

RESEARCH

The representation of syntactic action at a distance: Multidominance versus the copy theory

Brooke Larson

Department of Linguistics, The University of Iowa, US
brooke-larson@uiowa.edu

It is a common understanding that Merge (Chomsky 1995) effectively explains the preponderance of displacement in language. That is, at least since Chomsky (2001), the mechanism that captures displacement (Internal Merge) has been recognized as something that comes ‘for free’ along with Merge. However, the particular representation of that displacement has been subject to disagreement with some researchers assuming a copy-theoretic view and others a multidominance view. In this paper I offer arguments that support the copy theory of movement over that of multidominance. Multidominance demands that the grammar operate over positions instead of terms, which is incompatible with a Merge-based approach to structure building, and the copy theory demands no such thing. I also argue that the discontinuous interpretation of moved elements can be seen as evidence in favor of the copy theory.

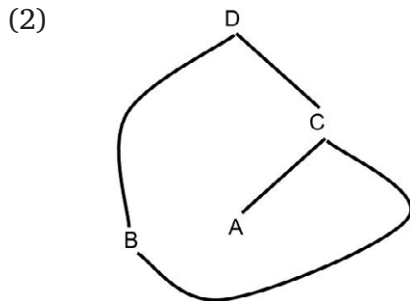
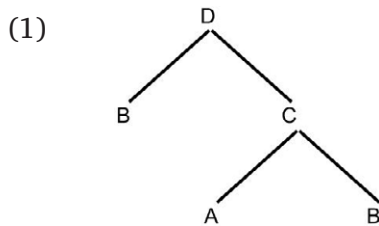
Additionally I note that previous arguments comparing the two representations fail on one of two counts. They either 1) rely on interface-dependent notions about which too little is known to be used to distinguish the two or 2) depend on issues of mathematical power that are not a priori relevant. The new arguments presented here rely on syntax-internal notions and interface notions that are on more solid empirical footing.

Keywords: movement; copy theory; multidominance

1 Introduction

Merge – as introduced by Chomsky (1993; 1995) – has been argued to explain the ubiquity of displacement in human language. Prior to the postulation of Merge, displacement was seen as something surprising and superfluous (Chomsky 1997). However, if Merge is the sort of thing whose output can also be its input, then it should be possible to merge an element a second time to one of the subsequent outputs of that operation. The result of this second merger is the displacement. So much is clear and in this section I lay out the ways this can be represented.

There are two logically possible ways of representing the way in which Merge effects displacement. It can create an entirely new term in what is commonly called the Copy Theory of Movement (CTM) and is discussed in Chomsky (1995) and Nunes (2001; 2004) among others. This is sketched in (1) below. The other logical possibility is that no new instance of the moving element is created, but rather the displaced element is introduced into a new position without vacating its old one. This is represented in (2) below and represents Multidominance (MD) accounts as discussed by a large number of researchers including Epstein et al. (1998), Starke (2001), and Gärtner (2002) among others for simple, ‘upward’ movement. In (1) we find two distinct instances of the displaced element whereas in (2) we find essentially the same structural relations represented using but one instance of the element in question.



The issue at hand is: does (1) or does (2) better represent displacement? Does displacement in the form of re-merge result in multidominance or copies?

What the distinction between CTM and MD boils down to is how to link positions across structural distance. This was once done via movement *chains* (see Chomsky 1995), but given minimalist methodological strictures, we want to see if we can do without those in the syntax. The MD approach is to subsume the chain relation into the Merge operation. In (2) the chain and the displacement are the same thing. The CTM approach exorcises the chain relation from the narrow syntax altogether and leaves it up to the interfaces to do the linking. Collins and Stabler (2016) show that this sort of linkage can be taken care of via (co-)indexation without chains. They note that this approach is compatible with MD representations, but not contingent upon them. In effect the representations demanded by the copy theory should be like that in (3) below:

$$(3) \quad {}_D[B_1{}_C[A B_1]]$$

In short, what MD represents explicitly in the syntax, CTM leaves implicit to be captured by the interfaces, though further discussion of this possibility lies outside the narrow scope of this paper. Each requires stipulations to work correctly, but I argue that adding explicit chains to the syntax causes problems. Leaving the chain relations to the interfaces to deal with avoids those same issues (though perhaps leading to unknown additional problems). This sort of methodological step is a core tactic of minimalism: attributing relations and operations to the interfaces when they cause complications (either empirical or conceptual) in the narrow syntax.

Also, in attempting to answer the questions above, I argue that comparing the resultant shape of the interfaces given one representation or the other is not on its own a viable means to adjudicate between them. Though much is known about the interfaces (Fox 2000; Selkirk 2001; and many others), to this author's awareness there does not exist any work that makes arguments about the representation of displacement by means of interface-related facts. Instead, adjudication between the two must take one of two forms. One, it must be shown that one representation requires a theoretical impossibility, not merely something ungainly. Here I argue that MD analyses require that Merge not operate over terms, but rather positions, which is not possible if one assumes a Merge-based approach to structure building. Two, it must be shown that, given empirically established interface effects, one approach cannot capture a paradigm without recourse to ad hoc stipulation. Without arguments like these, our understanding of the interfaces is not such that it can

at this point serve as a means of arbitration. It is of course difficult to prove this negative and I do not take it as a priori that no such arguments could be made.

Finally, In addition to the authors cited above who focus their attention on ‘upward’ MD displacement that results in c-command, there is an even larger number of researchers who focus on MD that works akin to the sideward movement of Nunes (2001; 2004). That is, MD where the two positions of the element are not in a c-command relation to each other. These works include but are not limited to McCawley (1982); Ojeda (1987); Goodall (1987); Blevins (1990); Wilder (1999); Chung (2004); Guimarães (2004); Citko (2005); de Vries (2005); Chen–Main (2006); van Riemsdijk (2006); Gracanin–Yuksek (2007); Bachrach & Katzir (2009); and Citko (2011). I do not engage with the particular empirical motivations for these approaches because they do not hinge on MD in the first place to a large degree. That is, the empirical effects targeted in the works cited above do not crucially rely on nodes having more than one immediate mother, but rather that displacement be able to result in two elements that do not c-command each other (see Larson 2015 for discussion).¹ In this article, I’m concerned with the problems that result after ‘sideways’ MD applies, not the empirical motivations for positing sideward displacement in the first place.

2 How not to adjudicate

In this section I discuss differences between the two approaches but note that these basic differences are an insufficient basis to choose between the two approaches. First I discuss how interface mechanisms deal with the two approaches. Second I discuss the differences in computation power required by the two approaches. The interface mechanisms that each approach requires are different, but these differences hold at a level of representation where there is little firm methodological footing. The differences in power, while provable are not necessarily relevant as has been shown repeatedly in the past.

