Samuel D. Epstein, Hisatsugu Kitahara* and Daniel Seely

Phase cancellation by external pair-merge of heads

DOI 10.1515/tlr-2015-0015

Abstract: As Chomsky (2004, 2005) notes, a theory with set-Merge allows this operation to apply in two different ways, externally (to two separate objects) and internally (one object contained within the other). Here we extend Chomsky's form of argument to pair-Merge; i. e. in the absence of some stipulation preventing it, it too can apply in two ways: internally and externally. We will argue that external pair-Merge of heads overcomes a paradox concerning bridge verb constructions. In the final section we note that external pair-Merge of heads is, in effect, a "presyntactic" morphological ("word formation") rule entailed by current syntactic theory. The extent to which the standard theory of morphological operations can be subsumed by external pair-Merge of heads, further unifying syntax and (aspects of) morphology is left for further research.

Keywords: merge, pair-merge, minimalism, phase, copies

1 Introduction

Adopting the general framework outlined in Chomsky (2013), Chomsky (2015) discusses a long-standing problem regarding unvalued phi-features (uPhi) on bridge verbs (such as *think*).[1] Consider the following structure for the matrix v*P

1 The problem associated with bridge verbs has taken a number of different forms in the course of the development of the theory of generative grammar. For GB theory, for instance, the question was whether embedded CPs required (and could receive) Case from the bridge verb (if the verb in fact had Case to assign). Thus, in *I believe the report* vs. *I believe the report is flawed*, does (or can) *believe* assign Acc Case to both the DP and the CP object? See, among others, Chomsky's (1986) visibility analysis, whereby all arguments need Case in order to receive a theta role. But see Epstein (1990) and Davis (1984) for arguments against the visibility analysis. See also Safir (1985) for relevant discussion.

***Corresponding author: Hisatsugu Kitahara,** Keio University, Tokyo, Japan,
E-mail: kitahara@icl.keio.ac.jp
Samuel D. Epstein, University of Michigan, Ann Arbor, MI 48109, USA
Daniel Seely, Eastern Michigan University, Ypsilanti, MI 48197, USA

phase of *John thinks that he will win* (where R is the root THINK, v* is a phase head and a verbal categorizer, and EA is the external argument):

(1) [EA [v* [$_{\alpha}$ R [$_{\beta}$ C...]]]] *(John thinks that he will win)*

In (1), R inherits uPhi from v*, but the search domain of R, namely β, contains no element that agrees with R (and moves to Spec-R).[2] Under minimal search, the only element syntactically accessible in β, given the phase-impenetrability condition (PIC), is C.[3] Thus, two problems with (1) emerge: First, there is a valuation failure; uPhi of R never gets valued, which induces formal crashing at the interfaces. Second, there is label failure for α, see Chomsky (2015); α is not identified, and is thus predicted to be uninterpretable, at the interfaces.[4] It would appear that we have a widespread undergeneration problem arising in all cases of bridge verb constructions ("V taking a CP complement" descriptively). Although Chomsky (2015) reveals the problem and outlines a solution to it, we suggest below that it is potentially problematic and propose an alternative that is entirely consistent with leading tenets of the Chomsky (2015) framework.

Our analysis appeals to no new mechanism beyond freely applied Merge that includes both set-Merge and pair-Merge (Chomsky 2004, 2005). The analysis we propose here also resolves a paradox that arises between Chomsky's (2015) analysis of simple transitive sentences and of bridge verb constructions. Furthermore, it resolves another long-standing problem concerning the phase-

2 See below for comparison with a typical DP object structure like *John read the book*.

3 If there were a DP in the Spec of the embedded CP, then in principle, R could agree with such an element on the edge. For analyses of this phenomenon, not addressed here, see among others, Kayne (1989) for French, Van Urk and Richards (2015) for Dinka, Carstens (2005) for Kilega, and Obata and Epstein (2012a, 2012b) for English. Also, as a reviewer points out, in some languages there is agreement with and shifting of a CP complement, as in McCloskey's (1991) example: *That he'll resign and that he'll stay in office seem at this point equally possible*. As has been noted, the agreement in such cases seems to us weak in the sense that *seems* appearing in this context is perhaps acceptable, and clearly better than e. g. *They seems happy*.

