch,* is also a candidate for scrambling, which applies in (15b). Thus, the ill-formedness of the resulting (15c) might be attributable to the nonlocal scrambling of β in (15a) out of the category α that dominates β. This movement is nonlocal since the closer element α could have undergone scrambling. The hypothesis that such dominance results in movement locality intervention effects is supported by the fact that if α is a topicalized phrase that does not undergo scrambling (rather, it undergoes topicalization movement in (16b)), scrambling of β (16a) is licit, leading to the well-formed (16c).

(15) a. daß keiner [$_\beta$ das Buch] [$_\alpha$ (das Buch) zu lesen] versucht hat

b. daß [$_\alpha$ (das Buch) zu lesen] [$_\beta$ das Buch] (das Buch zu lesen) versucht hat

c. *daß [(das Buch) zu lesen] keiner [das Buch] (das Buch zu lesen) versucht hat
   that to read no-one the book tried has

(16) a. hat keiner [$_\beta$ das Buch] [$_\alpha$ (das Buch) zu lesen] versucht

b. [$_\alpha$ (das Buch) zu lesen] hat keiner [$_\beta$ das Buch] (das Buch zu lesen) versucht

c. [(das Buch) zu lesen] hat keiner [das Buch] (das Buch zu lesen) versucht
   to read has no-one the book tried

Note that MutCom, PathQuan, and SubPath each account for (15)–(16) without modification. For any movement candidate K, movement of a category L that dominates K will clearly create fewer mutual c-command relations and a smaller path than would movement of K. Furthermore, this path will be a subset of the path that would result from movement of K. However, Com contains no dominance restriction, and so as stated it cannot account for the effects seen in (15)–(16). But while constraints like Com that define closeness in terms of c-command fail to predict A-over-A effects, Dom, which utilizes dominance rather than c-command, has no trouble with these facts.[15]

3.2 Predicting the Landing Site of Movement

Richards (1999) argues that movement locality should constrain both the choice of attracted category and the landing site of that category. He presents compelling evidence that when a head H attracts multiple categories to specifier positions of H, each movement operation inserts the moved category in the specifier position closest to H, a phenomenon dubbed ‘‘tucking in.’’[16] Thus, the derivation of the Bulgarian (17c) proceeds as in (17a–b). Richards argues that movement of *kogo* to the outer specifier of C_{+wh} as in (17d) is illicit, as evidenced by the ill-formedness of (17e).

[15] I thank an anonymous *LI* reviewer for pointing this out.

[16] It may be that tucking in only arises 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 nuanced treatment of the locality of landing site choice.

(17) a. [$_{CP}$ koj [$_{C'}$ vižda (koj) [kogo]]]

b. [$_{CP_1}$ koj [$_{CP_2}$ kogo [$_{C'}$ vižda (koj) [(kogo)]]]]

c. Koj kogo vižda?
who whom sees

d. [$_{CP_1}$ kogo [$_{CP_2}$ koj [$_{C'}$ vižda (koj) [(kogo)]]]]

e. *Kogo koj vižda?
whom who sees

While AC approaches restrict only the choice of the attracted category, restricting only the *structural description* of Attract (1), the tucking in of *kogo* in (17b) requires attention to the *structural change* of Attract. Consequently, AC approaches alone cannot account for tucking in. Were such an approach to be adopted, landing site phenomena such as tucking in would have to be dealt with separately.

MutCom, PathQuan, and SubPath all account for Richards's tucking in phenomenon without modification. Movement of *kogo* to the outer specifier of C_{+wh} as in (17d) will result in the creation of at least one more mutual c-command relation, *kogo–koj,* than does movement to the inner specifier of C_{+wh} (17b). Thus, movement as in (17b) is predicted by MutCom. Similarly, the movement path P of *kogo* in (17d) contains at least one more element, CP_2, that is not contained in the path P′ of *kogo* in (17b). Thus, P′ is smaller than P and (17b) is preferred by PathQuan. Since $P' = P \cup \{CP_2\}$, it follows that $P \subset P'$, and so (17b) is also preferred by SubPath. Therefore, to the extent that tucking in should be accounted for by movement locality, and not independently of the selection of the element to be moved, AC approaches are empirically deficient, while some relation-quantifying and path-theoretic approaches fare better (but see footnote 16).