2.1 Interface mechanisms

Let us recall what displacement results in. For both approaches, CTM and MD, displacement results in two (or more) separate yet derivationally related structural positions that each have pronounceable material in them. However, for each of these approaches, only one of the related structural positions ends up being the one with something pronounced in it. The challenge for CTM and MD is to capture how it comes to be the case that only one position has the privilege of pronunciation.

Similarly, for both approaches the same two (or more) positions wind up with semantically interpretable material in them. Moreover, the semantically interpretable material in these positions is identical as far as Merge is concerned. Merge does not alter the semantics of that which it works over. The challenge for CTM and MD here is that these

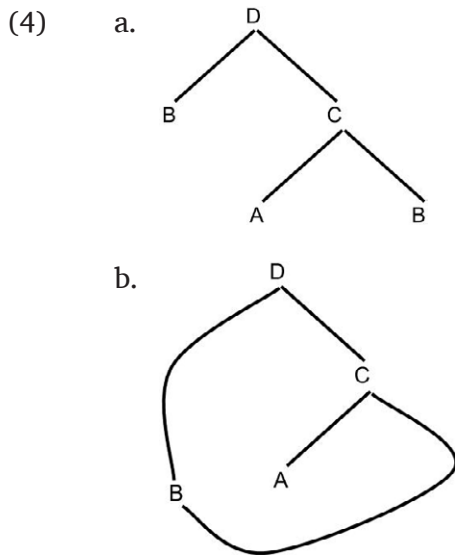
¹ For example, a multidominance analysis of right node raising (like *Jane bought, and Becky sold, used bicycle parts*) maintains that the shared internal object is simultaneously in two conjuncts. This has been argued to explain the fact that right node raising is not subject to island constraints, first noted by McCawley (1982):

i. Jane met the man who sold, and Becky met the man who bought, used bicycle parts.

However, a sideward movement account can do the same since that type of movement is not subject to islands according to Nunes (2001; 2004). And this state of affairs holds generally for non-upward movement. What can be specially attributed to MD can also be invoked as a special ability of sideward movement. The point is basically that, while we know a lot about upward displacement as caused by merge, less is known about sideward relations and it seems that what can be said of sideways MD can also be said of sideways copy movement. Both theories are unconstrained to a degree that both can say, for whatever the facts may be, that that’s simply a property of sideward MD/CTM. Because of this theoretical lability, I don’t think sussing these two apart is possible.

two positions must somehow be distinguished semantically given the obvious fact that displaced elements frequently serve distinct semantic purposes in their different positions.

To see how this is achieved in action, take the toy cases in (4) below, repeated from above:



A descriptively adequate interpretive mechanism for (4a) would need both a means and a motivation to disregard the lower instance of term B as far as the pronunciation is concerned. One well worked-out approach can be found in Nunes (2001; 2004). This work posits that that pronunciation is determined by the number of checked features on a copy. In English, the highest copy is generally going to be the one with the most checked features and as such it will be spelled out. Crucially, this requires a mechanism that cannot allow the pronunciation of unchecked features.

For MD, the interface mechanism needs to be able to distinguish the structurally lower position of B from its higher one and ignore it. Kural (2005) as well as de Vries (2009) posit a means to achieve this goal via ‘graph traversal’, a common type of algorithm from computer science. Under this approach the sketch in (4b) (interpreted as derived graph) is traced by an algorithm that determines the linear order of the terminals based on which ones were encountered first. In the case of *wh*-movement in English, the tree traversal algorithm would, by stipulation, hit upon B in its higher position prior to any other. Crucially, this requires a mechanism that must somehow know to ignore that lower position of B in terms of pronunciation (B is somehow marked as already pronounced upon each subsequent encounter along the traversal.).

These two approaches are not necessarily the only possible approaches to the challenge, but they are representative and embody the minimally conceptually necessary characteristics of a successful approach. That is, they differentially treat either a term or a position. However, we are not in a position to adjudicate between CTM and MD based on a comparison of these interface strategies. Each of these mechanisms does a perfectly adequate job of meeting the fundamental empirical challenge: either via unchecked features or already-traversed elements, pronunciation will only happen once. As such, they are indistinguishable based on the basic facts.

There could be an argument that one type of mechanism or other aligns better with the syntactic structure that feeds it. This would take a form like: tree traversal works better with MD, therefore MD. Any gesture in that direction would simply be begging the question since it is exactly the nature of the syntactic structure that is in question.

So, empirical grounds and syntactic concerns are not viable venues to tackle this issue. Instead, a comparison must be made based on whether one approach or the other is more plausible *as a PF-interface mechanism*, however, there is no current argument that suggests one mechanism is more plausible than another.

The same concern holds for the meaning side of things. The challenge that faces the CTM account is that two terms that are non-distinct in the syntax must somehow bear differing denotations at the semantic interface. Take the concrete example in (5). The higher instance of *which person* must be interpreted as an operator and the lower instance of it must be interpreted as a variable.

(5) Which person did you see < which person > ?

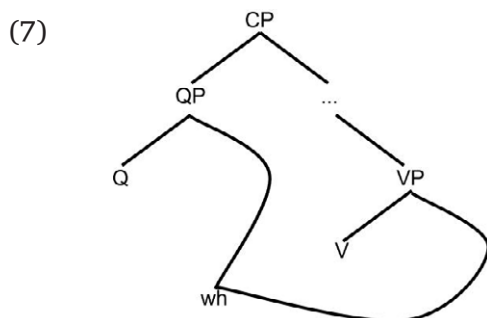
Given that there are two different terms in the CTM version, it is logically possible to treat them different at the LF-interface. One way of doing this has been explored by Fox (2002) as well as Sauerland (2004) and has been dubbed ‘trace-conversion’ (though it of course works with copies just as well as traces). This proposed interface operation takes the structure implicated in (5) and converts the lower version into a definite description bearing a bindable variable as shown in (6) below (after λ -abstraction):

(6) Which person λx [you saw the person x]

The problem of differential semantic interpretation in different positions is on the surface much more difficult for MD approaches. Any approach that follows the intuitions of trace conversion would require the self-same element to be interpreted differentially.²

Johnson (2012) avoids this issue by denying that the displaced element needs to be so dually interpreted. Instead he posits that the *wh*-phrase functions solely as the variable. By analogy with the approach sketched for CTM, under Johnson’s analysis the *wh*-phrase functions like the lower copy in (6) and something else serves as the operator. The operator rule is fulfilled by a Q-head of the sort hypothesized by a variety of researchers among them Hagstrom (1998), Kishimoto (2005), and Cable (2007; 2010).