4 Recall that for Chomsky (2015) R in English is 'weak' and hence can't serve as a label on its own. Regarding label failure, as Chomsky (2013: 43) notes, "For a syntactic object SO to be interpreted, some information is necessary about it: what kind of object is it? Labeling is the process of providing that information." Note this does raise an interesting question, one we will not pursue here, regarding the exact relation between "(non)labeled" and "(un)interpretable." For example, **(I think that) will Bill read the book* where *Bill* has not raised from spec of v*P involves label failure and yet it is not entirely clear in what sense it is "uninterpretable." For our purposes here, it is sufficient to adopt Chomsky's hypothesis that "object identification" (i. e. "labeling") is required for interpretation at the interfaces (leaving the precise status of "interpretability" open). See Epstein (2007), and Epstein et al. (2010) for discussion of different kinds of (un)interpretability.

head status of little v* and eliminates the stipulated "strong vs. weak" distinction regarding the notion "phase." The paper is organized as follows. Section 2 reviews how Chomsky (2015) analyzes simple transitive sentences like *John likes the dog* and bridge-verb constructions like *John thinks that he will win*. We then point out a paradoxical situation with respect to the visibility of copies left by pair-Merge. Section 3 resolves this paradox, presenting a new analysis of *John thinks that he will win*, while keeping intact Chomsky's (2015) analysis of *John likes the dog*. Section 4 extends the proposed analysis to Icelandic dative subject constructions, and resolves the long-standing problem concerning the phase-head status of little v* and eliminates the "strong vs. weak" distinction regarding the notion "phase." Section 5 notes that the recognition of external pair-Merge of heads – allowing our solution to the specific bridge verb problem noted above – as entailed by the current theory, would appear to clear the way for significant unification of (aspects) of "presyntactic" morphology ("word formation") and syntax along the lines originally suggested in an earlier framework by Marantz (1997) and much subsequent work.

2 Background theoretical and analytical assumptions

One recent development in minimalist theory (Chomsky 2013, 2015) is that Merge, formulated in the simplest form, applies freely as long as it conforms to third factor principles such as the condition of inclusiveness: "no new objects are added in the course of computation apart from arrangements of lexical properties" (Chomsky 1995: 228), and the no-tampering condition: "Merge of X and Y leaves the two SOs unchanged" (Chomsky 2008: 138). Merge, by hypothesis, is no longer operating only "in order to" create a configuration that allows interface illegitimate features to be checked; rather, Merge optionally applies, and so crashing happen.[5] Chomsky (2015) assumes that "operations can be free, with the outcome evaluated at the phase level for transfer and interpretation at the interfaces."

As for the formulation of Merge, there are two kinds: (i) set-Merge that takes X and Y, and forms the two-membered set {X, Y}, and (ii) pair-Merge that takes X and Y, and forms the ordered pair <X, Y>. Informally, as Chomsky (2004, 2005) notes, these are descendants of substitution and adjunction in earlier theories.

5 That crashing happens is contra, for example, Frampton and Guttman (2002); see also Putnam (2010) for discussion of crashing vs. "crash proof" grammar.

We propose that, unless stipulated, set-Merge and pair-Merge apply optionally and if applied are free to apply in any order, and we take this free rule-ordering in the optional application of Merge (including set-Merge and pair-Merge) to be the null hypothesis.

Given these assumptions, let us first review how Chomsky (2015) analyzes simple transitive sentences. Consider the following structure for the v*P phase of *John likes the dog* (where R is the root LIKE, and *t* represents a full copy of the internal argument IA):

(2) [EA [<R, v*> [$_{\alpha}$ IA [R *t*]]]] *(John likes the dog)*

In (2), (i) set-Merge externally forms {R, IA}, (ii) set-Merge internally merges IA to Spec-R,[6] (iii) set-Merge externally introduces v* and then EA (cyclically) into the derivation, yielding the v*P phase, (iv) R inherits uPhi from v,* (v) R agrees with IA,[7] valuing Case, (vi) α is labeled <ϕ, ϕ> under minimal search,[8] (vii) pair-Merge internally forms <R, v*> (= R with v* affixed),[9] (viii) v* becomes invisible (and thus is no longer the phase-head), (ix) the phase-head status is activated on the copy of R, and (x) the complement of R, namely *t*, gets transferred. Importantly, notice that *in order to function as the "derived" phase-head, the copy of R in α, left "behind" by pair-Merge, must be visible to minimal search.*

6 We assume cyclic rule application of IM moving IA from R-Complement to Spec-R, as does Chomsky (2015). A reviewer points out that such movement is inconsistent with antilocality analyses (Bošković 1994, 2015; Abels 2013; Grohmann 2003; Saito and Murasugi 1999; among many others). In this paper, we will not adopt the antilocality constraint, and continue to assume (with Chomsky [2015]) that Merge applies freely as long as it conforms to third factor principles such as the condition of inclusiveness and the no-tampering condition.

7 Note that although *t*, the full copy of IA, does not count as being in the domain of the set {R, *t*} (since not all occurrences of IA are within that set), both IA and *t* are in the search domain of v*, which transmits its features to R. If feature-transmission is taken to be part of Agree, then the search domain of Agree may be extended to the set {v*, {IA, {R, *t*}}}. In this paper, we assume that R agrees with IA, valuing IA, while leaving a precise implementation of Agree open.

