Just Pair-merge

Ken Safir, Rutgers University

221 North 2nd Avenue

Highland Park, NJ 08904

USA

Abstract: Two structure-building operations are currently posited in minimalist theory: an operation forming sets (set merge) and an operation forming ordered pairs (pair-merge). I argue that pair-merge is sufficient to generate syntactic relations so set merge, also called simple merge, should be eliminated from syntactic theory on grounds of simplicity. This conclusion requires re-evaluating the relationship between structure-building and labeling of constituents for the expression of syntactic relations, since labeling plays a crucial role in this comparison of theories according to the simplicity metric. An existing labeling hypothesi, specifically Chomsky (2013) and (2015), is shown not to have any advantage claimed for it by comparison with the just pair-merge theory proposed here. An advantage of the only pair-merge hypothesis is that it provides a more principled origin for the inherent asymmetry in the c-command relation that does not follow from a theory that includes set merge.

1 Introduction

      In current minimalist theory, there are two structure-building operations, simple merge, which forms sets, and pair-merge, which forms ordered pairs. Chomsky (2004) proposes that syntactic adjuncts are so-defined by the fact that they are always introduced by pair-merge. However, it is intuitively appealing and consistent with minimalist aspirations to see if linguistic

theory can do with just one rather than two merge operations. I will posit that only pair-merge is necessary to generate the full range of syntactic configurations that grammar requires. I argue for the following propositions, with succeeding sections of this essay devoted to each point.

I. A theory with just pair-merge and no simple (set) merge is a simpler theory more consistent with minimalist assumptions.

II. A theory with just pair-merge and no simple merge is just as adequate to generate all relevant syntactic relations.

III. Apparent advantages of the two merge operations hypothesis, including the labeling algorithm, are illusory. (The existence of agreement is not derived, the EPP is not derived).

IV. Only the just-pair-merge hypothesis derives the inherent asymmetry of c-command.

The claim that all merge is pair-merge has been proposed by a number of others before, most notably, Kayne (1994) and Zwart (2005, 2009, 2011) a. o., though I crucially do not adopt the claim that c-command translates into any commitment as to linear ordering of the pair members. In a final section, before concluding, I will discuss other "just pair-merge" proposals in the literature and further distinguish them from what I propose here.

The revision I propose, though it simplifies the internal mechanisms of the minimalist theory of UG, does not have important empirical consequences, and so some would doubt the utility of the enterprise. However, if we are to take Hauser, et. al. (2002) (henceforth, HCF) seriously, and I am among those who do, any explanation of our linguistic ability must be

consistent with how such an ability could have evolved. HCF contend that linguistic ability

appears to have emerged in the history of homo sapiens rather recently and constituted a

relatively abrupt change. They reason that our capacity to generate linguistic structures was not

built incrementally by genetic change, but rather emerged quite suddenly when one new

capacity, added to existing cognitive capacities, permitted UG to emerge much as it is now. HCF

proposed that the ability to build recursive hierarchical organization of thought, the simple and

sudden new capacity in question, is the Merge operation. For the purposes of this discussion, I

am taking HCF as seriously as I can. However, Chomsky (2004), while building a theory intended

to be as consistent as possible with the HCF evolutionary scenario, actually introduced two

separate merge operations, simple merge and pair-merge. If we are to adopt the strongest

theory consistent with the evolutionary scenario and if the introduction of Merge was the key

genetic innovation, then there should be only one merge operation.[1] I have couched my

argumentation on the grounds of seeking a logically simpler theory, but the ultimate motivation

is not uniquely an appeal to Occam's Razor, but rather to build a theory consistent with the HCF

evolutionary scenario.[2]

---

[1] If there are two separate changes introducing the two merge operations, then the question would be if there was

ever a version of homo sapiens who had one operation and not the other. On incremental theories, the homo

sapiens with just one of these operations would have had to be outcompeted for survival by those who had both

operations. It is hard to imagine a more improbable scenario.

[2] The HCF hypothesis that only merge is the key factor in the advent of UG faces many challenges concerning

apparently syntactic phenomena such as binding, case, agreement and thematic assignment, etc., that the role of

Merge does not currently address (pace Hornstein 2000, 2009, see Safir 2008). The future analysis of those

2 What is simpler?

In order to compare two theories and rigorously determine which is simpler, one must have a measure of simplicity that is quite specific. The metric I will use here is based on a subset relation. If theory A has only a subset of the necessary assumptions of theory B, then theory A is simpler than theory B. If theory A has necessary assumptions that are not in theory B, however, then the metric does not decide between them. In this section I argue that a theory assuming the Just Pair-merge Hypothesis (JPMH) is a simpler theory than one that includes Chomsky's Two Merge Operations Hypothesis (TMOH).

TMOH, based on Chomsky (2013, 2015) has the following assumptions.

1a) **Simple Merge** forms sets, e.g., X and Y merged form {X, Y} with no linear order specified.

 b) **Pair-merge** form ordered pairs, such that X adjoined to Y forms <X, Y> with no linear order specified.

 c) **Labeling Algorithm (LA)**[3]

---

phenomena from the minimalist point of view may ultimately rely on richer component-boundary conditions – obviously a promissory note. This is one reason why there is a minimalist program, rather than a single minimalist theory. In this paper, binding, case, agreement and thematic assignment will be bracketed so that discussion is focused on the single question of whether or not more than one merge operation is necessary or desirable.

[3] LA locates the nearest syntactic object by a minimal distance search, which is unnecessary for labeling for a theory committed to JPMH, However, minimal search is a natural least-effort computational feature of any system searching in a hierarchical structure and so would presumably be available to computations over structures generated by JPMH. That JPMH does not appeal to minimal search for labeling is therefore not an argument

i. If X is a head and Y is a non-terminal[4], then $\{X, Y\}_X$

ii. If both X and Y are non-terminals, then they only project a label if they "agree"

iii. If only one of two non-terminals {A, B} is visible to LA, then the label for {A, B} is the label of the member of it that is visible.

iv. No label can be applied to {X,Y} if none of these conditions are met.

d) lower copies are invisible to LA.

e) Adjuncts formed by pair-merge are invisible to further syntactic operations. If A is pair-merged to B, then A is invisible to LA.

f) **Labeling Filter:** Every constituent must have a label.

Where A and B are non-terminals, the set formed by merge of A and B would not have a label because LA cannot choose the label of just one of two non-terminals to assign a label to {A, B}. The rationale for (1d) is that movement of one of two non-terminals of {A, B} leaves a copy invisible to LA (e.g., a copy of A represented by italicization), with the consequence that the label of B can be assigned to {*A*, B}. This provides a motivation for movement derived as a consequence of LA, a matter to which we will return.

---

against TMOH. Chomsky assumes minimal search is crucial to Agree in order to get minimality effects, for example. Agree is not at issue in this theory comparison, so if JPMH also has Agree, it also appeals to minimal search.

[4] My use of "terminal" and "non-terminal" correspond to "simplex syntactic object" and "complex syntactic object" for those who would prefer to use non graph-theoretic terms. The term "constituent" can also be taken to mean "syntactic object." A reviewer points out that Zyman (to appear) makes a point of distinguishing set-theoretic characterizations of syntactic objects from graph-theoretic ones, a point that will be orthogonal to my discussion. Zyman does not address issues involving pair-merge or what sort of object results from it.

Chomsky (2004) proposed (1e) to account for the purported generalization that extraction from adjuncts is generally banned.[5] This assumption has the consequence for LA that adjuncts are also invisible to LA. Thus (1ciii) Is satisfied when pair-merge forms the ordered pair <A, B>, because if A is adjoined to B, then A is invisible.[6]

JPMH, as proposed here, has the following assumptions.

2a) **Pair-merge** forms ordered pairs such that X adjoined to Y forms <X, Y> with no linear order specified.

b) **Pair-merge Labeling (PML):** If X is adjoined to Y, then $<X, Y>_Y$

c) **Labeling Filter:** Every constituent must have a label.

The asymmetry of pair-merge is evident in the labeling of the output. No constituents are stipulated to be invisible to PML, i.e., both adjuncts and copies are visible to subsequent syntactic operations, hence they are visible to PML.

