MR. YUSHI SUGIMOTO (Orcid ID : 0000-0002-7980-5770)

Article type : Original Article

LH: Yushi Sugimoto

RH: *Labeling defective T*

LABELING DEFECTIVE T*

* This paper is a revised version of a paper that I presented at GLOW39 (Göttingen) and CoSOLE24 (York). For valuable comments and discussion, I am indebted to Andreas Blümel, Noaki Fukui, Nobu Goto, Takaomi Kato, Alan Hezao Ke, Andrew McInnerney, Terje Lohndal, Hiroki Narita, as well as the audience at the two conferences (GLOW39 and ConSOLE24). I'm also very grateful to two anonymous reviewers for comments and suggestions. I also thank Acrisio Pires, Daniel Seely and Hisatsugu Kitahara for sharing their ideas and giving me very helpful suggestions to improve this paper. Finally, this paper is dedicated to Samuel Epstein and Roger Martin, who both unfortunately passed away last year. Without their encouragement, this paper would not be published here.

Yushi Sugimoto

Abstract. This paper proposes a new analysis of so-called 'defective T' (a postulate originally discussed in Chomsky 2000, 2001) and seeks a label-based theoretical explanation of phenomena associated with the distribution of subjects in infinitival clauses. I argue that we do not need to postulate 'defective T' as an element of the lexicon. Following Chomsky's (2013, 2015) labeling approach, which investigates only finite clauses, I argue here that the subject of infinitive-*to* is pair-merged structure <T, C> and the associated predicate forms an {XP, YP} structure, which is unlabelable. To attain labeling, the subject position ([spec, TP]) in the infinitival clause moves out and the label of the resulting structure is the amalgam <T, C>, which has no visible Spec (violating EPP). As a consequence, we eliminate the Merge-over-Move principle.

1. Introduction

Chomsky (2013, 2015) provides a new explanation for the EPP effects. Chomsky (2013, 2015) reduces such an effect to a labeling problem, assuming that T cannot be a candidate for a label of syntactic objects by itself in the case of subject raising. Although the satisfaction of the EPP in subject finite clause is addressed, non-finite clauses are not analyzed in Chomsky (2015).

A critical component of Chomsky's analysis concerns Phase Theory. Chomsky (2015) proposes that the phasehood of C is inherited by T after C is deleted at the phase level, while the phasehood of *v**is inherited by R after the head raising. To extend this analysis, we can generalize it in the following way; the phase head status can be inherited by non-phase heads along with uninterpretable features (such as φ-features). As a result, non-phase heads can be 'derived phase heads' (cf. Epstein, Kitahara and Seely 2016, henceforth EKS 2016), which are activated when non-phase heads are pair-merged with phase heads.[1] EKS (2016) show that bridge verbs are an externally pair-merged amalgams (i.e., <R,*v**>) to solve a lack of valuation problem (see section 3). To extend EKS's (2016) proposal, this paper investigates the possible application of external pair-Merge of T to C (<T, C>), rendering C invisible, and its consequences, which provides an alternative analysis of defective T in non-finite clauses, improper movement and the EPP requirement, in terms of the labeling and phase theories developed in the framework of Chomsky (2015).

This paper is organized as follows. Section 2 and Section 3 provide background and specify the framework adopted in this paper. Section 4 puts forward the proposal that

[1] In this paper, we will discuss pair-Merge of heads (EKS 2016), but not pair-Merge of phrases (cf. Chomsky 2004, Richards 2009).

external pair-Merge of heads T and C is an automatic, unnoticed consequence of free unconstrained (set and) pair Merge; Section 4 also explores the consequences of external pair-Merge of heads, including: overcoming a {XP, YP} labeling problem in infinitival clauses, and eliminating a generalized EPP in terms of labeling. Section 5 concludes.

2. Subject position under labeling theory

Assuming the framework of Chomsky (2013, 2015), I will discuss infinitival clauses in terms of labeling. The example in (1) is a case of Exceptional Case Marking (ECM) in English.