8 This is an instance of the agreeing shared-prominent-feature option of the labeling algorithm of Chomsky (2013). Thus, in {XP, YP}, XP, YP non-heads, minimal search finds both heads X and Y equally. This "ambiguous" search result can be resolved, according to Chomsky (2013), only if minimal search can find at least one agreeing shared feature borne by both X and Y. Intuitively, if, say, phi of Y values phi of X, then there is a single feature, namely, phi, that is found by minimal search, even though minimal search finds X, Y; there is still a single disambiguating feature, or feature-set, namely, phi, that counts as the object identifier (serving as the interface-relevant information "I am a phi object"). See also Bošković (2015) for discussion of additional motivation for labeling within a different set of assumptions.

9 For earlier analyses in which the mover projects in its landing site, see Epstein (1998) and Donati (2006).

To be "visible" is to be within a relevant domain. Chomsky (2013) assumes γ to be in domain D iff every occurrence of γ is a term of D. Suppose an occurrence of γ is a sister-category merged to γ by set-Merge. Then, there is only one occurrence of R in (2), namely, the sister-category β (i. e. the IA merged to R by set-Merge). Given this interpretation of visibility, minimal search will identify the copy of R in α, left by internal pair-Merge of R to v*, as the only visible R head in the representation.[10]

Let us now compare how Chomsky (2015) analyzes bridge-verb constructions. Consider the following structure for the matrix v*P phase of *John thinks that he will win*:

(3) [EA [<R, v*> [$_{\alpha}$ R [$_{\beta}$ C...]]]] *(John thinks that he will win)*

In (3), (i) set-Merge externally forms {R, β}, (ii) set-Merge externally introduces v* and then EA (cyclically) into the derivation, yielding the v*P phase, (iii) pair-Merge internally forms <R, v*> (= R with v* affixed), (iv) v* becomes invisible (and thus is no longer the phase-head), but here, unlike the analysis of (2), the copy of R in α, left by internal pair-Merge, is assumed to be invisible. This assumption is seemingly necessary so as to allow α to be labeled. Given that there is no DP shifted to Spec-R, a <phi, phi> label (interpretation) for α, under phi agreement, is impossible (and recall R is too weak to serve as a label by itself). With the copy of R invisible, labeling by β (as C) is possible. Although it does not consider all the details, this is Chomsky's (2015) proposed solution to the bridge verb problem; the copy of R in α is invisible and that solves the label failure problem. In addition, since v* is rendered invisible by internal pair-Merge of R to v*, this solves the problem that there is uPhi on v*, which finds no DP to probe and agree with.[11] *Thus, for α to be labeled by β, the copy of R in α, left by internal pair-Merge, must be invisible to minimal search*. Here's the paradox:

10 Note that we do not consider the nature of 'occurrence' with respect to pair-Merge, and interesting questions arise. For example, if there is only one occurrence of R in (2), as claimed, then what is the element represented as "R" in <R, v*>. The little v* is not an occurrence of R (since <R, v*> was not created by set-Merge). So, what is it? Note further that if there is only one occurrence of R in (2), then the following situation arises: Derivationally R was internally Merged to v*, but representationally there is only one occurrence of R in the resulting structure; so this is movement (derivationally) but non-movement (representationally).

11 A reviewer raises the issue of uPhi inheritance from v* to R. If R bears uPhi after inheritance, then the invisibility of v* with respect to these unvalued phi features would be irrelevant; uPhi of R in <R, v*> would remain visible and unvalued, and that should induce crashing at the interface. One possible way out of this problem is that uPhi does not get transmitted to R (feature inheritance is optional). See also Nomura (2015) for relevant discussion.

within Chomsky's (2015) analysis, it is (implicitly) presumed that R left by internal pair-Merge both is and is not visible, specifically it is visible in (2), crucially to allow the copy of R to serve as the "derived" phase head, and it is invisible in (3) crucially to avoid label failure.[12]

How can we resolve this apparent paradox? In the next section, we suggest that pair-Merge of R to v* invariably leaves a visible copy of R. Thus, we keep Chomsky's (2015) analysis of (2) intact and provide an alternative analysis of bridge verb constructions like (3).

3 A solution with no new assumptions

To resolve this paradox, let us begin with an overview of the four types of rule application that have been recognized in minimalist literature; namely, external and internal set-Merge, and external and internal pair-Merge. Set merge of X, Y is widely recognized to have two forms of application: (i) external (X and Y are separate) and (ii) internal (one of X, Y is contained within the other). To the best of our knowledge, pair-Merge was first introduced by Chomsky (1995),[13] and developed in some detail by Chomsky (2004), to represent adjunction. Although it was not referred to as such, it is external pair-Merge of phrases or phrase-to-head that is employed to create adjunction structures in Chomsky (2004). Richards (2009) recognizes external and internal set-Merge, and external pair-Merge, and was perhaps the first to note that *internal* pair-Merge is a logical possibility, and then goes on to explore that possibility. And internal pair-Merge of heads plays an important role in Chomsky (2015). Also note that external set-Merge of heads has been employed in Saito (2012, 2013, 2014). What we would like to stress here is that in the absence of a stipulation forbidding it, there is a new type of rule application, namely, external pair-Merge of heads,[14] and all

12 One might argue that there is no paradox; rather, pair-Merge of R to v* has an option of leaving a copy visible or invisible to minimal search. Leaving aside the unclarity of how this option is to be implemented, the system faces (at least) a serious overgeneration problem. This option in effect allows IA to stay in situ as in e.g. **I believe to be someone in the room*. Notice that R (BELIEVE) inherits uPhi from v*, R (BELIEVE) and IA undergo Agree, and pair-Merge of R (BELIEVE) to v* leaves a copy of R (BELIEVE) invisible to minimal search. Thus, the system would overgenerate this derivation. If the option is eliminated, the paradox remains.