3.3 Locality of Head Movement

Under an approach to c-command such as the one taken here, a head adjoined to another head through movement does not c-command its departure site.[17] AC approaches predict the expected head movement locality effects despite this fact. In (18) *could* cannot be attracted since *would* c-commands it (and the mother of *would* dominates *could*).

(18) a. C_{+wh} [$_{TP}$ John would think [$_{CP}$ that Bill could see who]]

b. [$_{CP}$ who [$_{C'}$[$_{C^0}$ would–C_{+wh}] [$_{TP}$ John (would) think [$_{CP}$ that Bill could see (who)]]]]

c. *[$_{CP}$ who [$_{C'}$[$_{C^0}$ could–C_{+wh}] [$_{TP}$ John would think [$_{CP}$ that Bill (could) see (who)]]]]

[17] Unless a more stipulative definition of c-command is used, as in Kayne 1994 or May 1985.

However, approaches to head movement locality that rely on look-ahead (e.g., MutCom, PathQuan, SubPath) do not predict a difference between (18b) and (18c). The only new mutual c-command relation formed by pairing two heads H and K is H–K. In (18) movement of *would* and movement of *could* both result in one new mutual c-command relation, *would*–C_{+wh} and *could*–C_{+wh}, respectively. Thus, relation-counting approaches that rely on c-command fail to make locality distinctions between any two head movement operations.

Similarly, with *path* defined as in (12b), all movement paths resulting from head movement $= \emptyset$, and so all head movement is equally local. For example, *would* has occurrences {C_{+wh},[think . . .]} in (18b), from which we compute the path $P = \emptyset$; the set of terms of C_{+wh} that dominate [think . . .] is $\emptyset$.

To circumvent this problem, we might modify our definition of *path* to resemble more closely the approach taken by Collins (1994).[18]

(19) *Path* (Collins 1994:56)
Let P_1 and P_2 be two categories in a tree. Let S_1 be the set of categories dominating P_1 and let S_2 be the set of categories dominating P_2. The path between P_1 and P_2 is defined as follows: $\text{Path}(P_1,P_2) = (S_1 \cup S_2) - (S_1 \cap S_2)$.

Here P_1 and P_2 are the relevant sisters of a moved category K. It is important to note, however, that path-theoretic approaches treat the notion of path as something that is simply defined. Thus, a priori, there does not seem to be any reason to prefer (19) to (12b), and the fact that relation-based approaches such as Com and Dom predict Head Movement Constraint effects without modification or the creation of additional arbitrary theoretical constructs would tip the scales in their favor, all else being equal.

3.4 Locality in the Absence of C-Command

Consider a case such as (20a), where we find two potentially attractable categories *where* and *what,* but neither *wh*-phrase c-commands the other. As first noted by Fiengo et al. (1988), (20b) is, perhaps surprisingly, well formed. Though movement of *what* is ‘‘longer,’’ in the sense of MutCom and PathQuan, than would be movement of *where,* this operation is nevertheless permitted.

(20) a. C_{+wh} [$_{TP}$[$_{DP}$ people from where] [$_{VP}$ try [$_{TP}$ PRO to buy what]]]
b. What did [$_{TP}$[$_{DP}$ people from where] [$_{VP}$ try [$_{TP}$ PRO to buy (what)]]]?
c. cf. *What did [$_{TP}$ who try [$_{TP}$ PRO to buy (what)]]?