Under this approach, the displaced *wh*-phrase does not find itself multiply dominated in its base position and *spec,CP*. Instead, the Q-element heads up a QP phrase in *spec,CP* and the *wh*-phrase is multiply dominated in its base position and as the complement of the QP. To clarify, this structure is given below in (7).



² This would be possible only if the element that is represented a single object in the syntax is ‘copied’ in some sense in at the LF-interface. This is of course possible, but it is not without its intuitive problems for the task at hand here. For example, as an argument for MD in the syntax, this sort of mechanism would involve some sort of bizarre question begging. Instead of presuming the conclusion, it would require presuming a conclusion that does not follow from the argument.

Again, we find two approaches that are equally well-positioned to capture the driving empirical fact: the displaced *wh*-phrase is involved in an operator-variable relationship. This being the case we are again in a position where it is not possible to argue for CTM or MD over one another on empirical grounds. Further, we are once again in a position where arguing for CTM or MD based on how well these approaches align with those approaches begs the question.

Given the facts as they stand here, the only legitimate way to compare CTM and MD is whether the interface-mechanisms required of these approaches are more plausible again as *LF-interface mechanisms*. The CTM approach requires something like trace conversion and the MD version, as admitted by Johnson requires “a theory of the syntax/semantics interface that does not force sisters to systematically combine” (Johnson 2012: 20). Much like with the PF-interface question, there is no argument that the nature of the LF-interface is better suited to one or the other of two mechanisms.

In short, CTM and MD approaches to displacement do not in and of themselves solve the displacement problem better than their competitor. They both require interface mechanisms to capture the basic empirical facts. As we have seen, it is possible to devise such mechanisms that are both coherent and empirically adequate. However, arguments in favor of CTM or MD cannot be adduced based on these mechanisms. In fact, the logic goes the other way. If it were possible to somehow conclude that CTM or MD was correct, then it would be possible to make interesting claims about what the nature of the interface mechanisms would need to be. Before that, in the next subsection I will discuss another possible means of arguing for CTM or MD, one that also ends up insufficient.

2.2 Power issues

Another means to potentially distinguish CTM from MD and in turn favor one or another is via formal computational power comparisons. In this subsection I discuss some previous instances of this sort of reasoning and argue that they too are insufficient. Much like with the previous discussion with interface mechanisms, this is not to deny the results or import of these investigations. Rather, I intend to deny that the results ought to be used to determine whether CTM or MD is the correct representation. To the contrary, the results of investigations into the relative power of CTM and MD serve to make clear the unavoidable consequences of the correct representation.

As a guide to our current discussion, it is helpful to return to some issues that were first broached previously in the field. Before Chomsky argued in favor of transformations (in Chomsky 1957), he used computational power to arbitrate between two conceptions of grammar. Finite state machines were shown to be fundamentally incapable of expressing the sort of patterns found in human language. In light of this result, it was necessary to posit a conception of the grammar that was at least as powerful as a phrase structure grammar. The upshot of this story is to note that this sort of reason is clear of ambiguity or dispute. If there are patterns of human language that are inexpressible assuming CTM or assuming MD, the decision between the two will be simple.

On the other hand, the decision between CTM and MD conceptions of displacement may end up like the comparison between the trace theory of displacement (as in Chomsky 1981) versus the CTM that supplanted it. Compare the representations in (8). In (8a) there is an inherently silent trace in the direct object position whereas in (8b) there is an unpronounced copy of the moved term.

- (8) a. What_i did you see t_i
 b. What_i did you see <what_i>

It is trivially the case that these two options are of equal expressive power: any position that movement could have stemmed from could equally leave a trace or an unpronounced copy. The way to distinguish these two options is via empirical evidence and theoretical concerns. Reconstruction effects are the empirical evidence in favor for CTM (as argued in Chomsky 1993). On the theoretical side, the Inclusiveness Condition (Chomsky 1995) serves to eliminate trace theory as a viable option. Again, even if two options have the same expressive power, they can still be theoretically and empirically distinguished.

To recapitulate the previous discussion: One, equality in terms of expressive power does not entail that linguistic distinctions (empirical or theoretical) cannot be made. Two, insufficient expressive power is indeed dispositive.

There indeed have been some mathematical investigations into CTM and MD. One notable investigation is that of Kracht (2008). He argues CTM and MD are very similar and while not identical, they are “identical for all linguistic purposes” (Kracht 2008: 527) and shows this by proving various mathematical mapping relations between the two. As it turns out, the formal mathematical properties of CTM and MD do not relevantly distinguish the two. Anything that can be stated in one can be stated in the other. Because of this, he argues, they are linguistically identical.

This should remind the reader of the trace-theory versus CTM discussion above. There too it is the case that anything that can be expressed in one can be expressed in the other. In other words, the argument that CTM and MD are linguistically identical is not necessarily true. As a general point, it should be well known that when doing cognitively- and biologically-minded linguistics mathematical considerations are not necessarily the final word. Berwick and Weinberg (1982: 177) argue that “mathematical relevance need not imply cognitive relevance.” However it is this implication that many see Kracht as making. This is not to dispute the basic results of this work (though see Kepsner 2010 with some criticism), but to maintain that they are not necessarily the sort of results that help us choose between CTM and MD. This paper has not been concerned with formal computational equivalence or lack thereof between CTM and MD. Instead, it has been concerned with whether a particular choice of encoding is ‘better’ with respect to syntactic theory and empirical considerations.

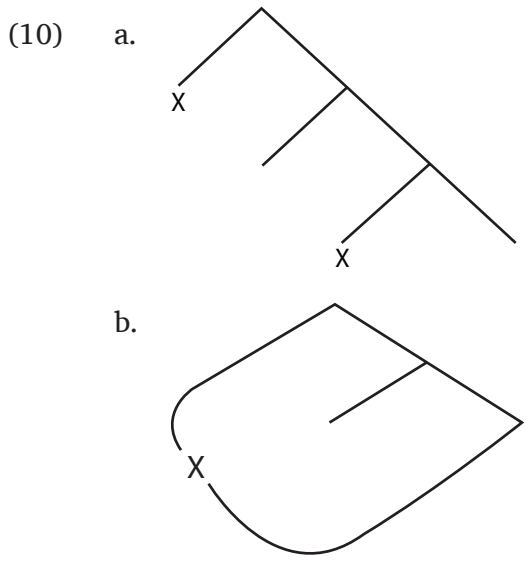
3 How to distinguish MD from CTM

It is not easy to determine whether CTM or MD is the correct representation of displacement. Each approach has the capacity to capture the basic empirical facts: Overt elements can be interpreted far from where they appear phonetically. But we can settle this issue by focusing on theoretical arguments or arguments involving more complicated empirical effects. In this section I provide two means of distinguishing between the two, a theoretical argument and an empirical one. The first involves fundamental notions of what the target of the Merge options is and the second concerns the possibility of, and constraints on, ‘scattered’ interpretation.

