Minimal search as a restriction on Merge

1. Introduction

One of the most prominent and difficult puzzles in linguistics is: Why does the faculty of language exist in the form that it does? From an evolutionary perspective the question takes on an additional modification: How could that form have arisen in the relatively narrow time-frame that we surmise that is must have (following thoughts from Tattersall 1998 and Bolhuis, Tattersall, Chomsky, and Berwick 2014). This puzzle has been deemed Darwin's problem (see Hornstein 2009 and Boeckx 2011 among others) and given much current linguistic theory, this puzzle is especially important.

Over the years a thick skein of rules and representations has accrued in linguistic theory. The adoption of these mechanisms was driven by the desire to capture the facts as they exist and ultimately explain how a child could arrive at their adult state absent clear, unambiguous data in their input. However, the baroque genetically encoded knowledge that these theories seemed to require created a well-known tension. The more the acquisition problem is mitigated by genetically encoded guidance, the more difficult it is to conceive of a plausible explanation of how that guidance came about originally.

One logical way to extricate ourselves from this tension is to relieve genetics of some of its explanatory burden. This strategy is most strongly endorsed in Chomsky 2005. In this work, Chomsky outlines a research agenda wherein linguistic attributes heretofore lumped into the genetic endowment are to be analyzed as arising from forces independent of genetics. Such so-called "third factor" forces could take a variety of forms, but for the sake of the current point, I will focus on the notion of “minimal search”, namely, the mechanism by which elements to be Merged are determined.

Chomsky posits the idea that operations over linguistic elements in large part take the shape that they do because factors collude so as to constrain the domain of possible operands. Instead of operating over a vast sea of possible terms, linguistic operations are limited to small pools. For example, Chomsky argues for the notion of phases (Chomsky 2001, 2008) because of their capacity to reduce to the set of possible targets of movement via the Phase Impenetrability Condition. Without phases as interpreted here, linguistic elements at arbitrary structural depths could be potential targets for movement: a very open search indeed and importantly not consistent with the facts.

With this as background, let us now get into the particular point to be made here. In this paper I argue that a simple conception of minimal search serves to restrain the application of Merge. I argue that the search for possible terms of Merge is in this sense optimally economical. The domain of Merge defaults to the smallest search domain possible and only when that domain is exhausted or otherwise inapplicable is Merge across a wider domain allowed. In effect, this will entail that internal Merge is the favored default since the domain of possible Mergees is the most constrained. This idea has most recently been mentioned Chomsky (2013, 2014) and it was first explored in Shima 2000. Shima briefly notes the notion of search as a motivation for a move-over-merge system and then presents some empirical arguments. Only when internal Merge is impossible will Merge between roots across workspaces (external Merge) be allowed. Finally, when neither internal nor external merge is allowed, only then will so-called parallel Merge (a term borrowed from Citko 2005 referring to the Merger of root-internal elements across workspaces) be allowed.

This enforcement of an ever-widening search space will have three effects given current approaches to syntax (as in Chomsky 2013). First, it will make failure to escape a phase impossible. Second, it will actually serve to preclude parallel Merge as a viable option without explicitly barring it, which will have positive empirical repercussions. Third, while still barring parallel Merge, it will explicitly capture the long-assumed, but un-supported notion that Merge alone entails displacement.

In the next section I outline the Merge operation and discuss how, all else being equal, it forces us to predict a very free conception of Merge. In section 3 I discuss how minimal search allows us to make conceptual progress with respect to phase escape. Following this in section 4 I explore the predictions that this restriction yields with respect to sideward movement. Section 5 concludes the paper.

2 Restricting Freedom

In this section I rehearse notions of freedom surrounding Merge and in turn displacement. Merge on its own is a potentially very powerful operation but in the past there have been reasons both empirical and conceptual to limit it. This limitation, in the form of phases, has served to make the search domain for displacement (or internal Merge) smaller. However, this sort of limitation on Merge can be further generalized as will be shown later.

Merge itself can be seen as a means of massaging Darwin's problem. A well known descriptive fact is that there is a preponderance of displacement phenomena in human languages. Prior to the postulation of Merge, the genetic endowment was presumed to somehow encode the possibility for displacement operations separate from structure building operations (viz. X-bar theory and move alpha). However, given the current conception of Merge both structure building and displacement are the result of a single operation. In short, if it is possible to take linguistic objects X and Y and create linguistic object alpha as in (1), it should be able to take linguistic objects Y and alpha and create linguistics object beta as in (2).

(1) Merge(X,B) → $_{\alpha}$[A B]

(2) Merge(B,C) → $_{\beta}$[B $_{\alpha}$[A <B>]]

It is important to note that Merge itself does not *necessarily* entail displacement; it requires the ability to search inside pre-made structure for displacement to be a possibility. That displacement does not come for free with Merge as currently conceptualized has long gone underappreciated and this paper seeks to actually make explicit a plausible means for Merge to entail displacement.

Nevertheless, in (1) we find the simplest case of structure building. In (2) we find the simplest instance of displacement. Element B is now also in a position separate from its original position. As mentioned earlier, displacement is a extremely common phenomenon. That is, it is often the case that a linguistic element shows the effects of being in a position without being there overtly. In (3), we see that the term *carrots* appears sentence initially despite bearing an intuitively very tight relation with the verb *like.*

(3) Carrots, I like.

A Merge-based conception of displacement captures this intuitive verb-object relation in a very simple way: *carrot* is in some sense still there composed with the verb as sketched in (4).

(4) Carrots, I like <carrots>

This is rightly taken to be a positive theoretical advancement, but as presented here, there is still more to be said. Note that the second Merge of B in (2) is with an element alpha that contains B as a subpart. But we cannot limit that second Merge of B only to elements containing B. That is, we cannot limit B to 'upward' displacement: We know independently that that Merge must be able to operate over elements in separate workspaces.