Here MutCom and PathQuan both predict that movement of *what* is illicit since there is an element, *where,* that is closer, in the sense that movement of *where* would create fewer mutual c-command

[18] Alternatively, one might embrace MutCom, PathQuan, or SubPath and argue that this failure to predict head movement locality supports Chomsky’s (2001) proposal that head movement should not be included in narrow syntax. Rather, Chomsky argues, head movement is a phonological operation applying after Spell-Out. However, see Fitzpatrick 2001b for discussion of several problems with Chomsky’s approach and Matushansky 2001 for discussion of head movement in a minimalist framework.

relations (MutCom) or a shorter path (PathQuan) than does movement of *what.* In fact, were *what* to be more deeply embedded than *where,* the predictions of MutCom and SubPath would be reversed, only allowing movement of *where.* AC and SubPath, on the other hand, do not block movement of *what* in this case. Movement of *what* is not ruled out by AC approaches since there is no closer *wh*-phrase that intervenes between *what* and the attracting C_{+wh} head. SubPath fails to rule out movement since no category K exists such that movement of K would result in a path that is a proper subset of the path that would be formed through movement of *what.*

However, though this suggests that MutCom and PathQuan are empirically deficient, note that movement of *where* rather than *what* (21) is illicit, presumably due to the ‘‘Subject Condition’’ or whatever aspect(s) of the grammar ultimately rule(s) out extraction from within Spec,T. For example, in Nunes and Uriagereka's (2000) approach to the Condition on Extraction Domain (CED), *where* is simply not accessible in (21).

(21) *Where did [$_{TP}$[$_{DP}$ people from (where)] [$_{VP}$ try [$_{TP}$ PRO to buy what]]]?

Clearly, then, the importance of these effects for our formulation of movement locality constraints depends on other assumptions. However, if *where* does remain accessible in (20)–(21), one might suggest that we should modify MutCom and SubPath to allow nonlocal movement in a case such as (20), where movement of the closer element *where* would lead to ill-formedness due to some other principle. However, further consideration shows that this approach may be problematic. For example, though movement of the superior *who* in (22a) would result in a *that*-trace effect, movement of the more distant *what* is not licit (22b).

(22) a. *Who did you say that (who) bought what?
b. *What did you say that who bought (what)?

Turning to (23), we note that movement of *where* in (23a) is prohibited since *wonder* requires a *wh*-element in the specifier of its CP complement (see Epstein 1992). However, movement of the lower element *what* is nevertheless illicit.

(23) a. *Where do you wonder (where) John saw what?
b. *What do you wonder where John saw (what)?

Though these conclusions are contingent on present assumptions regarding the CED,[19] it appears that MutCom and PathQuan should not include a blanket ‘‘last resort’’ clause allowing nonlocal movement in cases where the most local movement would violate some other principle. While RQ and PathQuan are unable to account for the effect in (20), AC and SubPath accounts predict that when two categories are unrelated to each other by c-command or dominance, either category may be attracted. Further restrictions may of course arise due to other principles of grammar.

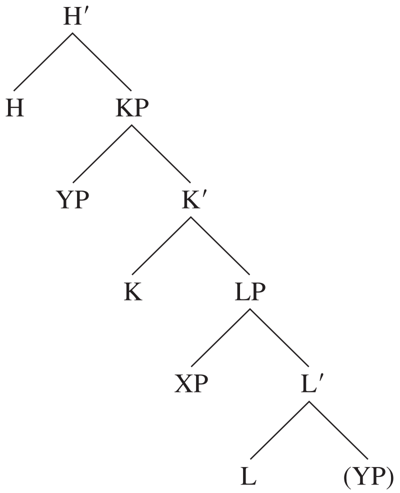
3.5 Mutual C-Command and Unattractable Categories

Recall that under the assumptions of copy theory outlined in section 1, XP and YP c-command each other in (24a), even under a representational construal of c-command. Derivationally, XP