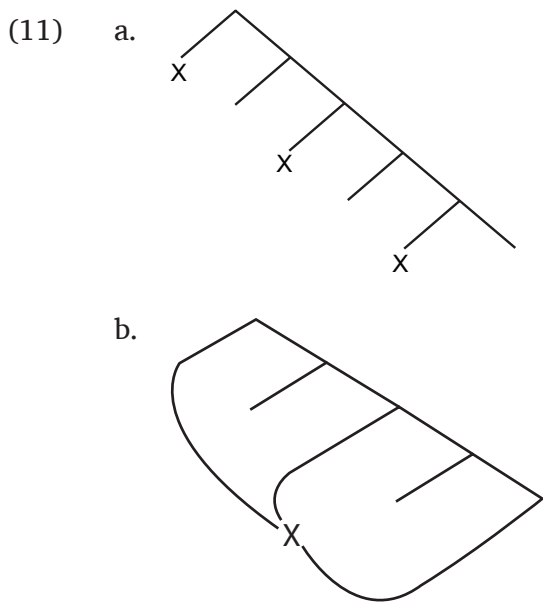
3.1 The target of successive merges

Prior to displacement, as in a structure like that in (9) below, there is only one token instance of X and as such it is at this point that CTM- and MD-style representations are identical. It is only after displacement (second Merge) has occurred that the representations differ along the lines shown in (10).

(9) [... X ...]



In (10a), there are now two token instances of X whereas in (10b) there is still but one. Of course displacement does not necessarily halt after one hop. Rather, displacement can happen again and again in an iterated fashion as argued by McCloskey (1979), Torrego (1984), Henry (1995), and McCloskey (2001) among others. That is, if X were to continue to be Merged into higher positions, we would be left with representations like those in (11).



However, in (11b) it is always the case that X is *still there* in its base position even though it also finds itself much higher in the structure as well. This contrasts with CTM where there exist higher tokens of X structurally independent of the base position copy. This distinction lies at the heart of the differing interface strategies outlined in the previous section. In those instances, work went into trying to capture how CTM or MD worked with grammatically licit structures and as we saw, there was little to distinguish them by. That is, each system was powerful enough to capture the licit structures as they pertain to interface interpretation. In this section, I employ a different method.

What other discussions neglect is how the differing interpretations of displacement fare when dealing with grammatically illicit structures and interface-independent notions. In this section, I will argue that MD is not in a position to explain certain ungrammatical sentences based on these narrowly syntactic concerns. In fact, the syntax-internal wedge that

will serve to distinguish the two approaches will be Merge, the operation that prompted this discussion of CTM and MD to begin with.

The logic will be as follows. In order to capture the ungrammaticality of certain island violations, MD analyses require something that the system does not allow: the targeting of non-terms by Merge. Because Merge necessarily works over terms, this means that MD representations are incapable in principle of being used to capture island effects.

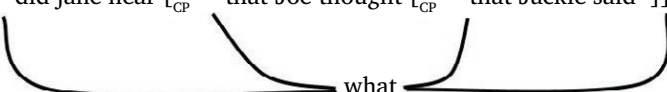
It has long been known that displacement cannot occur over an arbitrarily great structural distance in a single derivational step. To take a classic example from Ross (1967), displacement cannot occur across a complex noun phrase like in (12).

(12) *What did Jane hear the rumor that Joe saw?


Chomsky (1973 et seq.) analyzed this constraint as one of moving too great a distance. The precise formulation of the short-steps constraint has changed over the years, but the basic notion still holds: displacement cannot take place over an arbitrarily great distance.³

However, under MD accounts and assuming a representational approach to modeling syntactic relations, movement relations in fact *do* hold over arbitrarily great distances. An MD-style representation of the sentence in (13) requires that there be a Merge-derived relation between the base position of X and its final position. This is abstractly sketched in (14).⁴

(13) What did Jane hear that Joe thought that Jackie said?

(14) [_{CP} did jane hear [_{CP} that Joe thought [_{CP} that Jackie said]]]


In (14) there is in effect a movement relation that holds across a very great distance. Contrast this with the CTM representation in (15) where there are only local relations.

(15) [_{CP} what did jane hear [_{CP} <what> that Joe thought [_{CP} <what> that Jackie said <what>]]]


This distinction alone is not sufficient to prefer CTM over MD as it could easily be argued that somehow in virtue of the myriad shorter dependencies in (14) the very long distance dependencies are rendered licit. Perhaps the fact that a series of short hops are instantiated in (14) is sufficient. That is, in virtue of there being a licit 'route' from base position to final derived position, the representation is grammatical. At each step in the derivation

³ A contemporary interpretation of the prohibition on movement over too great a distance is found in phase-theory (Chomsky 2001;2008). Under phase-theory, lower structure is periodically rendered inaccessible for grammatical operations. This causes independent problems for MD analyses as it would require that the moving element be rendered inaccessible to further operations as well: an unwanted consequence. If the moving element were somehow not rendered inaccessible, this would force a nonsensical state of affairs wherein a VP would be rendered inaccessible, but not its constituent parts. To the degree that a lower position of a moving element in an MD theory can be spelled out, but a higher position not, this would just be incorporating copies back into theory. If MD structures are rendered into copies upon phase-based spell out, the advantages of reifying the chain in the syntax is lost.

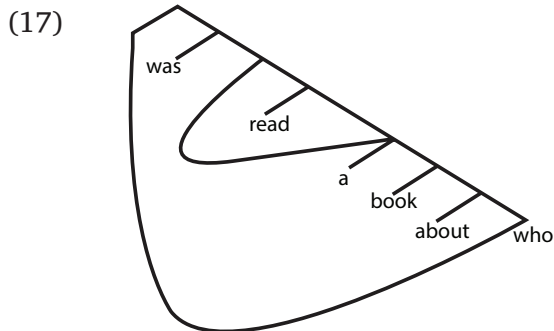
However, a reviewer points out that phases cause trouble for CTM analyses as well since successive cyclic movement is predicated on the fact that the moving element bear an uninterpretable feature. However, the lower copies that spurred the movement will not get their features checked and, when spelled in a phase-based fashion, may cause the derivation to crash.