(1) a. They expected John to win.

b. $\{\text{they}, \{_{\gamma} \langle \text{R}, v^{*} \rangle, \{_{\alpha=\langle \varphi, \varphi \rangle}$[2] $\text{DP}_{i}, \{\text{R}(=\text{expect}), \{_{\beta}\ \text{DP}_{i}\ (=\text{John}) \ldots\}\}\}\}\}$

Under Chomsky's (2015) analysis involving "raising-to-object" (cf. Postal 1976, Lasnik

[2] I use this notation just for expository purposes. This notation indicates that the label of α is $\langle \varphi, \varphi \rangle$.

and Saito 1992), the derivation of the matrix *v** (“expect”) phase in (1) proceeds as follows: (i) Form {R, β} by External (set-)Merge (EM)[3], (ii) Internal (set-)Merge (IM) of DP in forming α, (iii) Merge *v**, reaching the phase level, (iv) Feature Inheritance from *v** to R, (v) Labeling of α as <φ, φ>, (vi) R raises to *v** forming R with *v** affixed (via pair-Merge[4]), i.e., <R, *v**>, hence *v** is invisible and so phasehood is activated on the lower copy of R[5]), (vii) Transfer of β. In this derivation, R is a ‘derived’ phase head, since phasehood is activated on R when *v** becomes invisible as a result pf pair-Merge. Further, Labeling theory plays an important role. Assuming Simplest Merge (following Chomsky

[3] Whether a root takes its complement or not is controversial in the recent literature. See Alexiadou and Lohndal (2017) for relevant discussions. Chomsky (2015) assumes here that a root can take a complement though the root moves (via head movement) to the categorizer to get categorized.

[4] Chomsky (2015) assumes that a mover becomes the host of the structure rather than affixed to the upper head, which means that R, in the case of (1), becomes the host of the structure and *v** is affixed to the root. See also Epstein (1998) for an analysis in which V raises to AGR, and the moving V ‘projects’ and renders AGR invisible.

[5] Chomsky (2015: 12) argues that “[a]ccordingly, raising of R to *v** yields an amalgam with *v** adjoined to R, and the affix is invisible to the labeling algorithm. Note that although R cannot label, the amalgam [R, v*] can.” However, the following point may be controversial: when a root merges with a categorizer, it seems that the categorizer becomes the label of the syntactic object. One reviewer points out that there is another possibility; the root can be the labeler, since it is visible according to the reviewer. In this paper, however, I adopt Chomsky’s (2015) idea that the amalgam becomes the labeler. If the root can be a label and roots do not include categories (Borer 2013, Marantz 1997), roots seem to include, at least, concepts, but roots do not show whether they are predicates or nominal, which should be indicated in the narrow syntax, if we adopt Borer (2013) and Marantz’s (1997) approach. As for pair-Merge of T and C see the following discussion in the paper. See also Alexiadou and Lohndal (2017).

(2013, 2015): Merge (α, β) = {α, β}), there are no projections or syntactic labels. However, it is assumed in Chomsky (1995, 2013, 2015) that labels are necessary to interpretation at the interfaces. Minimal search finds the label of a syntactic object. A simple Labeling Algorithm (LA) applies in a top-down fashion and returns the most prominent (lexical) item it finds. Chomsky (2013) proposes that this LA yields descriptively different outcomes when applied to three distinct structural configurations; (i) {H,XP} cases, (ii) {XP, YP} cases without agreement, and (iii) {XP, YP} cases with agreement.

The trivial case is {H, XP}, where minimal search unambiguously finds H as the label. A problematic case is {XP, YP}. When minimal search finds heads in an {XP, YP} structure, it finds both the head of XP and the head of YP simultaneously, and thus the label is not determinable as a unique head.[6] One solution is that one of the syntactic objects moves out of this structure. It is assumed that the copies of syntactic objects are invisible for minimal search except for the highest copy (Chomsky 2013). Then, for example, if XP moves out, then minimal search on {XP, YP} does not find XP, and thus Y becomes the uniquely identified label.. As for the second solution to the ambiguous {XP, YP} structure, consider a case of subject raising as in (2).