A simple instance of this is seen in the introduction of complex left branches as in (5) where the derived structure *the boy* is merged with the structure *hit the ball*.

(5) Merge([the boy],[hit[the ball]]) → [[the boy] [hit [the ball]]]

Since Merge allows us to manipulate sub-objects like B in (2) and since Merge allows us to operate across workspaces like in (5), it would require independent stipulation to rule out Merge of a sub-object across workspaces like in (6).

(6) Given $_{\alpha}$[A B] and Z, Merge(B,Z) → $_{\alpha}$[A <B>] $_{\beta}$[B Z]

The argument above was made most convincingly by Citko (2005) and in effect makes it such that "sideward" instances of displacement should be predicted in the same way that upward ones are (again I follow Citko in dubbing this "parallel

Merge", but refer to its effects as "sideward displacement"). That is, this simplest version of Merge predicts both upward displacement like in (7) and sideward displacement like in (8).

(7) [B…[A <B>]]

(8) [[B Z] …[A <B>]]

This conception of Merge has been widely exploited to analyze where sideward displacement seems possible. Going under a variety of names (multidominance (McCawley 1982, Ojeda 1987, Blevins 1990, Wilder 1999, Chen–Main 2006, Johnson 2007, Bachrach & Katzir 2009), inter-arboreal movement (Bobaljik 1995 and Bobaljik and Brown 1997), parallel Merge (Citko 2005) sideward movement (Nunes 2001, Agbayani and Zoerner 2004, and Fernández-Salgueiro 2008), grafting (van Riemsdijk 2006), sharing (Guimarães 2004, Chung 2004, de Vries 2005, Gracanin–Yuksek 2007), etc.), the notion that a single syntactic element can directly compose in two positions that are not in a c-command relation with one another has been often explored despite being variously construed as movement or as a static structural relation. And it is for good reason that this exploitation has occurred: all else being equal we should expect structures like (8). Again, it would take an independent force to defuse (8) as a predicted result of Merge.

However, in this paper I argue that such syntactic relations are effectively ruled out. Moreover, ruling out such relations results in an advantageous empirical perspective and not much is lost despite the notion itself being clearly fecund as evidenced by the above multitude of publications. This will have larger consequences that will not be explored here. If it is the case that parallel Merge is ruled out, the constructions targeted by the analyses listed above still demand explanations. Further, if Merge writ large is ruled out as part of the source of those constructions, the explanations must effectively lie outside of syntax proper and in turn outside the scope of this article. As much has already been noted by Larson (2013a) in the course exploring the difficulties facing those constructions analyzed using parallel Merge.

Yet limiting Merge so as to rule out parallel Merge finds analogues elsewhere. An instructive lesson can be found in limitations on internal Merge. Much like with sideward displacement, the simplest form of Merge does not on its own rule out displacement of an element from an arbitrarily deep position in a structure. That is, there is nothing in the definition of Merge that would serve to preclude it from operating over the root node A and the arbitrarily deeply embedded element Z in the structure in (9).

(9) [A [B [… Z]]]

However, all else is not equal in this case and we have evidence of long-distance movement taking place over short, punctuated steps (McCloskey 1979, Torrego 1984, Henry 1995, and McCloskey 2001). It seems that it is to the contrary *not* licit to Merge A and Z over long distances. but rather Z must first have moved into a position sufficiently close to A. This may involve multiple steps as sketched in (10).

(10) [A [**Z** [B [**<Z>** [C [… **<Z>**]]]]]]

This sort of cyclicity is the empirical basis for attempts to re-analyze successive cyclic movement as being a forced necessity somehow. Previously, cyclicity was enforced extrinsically via subjacency (Chomsky 1973, 1986). Movement could in principle be as far as possible: There were no restrictions on how deeply the system could 'look' into a structure. However, if movement took place over too great a distance, it would result in structures that violate various grammatical principles.

Currently, cyclicity is the empirical basis for the 3rd factor minimal search constraint embodied by phases (Chomsky 2001, 2008). Phases are a means to delimit the domain of Merge operations via the Phase Impenetrability Condition. In this way, it is no longer possible to move over too great a distance because elements that are too far away are not visible. Merge between A and Z in (9) simply is not stateable.

This is a potentially subtle distinction. In both phase-based and subjacency-based analyses, the relevant limiting nodes are chosen to fit the facts and in this way it might seem that cyclicity is extrinsically encoded. However, phase-based approaches capture cyclicity by analyzing its violation as *impossible* while subjacency-based approaches analyze cyclicity violation as *possible but ungrammatical*. With third factor considerations in mind, it is clearly preferable to attempt to couch things in terms of linguistic impossibility instead of grammatical violations. That is, phases may indeed stipulate the relevant bounding nodes and this was also true of subjacency-based theories. The advantage to a phase-based analysis coupled with the PIC Grammatical is that they render certain domains invisible to Merge (namely domains that have already been spelled-out). This differs from previous analyses that were radically free to operate over any constituent (move alpha) in a way that might lead to ungrammaticality. In such cases, this freedom needed to be ruled out as grammatically stateable, yet deviant. The differences being that violations require a means to assess the violations, things that are impossible do not. To the extent to which we can ascribe impossibility to the ungrammatical, we have made conceptual progress. In the rest of this section, I posit that we can do this with Merge more generally.

In the next section I make an argument that the application of Merge in general should be constrained based on search domain considerations in a similar way to phases. That is, we should extend the lesson of phases to Merge in general.

3. Third Factor restrictions on Merge

3.1 Initial considerations

Phases, as described above, employ the notion of minimal search to constrain the freedom of Merge to look deep into a given structure. By limiting the domain of Merge, we can not only capture the facts, but do so in a manner more evolutionarily plausible. But the innards of a structure is not the only domain of Merge. As we noted above, all else being equal we should be able to Merge elements across structures: either in a simple external Merge fashion or in the sideward movement fashion. These applications of Merge require different search possibilities and as such can be assessed on that basis. The notion that Merge can be restricted based on minimal search can apply here as well. Below I argue that doing this can lead to some explanatory progress.