⁴ I ignore here landing sites other than Spec, CP for expository purposes.

there was a position that X was in that was accessible for further movement, even though there were other positions that X was in which were inaccessible. Call this the ‘whatever works’ interpretation.

This ‘whatever works’ interpretation of MD movement runs into immediate problems when looking at freezing effects like those in Wexler and Culicover (1980). Such effects are found when movement is attempted out from within an element that itself has already moved. This can be seen in (16) where the subject of the sentence has A-moved from its base position as the internal object of the verb. From within this derived subject, a wh-word has been extracted, leading to ungrammaticality. The sentence in (16) as depicted under MD would be like that in (17).

(16) *Who was a book about read (by Jane)?



Note that the ‘whatever works’ approach here makes the wrong prediction. Moreover, it makes the wrong prediction for CTM as well. There in fact exists a licit ‘route’ between the base position of *who* and its final position: via the base position of the derived subject. As seen in (18), movement from within the noun phrase in object position is perfectly licit. For some reason, it is only the most recent position that the moving elements are in that matters, for both MD and CTM. This will become more relevant presently.

(18) Who did Jane read a book about?

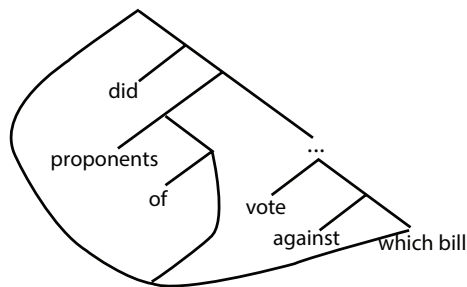
As such, the ‘whatever works’ approach does not seem tenable. Another plausible approach would be to restrict operations to just between the root of the tree and ‘structurally closest’ instantiation of the to-be-moved element. This is a MD version of Shortest Move (Chomsky 1995): create the shortest link. For the representation in (17), the wh-word is effectively in two positions. It is within the larger DP in the lower, PP-complement position as well as within the DP in the higher, subject position. If we adopt the ‘structurally closest’ stricture, movement of the wh-word is necessarily assessed as if it came from the subject position. Since this closer position is an island for movement, the sentence is correctly ruled out.

However, even the ‘structurally closest’ approach cannot be viable. There are instances where extraction from the structurally closest position is also illicit. Take the following subject-internal parasitic gap sentence.

(19) Which bill did even proponents of decide they had to vote against?

Following the approach to parasitic gaps found in Nunes (2001; 2004), the above sentence is derived by the wh-word beginning its derivational life within the subject. It then side-wards moves into the complement position of *against* (crucially before the incipient subject is merged to the spine of the tree and thus becomes an island) and then makes its way to the left periphery of the matrix clause. The MD representation is sketched in (20) below.

(20)



In this sort of parasitic gap, the structurally highest position of the moving wh-word is an island for movement as seen in (21) below where there is no alternate position to assess:

(21) *Which bill did even proponents of insult each other?

If one were to adopt the ‘structurally closest’ approach, the assessment of whether the movement is licit would be restricted to the subject-internal position. Since that position is island-internal, the sentence should be ruled out, contrary to fact.⁵ Even if we were to posit, as a reviewer suggests, that the wh-word evacuates the incipient subject and moves to the spec, CP position before the merger of the subject to the spine, we’d still run into problems. In this scenario we are forced to assume a counter-cyclic merger of the subject to the spine after the wh-word has already moved above that position.

In sum, it’s not that any licit link rules in a sentence, nor is it the case that as long as the shortest link is illicit, the sentence is ruled out. Instead it seems that recency is what matters, not just for parasitic gaps, but for all upward movement. For CTM it is the most recent copy of the moving element that can be targeted; for MD it is the most recent position of the moving element that can be targeted.⁶

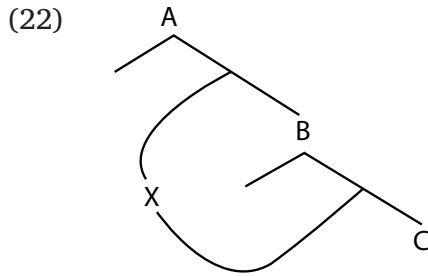
Herein lies the problem for MD accounts. Merge does not operate over positions, but rather terms in the sense of the constants and variables that constitute sentences (Chomsky 1995). In the minimalist program, terms are taken to be elements like *the* and *dog*, but not relations between them.⁷ One such relation between terms is that of sisterhood. Sisterhood can be used to identify a position, but neither the sisterhood relation or the position that it defines are terms. This will cause problems for MD accounts.

Take the MD representation in (22). Here, the only licit means for X to be Merged with A is for the position “sister of B” to be operated over by Merge. It cannot be the case that that X itself is targeted for Merge because X is also in its too distant base position. Recall that it is simply not enough to say that Merge with X is licit in virtue of there being some licit extraction site. As we saw with the freezing effects above, the existence of a licit extraction site is necessary but not sufficient. If Merge targeted X, it would in effect be targeting an inaccessible term. Only the most recent *position* can be operated over. The position (sister of B) is not a term and as such cannot be operated over by Merge.

⁵ It should be noted, as a reviewer points out, that the same sort of argument can be made in principle even if a sideward movement analysis of parasitic gaps is not assumed. Brody (1995) posits that in such constructions there is a certain type of chain between the moved element and its trace along the spine of the tree and a different type of chain between a null operator and that which licenses it. He captures the same facts by recourse to the reference of terms and not positions.

⁶ For CTM analyses, the notion that only the most recent copy is relevant for further Merge could be (and is) easily implemented in phase-theory wherein previous copies are rendered inaccessible throughout the derivation. In such a case, the most recent position can be determined and the term in that position can enter into the Merge operation. However, a review points out that if in situ copies of moved elements retain unvalued features that are illegible at the interfaces, CTM will have problems with phase theory as well.

⁷ That operations target terms in positions is not a novelty of Merge. Rather, terms have always been that which grammatical operations target going back through move alpha (Lasnik and Saito 1994) on through to Chomsky (1955).



Because MD requires Merge to operate over non-terms for parasitic gap constructions, it cannot be maintained as a possible representation of displacement. For MD accounts to work for these constructions, it must be the case that Merge be altered so as to be able to operate on more than just grammatical terms.

