**If all merge is pair-merge**

**Ken Safir – Rutgers University**

Current versions of minimalist theorizing assume two structure-building operations, simple merge and pair-merge and a means of providing a syntactic label to the output of simple merge. Simple merge takes two syntactic objects (SOs) and includes them in the same set to form a single SO. The labeling algorithm (LA) of Chomsky (2013) assigns the label of a head to the SO formed by simple merge with a non-head. When two non-terminals are simple-merged, there is no determinate label possible, unless the two SOs simple-merged share features that can constitute a label. Pair-merge forms ordered pairs of SOs, with the asymmetry that the first member of the pair SO1 is adjoined to the second SO2 and the label of SO2 becomes the label of the single SO that is formed. In addition, the theory requires a filter that excludes the interpretation of structures that lack a label. I propose that if we assume just a simplified version of pair-merge, both simple merge and LA are unnecessary. If there is only pair-merge, then there is no need for a separate labeling algorithm such as Chomsky’s LA because every output of merge will bear the label of the second member of the ordered pair. I will demonstrate that the simpler only-pair-merge (OPM) theory permits all necessary syntactic constructions and labeling permitted by the two merges plus LA (TMPLA) theory. A descriptively adequate theory with only pair-merge is superior to the TMPLA theory by Occam’s Razor. However, there are also properties of the OPM theory that represent an advance over TMPLA beyond offering a simpler account.

The claim that all merge is pair merge, directly producing an ordered pair output, is explicitly proposed by Kayne (1994) and more recently, in the context of more recent minimalist, assumptions, by Zwart (2009, 2011) and de Belder and van Craenenbroeck (2015). Zwart, like Kayne (1994), sees a consequence for linear ordering based on the ordered pair asymmetry, which is not assumed here. Some of the differences between these theories and mine (discussed towards the end of the squib) are not the focus of this article. Rather the main point raised here is that current syntactic theory has two algorithms, LA and the labeling of pair merge; I show that LA is unnecessary if all merge is pair merge, a position fundamentally consonant with Zwart’s view that all merge is pair merge.

An important point to keep in mind in this discussion is that I am not comparing the relative simplicity of pair-merge vs. set merge. In isolation, set merge is simpler. However, set merge is not viable without LA, and so it is not at all clear that set merge plus LA is simpler than pair-merge. Moreover, the argument is not based on the simplicity of pair-merge vs. set merge plus LA, but rather that no theory should have all three and that the best reduction is OPM.

It is natural to simplify the theory so as to avoid stating two merge operations. An alternative strategy would be to eliminate pair-merge in favor of simple merge plus LA, if nothing more than LA is needed (see Oseki, 2015). For this reason, we must question what the motivation for pair-merge is. Chomsky (2004, 2008) suggests that pair-merge can offer an account of anti-reconstruction and of restrictions on extraction from adjuncts on the assumption that pair-merged (adjoined) elements are on a different plane, invisible to further manipulation (requiring an otherwise unnecessary simplification operation, SIMPL, to reintegrate adjuncts into interpretation). There are now other accounts of anti-reconstruction that do not rely on the different-plane assumption and that are better integrated into other aspects of minimalism (see, e.g., Sauerland, 1998, Takahashi and Hulsey, 2009 and Safir, 2019). For the latter accounts, however, pair-merge labeling for adjuncts as opposed to LA labeling for complements remains an underlying assumption. That assumption still motivates the need for pair merge (while pair-merge can subsume complementation labeling, as will be shown). Moreover, as Oseki shows, the restrictions on extraction from adjuncts is not as absolute as Chomsky’s different-plane theory would predict. However, the other feature of pair-merge that concerns us here, the one most closely associated with adjunction historically, is that two non-terminal SOs can be joined by pair-merge and that their label is determined by the operation. Chomsky (2004:117-118), where he introduces pair-merge, begins with the pronouncement, stated as an observation that “it is an empirical fact that there is also an asymmetric operation of adjunction”. This still appears to be necessary or adjuncts could not easily be introduced into structure at all – they do not require agreement to be licensed. It is this aspect of pair-merge that I propose is general to all structure-building operations.[[1]](#footnote-1)

To illustrate the need for pair-merge, what I will call, purely for orthographic reasons, “hyphenated adjuncts” which are internally phrasal, but are introduced as modifiers.