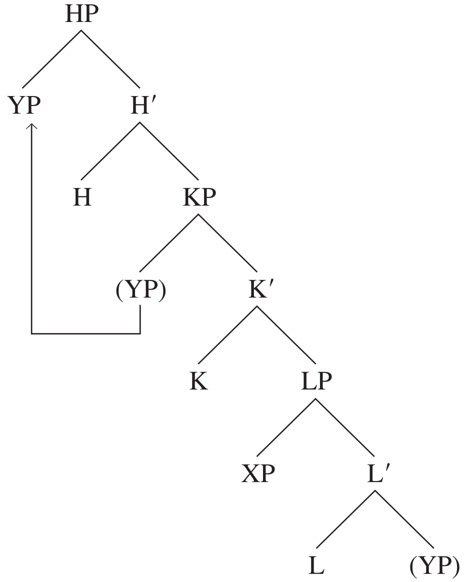
[19] My thanks to an anonymous *LI* reviewer for pointing out this complication.

and YP each c-command all and only the terms of the categories with which they were paired by Merge or Move in the course of the derivation. In forming (24a) XP was paired with L′, one term of which is YP, and YP was paired with K′, one term of which is XP. Thus, XP c-commands YP and YP c-commands XP.

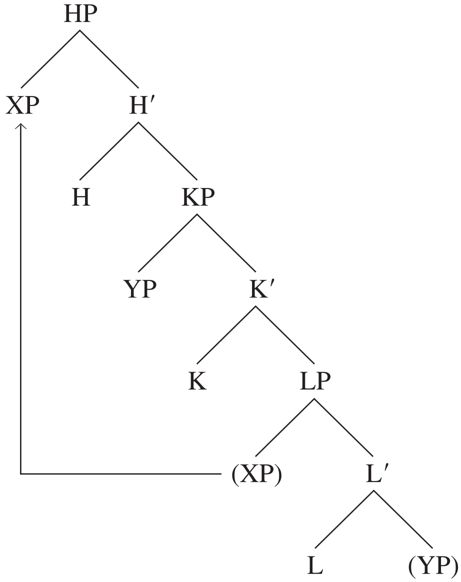
(24) a.



b. *Move(YP,H′)*



c. *Move(XP,H′)*



Assume that H must attract an element from KP and that both YP and XP bear the requisite feature. As the reader can easily verify, all approaches discussed in section 2, with the exception of Com and Dom, predict that YP is attracted rather than XP. However, some AC approaches predict that neither XP nor YP may be attracted (but see footnote 24 for AC approaches that allow movement in this case). XP cannot be attracted since there is an element YP that c-commands XP and bears the requisite feature F. YP cannot be attracted since there is an element XP that c-commands YP and bears the requisite feature F. The same holds for Dom, mutatis mutandis. XP and YP are ''frozen'' in place.

Surprisingly, we may find this sort of freezing effect in certain cases. (25), which appears to be a Superiority violation, is well formed in Spanish (see Jaeggli 1982).[20] If movement locality of the type investigated here is indeed an invariant principle of grammar, this movement must be due to a feature not borne by *quién.* One possibility is that this movement is due to a focus feature on *qué* (but not *quién*) that is checked following attraction to Spec,Foc (see Bošković 1999). However, the exact nature of this feature is not crucial here, and I will label this unknown attracting head *Z.*

(25) $[_{ZP}$ qué Z $[_{IP}$ compró $[_{VP}$ quién $[_{V'}$ (compró) (qué)]]]]
what bought-3SG who
'What did who buy?'

As with XP and YP in (24), the *wh*-phrases *qué* and *quién* c-command each other in (25). Now note that further movement of *qué* is barred: embedding of (25) and extraction of *qué* as in (26) is illicit.

(26) $*[_{CP}$ qué C_{+wh} crees $[_{ZP}$ (qué) que $[_{IP}$ compró $[_{VP}$ quién $[_{V'}$ (compró)
what think-2SG that bought-3SG who
(qué)]]]]]
'What do you think who bought?'

Interestingly, this is exactly the freezing prediction of Com and Dom.[21] *Qué* may not be attracted since there is a *wh*-phrase *quién* that c-commands it. Conversely, *quién* may not be attracted since there is a *wh*-phrase *qué* that c-commands it.