It is perhaps an option to amend Merge such that it can operate over positions (defined in terms of sisterhood: {X,Y}, X and Y being sisters) instead of or in addition to terms. However, given Bare Phrase Structure, there is no narrow syntax distinction between terms and their positions. Previously, under X-bar-theoretic conceptions of phrase structure (Chomsky 1970 and Jackendoff 1977), structural positions were in fact reified things (Spec, XP, for instance, existed independent of the term in that position.). As such, it would be theoretically possible to target a position as such. This is not the case under Bare Phrase Structure. The Inclusiveness condition (Chomsky 1995: 225) prohibits novel entities like positions to be syntactically reified and because of this, Merge cannot be amended so as to be able to target them.⁸

3.2 Scattered interpretation

Above I presented a theoretical reason why CTM is to be preferred over MD. In this section I present what is at heart an empirical case. As seen above, such arguments are not easy to make in any decisive way. Here an interface-related argument that is not only sound, but one that could serve as a template for other interface arguments is presented. The argument relies on developing an empirical generalization and then determining how the two approaches might handle that generalization.

The basic empirical point is found in Chomsky 1993. He notes that in a sentence like (23) there is a correlation: when *Bill* is antecedent of the reflexive, the sentence is ambiguous between an idiomatic reading in which *take a picture* means *photograph* and one where it means *take possession of*. In contrast, when *John* is the antecedent the sentence only has the second non-idiomatic interpretation.

(23) John wondered which picture of himself Bill took.

⁸ A reviewer notes that in Gärtner’s (2002) theory of MD it is possible to arbitrarily define representational constraints to effectively capture apparent reference to position. This is true, but this hides some problems. In Gärtner’s approach, the parasitic gap examples from above are ruled in by allowing displacement to be licit when it stems simultaneously from two non c-commanding positions (Gärtner 2002: 128). There are problems with this.

For instance, the structural description of when across-the-board or parasitic gap-like displacement is licit predicts that such displacement should be licit *so long as* the two relevant positions do not c-command each other. But this is false because across-the-board extraction from within a subject or adjunct is still ungrammatical:

- (i) a. *What did Jane arrive after buying and before selling?
- b. *What did supporters of and opponents of eat lunch?

The above sentences could surely be ruled out under Gärtner’s approach via further specifications on the configurations where such types of displacement are licit. But to capture the recency effects without explicit reference to recency demands the stipulation of specific constructions where movement constraints are specially obviated.

Chomsky accounts for the contrast by positing differing LF interpretations of the string in (23). For us, the crucial interpretation is the one where *John* is the antecedent and the idiomatic reading is not possible. For such an interpretation, the LF is like that in (24).

- (24) [John_i wonders [CP which picture of himself_i [Bill took ~~which picture of himself_i]]]]~~

Chomsky argues that the idiomatic reading is ruled out in such an instance because of three things. One, the reflexive must move above *Bill* for the relevant construal with *John* to arise. Two, it is not possible to interpret a copy of *picture* without interpreting the copy of *himself* that it dominates. That is, there can be no scattered interpretation of *picture* and *himself*. Three, in the framework at hand, DS no longer exists and as such the requirement that idioms be constituents at that level is reanalyzed as holding of LF instead. Because of this, the idiomatic reading requires that *picture* (and *himself* in turn) be interpreted in its base position. This precludes the simultaneous idiomatic interpretation and matrix subject construal of the reflexive.

Chomsky is not entirely technically explicit here, but the basic interface-theoretic notion is clear: The interpretation of a given moved syntactic term at LF necessitates the interpretation of that which it dominates. When the syntactic term is interpreted low, so too should its dependents. It should be clear from the previous discussions that MD approaches can easily capture this interface generalization. As far as the reflexive is concerned, one of its positions is sufficiently local to *John* and one to *Bill*. Further, one position of *picture* feeds an idiomatic interpretation and one does not. We can encode the interpretive dependences in the MD idiom as in (25).

- (25) If position 1 of element X is interpreted, then all elements Y that X dominates are also interpreted in position 1 unless elements Y displace to a yet higher position.⁹

Something like (25) prevents the scattered interpretations of constituents for MD, mimicking the analogous prohibition against scattered deletions discussed above in the context of the CTM. More particularly, (25) would prohibit interpreting the higher position of the *himself* and the lower position of *picture* in (23).

Again, we find ourselves in a position where CTM and MD are equally able to account for the data, given certain assumptions: If a moving syntactic term is interpreted in a particular position, so too must be its complement. However, something like what we have in (25) cannot be the last word. Sometimes scattered interpretations are in fact required. An example of this can be found in what are commonly known as Lebeaux-effects (Lebeaux 1988; 1991; Chomsky 1995). That is to say, that for the following sentences, it is still possible to get the idiomatic readings. Nevertheless, the adjunct to the wh-phrase is not interpreted low down in the idiom: There are no Principle C effects.

- (26) a. Which picture that Bill_i hated did he_i say that Mary took?
b. Which habit that Bill_i hated did he_i say that Mary finally kicked?

CTM accounts can rely on the late adjunction of the relative clause to capture these anti-reconstruction effects (as in Stepanov (2001) and others). After movement of the wh-phrase to its non-base position, an adjunct can Merge with it and not the copy left in situ.

⁹ This definition allows for the quantifier raising of a possessor DP to the exclusion of the possessee as in *Bill took everyone's picture* assuming a structure like [[everyone]_{DP1}] ['s [picture]_{DP1}] DP2] subject to independent covert movement constraints and similar things can be said for in-situ wh-phrases. In sentences like *which picture did Jane take* I follow Kayne (1994) and Cable (2010) and take the the moved phrase *which picture* to be decomposable into a DP with a DP specifier. The DP-specifier is what drives the movement, but there is obligatory pied piping. In essence, the specifier is interpreted high as the operator, but per (25) that does not demand that *pictures* be interpreted high since *pictures* is not dominated by the specifier of the DP.

The higher copy can be manipulated independently of the lower one and if the relative clause is appended to the higher copy only, no Principle C effects are predicted. MD accounts can surely also resort to late adjunction, but this is going to have an unwanted effect: the adjunct is late-adjointing low as well as high and the lack of Principle C effects is not predicted.

Perhaps the injunction against scattered interpretations in (25) could be altered so as to make reference to segments and the non-standard assumption that the relative clause adjoins to the DP ‘which picture’ not the NP ‘picture’. In this case, the moving DP will not *fully* dominate the relative clause adjunct. But even then, we could further embed the adjunct such that it would be fully dominated by something:¹⁰

(27) Which picture of [the dog that Bill_i hated] did he_i say that Mary took?