---

[5] Chomsky (2008: 146-7) remarks: "The adjunct-island subcase follows if an adjunct is not in the search domain of the probe."

[6] Chomsky (2004) assumes that adjunction never affects what it adjoins to, so the adjoinee's label remains the same. However, this amounts to Pair-merge Labeling in (2c) insofar as the ordered pair that results must have the adjoinee's label. Chomsky (2004:118-119) also has to introduce a rule he calls SIMPL that restores adjuncts (which exist on a different plane) to visibility in interpretation and phonology as it converts ordered pairs to sets. SIMPL is hypothesized to be part of Spell-out, which is the final line of a syntactic derivation. Subsequent surface ordering must know If the set reduced by SIMPL was formed by pair merge (p.119). Ordering of SIMPL as part of Spell-out ensures that the adjunct is invisible for preceding steps of the derivation. If all merge is pair-merge, then SIMPL and some of these intricate ordering assumptions are unnecessary.

As it stands, the two hypotheses are not in a subset relation, since (2b) is not in TMOH and (1a, c, d, e) are not JPMH. Both hypotheses assume pair-merge and the Labeling Filter. If we leave matters here, the subset simplicity metric cannot be applied. However, if (1e), the proviso that adjuncts are invisible, can be shown to be untenable empirically, then TMOH will not have a way of labeling the output of pair-merge when two non-terminals are pair-merged. If (1e) is then just a stipulation that adjuncts are invisible to Labeling Algorithm, then it has the same force as adding Pair-merge Labeling to TMOH. On this reasoning, (1e) is the equivalent of PML, in which case JPMH has a subset of the assumptions required by TMOH. The metric then apples: A theory containing JPMH is simpler than one that instead includes TMOH.

2.1 Evidence that adjuncts are visible to syntactic operations

Cattell (1976) and Huang (1982) have noted that extraction from adjuncts is problematic and this has led to the generalization, encompassed by the Adjunct Island Condition, that extraction from adjuncts is blocked. Chomsky (2004) takes this generalization at face value and treats it as a very strict prohibition, namely, as above, that adjuncts are invisible to syntactic operations such as extraction (Stepanov, 2001: 117 asserts that extraction from adjuncts is universally blocked). Similarly, if LA is a syntactic operation, then adjuncts are invisible to labeling. However, here is ample evidence, that at least some adjoined constituents are indeed visible to syntax in that their contents are accessible for internal merge to relate them to adjunct-external positions.

Chaves (2021), which reviews the history of noted exceptions to the adjunct condition, points out, with numerous examples from numerous sources, that such extraction is widely attested even though it is often restricted. Culling from the literature, Chaves (2021: 691 and

references cited there) presents examples of extraction the following sorts of adjuncts, including *without* gerunds, bare adjunct gerunds, *in order to* adjuncts, temporal adjuncts and conditionals (and the conditional-like example in (3g)).[7]

3a) That's the symphony that Schubert died without finishing.

b) What did he drive Mary crazy trying to fix?

c) Who did you go to Girona in order to meet?

d) This is the house that Mary died before she could sell.

e) This is the watch that I got upset when I lost.

f) The person who I would kill myself if I couldn't marry is Jane.

g) Which celebrity did you say that the sooner we take a picture of, the quicker we can go home?

One way to salvage the ban on extraction from adjuncts would be to argue that the cases in (3) are not really extractions at all. For example, Cinque (1990) argued that extractions of the sort in (3) might be cases where there is A'-licensing of a pro inside the adjunct. To support this proposal, he shows that PPs cannot be extracted from adjuncts but adjuncts can have a nominal gap. He argues that pro is only a proform for nominals, so the absence of PP gaps in adjuncts is because there is no proform for PPs in Italian. I don't have an answer for why PP extraction is blocked out of adjuncts, however, it is generally assumed that pro does not allow reconstruction.

4a) Which of his$_i$ responsibilities did every congressman$_i$ leave town without/before taking care

---

[7] A reviewer alerts me to McIinnerney (to appear), who argues that whether or not PPs can be extracted from is orthogonal to their status as adjuncts or arguments, but under the right conditions, extraction from adjunct PPs is possible.

of?

b) How many of his$_i$ incompletes can any student$_i$ graduate from here without taking care of?

c) Which of his$_i$ creditors would you expect any gambler$_i$ to skip town (in order) to avoid

The bound reading of the pronouns in these cases is taken to be possible in these contexts because *every congressman* c-commands a pronoun in the extraction site (on the copy theory of movement, inside a copy). If the extraction site contains a copy with internal structure that includes the pronoun *his*, then *any student/gambler* will c-command the pronoun, but not otherwise. The reconstruction effect shows that at least these cases must involve internal merge leaving a copy in current minimalist terms. Thus, these cases have to be extractions from adjuncts, at least for English.

The extraction from adjuncts is not restricted to English. Kush et. al. (2019:399) illustrate extraction from adjuncts in Norwegian, as in (5a) (see also Bondevik, et. al., 2021). Biskup & Simik (2019) present a variety of Czech examples, primarily extraction from left peripheral adjuncts by relativization, as illustrated in (5b).

5a) Bakdøren   blir   han   nervøs om   de   lar   stå   ulåst.

back.door.DEF gets   he   nervous if   they   leave stand   unlocked

'The back door he gets nervous if they leave unlocked.' (Topic-Island, Norwegian)

b) To   je   řečník,   kterého   [aby   nalákali t],   museli

That   is   speaker   which.ACC   in.order   attract   must

by   mít   peníze.

SBJV   have   money

'That's a speaker such that they'd have to have money in order to attract him.'

9

Chaves concludes that the restrictions on extraction of adjuncts are not uniquely syntactic in nature, as they seem sensitive to context. Truswell (2007, 2011) analyzes extraction of adjuncts as possible when there is a particular temporal relation between events, but as Chaves points out, extractions from tensed adjuncts, such as those illustrated above, do not seem consistent with that approach. Biskup & Simik take a more syntactic approach based on the leftmost or rightmost position of the adjunct. Bošković (2018) argues that extraction of adjuncts, which he illustrates with Serbo-Croatian examples, is possible if base generated syntactic objects are at the edge of adjuncts. However, the best account of why extraction of adjuncts is possible in some contexts and not others is not at issue here. What is pertinent for my argument is that extraction of adjuncts is clearly possible and thus a categorical syntactic ban on extraction of adjuncts is simply not empirically viable.

Finally, there is potential evidence that depends on a particular approach to multiple interrogation, namely that of Bobaljik (2002). On Bobaljik's theory, internal merge applies in the syntax to the in situ wh-phrases in (6a, b) though only the lower copy is pronounced.

6a) Which of the congressmen left town before talking to which portion of their

constituents?

b) Who crossed the room (in order) to avoid who?

An absolute restriction on access to adjuncts on the grounds that they are invisible to further syntactic operations would require abandoning Bobaljik's approach. If Bobaljik is right, however, then the posited invisibility of adjuncts to further syntactic operations is too strong and thus not viable on theoretical and empirical grounds. If Bobljik's approach is to be rejected, then these cases do not bear on the question at hand.

It is also worth pointing out that Chomsky's theory of wh-extraction proceeding by movement to intermediate positions at the edge of phases (cyclic domains) is not formulable as adjunction (though it always is) if his approach to pair-merge renders pair-merged adjuncts invisible to further operations. Once a wh-phrase is adjoined to a phase, the adjoined phrase would be invisible, so internal merge could not apply to it successive cyclically.[8] For this account of wh-movement, adjuncts must be visible to syntactic operations.

If adjuncts are visible to syntax, then there is no reason to say that they are not visible to Chomsky's LA. LA is faced with indeterminacy for constituents created by pair-merge unless another form of abstract agreement is posited to allow for shared-feature labeling determined by the features of the adjoinee and the adjunct. Inventing such an agreement, or perhaps some other stipulation that will feed the LA to get the right result, must be independently motivated. In other words, the purported invisibility of adjuncts to just labeling theory, not merge, is then just an additional stipulation, equivalent to Pair-merge Labeling, to ensure that the output of pair-merge gets the label of the adjoinee. In short, both TMOH and JPMH stipulate PML or its equivalent.