1a) Don’t give me that what-am-I-supposed-to-do look.

b) I just love his off-to-the-races enthusiasm.

c) A please-don’t-bother-me, disdainful, I-am-too-far-above-you attitude will not serve you well in our workplace.

d) Your too-long-in-the-oven crumpled black croutons are a disgrace to our restaurant.

e) The newly painted, please-don’t-call-them-taupe, now sound-proof, dividing walls are

sure to appeal to our new clientele.

These hyphenated-adjuncts act just as adjectives do, not as compounds, and they are easily accommodated by structure-building using pair-merge indifferent to the category of the first member of the pair, but there is no obvious way in which sentences and PPs are agreeing with the head noun. Whatever their category, hyphenated adjuncts combine with a nominal to yield a modified nominal (or when adjoined to an adjective, a modified adjective, e.g., *For his new sportscar he chose a welcome-to-my-midlife-crisis* *red*).[[2]](#footnote-2)

What is needed to capture such structure building is an operation, pair-merge, which takes two syntactic objects in an ordered pair <Xw, Yz> and forms [z Xw Yz]. We describe this as “X is pair-merged to Y” such that the label of Y (its subscript, for presentation) becomes the label of the SO that is the output of pair-merge. Historically, the two merge operations have their roots in the contrast between substitution transformations and adjunctions. LA has its historical roots in X’-theory, where adjunction was an accretion on the labeling established by X’ endocentricity, repeating both the label and the bar level of the adjoined-to constituent. I am assuming that linear ordering of SOs is not established by pair-merge, but is rather determined independent of structure-building (narrow syntax).[[3]](#footnote-3)

If only pair-merge builds structure, this predicts four cases of external pair-merge for syntactic objects that are either terminals or non-terminals.

2a) A non-terminal X is pair-merged to a terminal Y and the label of Y projects.

b) A terminal X is pair-merged to a non-terminal Y and the label of Y projects.

c) A non-terminal X is pair-merged to another non-terminal Y and Y projects.

d) A terminal X is pair-merged to another terminal Y and Y projects.

(2a) forms complementation structures. This is identical to the output of set merge, especially after labeling if Langendoen (2003) is right, and so in a theory with both set merge and pair-merge, eliminating set merge reduces redundancy in the theory. The need for LA to address this case disappears. (2b) forms what many have assumed for adverb, pronominal clitic or quantifier particle structures, that is, where a nominal head adjoins to a non-terminal vP or TP or an adjective adjoins to an NP. Set merge and LA gives the wrong result for such cases (e.g., the A would contribute the label for set-merged [A NP]), so pair-merge is independently necessary. The addition of adjectives in this way does not require that adjective modifiers be terminals or non-terminals, though other considerations may determine that they must be adjectival. With respect to (2c), there is no indeterminacy in the OPM theory since whatever is adjoined to projects its label. This means that [XP YP] formed by pair merge <XP, YP> will be a YP and [XP YP] will be an XP if it is formed by pair merge <YP, XP>. Thus there is no crisis of indeterminacy leading to movement, so this prediction of LA is lost. However, we will see that the LA is not necessary motivate movement. Once again, pair-merge is independently necessary for adjuncts of the form [YP XP YP] to be formed, as in many accounts of relative clauses or adjunction of non-terminals to verb phrases. Finally, with respect to (2d), a head X can be adjoined to a head Y before these heads are merged to any non-terminal. Both set merge and pair-merge permit heads to merge with heads, but pair-merge does not need LA to label the outcome. If both heads have a label, LA requires one of the heads to move (a bad prediction, it would appear). If two roots do not have any category, pair-merge will not project any label, but if every non-terminal must have a label (a version of the labeling filter), any category-less head will have to be adjoined to something that projects a category label.