We can turn to the idiom *to take advantage* to test this. In example (28) below we have a discontinuous idiom with *advantage* and its complement being part of a degree question. Within the complement of *advantage* we have a noun that can take either an adjunct relative clause or a complement clause. As can be seen, the example with the relative clause does not effect a Principle C violation whereas the example with the complement clause does.

(28) a. How much advantage of the claim that Bill_i made did he_i end up taking?
b. *How much advantage of the claim that Bill_i died did he_i end up taking?

What injunction could be made such that an MD account could handle these anti-reconstruction effects? Perhaps a “whatever works” sort of approach that would absolve any Binding Principle violation if at least one occurrence obeyed them. This would falsely predict that non-adjunctions would also show anti-reconstruction effects:

(29) *Which picture of Bill_i did he_i say that Mary took?

One could claim that idiom-internal adjuncts are immune to binding theory strictures. Though this would also make the incorrect predictions:

(30) *He_i took some pictures that Bill_i hated.

One final option, as offered by a reviewer, is to tie the assessment of at least Principle C violations to derivational timing. That is, each time a nominal is (re)merged into the structure, it is determined whether that nominal c-commands a co-indexed R-expression. Once a nominal reaches its highest A-position, it is no longer considered a potential antecedent. If an adjunct containing a co-indexed R-expression late-adjoints to a position in its c-command domain, that nominal will not be reassessed and there will be no Principle C violation.

This approach runs into problems because late-adjunction doesn’t depend on displacement. If we assume both this sort of Principle C assessment and the possibility of late-adjunction, this evasion of Principle C effects should arise independently of whether the adjoined-to element has moved or not. That is it should be possible to late-adjoin an element with a R-expression to an in situ element in the c-command domain of a co-indexed element and not effect a violation. But this is not the case:

(31) *He_i picked out a color that Jake_i liked.

¹⁰ A reviewer notes that this test might not be entirely useful since Nunes (2001) argues that reconstruction-driven Principle C effects are not as strong when the R-expression is embedded more generally, even within nominal complements which should obligatorily reconstruct with their host. However, it is hard to construct the correct test with the picture idiom since the picture would need to be of a noun that takes a complement clause (like ‘rumor’ or ‘claim’) and such nouns are not generally things that can be photographed.

Of course one could further stipulate that late adjunction only applies to displaced elements, but this just puts off the same problem:

(32) *He_i knew which of the colors that Jake_i selected would be most popular.

To make such an approach empirically viable, it would be necessary to stipulate that late adjunction of elements containing R-expressions is only possible if the element they adjoin to is not in the c-command domain of a co-indexed expression. That is, it would be necessary to simply redescribe the empirical facts.

In short, the scattered interpretation facts can be cleanly captured under CTM but cannot be so captured by MD theories. CTM allows for the independent manipulation of terms in a way that cannot be said of MD. That is, late adjunction of adjectives allows for anti-reconstruction effects in certain instances but obligatory reconstruction for certain complements. The same cannot be said for MD. Again, the distinction comes down to dealing with terms or positions. In this case, we assume that adjuncts attach to terms and not to positions. If the adjuncts above could adjoin to positions to the exclusion of the term in them, then the problem for MD would not arise. However, this is not possible in a Merge-based system the deals in terms. This is not to say that MD accounts cannot in principle somehow capture this data, but much like Chomsky's argument from transformations, that approach would be less enlightening.

4 Conclusion

Merge as an explanation of displacement is an important conceptual step forward in the explanation of grammatical properties. The otherwise problematic fact of syntactic action at a distance is rendered much tamer by this re-interpretation of structure building. Coupled with other current guiding principles in syntax, it has led to a question about the result of an object undergoing Merge more than once.

The two logically possible options, each reasonable in its own right, have been assumed, adopted, and argued for in the past, but here we have seen that there has been little clear basis for those moves. The clear differences between the copy theory and multidominance do not readily translate into strong arguments in favor of one over another.

The answer to these questions of which representation is correct will have repercussions elsewhere. The obvious import of the answer lies in the fact that the merge-based explanation of displacement is one of the major theoretical innovations of modern syntactic theory and answering the above questions will of course be inherently important for that very reason. Additionally, the choice between the two representations will be practically important in terms of the demands that the competing representations make on other parts of the grammar. CTM and MD each require different things of the interfaces and opting for one representation over another forces the interfaces to have certain properties.

Instead I have argued that to adjudicate between the two requires discussion of interface-independent notions (such as what is targetable by Merge) or interface generalizations that enjoy some empirical support. When comparisons of that sort are made, it becomes possible to argue for one approach over another. In the instances relayed above, a copy-theoretic conception of displacement is preferable.

Acknowledgements

I would like to thank Norbert Hornstein, Dave Kush, Barbara Citko, Edith Aldridge, and Dennis Ott for helpful criticisms of this work as well as the tireless and very fruitful reviews of the anonymous Glossa reviewers who worked to improve this article.

Competing Interests

The author declares that she has no competing interests.

References

- Bachrach, Asaf and Roni Katzir. 2009. Right node raising and delayed spellout. In Kleantes Grohmann (ed.), *InterPhases: Phase-theoretic investigations of linguistic interfaces*, 283–316. Oxford: Oxford University Press.
- Berwick, Robert C and Amy S. Weinberg. 1982. Parsing efficiency, computational complexity, and the evaluation of grammatical theories. *Linguistic Inquiry* 13. 165–191.
- Blevins, James. 1990. *Syntactic complexity: Evidence for discontinuity and multidomination*. Amherst, MA: University of Massachusetts dissertation.
- Brody, Michael. 1995. *Lexico-Logical Form: A radically minimalist theory* (Linguistic Inquiry Monograph 27). Cambridge, MA: MIT Press.
- Cable, Seth. 2007. *The grammar of Q: Q-particles and the nature of Wh-fronting, as revealed by the Wh-questions of Tlingit*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford: Oxford University Press. DOI: <http://dx.doi.org/10.1093/acprof:oso/9780195392265.001.0001>
- Chen–Main, Joan. 2006. *On the generation and linearization of multi-dominance structures*. Baltimore, MD: John Hopkins University dissertation.
- Chomsky, Noam. 1955. Logical structure of linguistic theory. Ms. Parts of 1956 revision published in New York: Plenum, 1975, and Chicago: University of Chicago Press, 1985.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Walter de Gruyter.
- Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs and Peter Rosenbaum (eds.), *Reading in English transformational grammar*, 184–221. Waltham: Ginn.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart, and Winston.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale and Samuel J. Keyser (eds.), *The view from building 20*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1997. *New horizons in the study of language*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press. DOI: <http://dx.doi.org/10.7551/mitpress/9780262062787.003.0007>
- Chung, Dae-Ho. 2004. Multiple dominance analysis of right node raising constructions. *Language Research* 40. 791–812.
- Citko, Barbara. 2005. On the nature of merge: External merge, internal merge, and parallel merge. *Linguistic Inquiry* 36. 475–496. DOI: <http://dx.doi.org/10.1162/002438905774464331>
- Citko, Barbara. 2011. *Symmetry in syntax: Merge, move and labels*. Cambridge: Cambridge University Press. DOI: <http://dx.doi.org/10.1017/CBO9780511794278>
- Collins, Chris and Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19. 43–78. DOI: <http://dx.doi.org/10.1111/synt.12117>
- de Vries, Mark. 2005. Merge: Properties and boundary conditions. *Linguistics in the Netherlands* 22. 219–230. DOI: <http://dx.doi.org/10.1075/avt.22.20vri>