Thus, the extraction evidence that is supposed to justify the claim that adjuncts are invisible to syntactic operations in fact gives evidence for the opposite conclusion: Adjoined

---

[8] Stepanov (2001:117-118) notices this consequence of treating adjuncts as invisible but does not attempt a solution consistent with his assertion that adjuncts are invisible to further syntactic operations. A reviewer suggests that one can always appeal to a theory of multiple specifiers at the cycle edge to avoid this consequence. The result would be a much more porous theory of cyclic domains. I have limited specifiers to the first merge to a head and its complement. In any case, an additional specifier position is not what Chomsky proposed.

constituents must be visible to syntax or extraction from them would be impossible. Once it is established that adjuncts must be visible to syntax, Chomsky's LA is not sufficient to provide the right labels when pair-merge applies to form a constituent of two non-terminals. If PML is part of TMOH, then label indeterminacy is avoided. However, the upshot of adding PML to TMOH is to render JPMH assumptions a subset of those in TMOH. By this metric, JPMH is straightforwardly simpler than TMOH.

3. JPMH is adequate to generate all known relevant syntactic relations

Even if theory A is simpler than theory B by the subset metric, theory A is not attractive if it is insufficient to characterize the relations that are important for explaining the phenomena that the theory is supposed to address. JPMH, if it is to be a viable theory of natural language syntax, must be able to generate all of the syntactic hierarchical relations that are necessary for any adequate theory of grammar, or at least those that are currently claimed to be necessary. Among those relations, linguists appeal to the following:

7a) Categorization – combination with a root and a labeled head

  b) Complementation – combination of a labeled root+head and a non-terminal

  c) Non-complement argument assignment – "specifiers"

  d) Adjunct relations – non-argument subordination

  e) Head-head relations (and morphological complexity)

Pair-merge is indeed adequate to generate structures that these hierarchical relations can be stated on. The ordered pairs in (8) generated by pair-merge illustrate how the purely syntactic relations in (7) can be represented in a system without set merge (I postpone discussion of

(7e)). In the examples below, the subscript indicates which member of the ordered pair

contributes the label for the whole constituent.

8a) Categorization $<root, X>_X$

  b) Complementation $< YP, <root, X>_X>_X$

  c) Specifier $<EA, <YP, <root, X>_X>_X>_X$

  d) Adjunct $<ZP, <EA, <YP, <root, X>_X>_X>_X>_X$

Tree representations only represent dominance relations and do not distinguish between

constituents that are sets and those that are ordered pairs. Let us assume for the purposes of

presentation that the tree in (9) (which abstracts away from linear relations) has been

generated entirely by pair-merge. I assume that the relevant syntactic relations in (7) are

conditioned and determined at the point in the derivation where the hierarchical relations are

constructed.

9)
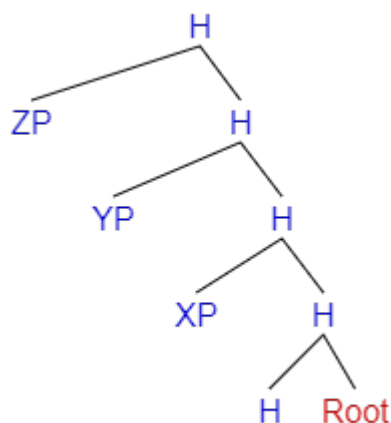
Since JPMH includes a labeling filter and a morphological root lacks a label by hypothesis, as proposed by Marantz (1997) and Borer (2005), the root must adjoin to a categorizing head H to produce a categorized constituent, namely, a head projecting H, which corresponds to the lowest pair of the tree. XP pair-merged to a categorized head yields a complementation relation. Given PML, it matters which of the pair is adjoined to the other. I assume that PML applies after each instance of pair-merge. If the categorized head were pair-merged to the complement, the label for the whole would be XP, which is not a complementation relation to categorized head H. The structural condition for being what we descriptively call a "specifier" is that it be the first constituent to adjoin to a categorized head X and its complement. The output bears the label of the pair <complement, categorized head> and it is where I assume an external argument (EA) can be assigned.

However, the syntactic relations in (7) are not sufficient without selectional conditions. Consider the role that selection plays in the definition of "specifier". There are, after all, unaccusative predicates that do not appear to select for a specifier, so if $<_{XP}$ complement, X> is adjoined to a head (which takes the former as its complement), then there simply is no specifier of XP. However, suppose non-terminal is pair-merged to an unaccusative predicate, on the assumption that these predicates take their single argument as a complement that raises to subject position. For example, in a sentence like *The new century arrived without fanfare*, the non-terminal *without fanfare* modifies the arrival, not *the new century* directly. *Without fanfare* is interpreted as an adjunct. Since the predicate in question does not select an external argument, pair-merge of the non-terminal *without fanfare* constituent to <complement, head> will not meet the selectional condition (*arrive* does not select for a manner EA), only the

syntactic one, so it needs to be characterized as an adjunct when it enters the derivation. In other words, the first non-terminal pair-merged to $<_{XP}$ complement, X> is only a specifier if X selects for a non-terminal in that position. This is, in fact, the classic version of the EPP stipulation about T, namely, that it requires there be a specifier for T. Given these concerns, I define the central syntactic relations as follows, where structure is a condition in the definitions of "specifier" and "complement".

10a) Categorization is achieved by pair-merge labeling in the configuration <root, X>$_X$

b) A non-terminal, YP, is a complement in the structure < YP, <root, X>$_X$>$_X$ if and only if <root, X>$_X$ selects for YP.

c) ZP is a specifier in the structure <ZP, <YP, <root, X>$_X$>$_X$>$_X$ if and only if <root, X>$_X$ selects for ZP and YP is a complement.

d) If a syntactic object in structure is not defined by (a), (b) or (c), then it is an adjunct.

This accommodates an unaccusative predicate modified by a non-terminal adjunct and allows for the possibility that a terminal node could also be an adjunct, e.g., adverbs and/or internally merged clitics. This may or may not be the best way to characterize adverbs or clitics, but these are not issues I will resolve here, rather, I merely point to the fact that these possibilities are instantiated in existing theories and one does not need to contrast set merge and pair-merge to treat these elements as adjuncts. [9] All of these relations can be established at the moment they

---

[9] Analytically speaking, it is not clear whether there is any reason to ensure that only a non-terminal can be selected as a specifier, so I have not characterized ZP as a non-terminal, but it could easily be stipulated that Z in

are formed. The moment that a non-terminal is merged to <complement, head>, it is characterized either as a head or adjunct based on whether or not there is a selection relation. No later inspection of the tree is required. [10]

It is also necessary to account for the case of unergative intransitives, which have no overt complement and assign their single thematic role as an external argument. This theory does not allow a constituent to meet the structural condition for being a specifier if a non-terminal cannot be merged to <complement, head>.[11] If however, the EA is always introduced by VoiceP, as in fn.11, then the vP containing any sort of verb will be a complement to Voice, thus always permitting the EA to meet the syntactic condition on specifiers.

By providing descriptive definitions of what counts as a specifier, complement or adjunct, I am not adding anything new to the theory, as assignment of thematic roles has always been contingent on the position of the syntactic constituent that counts as an argument or adjunct for interpretation. With respect to adjuncts, however, JPMH posits no special

---

(10c) must be non-terminal. The outcome of such an adjustment has no consequence for the reasoning presented here comparing JMOT and TMOH.

[10] It is not apparent that there is any demonstration that elements interpreted as arguments cannot be introduced by pair-merge except by stipulation. For example, Rubin (2003), who assumes set merge generates complementation and specifier relations, attempts to justify the introduction of pair-merge in a structure by means of a head that stipulates the adjunct is a modifier.

[11] If it is assumed that the EA is introduced in the specifier of VoiceP, following Kratzer (1996) and most minimalist work since, it is still assumed that Voice takes a vP complement, so the structural condition for specifiers is satisfied for the introduction of EA by VoiceP.

operation that defines them and contrasts with a distinct operation that introduces syntactic

constituents that can count as arguments.

From this perspective, TMOH seems to have unnecessary descriptive power, since

TMOH allows all the syntactically relevant hierarchical relations in (9), as elaborated for pair-

merge in (10), to be generated by simple/set merge as well. The key syntactic relations as

generated by set merge are described in (11).