As Richards (2009) points out, if internal set merge is an automatic consequence of set merge, then internal pair-merge should be an automatic consequence of pair-merge. This predicts three or four cases of internal pair-merge, assuming Extension (Chomsky, 1995: 190), whereby each instance of merge applies to the root node.[[4]](#footnote-4)

3a) A non-terminal Y is internally pair-merged to a terminal X and X projects.

b) A terminal Y is internally pair-merged to a non-terminal X and X projects.

c) A non-terminal Y is internally pair-merged to another non-terminal X and X projects.

d) A terminal Y is internally pair-merged to another terminal X and X projects.

(3a) is not really possible, but insofar as it is, it is identical to (2a) because Extension is only satisfied if the structure created is bigger at its apex than the two syntactic objects combined. Internal merge of a non-terminal to a subconstituent terminal within XP would not extend XP.

(3b) is essentially a clitic structure, where an N or D-head originating in the extended VP adjoins to VP, for example. (3c) is instantiated by structures where a QP originating in a thematic position adjoins to IP to establish scope. (3d) is head-to-head movement, insofar as the theory of Extension allows it (see fn.4).

If both pair-merge and simple merge are in the theory, the labeling indeterminacy theory is undercut. Consider that if intermediate A’-movement, e.g., adjunction of a wh-phrase to vP, is pair-merge, the reason for further movement to satisfy the TMPLA labeling filter (conceived of as a precondition for interpretation) is lost. In a theory that allows pair-merge, the structure will have a label.

4a) [vP …[wh-XP]]

b) [vP [*wh-XP*] [vP …*wh-XP*]] - Formed by pair-merge <[wh-XP, vP]>

Thus if a wh-phrase adjoins to vP by pair-merge as in (5), then the resulting structure has a label, namely, vP. Since this is always possible, the claim that the LA motivates movement from intermediate landing sites has always been illusory, at least as long as pair-merge is in the theory.

Thus a different reason is needed as to why the intermediate stopping point required by the Phase Impenetrability Principle is not where the wh-movement simply stops raising. This is the free merge approach, where movement must end up where interpretation succeeds (see, e.g., Chomsky, 2013, and Safir 2010, 2019). If wh-XP must find the right configuration in which to be interpreted, namely in configuration with an interrogative C, then it will have to move for that reason. Except for the subject position (Spec,TP) this eliminates the need to appeal to LA failures to motivate evacuation from intermediate positions in movement derivations.

The only remaining motivation for the LA is that it purportedly provides a principled motivation for the existence of agreement. Chomsky posits that agreement exists to resolve labeling paradoxes. This must assume that there is independent reason to believe that labeling indeterminacy is needed. If all Merge is pair-merge, then there is no such motivation. Introducing the labeling indeterminacy to explain why agreement occurs is not then an explanatory move. All of the non-subject “agreement” relations are not about phi-agreement principally, but about what is descriptively[[5]](#footnote-5) Spec-head matching for semantic licensing, e.g., a topicalized constituent is the “Spec” of TOP, focus is so related to FOC, interrogative related to C+Q, possessive related to Spec,DP, etc. The only agreement that is not semantically motivated in this sense (and that purportedly resolves a labeling paradox) is agreement of what is descriptively Spec,TP and T. This is supposed to explain why subjects can reside in Spec TP.

Since T and its complement is a non-terminal [T vP] and the TP subject is a non-terminal, for subjects to remain in place only agreement can justify a label for the union of these two constituents.

This account faces, in part, the same objection that the TMPLA account of why movement cannot stop moving before reaching a destination where it is licensed. As pointed out, pair-merge could be applied instead of set merge for intermediate movement, eliminating any label indeterminacy. The same could be said for movement to subject position, that is, once pair-merge adjoins XP to T to yield a T label for [T T XP], then pair-merge adjoining the subject to [T T XP] will yield [T Subject [T T XP]] without any label indeterminacy.