13 See Chomsky (1995), Ch. 4, Section 4.3, and in particular, p. 248.

14 Although he does not say pair-Merge, Riqueros (2013) uses external Merge of two heads at X^0-level for deverbal nouns, arguing that deverbal nouns involve external Merge of a nominal head and a verbal head. We thank a reviewer for pointing out his work to us. See footnote 24 below for further comments.

these types of rule application are in fact entailed by the theory, as it now stands.[15] As we will now argue, the paradox noted above (the R copy both is and is not "visible") dissolves by recognizing and applying external pair-Merge of lexical items.[16] As noted, this is a form of "presyntactic"[17] morphology ("word formation") entailed by current theory, suggesting the possibility of significant unification of aspects of morphology and syntax.

---

15 In this paper, our focus is on external pair-Merge of two heads. Note that this is not the only unattested logical option (though see footnote 14 above), for note that if there are two operations (set-Merge and pair-Merge) each applying in two ways (externally and internally) and if there are three types of syntactic objects on which Merge can operate, namely, heads, sets and ordered pairs, then there are some 32 logically possible types of rule application. To the best of our knowledge the following (logically possible) cases have not be attested in the literature:

(i) External set-Merge {ordered pair, ordered pair}; i. e. {<x, y>, <w, z>}
(ii) External set-Merge {ordered pair, set}; i. e. {<x, y>, {w, z}}, though see below.
(iii) External set-Merge {head, ordered pair}; i. e. {H, <w, z>}
(iv) Internal set-Merge of α, β where one of (or both) α, β is an ordered pair; e. g. {<x, y>, {<x, y>, {w, z}}

And there are various options yet unattested for pair-Merge; for example pair-Merge of α, β where one of (or both) α, β is an ordered pair. Other unexplored logical possibilities (e. g. internal pair-Merge as in: {X, Y} => <X, {X, Y}> await future research. The important issue is to first identify the types of rule applications allowed by the theory as formulated and then to determine the empirical domain, or principled exclusion, of each.

16 To be clear, we assume that there are two structure-building operations: set-Merge and pair-Merge. Each has two instantiations, called "external" and "internal." As noted in important work by Richards (2009), external pair-Merge (though not referred to as such), is proposed in Chomsky (2004) for phrasal adjuncts. Our focus here is on external pair-Merge of heads.

17 By "presyntactic" we mean what in earlier frameworks is the morphological component that builds words and hands these words over to the syntax as atomic units. The (tentative) suggestion here is that external pair-Merge of heads can subsume aspects of this word-building work and since (external) pair-Merge (of heads) is a syntactic operation, the divide between morphology and syntax is potentially significantly reduced (perhaps eliminable). See Marantz (1997) for important discussion. As Marantz p. 205 states:

> *To imagine a theory in which the grammar constructs all words in the syntax by the same general mechanisms ("merge and move"; see Chomsky 1995) that construct phrases, it is useful to make the natural assumption that whether you get a "zero-level category" (word-like unit) or a phrasal category by merging two constituents is a function of the (categories of the) constituents involved, not of the "merger" operation itself. That is, there is no reason not to build words in the syntax via "merger" (simple binary combination) as long as there are no special principles of composition that separate the combining of words into phrases from the combining of morphemes into words.*

We concomitantly appeal to the null hypothesis that these types of rule application are freely ordered; that is, freely applied set-Merge and pair-Merge, regardless of whether external or internal, can apply in any order, with only certain choices converging.

As an illustration of this type of rule application (namely, external pair-Merge of heads), let us now examine how repeated Merge can generate a derivation for *John thinks that he will win*. Imagine that the CP *that he will win* has been built. Now, suppose R (THINK) and v* are each taken directly from the lexicon and externally pair-Merged as <R, v*> (as opposed to R's external set-Merger to CP, as in Chomsky [2015]). After this step of the derivation, suppose the newly created ordered pair <R, v*> is then externally set-Merged with CP. Under this ordering of independently motivated operations (an ordering allowed by free rule-ordering) then the phase-head status of v* is cancelled prior to set-Merge of <R, v*> to CP because pair-Merge of R to v*, recall, makes v* (including its uPhi) invisible.[18, 19] In this derivation, notice that there is no "raising" (i. e. internal pair-Merge) of R to v*. Consequently, there is no copy of R in the representation and hence it no longer needs to be stipulated that the copy of R is invisible in (3) but visible in (2), that is, the paradox is averted.