11a) Categorization $\{root, X\}_X$

b) Complementation $\{YP, \{root, X\}_X\}_X$

c) Specifier $\{EA, \{YP, \{root, X\}_X\}_X\}_X$
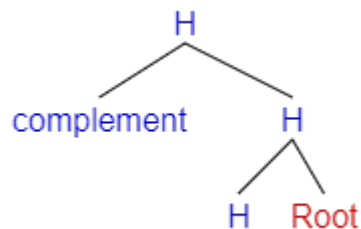
d) Adjunct $\{ZP \{EA, \{YP, \{root, X\}_X\}_X\}_X\}_X$

The redundancy extends further, however. TMOH also can generate mixed tree representations

that achieve identical dominance relations, as demonstrated in (12) for complements.

12a) <complement, $\{root, v\}_v{}^o$>

b) $\{complement, <root, v^o>_v{}^o \}$

13)

Altogether there are four ways to generate (13) for a theory committed to TMOH, but only one for a theory committed to JPMH. It is hard to see what justifies positing so much generating power if we can already generate the hierarchical relations in (13) with just pair-merge.

Finally consider (7e), which builds morphological structure. There may be cases where two morphemes both lack a label, as in a case where merge applies to two roots. Neither TMOH nor JPMH have anything explicit to say about these cases, but TMOH holds out the descriptive possibility, so far never exploited, that morphology could be constructed by either operation or even involve mixed trees that contain both ordered pair and set constituents. Again, this seems like unnecessary descriptive power. On the other hand, the Labeling Filter in either JPMH or TMOH would block such an output of pair-merge because neither subconstituent could provide a label. We could weaken the Labeling Filter to say that it does not apply to terminal nodes, so the non-terminal node formed by merge applying to two roots would not have a label, ruling out adjunction of two roots. However, the Labeling Filter so amended would not rule out adjunction of a root to a terminal node with a label (but see de Belder & van Craenenbroeck, 2015, for a different approach to these cases). It may be that joining two unlabeled roots into a larger root is not a syntactic operation, but in any case, it is not going to decide between JPMH and TMOH. Finally, neither account addresses the status of heads built from merge operations. In general, the assumption appears to be that heads are terminals or non-terminals that are evaluated as complex terminals (e.g., in Halle & Marantz, 1993: 116, where what I am calling heads are called "zero-level category nodes"). However unsatisfactory this description of heads may be, it does not distinguish TMOH from JPMH.

I turn now to the relevant syntactic relations that must be generated to characterize displacement structures. In keeping with Chomsky's treatment of internal merge as just an instance of merge, I assume that pair-merge, by its nature as a merge operation, allows for internal as well as external merge (as pointed out by Richards, 2009).

Internal pair-merge of a syntactic object X to a nonterminal Z that contains X is now the description of all such relations, including wh-movement and A-movement, traditionally treated as substitutions (generated in Chomsky's system by simple merge) as well as VP fronting, heavy NP shift, clausal extraposition, extraposition from NP etc. which are traditionally treated as adjunctions. Italics indicate the unpronounced copy left in the position from which the constituent was internally pair-merged.

14a) Which books did Mary read *which books*?

   b) The cat seems *the cat* to be out of the bag.

   c) Angry at Alfie though I was *angry at Alfie*, I still invited him.

   d) Selma will consider *any attempt to stop her from posting on Instagram* (to be) outrageous any attempt to stop her from posting on Instagram.

   e) It was apparent to us *that John would not be coming today* yesterday that John would not be coming today.

   f) We spoke to a Democrat *who was from Nashville* yesterday, who was from Nashville.

The ordered pairs of terms in (15a-f) correspond to the outputs in (14a-f) (for (15a-f), the square brackets simply avoid recapitulating the internal structure of the terms of the pairs).

15a) <[which books], [did Mary read *which books*]>

  b) <[the cat], [seems *the cat* to be out of the bag]>

 c) <[angry at Alfie], [though I was *angry at Alfie*]>

  d) <[any attempt to stop her from posting on Instagram], [*any attempt to stop her from posting on Instagram* to be outrageous]>

  e) <[that John would not be coming today], [was apparent to us *that John would not be coming today* yesterday]>

  f) <[who was from Nashville], [we spoke to a Democrat *who was from Nashville* yesterday]>

Notice also, that some adjunctions are to the left and others to the right, which is expected to be possible since pair-merge as posited here does not determine (or predict) linear order (nor does it in Chomsky, 2004). While it is unclear as to exactly where the (apparently) rightward-moved constituents land, it is quite generally assumed that they are adjoined where they land in the clausal architecture – as is necessarily the case for the JPMH. The diagram in (15) corresponds to (13e), where it is assumed that the CP adjoins in a position c-commanded by the subject, but outside the position of the temporal adjunct *yesterday*. The use of the AUXP label is for convenience of presentation only.

16)
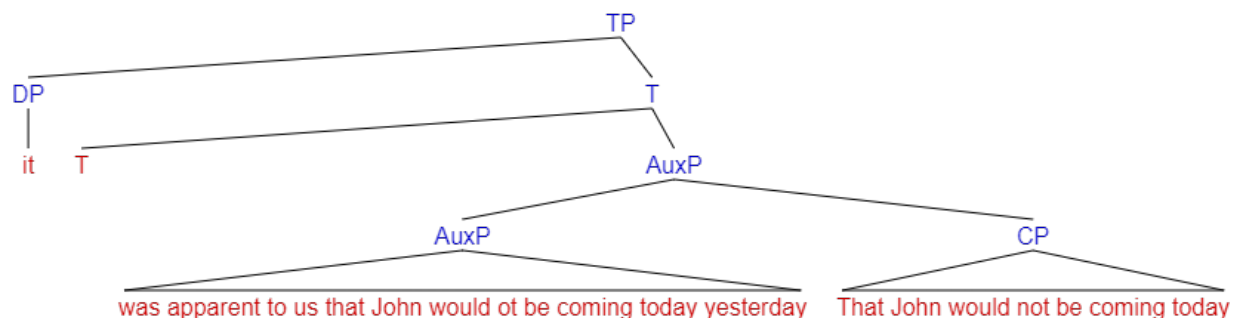
The instances of pair-merge in (15) are the only way to generate internal merge outputs assumping JPMH. TMOH can also generate all of (14a-f) with pair-merge, but it could also generate the same dominance relations in (14a-f) with simple/set merge.

The upshot of section 2 is that JPMH has the power to generate all the hierarchical relations that are required to express relevant syntactic relations, including the output of internal merge. Simple merge is not required to generate these relations, except by stipulation, added to TMOH, that certain syntactic relations must be generated by simple merge (e.g., complementation, argument structure). Thus, a theory assuming JPMH is adequate theory to express syntactic relations, in addition to being simpler than a theory committed to TMOH.

4 What is lost if we give up TMOH?

At least three advantages have been claimed to be consequences of TMOH with LA.

17a) The EPP can be derived.

   b) Label indeterminacy explains why intermediate movement does not halt short of its

      destination (the halting problem).

   c) The existence of agreement is motivated as a solution to label indeterminacy.

In this section I show that these advantages are only apparent. There is no principled explanation of them yielded by TMOH.

4.1 The Halting Problem – Is it a problem?

Suppose we accept that movement is not feature-driven, as suggested by Chomsky

(2013, 2015, 2020)[12] and Safir (2010, 2019) and that structure building is free (free merge).

Once we say that all structure-building is free, (free Merge), so too is internal merge. Thus

movement (=internal merge) can always take place, modulo island constraints, etc. If so, some

other factor or factors must determine that internal merge must take place. Triggers are

unnecessary to induce movement.

Chomsky (2008) and Safir (2010) suggested that movement outputs fail if constituents

do not reach positions where they can be interpreted. Rizzi's criterial positions (Rizzi, 2006) are

those destinations. [13] For example, a wh-phrase must ultimately be in a structural position

---

[12] Chomsky (2008: 144) posits free merge as a possibility but proposals in that essay still assume triggered

movement. See also Boeckx (2010).

[13] Recall I am assuming that linear order is determined at the morphophonological interface as well as free merge.