Note however that MutCom, PathQuan, and SubPath predict that once *qué* has moved to

[20] Several other approaches to this apparent Superiority violation have been proposed. See Bošković 1997 for discussion. My informants differed somewhat in their judgments of (25)–(29). However, several agreed on the judgments reported here. More research is clearly needed to determine the facts in this case, but regardless of the particular facts in this case, it is important to note that this *type* of pattern is predicted by certain AC approaches.

[21] One might wonder why *qué* may not move to a Spec,Z position of the matrix clause through the same mechanism as in (25). However, this mechanism seems to be clause bounded, as in (i), which appears to be a straightforward Superiority violation.

(i) ?*Qué dijo quién [que Juan compra (qué)]?
what said-3SG who that Juan buy-3SG

We might assume that the +Z feature on *qué* is erased in (25) and so further Z-driven attraction is impossible (see Sauerland 1999). If this is focus movement, we could account for (i) by stipulating that focus movement cannot originate from a Spec,C position (see Müller and Sternefeld 1993). It remains to be explained, however, why *qué* cannot be attracted by a matrix Z element directly from its VP-internal position without moving through Spec,C or Spec,Z of the embedded clause. A cyclic approach to Spell-Out as in Chomsky 2001 seems promising in this respect.

Spec,Z, it is superior to *quién* and may move on. Bošković (1997) exploits this fact and proposes a different account of (25), suggesting that non-*wh*-movement of a *wh*-phrase over a c-commanding *wh*-phrase may serve as an escape hatch for the lower category *qué.* Thus, assuming that Spanish subjects may remain in Spec,V overtly (see Suñer 1994), movement through Spec,Agr_O as in (27) makes the category superior to the subject by an approach like SubPath, and thus closer to C_{+wh} in (27a).

(27) a. C_{+wh} $[_{IP}$ compró $[_{Agr_OP}$ qué $[_{VP}$ quién $[_{V'}$ (compró) (qué)]]]]
     b. $[_{CP}$ qué C_{+wh} $[_{IP}$ compró $[_{Agr_OP}$ (qué) $[_{VP}$ quién $[_{V'}$ (compró) (qué)]]]]]

However, if this sort of movement serves as an escape hatch, *qué* should be able to undergo further movement to Spec,C of a higher clause. But contrary to this prediction, the resulting (26) is ill formed. Therefore, Bošković's approach, though suggestive, may be problematic.

Turning to further evidence of this sort, consider (28), from Rizzi 1982, which exhibits a well-known apparent exception to Relativized Minimality (Rizzi 1990) involving movement of the *wh*-phrase *a cui* over *che storie.*[22] As with the Spanish examples above, the mechanism of movement 2 remains largely mysterious.

(28) a. Tuo fratello, a cui mi domando che storie abbiano raccontato, era multo preoccupato.
        your brother to whom myself ask-I which stories told-they was-he very worried
        'Your brother, to whom I wonder which stories they told, was very worried.'
     b. Tuo fratello, [a cui] mi domando [che storie] abbiano raccontato (che storie) (a cui), era multo preoccupato.

We see in (28) that the *wh*-phrase *a cui* may move over the *wh*-phrase *che storie,* as long as neither phrase moves on. However, if movement over the c-commanding *wh*-phrase (movement 1 in (29b)) applies before movement of the superior element (movement 2), the result is ill formed.

(29) a. *Queste storie, che mi domando a cui abbiano raccontato, erano multo interesanti.
        the stories that myself ask-I to whom told-they were very interesting
        'The stories, which I wonder to whom they told, were very interesting.'
     b. *Queste storie, OP che mi domando $[_{DP}$ a cui] abbiano raccontato (OP) (a cui), erano multo interesanti.

Derivations like these were once ruled out by constraints like Pesetsky's (1982) Path Containment Condition, but I suggest that this may be the same phenomenon seen in (25)–(26). In (25) and (28) we saw that Spanish and Italian contain mechanisms that allow movement of a *wh*-phrase over a c-commanding *wh*-phrase.[23] However, since this operation results in mutual c-command

[22] The Spanish sentences (ia) and (ib) correspond to the Italian (28) and (29), respectively, and show the same effect. However, some of my informants found both (ia) and (ib) acceptable.