First to set the stage. In this section I will discuss different ways of hypothetically limiting Merge and after that I will show that progress can be made if we impose a general restriction on Merge outlined in (11). Call this the General Restriction on Merge (GRM).

(11) General Restriction on Merge: Merge can only apply to an object in a given search space if there is no possible Merge with an object in a more constrained search space.

The most constrained domain of Merge would be limited to just root-internal elements of a given root. The structure internal to the root is inherently more limited that allowing Merge between the root and elements in other workspaces which can be arbitrarily large. Root-internal elements are severely bounded by the PIC whereas the lexicon or other workspaces are not bounded whatsoever.[1] Call this most constrained purview of Merge the root-internal version:

(12) root-internal: The domain of Merge consists of only root-internal elements of a given root.

This is the most constrained search domain for Merge and given the GRM, it must be the default option. Nevertheless, it clearly presupposes already built structure to have any sort of theoretical force. A derivation driven by Merge could never possibly begin were it not for possible recourse to a wider domain of possible Merge operands. That is, it would prevent any first Merge to begin the derivation. External Merge must be a second option when root-internal Merge is not possible.

[1] Though numerations have indeed been theorized to consist of bounded sub-numerations (Chomsky 2000 and others), this notion seems to have been dropped in Chomsky 2013 where where external Merge involves already generated items and the lexicon, not any notion of numeration (Chomsky 2013:41).

The next less limited search space is the set of roots nodes in other workspaces. This domain is not limited by the PIC nor is it collected in a single structure. It is however limited in that only the root labels in these other workspaces are in its purview. The label of a root would serve to bar search inside of a structure, so it is larger than that of root-internal Merge but more limited than were it to include the innards of these other root nodes. Call this the root only version of Merge.

(13) root only: The domain of Merge consists of only root labels in other workspaces

This version of Merge is transparently insufficient on its own as it precludes any sort displacement. It can build structure via external Merge only.[2] But the GRM allows it as a second option when root-internal Merge is not possible. This allows for structure building to begin with and also for displacement: external Merge and internal Merge.[3]

In the above two versions of Merge, the operation is either barred from searching in other workspaces (root-internal) or barred from looking within a

---

[2] A disjunctive approach where Merge is root-internal *or* root only is descriptively sufficient, but as I argue below, we can effectively enforce the results of this disjunctive approach without explicitly encoding it.

[3] The difference between these two types of Merge in terms of search has already been noted by Chomsky: "If anything, [internal Merge] is simpler, since it requires vastly less search than [external Merge] (which must access the workspace of already generated objects and the lexicon)" Chomsky 2013:41. This observation is interesting because he immediately disavows the notion that one type of Merge is simpler or more preferable, saying that they are both freely available.

However, the notion that internal Merge requires less search can be seen as a fact and the mitigating notion that this does not matter is a theory-internal/aesthetic notion. Perhaps for his particular theory they are both free available, but the minimalist program allows for competing theories including differing notions of relevant 'third factors' (see Epstein, Kitahara and Seely (2012) for an exploration of a system in which root only Merge is the preferred, simpler version) . In fact, in this paper I accept that totally free Merge is the null hypothesis and argue that a more constrained view is advantageous.

structure (root only). The next least constrained search space for Merge would allow a root to not only search within itself and search for other roots, it would also allow search into the innards of other roots. This results in a very unconstrained version of Merge, call it Free Merge. Free Merge makes possible internal Merge, external Merge, and parallel Merge.

(14) Free Merge: The domain of Merge consists of only root-internal elements, other roots, and root-internal elements of other roots.

For the same reason as sketched above, Chomsky (2013:41) notes that internal Merge is a simpler operation than external merge. It requires less search. The General Restriction on Merge in (11) operationalizes this fact and requires external Merge (forced by a root only search space) to occur only when internal Merge (forced by a root-internal search space) is not possible. Further, this same restriction would require sideward movement (allowed by Free Merge) to occur only when external Merge is not possible.

3.2 Exploring the restriction

If we adopt the GRM (search narrowly, then wider, then widest), some nice things follow given current theory. Bluntly stated, search-based Merge restrictions effect a sort of 'move-over-merge' notion in the syntax (similar to that which is explored in Castillo, Drury and Grohmann 1999 and Shima 2000). Though counterintuitive in a field that once assumed a 'merge-over-move' syntax, given current theory this reverse conception is advantageous.

Currently (Richards 2007 and Chomsky 2013), it is only at the phase level that features that drive movement are introduced in a derivation. These phase heads carry with them the uninterpretable features that are rendered interpretable via displacement into their vicinity. As such, internal Merge will not be possible until the phase level and external Merge will be the next best option until that point. Crucially, once the phase level is reached and the relevant movement-inducing features are introduced, internal Merge is not only possible, but also the only option. Internal Merge *must* occur at this point because it involves the first choice smallest search domain. Note that for the argument presented here to go through, it is not necessary for *all* movement driving features to be introduced at the phase-level. Rather, it needs to be the case that features are introduced *at least* at the phase-level. It could be the case that internal Merge is forced at non-phase positions as well.

What this results in is obligatory phase-escape by the relevant elements. Recall earlier that we made conceptual progress with the PIC. The PIC makes it not merely ungrammatical to move from the innards of a phase, but impossible. The same sort of conceptual progress can be made here: The GRM makes it not merely ungrammatical to fail to escape a phase when it is an option, but impossible. At the point where movement to the edge of a phase is possible, it is not even an option for external Merge to introduce other elements to the root and thereby trap the phase-internal elements that could have otherwise moved. In short, we can reduce phase-escape to a notion of minimal search.