It is also dubious that LA plus simple merge succeeds at deriving the labeling paradox resolved by agreement. Clearly infinitives and gerunds lack phi-feature agreement with their subjects, but both require the establishment of subject-like arrangements. Both can have *there* subjects, which is usually the mark of an EPP effect. Yet agreement does not capture 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), especially as agreement with T does not involve the common features of *there* and T. For such cases, some agreement relation must also license the low subject position, but once again there is no overt morphological motivation for such a relation. It is always possible to create additional abstract agreement to render a theory symmetrical, and sometimes such idealizations lead to insight, but in this case, positing agreement relations that license positions that can appear on the surface serves no other function – it is a stipulation to save a theory.[[6]](#footnote-6)

A reviewer points out that the OPM theory will have to stipulate EPP rather than derive it, and this is true. This does not disadvantage the OMP with respect to the TEMPLA theory if the latter does not derive EPP either. The last paragraph gives some reasons why it doesn’t, in the case of *there* subjects in infinitives and gerunds, but there are additional reasons why the TEMPLA theory fails to derive EPP. For example, there are numerous instances where subjects are not nominals, but must be present. These cases, in addition to the non-agreeing nominals in infinitives (and in small clauses), show that there is no argument for agreement under LA as an account for EPP effects.

5a) Under the bed seems to be a good place to hide.

b) Is under the bed a good place to hide?

c) Under the beds 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. The same article shows small clauses must act as subjects (e.g. *Why shouldn’t workers ready to bargain be the goal of our tactics?*) The same questions can now be asked about gerunds and infinitives as subjects – how do we know they are nominal? They also raise and at least gerunds invert (*does talking to Bill seem (to be) difficult?*). Gerund and infinitive subjects have no overt phi-features that agree with T (at least in English). With less reason to believe that Case is necessary for a nominal (see, e.g., Mcfadden, 2004, among many others), there is no reason to assume that gerunds and infinitives are nominal subjects (at least in English). LA was designed to account for why subjects must agree, but when they are not nominal and have no apparent features, it is hard to see how LA makes an accurate prediction without a slew of ancillary assumptions. In short, LA does not account for the EPP as it stands, so OPM loses nothing in comparison with TEMPLA by failing to derive it.

As mentioned earlier, it is indeed necessary to distinguish DP specifiers from DP adjuncts, but this is easily handled by linking rules based on first merge, i.e., the first phrasal merge to X creates [X X Complement] and the first phrase merged to [X X Complement] is the Spec,X as others have proposed. Any subsequent adjunctions to [Spec,X [X X Complement]] will not be thematically assigned. Thus there is no reason to see the addition of the subject to structure by pair-merge as any more complex than the addition of a subject by set merge.

Moreover, LA does not explain the existence of phi-agreement because it does not account for most forms of agreement. It does not account for antecedent agreement, which may plausibly have a different source, but it also does not account for adjunct agreement, which is more constructional. The latter is an interesting case, since labels for adjunctions are provided by pair merge even in Chomsky’s system. In languages where all adjuncts to DP must agree with the DP head, e.g., for noun class in Bantu, the relations that are created by pair-merge should not have any need for agreement relations to introduce a label, yet agreement in Bantu is robust in such constructions. Example (6) is from Zulu, originally from Doke (1955:5), as cited in Baker, (2008:147) with the structure posited by Baker in (7).