(i) a. Tu hermano, a quién me pregunto cuáles historias le contaron, estaba muy preocupado.
    b. *Las historias, que me pregunto a quién le contaron, estaban muy bien.

[23] As discussed in Sauerland 1999, the interpretability (and thus erasability, following Chomsky 1995) of the feature

between the *wh*-phrases, following Com and Dom neither phrase may undergo further attraction by a head that c-commands both phrases, assuming this attraction is due to a feature borne by both phrases. Thus, (25) and (28) are well formed since no further *wh*-attraction applies. (26) and (29) are ill formed since movement of *qué* and OP, respectively, is nonlocal. To the extent that data such as (25)–(29) are attributable to movement locality, we have evidence for Com or Dom.[24]

4 Conclusion

We have seen five cases in which seemingly very similar approaches to movement locality diverge in their predictions. These empirical distinctions, along with certain formal differences discussed

---

involved in the movement in (25)–(29) may be important for considerations of locality. For example, assume that all movement operations shown in (28)–(29) are cases of *wh*-movement. If the *wh*-feature on *che storie* is uninterpretable and erased following movement 1 in (28), it is not surprising that *a cui* is attracted at a later point in the derivation since this phrase bears the only available unchecked *wh*-feature. However, if the *wh*-feature on *che storie* (28) is checked (and erased), the intervention effect observed in (29) would remain mysterious. It is, of course, possible that the feature may be checked but not deleted (Chomsky 1995, 2001) and thus still block movement of a lower phrase while not undergoing further movement itself. However, though this would allow an explanation of (29), it would leave the licit status of movement 2 in (28) unexplained. A closer look at the Italian facts suggests that the difference between interrogatives and relatives is a key factor. I will leave this issue for further research.

[24] This freezing effect would raise a further complication for standard phrase structures assumed in Chomsky 1993, 1995, where, as in (i), Subj and Obj undergo overlapping movement operations.

(i) a. $[_{\text{Agr}_\text{S}\text{P}}\ [_{\text{TP}}[_{\text{Agr}_\text{O}\text{P}}\ [_{\text{VP}}$ Subj $[_{\text{V}'}$ V Obj]]]]]

b. $[_{\text{TP}}\ [_{\text{vP}}\ [_{\text{vP}}$ Subj $[_{\text{v}'}$ v $[_{\text{VP}}$ V Obj]]]]]

Movement of Obj to Spec,Agr_O or to the outer specifier of v results in mutual c-command between Subj and Obj. If an attracting head is later merged, in some cases AC will predict that neither Subj nor Obj may be attracted. For example, if both Subj and Obj are *wh*-phrases, AC would predict that neither Subj nor Obj may be attracted by a $C_{+\text{wh}}$.

Some notion of equidistance (Chomsky 1993) or a modification of (i) along the lines of Richards's (1999) tucking in (ii) or Koizumi's (1995) split-VP hypothesis (iii) would solve this problem while maintaining the prediction of freezing in (25)–(29).

(ii) $[_{\text{TP}}\ [_{\text{vP}}$ Subj $[_{\text{vP}}\ [_{\text{v}'}$ v $[_{\text{VP}}$ V Obj]]]]]

(iii) $[_{\text{Agr}_\text{S}\text{P}}\ [_{\text{TP}}[_{\text{V}^\text{u}\text{P}}$ Subj V^u $[_{\text{Agr}_\text{O}\text{P}}\ [_{\text{V}^\text{l}\text{P}}$ V^l Obj]]]]]

Alternatively, the following revised versions of Com and Dom would eliminate freezing altogether.

(iv) *Com (revised version)*
G is the closest category in the sister of H iff there is no distinct category K such that K *asymmetrically* c-commands G and K bears a feature matching F.

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

Bruening (2001) provides an important argument that in some cases of mutual c-command freezing does not occur. In these cases mutual c-command is established by a single application of Merge, and either element may be attracted.