[6] As for an alternative approach to labeling, see Saito (2016) and Blümel (2017).

(2) a. John likes Mary

b. $\{C, \{_{\alpha}\ John_{i\ [j\varphi]}, \{_{\beta}\ T_{[u\varphi]}, John_{i}, \ldots$[7]

If the subject is raised to [spec, TP], the structure is $\{_{\alpha}$ DP, {T…}}, an unlabelable {XP, YP} structure. However, when X and Y share features (e.g., φ-features), these shared features become the unique label. Thus, the subject DP has interpretable φ-features and T has uninterpretable φ-features. So, the label of α becomes <φ, φ>. As for the label β, Chomsky (2015) stipulates that T in English is "too weak to serve as a label", whereas in rich agreement language such as Italian T can serve as a label. However, once it agrees with the subject in a spec-head configuration, T is strengthened according to Chomsky (2015).

In Chomsky (2015), derivations are totally cyclic and DP can IM at any point since Merge applies "freely." The only interpretable structure for interfaces is the case of subject raising to [spec, TP] in connection with labeling, an instance of minimal search. Thus EPP

---

[7] I use shading for expository reasons. It indicates that the element has been transferred.

is subsumed by labeling theory.

Chomsky's proposed derivation of EPP is promising, but it leaves a number of potential issues unexplored. One such case is the labeling of infinitival clauses. The example is ECM construction in English, as we discussed at the beginning of this section.

(3) They expected John to win

(=1b) $\{_{\gamma}$ <R, *v**>, $\{_{\alpha=<\varphi, \varphi>}$ DP_i, {R(=expect), $\{_{\beta}$ DP_i (=John)…}}}}

Chomsky (2015) argues that ϕ -less T is strengthened by agreement following inheritance of phi-features from C. But what happens with non-finite T, i.e., *to* in English? Finite TP is labeled as <φ, φ> by sharing φ-features in finite clauses, but this is presumably not possible in infinitives lacking such agreement. Thus, as things stand, {XP, YP} infinitives are predicted to be unlabeled and crash. In addition, we seem to predict that there is no way to label {XP, YP} infinitives. We saw in the example in (3)(=(1b)) that R is unlabeled if the object does not raise to [spec, RP].

(4) a. They expected the woman to win

b. {<R_j, *v**>{DP_i, $\{_{\alpha}$ R_j, $\{_{\beta}$ DP_i, $\{_{\delta}$ to, {be, win}}}}}

Here, DP moves to [spec, RP], strengthening R by agreement, so that α can be labeled. Further, movement of DP *out of* β makes the lower copy of DP invisible, resolving the {XP, YP} configuration of β. There is still a problem, however, how do we label δ since, recall, 'T is too weak to serve as a label'?

In the defective T cases, the label of β in (4b) is not determined because the structure is {XP (the woman), YP (to win)} and the feature sharing is not possible because non-finite T has no relevant features. The strategy of moving one of the SOs is available here. Thus, DP moves to another position, so that labeling failure does not take place. Before presenting this analysis in section 4, the next section introduces some necessary background.