6) Thin’ aba-kyulu si-ya-kuluma

We c1.1.pl-big 1ps-DISJ-be.speaking

‘We big ones are speaking’

7) [DP [DP we[1pl]k] [FAP FA [pl, \*1]k [AP big]]]

In this example, the functional adjective phrase (FAP) which takes AP as its complement is adjoined to the 1st person plural pronoun. Person agreement does not appear on adjectives, which is what Baker and Doke are illustrating (i.e., if ‘big’, *khulu*, had the 1st plural prefix *si-* it would be ungrammatical), but it also shows that adjuncts can agree for basic phi-features even in the absence of posited indeterminacy in the TMPLA theory.[[7]](#footnote-7)

Preserving simple merge alongside independently motivated pair-merge introduces a whole set of alternative derivations that can result in the same structures that pair-merge generates, as we have seen for intermediate movement, but notice also that certain kinds of relations will have to be defined twice. This is particularly true of the derivational definition of c-command, first introduced by Epstein (1999: 329).

8) 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.[[8]](#footnote-8)

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 reformulated adjunction as pair-merge), or for movement that results in adjunctions (which Hasegawa also noted and set aside). These complications reduce to the following definition once simple merge is eliminated and, following Chomsky (2004), movement is reduced to internal merge (internal pair-merge in the OPM theory).

9) C-command: If X is adjoined 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.[[9]](#footnote-9)

Furthermore, there is something to gain. 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.[[10]](#footnote-10) If that is true, then this theory has an explanation for it, namely, sisters created by pair-merge are, by definition, never symmetrical (see also Zwart, 2011). Is there ever an empirical reason to assume, for example, that a predicate must c-command something inside what it is predicated of? From the perspective of pair-merge, only the adjoiner can c-command something in the adjoinee. Consider, for example, the following two structures:

10a) [[*every man’s* mother] [praises *his* new spouse]]

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

In both structures, a quantifier is embedded in a constituent that c-commands a bound pronoun and a bound pronoun is embedded in a constituent that c-commands its quantifier. We can now capture this weak crossover contrast by saying:

11) A pronoun (or what dominates it) cannot c-command the quantifier that binds it.[[11]](#footnote-11)

In a theory that allows [*praises x’s spouse/mother*] to c-command the contents of the subject, the two sentences have the same violation, unless it is stipulated that the pronoun cannot be contained in a nominal that c-commands the quantifier that binds it. In the OPM theory (and in the other ordered pair output theories), the subject is adjoined to [ T XP], so it asymmetrically c-commands the contents [T XP] without further stipulation limiting c-command to nominals.[[12]](#footnote-12)

If we adopt pair-merge instead of set merge, then the binarity of syntactic trees (first proposed by Kayne, 1984) follows directly from how they are combined. Set merge is not inherently a binary operation, but rather other constraints or stipulations are required to insure binarity. Set merge would not be a different sort of operation if three syntactic objects were merged in a single operation. By contrast, positing an ordered triple, rather than an ordered double, introduces at least four potential asymmetries rather than one, that is, for the objects X, Y and Z in the ordered triple <X, Y, Z>, we have X>Y, Y>Z, X>Z, and of course X>Y>Z. The one way in which the ordered pair in pair-merge is asymmetric is with respect to label, but in an ordered triple, many asymmetric relations can be expressed in addition to labeling. However, no such relations are needed (e.g., if X, by virtue of participating in triplet-merge, could presumably have multiple, perhaps ordered, labels from both Y and Z). The point is that the difference between set merge of either two or three syntactic objects is not different with respect to the kinds of relations it implies, whereas any change of pair merge to n-tuple merge would be a rather drastic change implying a much richer set of relations. Thus pair-merge, unlike set merge, directly implements binarity, but does so in the most minimal way, introducing only one asymmetry (e.g., instead of \*{a, b, c} as a filter or restriction on the output of set merge)..