Let us see this in action. Given a pair of workspaces like in (15a), it is clear that root-internal, internal Merge is not a possibility. As such, a wider, root only search space is permitted. Merge is allowed between X and Y and the result is an element α like in (15b). Crucially, I assume that there is a certain motivation for the elements to Merge. For the sake of concreteness, let us say that X is motivated to Merge with Y via a '-Y' feature roughly following Pesetsky and Torrego 2007.[4]

(15) a. $X_{\text{-Y}}$ Y

b. Merge($X_{\text{-Y}}$,Y) → $_{\alpha}$[X Y]

Now, given the resultant structure in (15b), we now have the logical possibility of internal Merge (we could Merge Y and alpha), and given the GRM it may be that we are forced to move in this case. This would be a very negative repercussion as it would preclude the external Merger of any other element and the derivation would be doomed to consist of the iterated internal Merge of X and Y.

Luckily, according to contemporary theory, there is no motivation to do so. Following Abels 2003, the initial motivation for X and Y to makes movement of Y to such a structurally local position redundant and ruled out. As such, the search domain of Merge is opened up and external Merge is now viable. That is, Merger of Y or X and alpha should be the first thing considered, however movement will be precluded by a lack of motivation.

Were there more structure between Y and root and motivation for Y to move, the movement would not only possible, it would also be required before any other head were Merged.[5] Consider a phase-head Z with both a motivation to Merge with alpha (-α) and motivation to Merge with Y (-Y) as in (16a). There are two logically

[4] Throughout this section I present asymmetric motivations for Merge (one element as a featural need to compose with another). This is merely for expository ease and I take Merge to result in an entirely symmetric relation between the two terms once Merged. For example, specifically demands that Y be the mover in what follows, X could just as well have been the mover.

[5] This would jibe with notions of "agnostic movement" which posit that moving elements do so at every single available chance in a derivation. See Bošković (2002), Franks & Lavine (2006), Boeckx (2008), and Larson (2008) for discussion of this idea.

possible Merges: external Merge of Z and alpha as in (16b) or parallel Merge of Z and Y as in (16c).

(16) a. $Z_{-\alpha/-Y} \quad {}_{\alpha}[X\ Y]$

b. Merge($Z_{-\alpha/-Y}$, α) $\rightarrow {}_{\beta}[Z_{-Y}\ {}_{\alpha}[X\ Y]]$

c. Merge(($Z_{-\alpha/-Y}$, Y) $\rightarrow {}_{\beta}[Z_{-\alpha}\ Y] \quad {}_{\alpha}[X <Y>]$

The operation in (16c) is ruled out due to the GRM. The Merge of (16b) operates over roots across workspaces whereas the Merge of (16c) operates over a root and a root-internal element across workspaces. As such, it is the only possible next step from the state of affairs in (16a).

The resulting structure in (16b) not only has a built-up domain for for root-internal search, there is also motivation for an element within it to move since the phase-head desires a Y. Since this most constrained domain of search is viable, the GRM says that it *must* be used. This forces the Merge of Y and beta as in (17).

(17) Merge(Y,β) $\rightarrow$ $[Y\ {}_{\beta}[Z\ {}_{\alpha}[X <Y>]]]$

However, imagine the stage of the derivation prior to the internal Merge above. There is eventually going to be another phase-head H that could externally Merge with beta instead of Y. This would result in a structure like in (18) and trap Y within the phase headed by Z.

(18) Merge($H_{-\beta}$,β) $\rightarrow$ $[H\ {}_{\beta}[Z\ {}_{\alpha}[X\ Y]]]$

The step in (18) is not possible under the GRM. What's more, the comparison between (17) and (18) is not possible.[6] The operation in (17) is the only possible step given (16). It is in this sense that the GRM allows the failure to escape a phase to be rendered impossible without explicitly, extrinsically encoded prohibition.[7]

Additionally, parallel Merge of the sort allowed by Free Merge will only be allowed when external Merge is not a possibility. This restriction effectively limits us to internal Merge and external Merge because the possible points in a derivation where neither internal Merge nor external Merge is possible are severely limited. They are limited to such an extent that it seems plausible to say that they do not exist in practice. Let us see what situation would have to hold for parallel Merge to apply. There would need to be two workspaces A and B as shown below in (19). In workspace A, it must be the case that all movement to the phase edge is completed (internal Merge possibilities exhausted) and that there will be no more Merge to the root (external Merge possibilities exhausted). In workspace B, it must be the case that no more internal Merge can take place and that the only thing that can undergo external Merge with the root is the non-root phase-edge in workplace A.

(19) A: $_{\alpha}$[XP H] B: $_{\beta}$[H ZP]

That is, it needs to be the case that XP must Merge with beta. In the next section I show that there is no reason to suspect that such a state of affairs arises

---

6 A similar notion can be found in Shima 2000 concerning strict cyclicity and how it is forced by assuming move-over-merge.

7 A review notes that by making phase escape impossible, problems of phase-based accounts of island phenomena are starker. That is, the GRM makes it impossible to say that so called weak islands derive their deviancy from phase escape. It is not clear whether this is a positive repercussion or not. A syntactic approach to weak island phenomena is lost, but this might be advantageous (see Abrusán's (2014) semantic account of weak islands).

and that the instances where parallel Merge is fruitfully employed are empirically suspect. That is, there is neither theory-internal nor clear empirical reason for parallel Merge to exist. Given GRM, this is a predictable outcome.

This conception of an ever-expanding domain of Merge in the face of inapplicability also raises the spectre of what Grohmann et al. (2010) deem the “apex paradox”. If one assumes that the complement of a phase-head is only spelled-out upon the Merger of the next highest phase-head, how then would it be possible to spell-out the the root CP and its complement? It surely cannot be the case that that material simply is not spelled-out. Nor is it very explanatory to propose that a series special-purpose phase-heads are Merged so as to ensure the spell-out of the CP. Instead, a simply idea would be to simply spell-out the CP upon its completion. This is where the paradox arises as this would preclude the external Merger of a head that takes a CP as a complement (like *say*).