Notice with the theory's recognition of external pair-Merge of heads, there are now two ways to generate a representation of an ordered pair such as <R, v*>: in one derivation, R undergoes *internal* pair-Merge to v*, leaving a copy of R (this is the derivation appearing in Chomsky [2015]); in the alternative derivation proposed here, R undergoes *external* pair-Merge to v*, generating no copy of R.

Having illustrated the crucial first steps, we propose that the following derivation is the only convergent one assigned to *John thinks that he will win*. Consider the following structure for the matrix v*P phase of *John thinks that he will win*:

(4) [EA [$_{\alpha}$ <R, v*> [$_{\beta}$ C...]]] *(John thinks that he will win)*

In generating (4), (i) first, pair-Merge of heads externally forms <R, v*> (= R with v* affixed), and v* becomes invisible (and thus is no longer the phase-head),

18 Nomura (2015) independently points out that pair-Merge cancels the phase-head status of v* when v* selects a bridge verb, and develops a different derivational analysis in which internal pair-Merge of R to v* takes place prior to inheritance of features from v* to R.

19 Though further research is needed, we assume the selectional requirement between R and CP, and the theta relation between v* and EA are met in {<R, v*>, CP} and {EA, {<R, v*>, CP}}, respectively. This also raises the question: if v* is canceled (and "invisible") right from the start of the derivation, then is it required at all? Interestingly, it is required. As pointed out above, although R itself can't label on its own, the "amalgam" <R, v*> can label. Thus, if v* were never merged in, label failure would result. It remains to be determined in a principled way just why it is that <R, v*> can label while R itself cannot.

(ii) set-Merge then externally forms {<R, v*>, β}, and (iii) set-Merge introduces EA (cyclically) into the derivation. In (4), there is no visible phase-head, provided that external pair-Merge forming <R, v*> made v* (including its uPhi) invisible. This application of external pair-Merge of R and v* cancels the phase-head status of now invisible v*; hence, Transfer does not apply, and the derivation continues. Assuming with Chomsky (2015) that "although R cannot label, the amalgam [R-v*] can," it follows that neither labeling of α nor valuation of uPhi poses a problem, because the amalgam [R-v*] can label α, and v* that includes uPhi is invisible.[20]

There is no need to stipulate any rule-ordering of set-Merge or pair-Merge. Phase-cancellation by external pair-Merge of heads is possible only when there is no need to transmit uPhi to the head of the phase-head-complement for Case-valuation. So, for example, if external pair-Merge of heads cancels the phase-head status of v* in transitive sentences like (2), then Case on IA remains unvalued, causing the derivation to crash. In effect, given freely applied Merge, a 'natural ordering' emerges. External pair-Merge producing <R, v*> *before* set-Merge of the object DP results in crashing (the Case of DP fails to be checked); the only order that produces a licit result is external set-Merge forming {R, Obj} and, later, internal pair-Merge (of R) producing {EA, {<R, v*>, {Obj, {R, Obj}}. Only with this ordering (one of many allowed under free ordering) will the direct object Case get valued. However, with bridge verbs, as we have seen, only external pair-Merge forming <R, v*> *before* set-Merge of CP and this ordered pair, will produce a fully convergent, interface interpretable syntactic object. We continue to assume that freely applied Merge including set-Merge and pair-Merge, regardless of whether external or internal, can apply in either order, with only certain choices converging.

4 Eliminating the "strong vs. weak" distinction regarding the notion "phase"

The proposed analysis sheds new light on another long-standing problem concerning the phase-head status of little v*. In English (5), T agrees with NP_{SUBJ}, assigning Nominative Case (NOM) to it, and T cannot agree with NP_{OBJ} inside the phase-head-complement VP, given PIC. By contrast, in Icelandic (6), NP_{SUBJ} bears Dative Case (DAT), and T can agree with NP_{OBJ} inside the phase-head-complement

20 Note that this raises the interesting situation where the uPhi of v*, if transferred to one or the other interface, does not induce crash. See Epstein et al. (2010) for relevant discussion; we leave the matter open here.

VP but only if NP_{SUBJ} moves to Spec-T, rendering its in-situ copy invisible (an example discussed in Jónsson [1996], also cited in Bobaljik [2008]):[21]

(5) a. **John like these socks*
John.NOM like.PL these socks.PL.ACC
b. *... T.PL... [$_{v*P}$ v* [$_{VP}$ V NP.PL.ACC]]

(6) a. *Jóni líkuðu þessir sokkar*
Jon.DAT like.PL these socks.PL.NOM
b. ... T.PL... [$_{v*P}$ v* [$_{VP}$ V NP.PL.NOM]]