As mentioned, the view that all merge is pair merge in the sense closest to what is meant here is pioneered by Zwart (2009, 2011) and is picked up in bottom-up derivation theory by de Belder and van Craenenbroeck (2015) (henceforth, B&C). I agree with these authors that the output of every merge operation is an ordered pair, although I differ from these accounts as to how this comes about. Zwart proposes a top down derivation where the order of operations intrinsically creates the ordered pair asymmetry. Assuming the existence of a numeration (these authors use the term “resource” for “numeration”), which is a preselected set of syntactic objects that initiates the derivation, one member of the set, x, is split from the numeration N such that <x, N-x> is the output. Subsequent operations remove single members of the set N leaving the remainder of N as the right sister constituent to the one that is split from the set until the set is exhausted.[[13]](#footnote-13) Numerations are thus crucial to Zwart’s formulation as it is to that of B&C (see B&C’s definition of Unary Merge, p. 636). However, as proposed here, pair-merge makes no reference to numerations There is reason to believe that numerations (or any notion of a preselected set of items to be used in a derivation before it is built) should not be in the theory either. An alternative to numeration indices is proposed in the appendix to Safir (2019), and as a result, he argues, the need for initiating derivations with randomly selected numerations disappears. If Safir is right, then the Zwart and B&C versions of pair-merge are undefinable.

B&C are interested in deriving the grammatical status of roots as label-less and featureless from the ordered pair asymmetry (i.e., in the initial pair merge of a functional head with the empty set, where 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 (but they nonetheless assume pair-merge because it does not affect their argument). They are partial to the idea, originally from Jaspers (1998) that derivational history creates the asymmetry if each move 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 OPM as presented here, there is no necessary assumption that one item is added to an existing derivation, rather than assuming that one SO is merged to another ahistorically.

An additional assumption of both Zwart (2011) and B&C is that phrases constructed in a separate derivation (in another “workspace”) can be treated as members of numerations so that phrases, like the hyphenated adjuncts, can be treated like words. Insofar as OPM as proposed here cannot refer to (non-existent) numerations, the OPM account of hyphenated adjuncts depends on the asymmetry intrinsic to the pair-merge operation.

The key point, however, is that set merge plus LA are redundant if we also assume pair-merge. It remains to be seen whether an alternative approach to reducing the redundant power of pair-merge and simple merge can be achieved by eliminating pair-merge plus LA, but if such an approach is to be successful, it must not introduce auxiliary assumptions devoted only to permitting simple merge to produce adjunction outcomes. OPM as presented here diminishes pair-merge to its historical function of generating labeled adjoined structures. It eliminates set merge and redundant power to generate structures identical to pair merge. It eliminates the LA and the need for spurious agreement features (e.g., in infinitives with or without PRO) and semantically and phonetically null heads (as in Bruening’s account of hyphenated adjuncts) posited solely to justify it. In addition, OPM simplifies the derivational definition of c-command, and captures the binarity of syntactic structure in a minimal way.

**References**

Baker, Mark. 2008. *The syntax of agreement and concord.* Cambridge: Cambridge University Press.

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

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

Bruening, Benjamin. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language* 94:1-42.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge: MIT Press.

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

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

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

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

Doke, Clement. 1955. *Zulu syntax and idiom*. Cape Town: Longmans Southern Africa

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

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.

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

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

Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

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

Kayne, Richard S. 2011. Why Are There No Directionality Parameters? *WCCFL* 28: 1-23.

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

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. Ph.D dissertation, University of Pennsylvania.

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

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

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: MIT Press

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

Safir, Ken. 2019. The A-/A’-distinction as an epiphenomenon. *Linguistic Inquiry* 50.

Sauerland, Uli. 1998. The meaning of chains. Doctoral dissertation. MIT, Cambridge, MA.

Takahashi, Shoichi, and Sarah Hulsey. 2009. Wholesale late merger: Beyond the A/A¯ distinction. *Linguistic Inquiry* 40:387–426.

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 *Ocxford handbook of linguistic minimalism,* ed. Cedric Boeckx, 96-118. Oxford: Oxford University Press.