Pair-merge does not translate to surface linear order on my assumptions. How a language comes to be SOV, SVO,

VSO, etc. will require morphological features sensitive to linear order that must be satisfied at the syntax-

morphology interface. If movements do not take place allowing the morphological features to be satisfied, the

derivation fails. These are, in effect, indirect triggers for movement. It is not obvious that derivations with direct

triggers are less descriptive. The same can be said of those theories that stipulate underlying order is universally

SVO (e.g., Kayne, 1994) and include diacritics on heads to attract movements to produce the empirical range of

necessary results. I make no claim that free merge accounts of word order variation will be superior to stipulations

necessary in any other theory, only that there is no reason to believe that the stipulations necessary for a free

merge theory are likely to be more numerous or worse. In any case, the free-merge assumption is not one that

distinguishes between TMOH and JPMH.

where it takes scope over the question. Relations like these are translated in Chomsky's labeling theory as the position where appropriate agreement can take place (e.g., interrogative feature on a wh-phrase agrees with the interrogative feature on C). If it is criterial relations that indirectly drive movement, then the consequence of stopping in an uninterpretable position means only derivations that move to criterial positions will lead to successful interpretations. Since movement is bounded by phases, getting to the edge of a phase where a wh-constituent can be moved again in the next phase is the only chance a wh-phrase has to reach CP when escaping the vP phase. Free merge allows for movement to the edge of a phase and for subsequent movement to (descriptively) Spec, CP. There is no halting problem. A derivation of the relevant sort is illustrated in (18).

18) Who did John see?

   a) *Who* originates as object of *see.*

   b) *Who* moves to the edge of vP phase by free merge.

   c) Derivations where *who* stops at the vP phase edge will fail at LF.

   d) *Who* moves to Spec CP.

   e) Derivations where *who* lands in Spec CP will reach criterial positions and yield
      acceptable outputs.

Quite apart from the fact that an account based on label indeterminacy is not necessary to solve the halting problem, there is an additional issue with the halting solution Chomsky proposes. In his theory, wh-movement from object position stops at the vP phase edge ***as an adjunction***. Adjoined positions are invisible to further syntactic operations by hypothesis for TMOH. An internal pair-merge derivation for vP adjunction of wh-phrases according to TMOH

would freeze them in place in adjoined position, regardless of whether the <adjunct, vP> constituent is successfully labeled a vP or not. Now that would be a halting problem.

4.2 Labeling as the motivation for the existence of agreement

In many languages, agreement is vestigial or absent, which raises the question why it should exist at all. Chomsky's (2013, 2015) labeling theory motivates the existence of agreement as one of the repairs available when two non-terminals are sisters and neither one on its own can determine the label of their constituent. However, we must take a lot on faith to see this as a successful derivation of the necessity of agreement, since abstract agreement will be required in a wide variety of structures where it is not exponed.

In English and certain other languages, gerunds, infinitives and small clauses lack phi-feature agreement with their subjects, but both require the establishment of subject-like arrangements.

19a) [PRO eating fish] didn't bother me

  b) [PRO to be late] would be unfortunate

  c) [Workers angry about the pay] is not what we expected when we agreed to this.

The bracketed constituents in (18) can be inverted by subject-auxiliary inversion and raising to subject, which are tests for subjecthood.

20a) Why would eating fish ever seem to be dangerous?

  b) To eat fish would seem to be a good solution

  c) Why shouldn't workers ready to bargain be the goal of our tactics?

Gerund and infinitive subjects have no overt phi-features that agree with T (at least in English). With less reason to believe that (abstract) Case must be assigned to a nominal (see, e.g.,

McFadden, 2004, among many others), there is less reason to posit that gerunds and infinitives are nominal subjects bearing phi-features (at least in English).[14] LA was designed in such a way as to account for why subjects must agree, but when subjects have no apparent phi-features, it is hard to see how LA makes an accurate prediction. Even if one accepts *-ing* as the T equivalent in gerunds and *to* as T in infinitives, for the small clause case subject examples (19c) and (20c) there is no candidate for T and still no visible agreement.

The same sort of demonstration can be made for examples like those in (20).

21a) Under the bed seems to be a good place to hide.

b) Has under the bed been a good place to hide?

c) Under the floorboards is/?*are a good place to hide.

It seems that in certain circumstances, PPs can be subjects in that they raise and invert with auxiliaries (see Safir, 1983), but it is not obvious that there is anything in PP to agree with, including the prepositional object. How is weak T propped up by agreement here?

Of course, it is always possible to mint additional abstract agreement that could apply between the various subjects in (18-20) and their setmates/sisters. Sometimes positing abstract (non-exponed) relations to render a theory symmetrical is the sort of idealization that can lead

---

[14] Saito (2016) argues that these problems are resolved by assigning shared Case between T and the element acting as a subject. This might apply to gerunds, if the possessive marker *'s* is a determiner heading a gerund, perhaps but it would not obviously extend to PP subjects. The agreement in Case would have to extend to contexts where Nominative is not involved but the gerund acts as a subject, e.g., *For treating this malady with antibodies to be effective…* and *John considers arresting the demonstrators to be a poor solution*. What is agreeing to license the relation between two non-terminals as sisters in these constructions?

to insight, but in this case, positing agreement relations that license positions where they can appear on the surface only serves, without independent motivation, as a stipulation to save a theory (see also Zyman, 2017:13 fn.21).

Moreover, the most glaring problem for the justification of agreement was never solved. [15]

22a) There are many men in the room.

b)*Are many men in the room. (Declarative)

c)*There is many men in the room.

Agreement on T is not sufficient to license a null subject so the constituent consisting of T and its complement must have an unexponed feature to agree with one that must be stipulated to be on the *there* subject. However, T does have an exponed agreement with the *there* associate. This latter agreement does not provide an explanation for why the notional subject can remain where it does, e.g. *there are [[three men] [in the room]]* (as Epstein et. al. 2014: 466, fn.13 point out and set aside). The problem extends to the gerunds and infinitives which can have *there* subjects as well, e.g, *there being no one in the room would bother Bill, For there to be only two people in the room would be unfortunate.* Abstract agreement would have to hold between *there* and the infinitive/gerund independent of the postverbal subject.

---

[15] There have been attempts. Abe (2018), building on Sato (2008) proposes that *there* "and its associate form a constituent underlyingly" and then *there* moves to subject position where, unlike other DPs, *there*'s uninterpretable and unvalued phi-features are then satisfied within the assumptions and reasoning about Agree. Goto (2017) (written after Abe's paper) attempts to adapt the Sato-Abe approach to Chomsky's (2015) LA, but to do so, Goto invents a special head-head labeling relation in addition to LA.

In short, *there*-constructions are evidence that agreement, expressed in typical minimalist theories as Agree, must be posited for reasons independent of labeling. LA as it functions in TMOH does not provide a sufficient natural rationale for the existence of agreement relations.

4.3 EPP is not derived by LA and weak T

The EPP is a generalization to the effect that every sentence must have a structural subject, overt or covert. On some accounts, it does not hold in every language, and in others it is considered a universal. I will not consider whether or not there is a parametric difference. Instead, I will focus on whether or not LA and the stipulation that T is too weak to project a label without agreement are together sufficient to derive the EPP in the languages where it is assumed to hold. If the constituent consisting of T and its complement (T', for the purpose of presentation) has no agreement partner, it cannot project a label, since the label would have to originate on a head, and T, by hypothesis, cannot project a label unless it is agreed with. The intuition behind the analysis is that sister agreement enables labeling where there is a relation between two non-terminals. Thus, the need for agreement (to enable a label to be assigned) requires that {T, complement} always has a non-terminal sister, a subject, and that derives the EPP.

I have pointed out in the last subsection that subject positions can be filled with gerunds, infinitives, small clauses, and PPs in certain contexts and in all of these cases subject-auxiliary inversion shows them to be subjects. If T agrees with these subjects, then it is not obvious that T has an agreement partner that has any phi-features to agree with. There are a variety of stipulations that might solve this problem. For example, it could be assumed that T

27

has either phi-features or an anti-phi feature, that it agrees with phi-features when the subject

has phi-features, and that it agrees with subjects that lack phi-features by matching an anti-phi-

feature. This comes close to stipulating the EPP in the agreement system, but even if we were

to allow such a move, it would still run afoul of *there*-constructions.