- de Vries, Mark. 2009. On multidominance and linearization. *Biolinguistics* 3. 344–403.
- Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. New York: Oxford University Press.
- Fox, Danny. 2000. *Economy and semantic interpretation* (Linguistic Inquiry Monograph 35). Cambridge, MA: MIT Press.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33. 63–96. DOI: <http://dx.doi.org/10.1162/002438902317382189>
- Gärtner, Hans-Martin. 2002. *Generalized transformations and beyond: Reflections on minimalist syntax*. Berlin: Akademie Verlag. DOI: <http://dx.doi.org/10.1524/9783050074757>
- Goodall, Grant. 1987. *Parallel structures in syntax: Coordination, causatives, and restructuring*. Cambridge: Cambridge University Press.
- Gracanin-Yukse, Martina. 2007. *About sharing*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Guimarães, Maximiliano. 2004. *Derivation and representation of syntactic amalgams*. College Park, MD: University of Maryland dissertation.
- Hagstrom, Paul. 1998. *Decomposing questions*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Henry, Alison. 1995. *Belfast English and Standard English: Dialect variation and parameter setting*. New York: Oxford University Press.
- Jackendoff, Ray. 1977. *X-bar-syntax: A study of phrase structure* (Linguistic Inquiry Monograph 2). Cambridge, MA: MIT Press.
- Johnson, Kyle. 2012. Towards deriving differences in how Wh Movement and QR are pronounced. *Lingua* 122. 529–553. DOI: <http://dx.doi.org/10.1016/j.lingua.2010.11.010>
- Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monograph 25). Cambridge, MA: MIT Press.
- Kepser, Stephan. 2010. On monadic second-order theories of multidominance structures. In Christian Ebert, Gerhard Jäger, and Jens Michaelis (eds.), *The mathematics of language*, 117–128. Berlin: Springer. DOI: http://dx.doi.org/10.1007/978-3-642-14322-9_10
- Kishimoto, Hideki. 2005. Wh-in-situ and movement in Sinhala questions. *Natural Language and Linguistic Theory* 23. 1–51. DOI: <http://dx.doi.org/10.1007/s11049-004-6574-0>
- Kracht, Marcus. 2008. On the logic of LGB type structures. Part I: Multidominance structures. In Fritz Hamm and Stephen Kepser (eds.), *Logics for linguistic structures*, 105–142. Berlin: Walter de Gruyter.
- Kural, Murat. 2005. Tree traversal and word order. *Linguistic Inquiry* 36. 367–387. DOI: <http://dx.doi.org/10.1162/0024389054396890>
- Larson, Bradley. 2015. Minimal search as a restriction on merge. *Lingua* 156. 57–69. DOI: <http://dx.doi.org/10.1016/j.lingua.2014.12.009>
- Lasnik, Howard, and Mamoru Saito. 1994. *Move alpha: Conditions on its application and output*. Cambridge, MA: MIT press.
- Lebeaux, David. 1988. *Language acquisition and the form of the grammar*. Amherst, MA: University of Massachusetts dissertation.
- Lebeaux, David. 1991. Relative clauses, licensing, and the nature of the derivation. In S.D. Rothstein (ed.), *Perspectives on phrase structure: Heads and licensing*, 209–239. San Diego: Academic Press.
- McCawley, James. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13. 91–106.
- McCloskey, James. 1979. *Transformational syntax and model theoretic semantics: A case study in Modern Irish*. Dordrecht: D. Reidel. DOI: <http://dx.doi.org/10.1007/978-94-009-9495-9>

- McCloskey, James. 2001. The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37. 67–100. DOI: <http://dx.doi.org/10.1017/S0022226701008775>
- Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32. 303–344. DOI: <http://dx.doi.org/10.1162/00243890152001780>
- Nunes, Jairo. 2004. *Linearization of chains and sideward movement* (Linguistic Inquiry Monographs 43). Cambridge, MA: MIT Press.
- Ojeda, Almerindo E. 1987. Discontinuity, multidominance, and unbounded dependency in generalized phrase structure grammar: Some preliminaries in discontinuous constituency. *Syntax and Semantics* 20. 257–282.
- Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Sauerland, Uli. 2004. The interpretation of traces. *Natural Language Semantics* 12. 63–127. DOI: <http://dx.doi.org/10.1023/B:NALS.0000011201.91994.4f>
- Selkirk, Elisabeth. 2001. The syntax-phonology interface. In Bernard Comrie (ed.), *Linguistics section of international encyclopedia of social and behavioral sciences*, 435–483. Amsterdam: Elsevier. DOI: <http://dx.doi.org/10.1016/b0-08-043076-7/02958-2>
- Starke, Michal. 2001. *Move dissolves into Merge*. Geneva: University of Geneva dissertation.
- Stepanov, Arthur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4. 94–125. DOI: <http://dx.doi.org/10.1111/1467-9612.00038>
- Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15. 103–129.
- van Riemsdijk, Henk. 2006. Grafts follow from merge. In Mara Frascarelli (ed.), *Phases of interpretation*, 17–44. Berlin: Mouton de Gruyter. DOI: <http://dx.doi.org/10.1515/9783110197723.2.17>
- Wexler, Kenneth and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.
- Wilder, Chris. 1999. Right node raising and the LCA. In Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest (eds.), *Proceedings of WCCFL*, Vol. 18, 586–598. Somerville, MA: Cascadilla Press.

How to cite this article: Larson, Brooke. 2016. The representation of syntactic action at a distance: Multidominance versus the copy theory. *Glossa: a journal of general linguistics* 1(1): 39.1–18, DOI: <http://dx.doi.org/10.5334/gjgl.15>

Published: 07 October 2016

Copyright: © 2016 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 