1. Oseki proposes that pair-merge should be eliminated in favor of a revised LA and an additional condition on labeling. Rubin (2003) introduces a head to make modifiers complements that must be introduced by pair-merge, but no proposal made here will stipulate how a constituent can merge based on its internal structure. [↑](#footnote-ref-1)
2. A reviewer suggests that set merge could merge a hyphenated adjunct with a null head that gives the whole phrase the category that it fills the position of (e.g., N if it acts like a nominal) as suggested by Bruening (2018,6 fn.6) . Any features this head needs to agree with what it is adjoined to could also be stipulated. This is possible, but not explanatory. Any number of null heads could be added to the adjunct this way (e.g., [A[N[A[hyphenated adjunct]]]]) and presumably some sort of semantic filter would rule out the vacuity. But if so, why have any semantically vacuous null head added at all? In the OPM theory, it follows that the category of the adjoined hyphenated adjunct is irrelevant when the hyphenated adjunct is introduced as an adjunct, because only the second member of the ordered pair provides the label. The adjunct could pair merge with a null head before the result of that pair-merge is pair-merged to the constituent providing the label, but this would not change anything, and would presumably be disfavored or banned by the same sort of semantic vacuity condition needed to avoid recursive generation of semantically vacuous heads in the TEMPLA account. Thus the OPM account directly derives the absence of any theory-accommodating semantically null head. [↑](#footnote-ref-2)
3. Here I depart crucially from Kayne (2011) and Zwart (2009, 2011) who also have a pair-merge only theories, but they include linear ordering as part of it along the lines of Kayne (1994). Right adjunctions, as, for example, in relative clauses without reconstruction, appear to require otherwise unmotivated movement to get the right word order within Kayne’s approach. I will not discuss his theory here, but I briefly return to Zwart’s theory at the close of the squib. [↑](#footnote-ref-3)
4. Other assumptions about how Extension is formulated will potentially have different consequences, but this is not a matter that needs to be entered into here. Head-to-head internal merge is not permitted by a strict interpretation of Extension and some have argued that such structures must immediately be reconfigured (e.g., Matushansky, 2006) while others have reformulated Extension to allow for more structural possibilities, including head movement (e.g., Safir, 2019) [↑](#footnote-ref-4)
5. Descriptively, “Spec,XP” is the first constituent merged to X and its sister. [↑](#footnote-ref-5)
6. The same reviewer notes that this is not evidence that pair-merge is necessary if there is a way to save the set merge theory without it. However, this argument can be turned the other way; What evidence is there that set merge is necessary to generate structures that pair-merge does not generate? This would be an argument for less descriptive power for set merge if LA did not have to come with it. As mentioned at the outset, it is hard to see why a theory with set merge and LA is simpler, especially with all the spurious agreement features, heads and relations that TEMPLA requires. However, the point here is that TEMPLA with pair-merge as well is both redundant and more descriptive than OPM. [↑](#footnote-ref-6)
7. This could be taken a step further to say that adjunction preserves phi-features as well as categorial ones, which is why adjuncts to nominals agree if there is morphology in the grammar to express it. The absence of person agreement on adjectives would have to be accounted for separately, and Baker has a proposal. [↑](#footnote-ref-7)
8. This is Epstein’s preliminary definition, but the differences between this and his final version do not concern us here. A reviewer points out that this approach answers Brody’s (2002: 28) objection to derivational c-command based on the absence of asymmetry in the operation of structure-building. [↑](#footnote-ref-8)
9. Of course, if (7) turns out not to be the way (what comes to be treated as) c-command should be defined, nothing is lost, unless it is shown that c-command cannot be formulated in a theory with only pair-merge. [↑](#footnote-ref-9)
10. See Kayne (1994) for an explicitly asymmetric theory of c-command that is in part the model for this one, but see fn.2. [↑](#footnote-ref-10)
11. Safir (2004) treats this as a correlate of his Independence Principle. [↑](#footnote-ref-11)
12. See Safir (2004: 52) where this detail is observed and stipulated. [↑](#footnote-ref-12)
13. As mentioned in fn.3, I am not assuming anything about linear order that might follow from the asymmetry of ordered pairs, but this is a major concern for Zwart, which I do not address. [↑](#footnote-ref-13)