As we saw in the last section, agreement is overt in English *there*-constructions, but it is

not agreement with the structural subject, rather T agrees with the associate of *there* which it

c-commands. In that case T must have phi-features that agree, yet English still requires the

subject position to be filled with *there*, even though T agrees with the *there*-associate and not

with the subject. To save the LA account of EPP, it could be said that T has either phi-features,

or anti-phi-features, or it can optionally have both types of features. In the latter case, T agrees

with the phi-less subject in a *there*-constructions, but also agrees with the associate in phi-

features. However, all this does is to build the EPP into stipulations about agreement with

never-exponed anti-phi-features. This is admittedly my straw man addition to TMOH, but

without such a stipulative move, or one like it, the LA+weak-T account of the EPP is not viable

for even paradigm cases like *there*-constructions.

JPMH has nothing to say about the EPP, but if the reasoning in this section is right, then

TMOH does not provide an explanation that gives it any advantage over JPMH.

4.4 Upshot of section 4

JPMH does not provide a rationale for the existence of agreement nor does it derive the

EPP from independently motivated general principles. In this section I have shown that TMOH

does no better in this respect, despite claims to the contrary. It is possible that the halting

problem is not a problem at all in a free-merge theory – if the result of movement must satisfy

criterial relations. There is indeed a halting problem in TMOH, however, if adjuncts are invisible

to further operations and cyclic wh-movement proceeds by adjunction. In short, by adopting

JPMH, nothing is lost by abandoning the purported explanations of these three phenomena

that are supposed advantages of TMOH.

5 An advantage of JPMH - The derivational definition of c-command

As shown in Section 2, preserving simple merge alongside pair-merge introduces a

whole set of alternative derivations that can result in the same hierarchical structures that pair-

merge generates, Consider the derivational definition of c-command, first introduced by

Epstein (1999: 329).

23) Derivational c-command: X c-commands all and only the terms of the category Y with

which X was merged in the course of a derivation.[16]

At the time (the article was circulated in 1994), Merge and Move were separate operations and,

as Hasegawa (1996) points out, Epstein did not even consider additional issues that arise for

adjunction (before Chomsky, 2004, formulated adjunction as pair-merge), or for movement

that results in adjunctions (which Hasegawa also noted and set aside). Since adjunctions also

create c-command relations, a theory with both set merge and pair-merge must create a cover

term "Merge" that encompasses both set merge and pair-merge. These complications reduce

to the following definition once simple merge is eliminated.

---

[16] This is Epstein's preliminary definition, but the differences between this and his final version do not concern us

here.

24) C-command: If X is pair-merged to Y then X c-commands Y and everything Y

dominates.

No special provisos for other operations creating c-command relations need be referred to.

A consequence of (24) is that there is no such thing as mutual c-command, as there is in

a theory where merge operations form binary sets.[17] It is not obvious that there is any situation

where fully symmetric (mutual) c-command relations are empirically needed, yet that is what

set merge generates. Indeed Brody (2002: 28) objects to derivational c-command based on the

absence of asymmetry in the operation of structure-building. JPMH explains why: Sisters

created by pair-merge are, by definition, never symmetrical (see also Zwart, 2011).

In fact, there is good reason to reject the possibility of mutual c-command. For example,

mutual c-command would permit a predicate to c-command something inside what it is

predicated of. Consider the weak crossover contrast illustrated in (25a,b).

_____

[17] A reviewer objects that the naturalness of the asymmetry in c-command that blocks mutual c-command would

be just as natural for the two-merge theory if an operation reduces all merge outputs to sets and then c-

commands is defined on sets as in i.

   i.   A node A c-commands a node B and everything B dominates iff (a) A and B form a set and (b) B

            projects its label.

Apart from having to convert pair-merge outputs to sets, this proposal fails for just the reason proposed here. In

Chomsky's labeling system, if neither sister is a terminal, then labeling can proceed by finding matching features in

the two non-terminals. In such a situation, there should be mutual c-command by the definition in i. or no c-

command at all, depending on how the contribution of matching features is interpreted with respect to "projects

its label". This illustrates how mutual c-command can stubbornly re-emerge if the merged constituents are just

treated as members of a set.

25a) [[every man's new spouse] [praises his mother]

   b)*[[his new spouse] [praises every man's mother]]

A typical formulation of the configuration where weak crossover is found is stated in (25).

26) Weak crossover configuration

   A pronoun cannot be irreflexively dominated by a constituent that c-commands the

   quantifier that it is dependent on (i.e., that binds it).

Although not every configuration where (26) holds yields a weak crossover effect, (26) is

criterial for what counts as a weak crossover effect.[18] This way of putting the generalization

captures the fact that *every man*, though it does not c-command the pronoun, is contained in a

---

[18] Safir (2004, for example, treats (26) as a corollary of his Independence Principle. As stated, (26) abstracts away

from reconstruction phenomena, though most theories of reconstruction based on copy theory end up with

structural analyses for which (26) is true. For example, a reviewer raises i. attributed to Edwin Williams.:

   i.   [that he$_1$ might someday meet the queen] inspires [every British soldier]$_1$

   ii.   It inspires every British soldier that he might someday meet the queen.

If it can be argued that i. is derived from an expletiveless version of ii. by movement of the CP to subject position

then this is another case where reconstruction figures in the determination of when (26) is true (more challenging

examples come from equative structures, see Williams, 1986; 287), The statement in (26) has been challenged on

the basis of a wide variety of apparent counterexamples, almost always of the form that an expected violation is

not unacceptable, or not as unacceptable as it should be. A broad range of such cases is discussed in Safir (2017).

All the examples that are unacceptable and treated as WCO in the literature can be characterized by the

configuration described in (26). Indeed, acceptability of the bound reading in the (26) configuration is how

exceptions to weak crossover are defined. In other words, the asymmetry predicted by the derivational definition

of c-command holds for all the cases in the literature that are agreed to be WCO violations.

constituent that c-commands the pronoun. However, if the constituent T' (the particular label is only assumed for ease of presentation) is a sister to the subject, then *every man's mother* is c-commanded by T'. Thus, the bound variable pronoun is dominated by a constituent that c-commands the subject and the quantifier in it that the bound variable pronoun depends on. Consider the diagrams in (27a) and (27b) which correspond to (25a) and (25b), respectively.