Although it doesn’t solve this paradox, the GRM suggests an ineluctable forced end-point of the derivation. When there is no longer the possibility for internal, external, or parallel Merge, the sole syntactic structure-building operation is rendered useless which could plausibly leave nothing but spell-out as an option. This is still a novel motivation for spell-out, but it avoids the paradox in that it first demands that the search for an CP-embedding verb.

Finally, it should be noted that this GRM conception of Merge actually serves to explain displacement in general in a way that Merge alone does not. Merge, as discussed above, allows for two elements to be brought together. This does not necessarily give it the power to effect displacement which of course requires the ability to search within already constructed elements. The ability to search within structures is simply assumed, though left unconstrained it is a very powerful ability.

Merge, coupled with the GRM, allows us to finally not only truly predict displacement, it allows us to also predict the limited range of displacement, as will be discussed below.

4 Parallel Merge and c-command

In the previous section we saw that under current theory, there is essentially no reason to expect parallel Merge under GRM: the conditions for its application never arise. In this section I show that this is an advantageous result because it allows for two things: a deeper explanation of c-command requirements in movement and an explanation concerning the empirically shaky ground that much parallel Merge-reliant analyses face.

4.1 C-command as the result of Merge

Before exploring the past exploitation of parallel Merge, it is worth noting a very conspicuous absence. If it is indeed generally possible that Merge can hold between a root and the innards of another phrase, we should expect the effects of parallel Merge to be as ubiquitous as traditional upward movement. Just as it takes only a cursory empirical assay to find displacement of the sort in (20), so too should it be trivial to find evidence of parallel Merge.

(20) Carrots$_i$, I like t_i.

However, such forays are met with silence. There is a striking absence of unambiguous cases of sideward displacement like that in (21) in which it is clearly the case that the moving element did not move to a c-commanding position.

(21) a. *John saw t_i and what$_i$ did Jane buy t_i?

b. *Who saw $what_i$ and Jane bought t_i?

c. *Who ate $what_i$ before buying t_i?

d. *$John_i$'s mother loves t_i.

In (21a), if *what* can move from its base-generated position (the complement of *saw*) to a root-internal position across workspaces, then it should be able to move to become the complement of *buy* before moving up to spec,CP in the second conjunct. Or similarly with (21b) and (21c), the wh-word should be able to move to a second complement position and remain there in a multiple wh-question in English. Finally, the verbal complement in (21d) should have the capacity to move to the specifier position of a possessive DP before that DP is Merged as the subject of the sentence.

This utter absence has been already noted, albeit indirectly, in Nunes 2001, 2004. Nunes posits as a condition on the licensing of parallel Merge that the moving element must eventually c-command each lower copy of itself. That is, the examples in (21) are ruled out not because of illicit movement, but rather because the resulting representation does not involve a copy that c-commands all its derivationally previous locations.

So in order to capture the general lack of parallel Merge effects, an output condition that employs the structural relation of c-command is required. However, if parallel Merge were independently ruled out as under the GRM, no such condition would be required and the ungrammaticality of the sentences in (21) would be predicted.

Not only would the deviance of the sentences in (21) be predicted, the GRM serves to explain in a more fundamental way the generalization that movement must result in c-command. The GRM effectively (coupled with the extension condition) demands all Merge take place to the root either with an internal node or another root. C-command as a result of Merge falls out from this (just as in Epstein 1999) without being encoded. This contrasts with the Nunes approach which allows for parallel Merge but demands the reification of c-command as an explicitly encoded output condition.

This reification of c-command is not necessarily undesirable. It does however add explanatory burden in a way that the GRM account does not. With GRM and the extension condition, c-command is the unavoidable result of Merge and its ubiquity hence explained. Were parallel Merge allowed, the question remains to be answered as to why a concept like c-command should matter in the first place. GRM is thus a positive theory-internal step forward. In the next subsection I argue that it predicts the dubious empirical nature of most uses of parallel Merge.

In essence, I argue that a trade of sorts should be made. The GRM account of Merge allows for the simpler account of why parallel Merge effects are generally not seen. In return however, the instances and constructions where parallel Merge ostensibly captures the data demand an account. I show in the next section that this is an advantageous trade as the positive evidence for parallel Merge is light on the empirical ground.

4.2 Unexpected asymmetries in parallel Merge

Recourse to Merge between a root and a root-internal object in a separate workspace has been made by a number of researchers for a number of constructions. When one tests for more clearly syntactic phenomena, asymmetries abound, as I will outline here.

At heart, each of these requires something like (22) to hold:

(22) Given $_{C}$[A B] and Z(P), Merge(B,Z(P)) → $_{C}$[A <B>] $_{Y}$[B Z(P)]

For example, Right node raising examples like that in (23) have been argued to be derived via Merger between *some books* qua object of the verb *bought* with the external root *sold* (Wilder 1999, Abels 2004, Bachrach and Katzir 2009 and others). This has generally been couched in Multidominance terms, but a copy-theoretic implementation makes the same empirical predictions and for the sake of typographical ease, I'll work with copies here.

(23) John bought, and Mary sold, some books.

(24) Given [bought some books] and [sold] Merge(some books,sold) → [bought some books] [sold some books]

The representation in (24) captures the shared thematic relations between the object and the two verbs. However, as has already been shown to be true at least since Kayne 1994, this sort of symmetry does not hold generally for effects more clearly syntactic than thematic interpretation. One example from Larson 2013a shows a clear asymmetry with respect to subject-verb agreement. As seen in (25) agreement only holds between the second conjunct subject and the shared material, never the first:

(25) a. Bill is happy that Iris, and James is happy that his parents, {like/*likes} reading fiction.

b. Bill is happy that his parents, and James is happy that Iris, {*like/likes} reading fiction.