3. External pair-Merge of heads in verbal domain

Recent simplifications of the theory of UG perhaps allow Merge to be the only operation in the narrow syntax (NS). Merge guarantees unbounded human language expressions and it comes for free. Merge also can apply either externally or internally (Chomsky 2004, 2005). Chomsky (2007, 2008, 2013, 2015) shows that labels of the syntactic objects are determined by another mechanism that comes from the "third factor" in the sense of Chomsky (2005). Merge takes *n* (two) syntactic objects and combines them into an unordered set, without projected labels (see Collins 2002 for representations without labels and Seely 2006 for an explanation of representational labelessness). Simplest Merge does not generate order or projection. Merge application is unconstrained, i.e., is not feature driven or 'purposeful'. Merge just optionally combines two syntactic objects, without modifying either (The *No Tampering Condition*, Chomsky (2008: 138)) or introducing extraneous objects (*Inclusiveness Condition*, Chomsky (1995: 228)). In addition to (set) Merge, there is *Pair-Merge*, which takes *n* (two) syntactic objects and combines them into an ordered pair (Pair-Merge (α, β) = <α, β>). Pair-Merge is introduced by Chomsky (2000, 2004a), for descriptive adequacy (regarding adjuncts) arguably unattainable with just set Merge.[8] Thus, it is natural to think that there are four modes of Merge: internal/external set-/pair-Merge. EKS (2016) extends the idea of Chomsky's (2015) internal pair-Merge applying in a case of head movement to the external Pair-merge, without any additional assumptions. As they note, it would require a stipulation to prevent external pair-Merge of heads, essentially a morphological operation. To account for the long-standing problem of unvaluation in bridge verb constructions, EKS (2016) propose a new analysis.[9] The

---

[8] There is some evidence against the pair-Merge analysis of adjuncts. See Oseki (2015) for relevant discussions. However, in this paper, I will only deal with the X^0-level, namely, head movement and external pair-Merge of heads (EKS 2016). See Richards (2009) for XP^0-level internal pair-Merge.

[9] See also Nomura (2014, 2015) for an alternative analysis.

relevant example (an instance of a bridge verb) is analyzed in Chomsky (2015) that raising-to-object might be optional.

(5) John thinks that he will win

a. $\{_{\varepsilon=?} \{_{\delta} \text{C}, \ldots\}, \{\text{think (R)}, \{_{\delta} \text{C}, \ldots\}\}\}$

“raising of CP object to [spec, RP]”

b. $\{<\text{R}_j, v^*>, \{_{\varepsilon} \text{think (R}_j), \{_{\delta} \text{C}, \ldots\}\}\}$

“head movement” (by internal pair-Merge of R to *v**)

If the CP complement of *think* undergoes internal set-Merge to the specifier of R (*think*) in (5a) (by analogy to a DP-object internally merging to [spec, RP] in transitives), there are no relevant agreeing features between *think* and δ(=CP), resulting in a labeling failure of ε. Chomsky (2015) argues that if the syntactic object δ does not move to specifier of *think* in (5b), the copy of the Root is invisible, so the label of *ε* is the label of δ, which is C. However, *v** remains with unvalued features, which results in a crashing derivation, contrary to fact. To solve this problem, EKS (2016) propose that external pair-Merge of R to *v** (<R, *v**> makes *v** invisible and so the unvalued features on *v** are invisible for narrow syntax, thereby averting crash. The phasehood is canceled since the loci of the unvalued feature are (visible) phase heads including *v**. Furthermore, the distinction between strong phases and weak phases can now be derived by free (unconstrained) application of simplest binary set-/pair-Merge. Thus, the derivation EKS propose is as follows.

(6) a. form <R, *v**> by external application of pair-Merge (a morphological operation)

b. construct CP by repeated application of set-Merge

c. Merge <R, *v**> to δ=CP: {<R, *v**>, {$_{\delta}$ C, …}}

To sum up, EKS (2016) propose the following hypotheses: (i) Pair-Merge as well as Set-Merge can apply freely. e.g., <R, *v**>, (ii) external Pair-Merge of R to *v** makes uninterpretable features on *v** invisible, (iii) Pair-Merge of R to *v** prior to Set-Merge of R to its CP complement de-phases *v**. Since *v** and hence it φ-features are invisible, there is no unwanted crash due to unvalued φ-features on *v**.