27a)
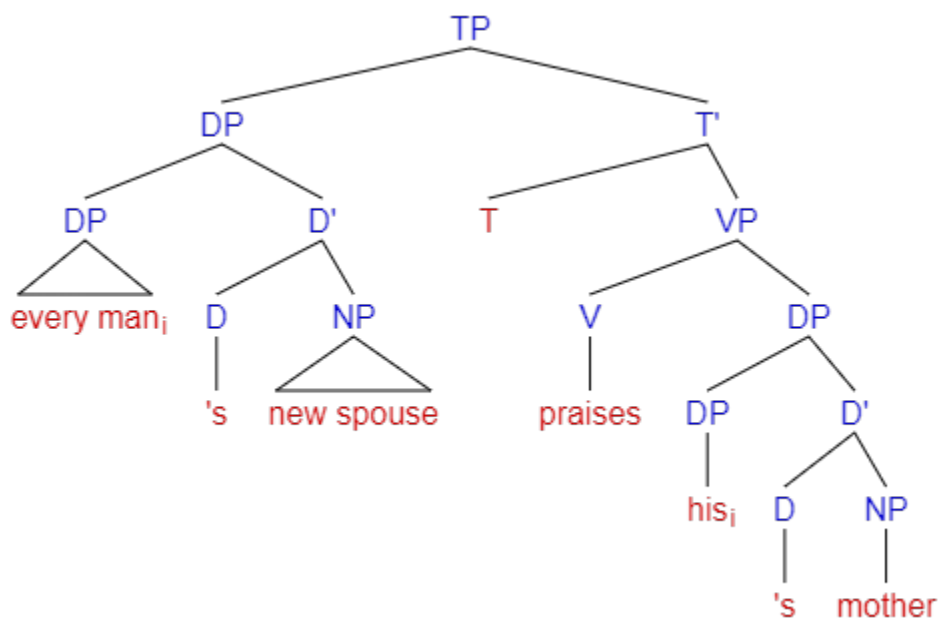
27b)
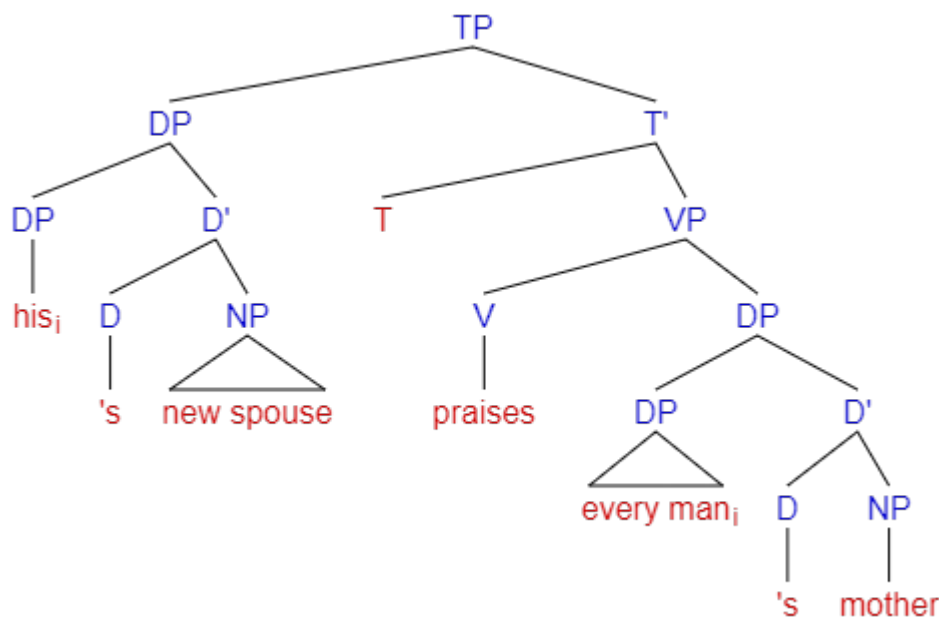
In a theory that allows [praises every man's mother] to c-command the contents of the subject as a result of set merge, the two sentences (25a,b) have the same violation, unless it is stipulated that the pronoun cannot be contained in a **nominal**[19] that c-commands the quantifier that binds it. For JPMH (and in the other ordered pair output hypotheses), the subject DP is adjoined to [T XP], so it asymmetrically c-commands [T XP] in (25a). No further stipulation about the subject being a nominal is necessary.

For a theory committed to JPMH, then, c-command could be recorded after each application of pair-merge, just as Epstein envisioned it.[20] The elimination of mutual c-command is a natural consequence of restricting all merge to pair-merge, which is inherently asymmetric.

---

[19] See Safir (2004: 52) where this detail is observed and stipulated.

[20] However, if c-command is calculated from representations, then these relations could be recovered from the values of the nodes in the tree, where each node corresponds to an ordered pair with a label. If the label of the

6.0 Previous just-pair-merge proposals

This essay attempts to show that a theory with just pair-merge is a simpler, adequate and advantageous by comparison to a theory that instead includes TMOH. As remarked in the introduction, the claim that all merge is pair-merge is not novel, though it is not mainstream. My approach owes a debt to previous accounts that have proposed that pair-merge is the only merge operation. However, most of the previous differ from what I have proposed for pair-merge and pair-merge labeling here in one way or another. The differences only matter for what I am proposing if the ways these approaches have formulated pair-merge necessarily conflict with what I have assumed about pair-merge here.

For example, several proposals, including Kayne (1994), Zwart (2005,2009, 2011), Fortuny (2008) and de Belder & van Craenenbroeck (2015) assume the asymmetry introduced by pair-merge is fundamental.[21] However, Kayne, Zwart and Fortuny assume that the determination of linear order is integral to pair-merge. Whether or not this is the right idea, for my purposes, the comparison of theories I have attempted would have to be much more complex because linear-order pair-merge introduces an assumption not included in TMOH. To

---

ordered pair A does not match the label of the node that immediately dominates it, then A is adjoined to its sister B, A c-commands B, and B does not c-command A. Other theoretical assumptions or proposals might turn out to result in differences between a representational approach and a derivational one, but that is beyond the scope of this essay.

[21] With respect to Kayne, this statement is anachronistic. Adjunction was not formulated as pair-merge when Kayne was writing. I am treating Kayne's view that the relation that arises between merged terms is asymmetric amounts to treating merge as an ordered pair forming operation.

try to refute the necessity of including linear order in pair-merge would go well beyond the scope of this essay and I will not attempt it.[22]

De Belder & van Craenenbroek are interested in deriving the grammatical status of roots as label-less and featureless from the ordered pair asymmetry. They argue that, in the initial pair-merge of a functional head with the empty set, the empty set is later replaced by a root. It is less important to them whether this asymmetry in output is part of pair-merge or is a consequence of something else (e.g., see Langendoen, 2003) that happens to structures created by set merge. They nonetheless assume pair-merge because it does not affect their argument. I have handled the categorization of a root differently, on the assumption that a category-less root can only be adjoined, because it cannot contribute a label. However, I have to give a pass to the root with respect to the Labeling Filter. On the other hand, it is not so clear that for the JPMH there is any need for the Labeling Filter, if pair-merge labeling is a reflex of pair-merge (which is not my current assumption) rather than a separate determination. I leave the matter open. De Belder & van Craenenbroeck are partial to the idea, originally from Jaspers (1998), that derivational history creates the asymmetry if each step in a derivation is an addition to something already built (formally realized in Zwart's pair-merge, where single items of a pre-existing numeration are sequentially broken off). Although it is not inconsistent with

---

[22] As Kayne has shown, a theory that introduces linear order by pair-merge requires a great deal of movement to achieve observed linear relations, particularly for the so-called right adjunctions. On the other hand, a free merge theory avoids one of the burdens of the linear approach to pair-merge because it is unnecessary to invent an enormous inventory of movement triggers. The structures built by free merge must, however, be interpretable. Overgeneration is then the challenge that such an approach would face.

JPMH as presented here, there is no necessary assumption that one item is added to an existing

derivation in time, rather than assuming that one syntactic object is merged to another without

regard to how the two syntactic objects came into being. In short, a proviso based on Jaspers'

suggestion is not an assumption that needs to enter into my demonstration JPMH is simpler

than TMOH.

7 Conclusions

As stated at the outset, a commitment to minimalist thinking, especially from the

evolutionary perspective proposed by HCF, certainly favors an approach with only one

structure-building operation rather than two. It would be natural to refute my account by

showing that a theory with only set/simple merge is superior to JPMH and TMOH, and some

have attempted such a simplification.[23] It would serve the interests of minimalist theory to

compare such theories with mine. Here I have developed an approach that rejects set merge in

favor of pair-merge. A more complete demonstration would show that no theory can be

empirically successful unless it includes pair-merge. I believe that case can be made, but to do

so would require many additional assumptions that would have to be balanced with

assumptions in TMOH for the simplicity argument to go through.[24]

---

[23] Oseki (2015) proposes that pair-merge should be eliminated in favor of a revised LA and an additional condition

on labeling.

[24] For example, head movement is often formulated as adjunction of one head to another head or to complex

heads also formed by adjunction. If that is indeed the case, then pair-merge is indispensable. Chomsky (2000) has

suggested the head movement is post-syntactic, partially on the grounds that it does not affect interpretation,

though that leaves open the question whether that operation is pair-merge of set merge. Matushansky (2006)

A syntactic theory with only pair-merge and no simple/set merge is simpler than one including TMOH and much closer to the minimalist ideal. A theory adopting JPMH is not only simpler, but it is also adequate to generate all the syntactic relations that are known to be necessary. I have demonstrated that the purported advantages of TMOH do not provide adequate explanations for the halting problem, the existence of agreement, or the existence of EPP effects. Insofar as my critique is successful, TMOH does not gain any advantage over JPMH with respect to those phenomena. JPMH provides an inherent advantage over TMOH in the formulation of c-command by eliminating the possibility of mutual c-command, since mutual c-command is not a crucial syntactic relation. Moreover, in certain circumstances, as argued in the last section, if mutual c-command is possible then it leads to applications of c-command with unfavorable empirical consequences.