This is a false prediction made by the parallel Merge account. Given the symmetry in the derived structure, the different options in (25) should be either equally grammatical or equally ungrammatical. Yet they are not. Numerous other examples redundantly making the same point can be found in Larson 2013a. Let (25) suffice as an exemplar of this sort of asymmetry and empirical evidence against parallel Merge.[8]

Another implementation of what is essentially parallel Merge is found in analyses of coordination-wh questions like that in (26). For the sentence below Gracanin-Yuksek (2007) provides the structural analysis in (27) (from Citko and Gracanin-Yuksek 2013:5). Without going into the details (for example, a linearization algorithm enforces the word order seen in (24)), we can see that there are numerous instances of structure-internal elements merged to formerly root-level nodes across a workspace (prior to joining at the final root).

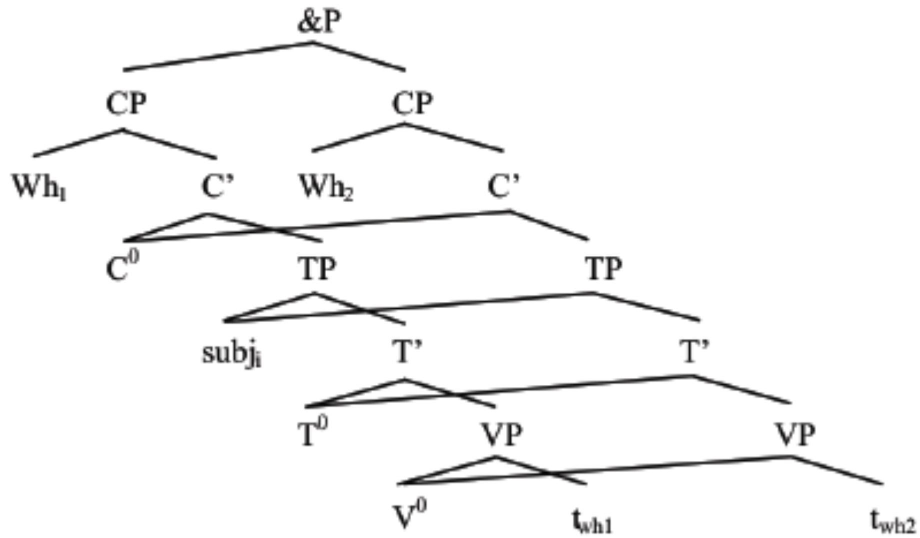
(26) What and when did John eat?

[8] Similar use of parallel Merge has been employed for ATB constructions like that in (i) that presume an underlying RNR structure (Fernández-Salgueiro 2008)

(i) What did John buy and Mary sell?

However a number of asymmetries can be found in these examples as well that militate against theories of ATB derived by parallel Merge (Larson 2013b and Parker and Larson 2013)

(27)



The structure, as Gracanin-Yuksek presents it makes an interesting prediction. Only verbs, like *eat*, which are optionally transitive are licit in such constructions in English. Why? The shared V node must be grammatical in the conjunct where there is no overt direct object. Verbs like *fix* should not be acceptable because they need an overt direct object in both conjuncts. This prediction is initially borne out as seen in (28).

(28) *What and when did John fix?

However, there is an asymmetry in acceptability to be found here as well. When the order of the wh-words is flipped around, the resulting sentence is acceptable, contrary to the prediction of the parallel Merge derived analysis. This can be seen in (29) and the judgements have been supported and replicated in a series of judgment studies (Lewis, Larson, and Kush 2012).

(29) When and what did John fix?

Again, this asymmetry is not predicted in the parallel Merge theory and as such we find herein empirical evidence against the sort of movement we ruled out theoretically in the previous section.

The hallmark example of parallel Merge is found in Nunes' 2001,2004 analysis of parasitic gap constructions like in (30).

(30) What did John eat shortly after buying?

Here, the wh-word moves from an adjunct internal workspace to the matrix clause workspace before those two sub-trees are merged. This is the paradigm case for which parallel Merge was designed (though see Chomsky 1986 and NIssenbaum 2000 for alternative analyses). As such, any analysis disavowing parallel Merge stands to lose empirical ground here. In virtue of the other advantages that such disavowal promises, this empirical loss is a fair price.

Furthermore, the positive empirical predictions of the parallel Merge analysis of parasitic gaps are not without their anomalies. For example, it is well known that the parasitic gap does not display the reconstructions effect that are found in analogous instances of movement (see Kearney 1983 and Munn 1993):

(31) a. Which books about himself did John file t before Mary read e?

b. *Which books about herself did John file t before Mary read e?

Additionally, Assmann (2012) has noted that there is an asymmetry in the strength of island effects between the matrix clause and the adjunct clause. That is, the potential wh-island formed by the embedded ‘how’ in the examples below shows differential effects depending on whether it arises in the adjunct clause or not. The adjunct clause displays less of an effect further suggesting that no movement has arisen from that position.

(32) a. *Welche Radios weisst du [$_{CP}$ wie man [ohne e zu reparieren] t verkauft]]?
which radios know you how one without to repair sells
‘Which radios do you know how to sell without repairing?’

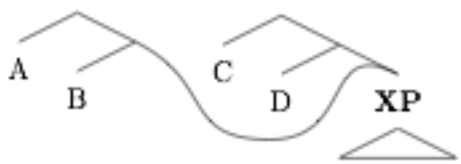
b. ?Welche Radios hast du [ohne zu wissen [$_{CP}$ wie man e repariert] t
which radios have you without to know how one repairs
verkauft]]?
sold
‘Which radios did you sell without knowing how to repair?’

One final construction that parallel Merge has been used for are so-called sentence amalgams (the term coined by Lakoff (1974). An example of such a construction can be seen below in (33). Here, there is a single element (underlined here) that plays a different role in each clause.