To allow T-NP_{OBJ} agreement in Icelandic (6), Chomsky (2001, 2004), contra Chomsky (2000), delayed the timing of Transfer under PIC until the merger of the next "higher" phase-head C. It was then crucially assumed that T inherently bears uPhi in its lexical representation. Thus, T could agree with NP_{OBJ}, exactly when T was merged with VP, and hence before the subsequent merger of C, which immediately triggers Transfer of the next "lower" phase-head-complement VP. Chomsky (2007, 2008), however, then proposed that T inherits uPhi from C. This feature-inheritance analysis creates a new problem for the analysis of T-NP_{OBJ} agreement: Before the merger of C to TP, T (lacking uPhi) cannot possibly agree with NP_{OBJ}, but immediately after the merger of C to TP, Transfer of VP is triggered. Consequently, even with Transfer of VP delayed until immediately after external merger of C, T cannot agree with NP_{OBJ}, because Transfer has already "removed" the phase-head-complement VP containing NP_{OBJ} upon external merger of C, hence before T has the opportunity to inherit uPhi from C.

Under external pair-Merge of heads as employed here, such T-NP_{OBJ} agreement is possible only when R (LIKE) is taken from the lexicon and externally pair-Merged with v*. This application of external pair-Merge forms <R, v*> and cancels the phase-head status of v* (prior to set-Merge of <R, v*> to IA). This is because external pair-Merge of R to v* makes v* (including its uPhi) invisible. Recall that phase-cancellation by pair-Merge of heads is possible only when

21 We leave aside how NOM gets realized on NP_{OBJ} in (6), since NOM can appear on NP_{OBJ} even when an intervening NP clearly blocks agreement between T and NP_{OBJ}, as in (i) (an example discussed in Schütze [1997], also cited in Bobaljik [2008]; see Nomura [2005] for discussion of nominative-assignment in cases such as (i)):

(i) *Mér virðist [Jóni vera taldir t líka hestarnir]*
Me.DAT seemed.SG Jon.DAT be believed.PL like horses.PL.NOM
'I perceive Jon to be believed to like horses'

there is no need to transmit uPhi to the head of the phase-head-complement for purposes of subsequent Case-valuation. This option is available for Icelandic (6) with dative subject, but not for English (5) with nominative subject. In Icelandic (6), external pair-Merge of R and v* cancels the phase-head status of v*, Transfer does not apply, and the derivation continues, leaving the door open for subsequent T-NP_{OBJ} agreement. In English (5), R agrees with NP_{OBJ}, valuing its Case, so the phase-head status of v* is active, and Transfer "removes" the phase-head-complement VP, thereby correctly blocking T-NP_{OBJ} agreement.

The proposed analysis also explains why T-NP_{OBJ} agreement becomes possible with passive/raising verbs, as in the following example discussed in Zaenen et al. (1985), also cited in Bobaljik (2008):

(7) a. *Um venturinn voru konunginum gefnar ambáttir*
In the winter were.PL the king.DAT given slaves.PL.NOM
'In the winter, the king was given (female) slaves.'
b. ... T.PL... [$_{v*P}$ v* [$_{VP}$ $V_{passive}$ NP.PL.NOM]]

Legate (2002) argued that a verbal phrase headed by a passive/raising verb is phasal. Chomsky (2001) stipulated that these are "weak" phases, which, unlike his original "strong" phases, elude Transfer application. Under the current assumptions, we can eliminate this *ad hoc* "strong vs. weak" distinction; i. e. there is just one v* in the lexicon and it is "strong," i. e. a phase head. In the event that this v* undergoes external pair-Merge creating <R, v*>, then the so-called "weakness" of v* follows from v* (in <R, v*>) being invisible. In (7), phase-cancellation by pair-Merge of heads is allowed because in passive/raising, there is no need to transmit uPhi to the head of the phase-head-complement for Case-valuation.[22] Thus, Transfer does not apply, and the derivation continues, leaving the door open for subsequent T-NP_{OBJ} agreement. As shown above, the effects of "weak phase" follow without postulating the notion "weak phase."

The proposed analysis thus predicts that phase-cancellation by external pair-Merge of heads takes place in verbal phrases with passive, raising, unaccusative and bridge verbs, but not in verbal phrases with transitive (taking a direct object) and intransitive (unergative) verbs if intransitive (unergative) verbs are hidden transitives in the sense of Hale and Keyser (1993) (see also Chomsky [1995]).[23]

22 We leave open whether this analysis might extend in part or in whole to middles (*John broke the vase* vs. *The vase broke*).

23 Both reviewers point out to us that the present proposal sheds new light on the wide-range of empirical issues discussed in Bošković (2014), and possibly derives the notion "phase collapsing" postulated there. We would like to return to these issues in future research.

5 Further implications

We argue in this paper that the theory in fact entails a new type of rule application, namely external pair-Merge of heads. This type of rule application was invoked here to resolve an apparent paradox in Chomsky (2015), namely that some R raising (internal pair-Merge of R to v*) crucially left no visible copy while other cases of R raising crucially *do*. To resolve the paradox, we generated the cases where we 'don't want' a copy of R (bridge verbs) by ("presyntactic," morphological) external pair-Merge (of R to v*) – a kind of word formation entailed by the existence of pair-Merge and of lexical access – which by definition creates no copy.