References

Abe, J. 2018. How to probe expletives. *Studia Linguistica* 72: 76-112.

---

does not assume head movement is an adjunction at all, but a version of simple merge to a specifier before morphology merges the head in the specifier position to the head of the phrase. Sifting through these proposals to determine whether there is a natural set merge alternative would require extensive argumentation that the conclusions I have reached so far do not require.

de Belder, Marijke and Jeroen van Craenenbroeck. 2015. How to merge a root. *Linguistic Inquiry* 46:625-655.

Biskup, Petr and Radek Šimik. 2019. Structure of conditionals and (cor)relative clauses: New evidence from locality. In Maggie Baird and Jonathan Pesetsky, editors*. Proceedings of the Forty Ninth Annual Meeeting of the North East Linguistic Society Vol. 1.* Amherst, MA: Graduate Linguistic Student Association, University of Massachusetts*.* 135-144.

Bobaljik, Jonathan. 2002 A-chains at the PF-interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory* 20:197-267.

Boeckx, Cedric. 2010. A tale of two minimalisms: Reflections on the plausibility of crash-proof syntax, and its free-merge alternative. In Michael T. Putnam, editor. *Exploring crash-proof grammars*. Amsterdam, the Netherlands: John Benjamins. 105-123.

Bondevik, Ingrid, Dave Kush and Terje Lohndal. 2021 Variation in adjunct islands: The case of Norwegian. *Nordic Journal of Linguistics* 44:1-32.

Borer, Hagit. 2005. *Structuring sense, Vol. 1: In name only*. Oxford: Oxford University Press.

Bošković, Željko. 2018. On movement out of moved elements, labels and phases. *Linguistic Inquiry* 49:247-282.

Brody, Michael. 2002. On the status of representations and derivations. In Samuel David Epstein and T. Daniel Seely, editors. *Derivation and explanation in the minimalist program*, editors, Cambridge: Blackwell. 19-41

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In Roger Martin, David Michaels, and Juan Uriagereka, editors. *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*.. Cambridge, MA: MIT Press. 89–155.

Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In Adriana Belletti, editor. *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*. New York: Oxford University Press. 104-131.

Chomsky, Noam. 2008. On Phases. In Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, editors. *Foundational Issues in Linguistic Theory*, Cambridge, MA: MIT Press. 133–166.

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

Chomsky, Noam. 2015. Problems of projection: extensions. In Elisa Di Domenico, Cornelia Hamann, and Simona Mattieini , editors. *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Philadelphia and Amsterdam: Benjamins. 1-16.

Chomsky, Noam. 2020. Puzzles about phases. In Ludovico Franco and Paolo Lorusso, editors. *Linguistic Variation: Structure and Interpretation*. Boston, MA/Berlin: De Gruyter Mouton. 163-168.

Epstein, Samuel David. 1999. Unprincipled syntax: The derivation syntactic relations. In Samuel David Epstein and Norbert Hornstein, editors. *Working Minimalism*. Cambridge: MIT Press. 317-345.

Epstein, Samuel David, Hisatsugu Kitahara and T. Daniel Seely. 2014. Labeling by minimal search: Implications for successive-cyclic A-movement and the conception of the postulate "phase". *Linguistic Inquiry 45*: 463-481.

Fortuny, Jordy. 2008. *The Emergence of order in syntax.* Amsterdam: Benjamins.

Goto, Nobu. 2017. How to label *there*-constructions. *Studies in English Literature*: Regional Branches Combined Issue 9: 33-43.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale and Samuel Jay Keyser, editors. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. 111-176.

Hasegawa, Hiroshi. 1996. Adjoin vs. Merge, and the concept of c-command. *English Linguistics* 13:15-39.

Hauser, Marc, Noam Chomsky and W. Tecumseh Fitch. 2002. "The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569-1579.

Hornstein, Norbert. 2000. *Move! A Minimalist Theory of Construal*. Oxford, UK: Blackwell.

Hornstein, Norbert. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge, UK: CUP.

Jaspers, Dany. 1998. Categories and recursion. *Interface* 12: 81-112.

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

Kratzer, Angelika. 1996. Severing the external argument from its verb. In J. Rooryck and J. Zaring, editors. *Phrase structure and the Lexicon.* Kluwer: Dordrecht. 109-137.

Kush, Dave, Terje Lohndal and Jon Sprouse. 2019. On the Island Sensitivity of Topicalization in Norwegian: An experimental investigation. *Language*: 393-420.

Langendoen, Terence. 2003. Merge. In Ander Carnie, Heidi Harley, and MaryAnn Willie, editors. *Formal approaches to function in grammar: In honor of Eloise Jelinek*. Amsterdam: John Benjamins. 307-318.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, editors. *Proceedings of the 21st Annual Penn Linguistics Colloquium*. Pennsylvania

Working Papers in Linguistics 4.2. Philadelphia: University of Pennsylvania, Penn Linguistics Club. 201–225.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37: 69–110.

McFadden, Thomas. 2004 The position of morphological case in the derivation: a study on the syntax-morphology interface. Doctoral thesis. Philadelphia, PA: University of Pennsylvania.

McInnerney, Andrew. To appear. The Argument/Adjunct Distinction Does Not Condition Islandhood of PPs in English. *Linguistic Inquiry*.

Oseki, Yohei. 2015. Eliminating pair-merge. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara, and Iris Chuoying Ouyang, editors. *Proceedings of the 32nd West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla. 303-312.

Richards, Mark. 2009. Internal Pair-Merge: The Missing Mode of Movement. *Catalan Journal of Linguistics* 8: 55-73.

Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In Lisa Lai-Shen Cheng and Norbert Corver, editors. *Wh-movement moving on.* Cambridge: MIT Press. 97-133.

Rubin, Edward. 2003. Determining Pair-Merge. *Linguistic Inquiry* 34: 660-668.

Safir, Ken. 1983. On small clauses as constituents. *Linguistic Inquiry* 14: 730-735.

Safir, Ken. 2004. *The syntax of (in)dependence*. Cambridge, MA: MIT Press.

Safir, Ken. 2008. Coconstrual and narrow syntax. *Syntax* 11: 330-355.

Safir, Ken. 2010. Viable syntax: Rethinking minimalist architecture. *Biolinguistics* 4.1. http://www.biolinguistics.eu/index.php/biolinguistics/issue/view/15/showToc.

Safir, Ken. 2017. Weak crossover. In Martin Everaert and Henk van Riemsdijk, editors. *Blackwell companion to syntax*. Oxford, UK: Blackwell.

https://doi.org/10.1002/9781118358733.wbsyncom090

Safir, Ken. 2019. The A/Ā Distinction as an Epiphenomenon. *Linguistic Inquiry* 50.2: 285-336.

Saito, M. 2016. (A) Case for labeling: labeling in languages without ϕ-feature agreement. *The Linguistic Review* 33:129-175.

Sato, Mai. 2008. Where is *there*? An approach to the expletive construction. Masters thesis. Tohoku Gakuin University, Sendai.

Stepanov, Arthur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4: 94-125.

Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117:1355–1377.

Truswell, Robert. 2011. *Events phrases and questions*. Oxford, UK: Oxford University Press.

Williams, Edwin. 1986. A reassignment of the functions of LF. *Linguistic Inquiry* 17: 265-299

Zwart, C. Jan-Wouter. 2005. Verb second as a function of Merge. In Marcel den Dikken and Christina Tortora, editors. *The function of function words*. Amsterdam: Benjamins. 11-40.

Zwart, C. Jan-Wouter. 2009. Prospects for top-down derivation. *Catalan Journal of Linguistics* 8:161-187.

Zwart, C. Jan-Wouter. 2011. Structure and order: Asymmetric Merge. In Cedric Boeckx, editor. *Oxford handbook of linguistic minimalism.* Oxford, UK: Oxford University Press. 96-118.

Zyman, Erik. 2017. P'urhepecha hyperraising to object: An argument for purely altruistic movement. In Patrick Farrell, editor. *Proceedings of the Linguistic Society of America, Vol. 2*, 53: 1-15. https://journals.linguisticsociety.org/proceedings/index.php/PLSA/issue/view/159