(33) Brooke bought I don’t know <u>how many books</u> at the store.

To capture the dual duty of the underlined material, a number of researchers have proposed parallel Merge-based accounts of the construction (Guimarães 2004, van Riemsdijk 2006, and Johnson 2012, 2013). In short, they take the underlined constituent above to be Merged both as the direct complement of *bought* and in the spec,CP position of the interrupting clause. This is roughly sketched in (34) (modified from Kluck 2014:28) where the XP is the underlined material from (33). According to the various theories promoted, the two roots in (34) are linearized in various ways so as to effect the word order in (33).

(34)



Again, we expect the XP to behave in a symmetric fashion with respect to both clauses. However, Kluck (2014) provides argumentation to the contrary. For instance, It is possible for there to be A'-movement of the shared material within the second, interrupting clause as seen in the example in (35).

(35) Bob kissed [how many girls]$_i$ you can't even begin to imagine t$_i$ at the party.

She notes that it is drastically unacceptable for the same underlined element to undergo A'-movement in the first, or matrix, clause:

(36) *[How many girls]$_i$ did Bob kiss you can't even begin to imagine at the part?

Kluck shows that the same facts hold for Dutch and proceeds to argue against the parallel Merge account of the construction in general, positing instead a sluicing account.

It should be noted that the successes of parallel Merge largely concern intuitions of compositional semantics. That is, by Merging an element across two internal workspaces is used mostly to capture the fact that the element is interpreted in each, despite showing up overtly in only one. While this is understandable motivation to group the two together syntactically, it is still at heart a semantic issue. The desire to ensconce a semantic notion like thematic composition in the syntax (at all cost) is perhaps a forgivable vestige of older, generative semantics-inspired theories. However, the more unambiguously syntactic facts tell a different story, one that suggests a need for a less syntactic-centric approach to the above constructions.

In short, the parallel Merge approach to parasitic gaps is not without its unexpected empirical shortcomings in the same way the right node raising and coordinated wh-questions constructions are. Complete empirical accuracy is of course not a reasonable criterion for any theory, but these qualms make the trade-off that was mentioned earlier, easier to bear. In exchange for the explanatory advantages of the GRM theory, all that needs to be ceded are a few (compared with the vast number of sentence types like those in (21)) unconventional constructions that themselves do not strongly support parallel Merge.

5 Conclusion

By adopting a minimal search stricture that requires the smallest search space possible be considered before all others, it is possible to derive a number of positive theoretical effects. A root node searching only within that root node (coupled with the PIC) allows for the most restrained search space in principle. If Merge must search within this space before anywhere else (effecting external merge), we are able to rule out unwanted phase-escape in a more explanatory fashion. Instead of phase-escape being possible yet ungrammatical, we render it wholly impossible. Using a wider search space and introducing a phase-inducing head via external merge before attempting internal merge can never occur. Further, if parallel Merge is only possible when external, root-to-root merge is not possible, there will effectively be no circumstance when parallel Merge is entertained. This is further advantageous since in the current theory, parallel Merge is not ruled out despite being empirically dubious. Finally, this approach allows us to enforce the notion the displacement must result in a c-command configuration without requiring any explicit extrinsic encoding of that requirement.

6 References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. Diss. University of Connecticut.

Abels, Klaus. 2004. “Right node raising: Ellipsis or across the board movement?”, in Moulton & Wolf (eds.), Proceedings of NELS 34, GLSA, University of Massachusetts, Amherst, pp. 45–59.

Abrusán, Márta. *Weak Island Semantics*. Vol. 3. Oxford University Press, 2014.

Agbayani, Brian, and Ed Zoerner. 2004. "Gapping, pseudogapping and sideward movement." *Studia Linguistica* 58.3: 185-211.

Assmann, Anke. 2012. "Deriving parasitic gaps by fission und fusion." *Proceedings of ConSOLE XIX* 49: 75.

Bachrach, Asaf, and Roni Katzir. 2009. "Right node raising and delayed spellout. "*InterPhases: Phase-theoretic investigations of linguistic interfaces*: 283-316.

Blevins, James. 1990. Syntactic complexity: Evidence for discontinuity and multi-domination. Amherst, MA: University of Massachusetts dissertation.

Bobaljik, Jonathan David. 1995. In terms of merge: Copy and head movement. MIT Working Papers in Linguistics 27, 41–64.

Bobaljik, Jonathan David & Samuel Brown. 1997. Interarboreal operations: Head movement and the extension requirement. Linguistic Inquiry 28, 345–356.

Boeckx, C. 2008. Understanding minimalist syntax: Lessons from locality in long-distance dependencies. Oxford: Blackwell.

Boeckx, Cedric. 2011. "Some reflections on Darwin's Problem in the context of Cartesian biolinguistics." *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*: 42-64.

Bošković, Ž. 2002. "A-movement and the EPP". *Syntax* 5:167–218.

Castillo, Juan Carlos, John Drury, and Kleanthes K. Grohmann. 1999. "Merge over move and the extended projection principle." *University of Maryland working papers in linguistics* 8.63: 103.

Chen–Main, Joan. 2006. On the generation and linearization of multi-dominance structures. Baltimore, MD: John Hopkins University dissertation.

Chomsky, N., 1973. Conditions on transformations. In: Anderson, S.R., Kiparsky, P. (Eds.), A Festschrift for Morris Halle, Holt, Rinehart, and Winston, New York, pp. 232–286.

Chomsky, N., 1986. Barriers, MIT Press, Cambridge, MA.

Chomsky, Noam. 1995. The Minimalist Program. MIT Press, Cambridge, Mass.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Step by step: Essays on minimalist syntax in honor of Howard Lasnik, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. "Derivation by phase. Ken Hale: A life in language, ed. by Michael Kenstowicz, 1-52.