Table 1
Empirical and formal differences among locality constraints

	A-over-A (§3.1)	Tucking in (§3.2)	Head movement (§3.3)	Lack of c-command/ dominance (§3.4)	Mutual freezing (§3.5)	Look-ahead	Quantification	Central relation/ theoretical device
Com	no	no	yes	yes	yes	no	no	c-command
Dom	yes	no	yes	yes	yes	no	no	dominance
MutCom	yes	yes	no	no	no	yes	yes	mutual c-command
PathQuan	yes	yes	no	no	no	yes	yes	path theory
SubPath	yes	yes	no	yes	no	yes	no	path theory

in section 2, are summarized in table 1.[25] Examining the table, we see that constraints that define ‘‘closeness’’ with c-command do not predict A-over-A effects. Moreover, neither c-command nor dominance approaches predict tucking in, a phenomenon that requires attention to both the structural description and the structural change of movement. Conversely, these AC approaches are well equipped to predict head movement locality. Relation-quantifying and path-theoretic approaches have trouble with head movement, unless the definition of path is modified. Though contingent on one’s particular assumptions regarding the CED, AC approaches as well as the SubPath path-theoretic approach make the expected predictions in cases where no c-command or dominance relation holds between attractable categories. Quantificational approaches, however, make somewhat bizarre predictions. Finally, as outlined in section 3.5, some AC approaches lead us to expect ‘‘freezing’’ in some cases, although modifications such as those outlined in footnote 24 may alter these predictions.

Next we see that path-theoretic and relation-quantifying approaches are nonlocal (in the sense of Collins 1997) in that they require look-ahead in order to determine differences in locality. Ideally, movement locality could be constrained locally—that is, without reference to later steps in the derivation and without interderivational comparisons.

Furthermore, counting, though by hypothesis not utilized in other parts of the grammar, is required by RQ and PathQuan approaches. Finally, note that addition of movement paths to our theoretical apparatus is required by the path-theoretic PathQuan and SubPath. This feature forces the definition of formal structures that are not utilized by other possible approaches, or elsewhere in the grammar.

Importantly, none of the locality metrics analyzed here account for all of the cases examined, and some have possibly important formal characteristics that should be kept in mind. The recogni-

[25] The first five columns of table 1 indicate the five phenomena discussed in section 3. *Yes* and *no* in these columns indicate whether the corresponding locality constraint accounts for that phenomenon. The final three columns summarize the formal differences discussed in section 2.

tion of these differences will clearly be important in further research. I believe that the bulk of the material presented here supports a theory of locality that utilizes simple relations like dominance and c-command (though the former seems to fare better here), and not more embellished approaches. Furthermore, this suggests that locality of attraction (i.e., choice of mover) should be divorced from choice of landing site. Though many questions remain, the central purpose here has been to reveal important empirical and formal distinctions between commonly used and seemingly similar approaches to the central concept of locality of movement. Though it seems likely that at most one of these approaches is correct, each poses its own problems.

References

Bošković, Željko. 1997. On certain violations of the Superiority Condition, AgrO, and economy of derivation. *Journal of Linguistics* 33:227–254.

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.

Bošković, Željko, and Howard Lasnik. 1999. How strict is the cycle? *Linguistic Inquiry* 30:691–703.

Bruening, Benjamin. 2001. QR obeys Superiority: Frozen scope and ACD. *Linguistic Inquiry* 32:233–273.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle,* ed. by Stephen R. Anderson and Paul Kiparsky, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger,* ed. by Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, Mass.: MIT Press.

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

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik,* ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language,* ed. by Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

Collins, Chris. 1994. Economy of derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25:45–62.

Collins, Chris. 1997. *Local economy.* Cambridge, Mass.: MIT Press.

Collins, Chris. 2001. Eliminating labels. MIT Occasional Papers in Linguistics 20. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL. [Also to appear in *Derivation and explanation in the Minimalist Program,* ed. by Samuel David Epstein and T. Daniel Seely. Oxford: Blackwell.]

Epstein, Samuel David. 1992. Derivational constraints on Ā-chain formation. *Linguistic Inquiry* 23:235–260.