4. Labeling in infinitival clauses

4.1 Non-finite clauses

Based on the proposal in EKS (2016) that external pair-Merge can apply freely, I argue that external pair-Merge of T to C (forming <T, C>) is available as a null hypothesis, and that such application of external pair-Merge yields de-phasing of C since external Pair-Merge of T to C makes (phase-identifying) uninterpretable features on C invisible. As for labeling, since the amalgam <R, *v**> serves as a label while R is too weak to serve as a label alone, I propose that non-phase heads (R and T) are too weak to serve as a label but the amalgam <R, *v**> and <T, C> can serve as a label if we assume the parallelism between C phases and *v** phases. I will assume that the amalgam <T, C> can serve as a label and that this amalgam is phonologically realized as *to* in English.[10] As for the status of the

---

[10] One of the reviewers points that there is another way to interpret the status of T. Therefore, although the agreed T in English is too weak to serve as a label (Chomsky 2015), the defective T in English can be a labeler; it does not agree with the subject since it does not have u ϕ, compared with the finite T. Also, as Chomsky (2015) mentioned, T in Italian has a rich agreement, thus it can serve as a label. Incorporating all of these, the reviewer's alternative approach suggests that Ts that show strong agreement or agreeless Ts are strong enough to become labels, whereas Ts that need something to their specifier are too weak to server as a label.

Although this could be an alternative approach, it seems that we add some complication with this proposal. First, u ϕ on T is transmitted from C (Chomsky 2007, 2008), so before T does not have a u ϕ, the finite T is labelable according to the reviewer's proposal. Then, the subject does not have to move to the [spec, TP], though a labeling problem occurs at the *v**P since the base-generated position of the subject ([spec,*v**P]) and the whole structure yield {XP, YP} problem. Suppose the subject moves to the [spec, CP], then C agrees with the subject in [spec, CP] since C did not transmit its u ϕ -features. Then, the derivation does not satisfy the EPP-effect at [spec, TP]. Another derivation is that T does not have u ϕ, but once feature inheritance applies, T needs a subject in its specifier since T becomes too weak to be a label.

Instead of adopting the reviewer's proposal, I adopt Chomsky (2015) and EKS's (2016) approach that R and T are not labelers and the amalgam becomes labeler such as [R,*v**]. If phasehood, when external pair-Merge applies to a phase head, it is de-phased, which derives weak/defective phases (i.e., there is no PIC effect or Transfer of the complement of a de-phased phase-head formed by external pair-Merge.). Assuming that unvalued features induce Transfer (strong phase effect), dephasing cancels Transfer.

Consider (7) below. As I mentioned before, the infinitival clause has a labeling problem. That is, the subject position in infinitival clause shows {XP, YP} problem. The labeling algorithm finds the head of X and Y simultaneously and the identification of label becomes impossible due to non-uniqueness/ambiguity.

(7) a. I expected John to win.

b. {$_{CP}$ …,{$_{\beta}$ <R$_j$, *v**>, {John$_i$, {R$_j$(=expect), {$_{\alpha}$ John$_i$ **<T, C>**(=to),…}}}}}

In the case of ECM constructions as in (10), the DP *John* moves to [spec, <T, C>P] and then to [spec, RP]. Then there is no syntactic object in [spec, <T, C>P]. The label of α becomes <T, C> and the label of β becomes <R, *v**>. Then how about the label of α ? The specifier of the <T, C> is filled by DP, and this creates the {$_{\alpha}$ XP, YP} problems. Here

---

we extend this approach, then T is an uniformly unlabelable object regardless of whether it's a finite or infinitival T. I believe this is the null hypothesis once we adopt Chomsky (2015).

agreement between X = D and Y = *to* is not possible due to lack of φ-features, though the movement out of this {XP(=*John*), YP} makes the DP *John*-copy invisible (Chomsky 2013), that is, the specifier of <T, C> is invisible.[11] The structure of α is {H, XP}, where H is <T, C>. The phasehood of C is cancelled here (C being invisible), thus the phase head C is inert. Note, this also suggests that the movement to [spec, <T, C>P] can alternatively skip this position since there is no PIC effect in this domain since there is no strong phase. The same analysis applies to the raising in (11)-(12). In the cases of raising, John moves first to the embedded [spec, <T, C>P] and then to the [spec, TP] in the matrix clause.