Chomsky, Noam. 2005. "Three factors in language design." *Linguistic inquiry* 36.1: 1-22.

Chomsky, Noam. 2008. On phases. in Foundational issues in linguistic theory:

Essays in honor of Jean-Roger Vergnaud, eds. Robert Freidin, Carlos Otero, and Maria Luisa Zubizaretta, 133–166. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2013. "Problems of projection." *Lingua* 130: 33-49.

Chomsky, Noam. 2014. Class Lecture. MIT. April 2, 2014.

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* 36.4: 475-496.

Citko, Barbara, and Martina Gračanin-Yuksek. 2013. "Towards a new typology of coordinated wh-questions." *Journal of Linguistics* 49.01: 1-32.

Epstein, Samual. 1999. 'Unprincipled Syntax and The Derivation of Syntactic Relations.' in Working Minimalism. Cambridge: MIT Press.

Epstein, Samuel David, Hisatsugu Kitahara, and T. Daniel Seely. "Structure building that can't be." *Ways of Structure Building* (2012): 253-270.

Fernández-Salgueiro, Gerardo. 2008. "Deriving the CSC and unifying ATB and PG constructions through sideward movement." *Proceedings of the 26th WCCFL*.

Franks, S. and Lavine, J. 2006. "Case and word order in Lithuanian". *Journal of Linguistics* 42: 239-288.

Goodall, Grant. 1987. Parallel Structures in Syntax: Coordination, Causatives and Restructuring (Cambridge Studies in Linguistics 46). Cambridge: Cambridge University Press.

Gracanin-Yuksek, Martina. 2007. "About sharing." *PhD diss., MIT*.

Grohmann, Kleanthes, Markus Pöchtrager, Tobias Scheer, Michael Schiffmann, and Neven Wenger. 2010. "The Apex Paradox: A Technical Issue for the

Explanation of Main Clause Phenomena" . GIST 2: Main Clause Phenomena. Het Pand, Ghent. (September 29 – October 1, 2010).

Guimarães, Maximiliano. 2004. Derivation and representation of syntactic amalgams. Doctoral Dissertation, University of Maryland.

Henry, Alison. 1995. Belfast English and Standard English: Dialect variation and parameter setting. New York: Oxford University Press.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge University Press.

Johnson, Kyle. 2007. LCA + alignment = RNR. Paper presented at the workshop on Coordination, Subordination and Ellipsis, Tübingen. [Eberhard-Karls-Universität Tübingen, 7–8 June 2007.]

Johnson, Kyle. 2012. Recoverability of deletion. Manuscript, University of Massachusetts at Amherst.

Johnson, Kyle. 2013. Licensing ellipsis. Manuscript, University of Massachusetts at Amherst.

Kayne, Richard. 1994. *The antisymmetry of syntax*. No. 25. MIT Press.

Kearney, Kevin. 1983. Governing Categories. Unpublished ms. University of Connecticut, Storrs, CT.

Kluck, Marlies. 2014. A sluicing account of amalgams. Manuscript, University of Groningen.

Lakoff, George. 1974. Syntactic amalgams. In Papers from the 10th regional meeting of the Chicago Linguistic Society, ed. Michael Galy, Robert Fox, and Anthony Bruck, 321–344. Chicago: University of Chicago.

Larson, Bradley. 2008. Agnostic Movement in Malagasy Focused Predicates. In *Proceedings of the 24th NWLC, 3-4 May 2008, Seattle, WA.*

Larson, Bradley. 2013a. The Syntax of Non-syntactic Dependencies. Doctoral disseration. University of Maryland, College Park, Maryland

Larson, Bradley. 2013b. "Not-so-across-the-board movement in Macedonian. "*Proceedings of Formal Approaches to Slavic Linguistics (FASL)* 21.

Lewis, Shevaun, Bradley Larson, and Dave Kush. 2012. "What and when can you fill a gap with something." *CUNY sentence processing conference*.

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.

McCloskey, James. 2001. "The morphosyntax of WH-extraction in Irish." *Journal of Linguistics*: 67-100.

Munn, Alan Boag. 1993. *Topics in the syntax and semantics of coordinate structures*. Diss. The University of Maryland.

Nissenbaum, Jonathan W. 2000. *Investigations of covert phrase movement*. Diss. Massachusetts Institute of Technology.

Nunes, Jairo. 2001. "Sideward movement." *Linguistic inquiry* 32.2: 303-344.

Nunes, Jairo. 2004. Linearization of Chains and Sideward Movement. volume 43 of Linguistic Inquiry Monographs. MIT Press, Cambridge, Massachusetts.

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.

Parker, Daniel and Bradley Larson. 2013. Two flavors of long distance dependency discerned through island effects. Talk presented at the CUNY conference on human sentence processing. University of South Carolina.

Pesetsky, David, and Esther Torrego. 2007. "The syntax of valuation and the interpretability of features." *Phrasal and clausal architecture: Syntactic derivation and interpretation* 262-294.

Richards, Marc D.2007. "On feature inheritance: An argument from the Phase Impenetrability Condition." *Linguistic Inquiry* 38.3: 563-572.

van Riemsdijk, Henk. 2006. Grafts follow from merge. In Mara Frascarelli (ed.), Phases of Interpretation (Studies in Generative Grammar 91), 17–44. Berlin: Mouton de Gruyter.

Shima, Etsuro. 2000. "A preference for Move over Merge." *Linguistic Inquiry* 31.2: 375-385.

Tattersall, Ian. 1998. The origin of the human capacity. New York: American Museum of Natural History.

Torrego, Esther. 1984. "On inversion in Spanish and some of its effects." *Linguistic Inquiry*: 103-129.

de Vries, Mark. 2005c. Merge: Properties and boundary conditions. Linguistics in the Netherlands 22, 219–230.

Wilder, Chris. 1999. "Right node raising and the LCA." *Proceedings of WCCFL*. Vol. 18.