Importantly, current syntactic theory then in fact incorporates the operation external pair-Merge of heads. Although external pair-Merge can operate on non-heads (creating phrasal adjunction, see Chomsky 2004), we have seen here that there is nothing to prevent external pair-Merge of two lexical items. In the latter case, this amounts to a morphological operation in the sense that it creates an "X^o-level" amalgam "presyntactically," i. e. unlike the X^o amalgams created by internal pair-Merge, i. e. syntactic internal pair-Merge of e. g. R to v*. Since neither of the two inputs nor the output of external pair-Merge of heads is a set, the operation can be said to be "morphological." What we mean by this can best be illustrated with an example.

Consider passive. In the standard analysis, there is a pre- and non-syntactic operation within the morphology that takes as its input, say, the active verb *arrest* and the passive morpheme pronounced *–ed/–en* and amalgamates the two creating the derived lexical passive item "arrested," which by hypothesis fails to value accusative Case on its internal argument. Under the current analysis, the same general effects are captured by the recognition of external pair-Merge of heads. Thus, R (ARREST) is externally pair-Merged with v*, which lexically bears uPhi, creating the ordered pair <R, v*>. In this resulting object, v* is invisible, hence uPhi on v* are invisible as well, and the passive for "arrested" is derived (unable to value accusative Case due to the invisibility of uPhi).

As Chomsky notes, a theory with set-Merge allows it to apply in two different ways, externally (to two separate objects) and internally (one object contained within the other). The recognition of these two forms constitutes Chomsky's normalization of displacement phenomena as nothing more than (internal) Merge, a form of Merge application allowed unless there is a stipulation preventing it. Concomitantly, Chomsky has eliminated the "PS-rule/deep structure vs. Transformational rule/S-structure" dichotomy present since the inception of generative grammar. Epstein et al. (2015) note the importance of this unification,

including the normalization of displacement; the unification cannot be stressed enough and is in our view underappreciated.

Correspondingly, we have extended Chomsky's form of argument to pair-Merge. In the absence of some stipulation preventing it, it can apply in two ways: internally and externally. External pair-Merge of lexical items would appear to subsume a certain class of morphological operations, including phase-cancellation as proposed here, and passive verb formation as briefly reviewed here. The extent to which external pair-Merge can capture such standard 'morphological' phenomena awaits further research.[24]

Acknowledgement: For helpful discussion of many of the ideas presented here, we thank Stefanie Bode, Noam Chomsky, Ángel Gallego, Masayuki Ike-uchi, Nobuaki Nishioka, Vitor Nobrega, Masashi Nomura, Masayuki Oishi, Mamoru Saito, two anonymous reviewers, the audience of the 150th meeting of the Linguistic Society of Japan, held at Daito Bunka University on June 20, 2015, and the participants of the intensive course on minimalist syntax, held at Kyushu University on June 30, 2015–July 3, 2015.

References

Abel, Klaus. 2003. *Successive cyclicity, antilocality, and adposition stranding*. Storrs, CT: University of Connecticut dissertation.

Bobaljik, Jonathan David. 2008. Where's Phi? agreement as a postsyntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*, 295–328. Oxford: Oxford University Press.

24 For example, see Nobrega (2015) for an analysis of NV compounding formulated in terms of external pair-Merge of heads. We thank Noam Chomsky for pointing out Nobrega's work to us and we also thank Vitor Nobrega for sharing and discussing his research with us, and for pointing out to us a second relevant example of external merge of heads, in this case External Merge of R to v*, a possibility suggested in Marantz 2013: 99 for conditioned allomorphs:

> "One possibility to explore in this context is that roots are always ("externally") merged to their category heads and never undergo head movement as part of word formation (although a given root might syntactically merge with different category heads). Other possibilities would involve deeper development of the theory of head movement and syntactic word formation."

We are however unsure if Marantz (2013) intended external SET Merge of heads or, (as we propose here and is proposed for different reasons in Nobrega (2015)) external PAIR Merge of heads.

See also Piggot and Travis (2013), who argue for the external adjunction of heads. To the extent that pair-Merge corresponds to adjunction, our proposal here is in the spirit of Piggot and Travis.