Epstein, Samuel David. 1999. Un-principled syntax: The derivation of syntactic relations. In *Working minimalism,* ed. by Samuel David Epstein and Norbert Hornstein, 27–43. 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.

Erteschik-Shir, Nomi. 1987. Right node raising. In *MIT working papers in linguistics 9,* ed. by M. A. Browning, Ewa Czaykowski-Higgins, and Elizabeth Ritter, 105–117. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.

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, 81–98. Stanford, Calif.: CSLI Publications.

Fitzpatrick, Justin M. 2001a. (In-)visibility. Ms., University of Michigan, Ann Arbor.

Fitzpatrick, Justin M. 2001b. Movement locality in a derivational theory. Ms., University of Michigan, Ann Arbor.

Gärtner, Hans-Martin. 1999. Phase-linking meets minimalist syntax. In *Proceedings of the Eighteenth West Coast Conference on Formal Linguistics,* ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 159–169. Somerville, Mass.: Cascadilla Press.

Groat, Erich M., and John O'Neil. 1996. Spell-Out at the LF interface. In *Minimal ideas,* ed. by Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, 113–139. Amsterdam: John Benjamins.

Jaeggli, Osvaldo. 1982. *Topics in Romance syntax.* Dordrecht: Foris.

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

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

Koizumi, Masatoshi. 1995. Phrase structure in minimalist syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Kracht, Marcus. 2001. Syntax in chains. *Linguistics and Philosophy* 24:467–529.

Martin, Roger. 1999. Case, the Extended Projection Principle, and minimalism. In *Working minimalism,* ed. by Samuel David Epstein and Norbert Hornstein, 1–25. Cambridge, Mass.: MIT Press.

Matushansky, Ora. 2001. Invitation to a beheading, or why head-movement doesn't exist. Presented at Ling-Lunch, MIT, Cambridge, Mass., Dec. 6, 2001.

May, Robert. 1985. *Logical Form.* Cambridge, Mass.: MIT Press.

Moltmann, Friederike. 1992. Coordination and comparatives. Doctoral dissertation, MIT, Cambridge, Mass.

Müller, Gereon. 1996. A constraint on remnant movement. *Natural Language & Linguistic Theory* 14: 355–407.

Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24:480–511.

Nissenbaum, Jonathan. 2000. Investigations of covert phrase movement. Doctoral dissertation, MIT, Cambridge, Mass.

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

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

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

Pesetsky, David. 1982. Paths and categories. Doctoral dissertation, MIT, Cambridge, Mass.

Phillips, Colin. 1998. Linear order and constituency. Ms., University of Delaware, Newark.

Richards, Norvin. 1999. Feature cyclicity and the ordering of multiple specifiers. In *Working minimalism,* ed. by Samuel David Epstein and Norbert Hornstein, 127–158. Cambridge, Mass.: MIT Press.

Rizzi, Luigi. 1982. Violations of the *Wh* Island Constraint and the Subjacency Condition. In *Issues in Italian syntax,* 49–76. Dordrecht: Foris.

Rizzi, Luigi. 1990. *Relativized Minimality.* Cambridge, Mass.: MIT Press.

Sauerland, Uli. 1999. Erasability and interpretation. *Syntax* 2:161–188.

Seely, T. Daniel. 2000. On projection-free syntax. Ms., Eastern Michigan University, Ypsilanti.

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

Suñer, Margarita. 1994. V-movement and the licensing of argumental *wh*-phrases in Spanish. *Natural Language & Linguistic Theory* 12:335–372.

Wilder, Chris. 1999. Right node raising and the LCA. In *Proceedings of the Eighteenth West Coast Conference on Formal Linguistics,* ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 586–598. Somerville, Mass.: Cascadilla Press.

Williams, Edwin. In press. *Representation Theory.* Cambridge, Mass.: MIT Press.

MIT
Department of Linguistics and Philosophy
E39-245
77 Massachusetts Avenue
Cambridge, Massachusetts 02139

jfitzpat@mit.edu