Bošković, Željko. 1994. D-structure, theta-criterion, and movement into theta positions. *Linguistic Analysis* 24. 247–286.
Bošković, Željko. 2014. From the Complex NP Constraint to everything: On deep extractions across categories. Ms., University of Connecticut, Storrs.
Bošković, Željko. 2015. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. Ms., University of Connecticut, Storrs.
Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language & Linguistic Theory* 23. 219–279.
Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, vol. 3, 104–131. Oxford: Oxford University Press.
Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22.
Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Mart Gärtner (eds.), *Interfaces + Recursion = Language?*, 1–29. New York: Mouton de Gruyter.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*, 136–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann, Simona Matteini (eds.), *Structures, strategies and beyond – studies in honour of Adriana Belletti*, 3–16. Amsterdam & Philadelphia: John Benjamins.
Davis, Lori J. 1984. *Arguments and expletives: Thematic and non-thematic noun phrases*. Storrs, CT: University of Connecticut dissertation.
Donati, Caterina. 2006. On *wh*-head movement. In Lisa Lai-Schen Cheng & Norbert Corver (eds.), *Wh-movement: Moving on*, 21–46. Cambridge, MA: MIT Press.
Epstein, Samuel David. 1990. Differentiation and reduction in syntactic theory: A case study. *Natural Language & Linguistic Theory* 8. 313–323.
Epstein, Samuel David. 1998. Overt scope marking and covert verb-second. *Linguistic Inquiry* 29. 181–227.
Epstein, Samuel David. 2007. On i(nternalist)-functional explanation in minimalism. *Linguistic Analysis* 33. 20–53.
Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2010. Uninterpretable features: What are they and what do they do? In Michael T. Putnam (ed.), *Exploring crash proof grammars*, 125–142. Philadelphia: John Benjamins.
Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely (2015). From *Aspects* 'daughterless mothers' (aka delta nodes) to POP's 'motherless'-sets (aka non-projection): a selective history of the evolution of Simplest Merge. To appear in Angel Gallego & Dennis Ott (eds.) 50 years later: Reflections on Chomsky's Aspects, 99–112. MITWPL *Aspects* volume.

Frampton, John & Sam Gutmann. 2002. Crash-proof syntax. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 90–105. Oxford: Blackwell.

Grohmann, Kleanthes. 2003. *Prolific domains: On the antilocality of movement dependencies*. Amsterdam: John Benjamins.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Jónsson, Johannes Gisli. 1996. *Clausal architecture and case in icelandic*. Amherst, MA: University of Massachusetts dissertation.

Kayne, Richard. 1989. Facets of romance past participle agreement. In Paolo Benincà (ed.), *Dialect variation and the theory of grammar*, 85–103. Dordrecht: Foris.

Legate, Julie Anne. 2002. Some interface properties of the phase. *Linguistic Inquiry* 34. 506–516.

McCloskey, James. 1991. *There, it,* and agreement. *Linguistic Inquiry* 22. 563–567.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). Article 14.

Marantz, Alec. 2013. Locality domains for contextual allomorphy across the interfaces. In Ora Matushansky & Alec Marantz (eds.), *Distributed morphology today*. Cambridge: The MIT Press.

Nobrega, Vitor. 2015. On Merge and Word Formation: the Case of Compounds. U. of Sao Paulo (USP) handout.

Nomura, Masashi. 2005. *Nominative case and Agree(ment)*. Doctoral dissertation, Storrs: University of Connecticut.

Nomura, Masashi. 2015. On pair-Merge. Unpublished manuscript, Chukyo University.

Obata, Miki & Samuel David Epstein. 2012a. Feature-splitting internal merge: the case of *tough*-constructions. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of Structure Building*, 366–384. Oxford: Oxford University Press.

Obata, Miki & Samuel David Epstein. 2012b. Feature-splitting internal merge: improper movement, intervention, and the A/A' distinction. *Syntax* 14. 122–147.

Piggot, Glyne & Lisa Demena Travis. 2013. Adjuncts within heads and complex words. In Raffaella Folli, Christina Sevdali & Robert Truswell (eds.), *Syntax and its limits*, 157–174. Oxford Studies in Theoretical Linguistics. Oxford University Press.

Putnam, Michael T. 2010. *Exploring crash proof grammars*. Philadelphia: John Benjamins.

Richards, Marc. 2009. Internal pair-merge: The missing mode of movement. *Catalan Journal of Linguistics* 8. 55–73.

Riqueros, José. 2013. *Spanish nominal(ization) patterns*. Doctoral dissertation. Storrs:University of Connecticut.

Schütze, Carson Theodore. 1997. *INFL in child and adult language: Agreement, case, and licensing*. Cambridge, MA: MIT dissertation.

Safir, Kenneth. 1985. *Syntactic chains*. Cambridge: Cambridge University Press.

Saito, Mamoru. 2012. Case checking/valuation in Japanese: Move, agree or merge? *Nanzan Lingusitics* 8. 109–127.

Saito, Mamoru. 2013. Conditions on Japanese phrase structure: From morphology to pragmatics. *Nanzan Lingusitics* 9. 119–145.

Saito, Mamoru. 2014. Case and labeling in a language without φ-feature agreement. In Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds.) *On peripheries: Exploring clause initial and clause final positions*, 269–297. Tokyo: Hitsuzi Syobo Publishing.

Saito, Mamoru & Keiko Murasugi. 1999. Subject predication within IP and DP. In Keith Johnson & Ian Roberts (eds.), *Beyond principles and parameters*, 167–188. Dordrecht: Kluwer.

van Urk, Coppe & Norvin Richards. 2015. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46. 113–155.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language & Linguistic Theory* 3. 441–483.