(8) a. John seems to be happy.

b. $\{_{CP}$ …,{John$_i$, {**<R, v*>**, $\{_{\alpha}$ John$_i$ **<T, C>**, John$_i$…}}}}

But notice again that under free Merge, movement to [spec, <T, C>P] is possible but not necessary. Of course, de-phasing can apply multiple times in single derivation; (9) shows

[11]Note EPP is satisfied after movement to [spec, <T, C>(=*to*)P], but after *John* moves to [spec, RP], the DP-copy in [spec, <T, C>P] is invisible, as if the EPP does not need to be satisfied in this representation. Informally, EPP is satisfied derivationally, and not representationally.

that de-phasing applies until the derivation reaches the matrix CP.[12,13]

(9) a. John$_i$ seems [$_{TP}$ John$_i$ to be likely [$_{TP}$ John$_i$ to win]](Lasnik and Saito 1992: 135 (144))

b. {$_{CP}$ … John$_i$, {**<R, v*>**, …, { John$_i$ **<T, C>**, …, { John$_i$ **<R, v*>**, …, {$_\alpha$ John$_i$ **<T, C>** , John$_i$, …

As for control structures, there is PRO in the specifier of <T, C>. In (10), the verb *hope* is also applied external pair-Merge (<R, *v**>) since the verb does not assign accusative case. *John* external set-Merge to <R, *v**>, then moves to the [spec, TP] in the matrix clause.

[12] This also suggests the possibility of the long-distance one fell swoop A-movement, which is compatible with labeling theory. See also Bošković (2002), Epstein and Seely (2006) and Epstein et al. (2014) for relevant discussions.

[13] One of the reviewers wonders about the availability of [spec,<T,C>]. The proposed analysis leaves this open. Thus, the [spec,<T,C>] position is available and it would be a landing site of successive cyclic movement. Regarding this point, the reviewer provides the examples below adopted from Chomsky (2008).

(i) *Of which car did [the driver t$_i$]$_j$ [t$_j$ cause a scandal] (Chomsky 2008: 147)

(ii) Of which car is [the driver t$_i$]$_j$ likely [t$_j$ to [t$_j$ cause a scandal]] (Chomsky 2008: 153)

As the reviewer mentioned, I agree that [spec, to] in (ii) is available, hence successive cyclic-movement via [spec, to] does not face any problem; only the lower copy of "the driver of which" remains there, but the labeling algorithm does not find it since it is invisible. Since the details regarding subject islands are not relevant in this paper, I will not go into them here.

(10) a. John hopes to leave.

b. {$_{CP}$ …,{John$_i$, {**<R, v*>**, {$_{\alpha}$ PRO$_i$, {**<T, C>**, PRO$_i$…}}}}

As for PRO, it moves from [spec, *v*P] in the embedded clause to the [spec, <T, C>P]. How to label α? It could be a<T, C> as well as raising constructions cases.[14,15,16]

---

[14] However, if PRO is interpretable, it is present at the CI interface, which might in turn create a problem for labeling. As a possible solution to this problem, perhaps it is the case that PRO and <T, C> might show some form of agreement (see Davis 1981 for early and important arguments that PRO, initially thought to be an "empty" category has relevant features). In this case, PRO is visible and EPP is satisfied by partial agreement/control (see e.g. Iatridou 1988 and Landau 2003). Furthermore, if viable, this approach would be an analog of Chomsky's (1995) null case theory within which there is a case-relation between PRO and "to," which in principle (assuming a checking type theory) could yield labeling as X and Y in {XP, YP} control structures do share a uniquely identifiable feature. See also Martin (2001). Another possibility is that PRO is completely invisible. In that case, PRO would stay in-situ (i.e., [spec,VP]) or move to [spec, <T,C>] position. I leave this important issue, not adequately addressed here, to future research.

[15]One of the reviewers wonders how to analyze "for-to" infinitival constructions such as

(i) For John to show up would be a surprise.

Since we propose the infinitival marker *to* in English is the amalgam <T,C>, which is a labeler, for-to infinitive clauses seem to be relevant here. There are two ways of analyzing this phenomenon: (a) *For John* is just a PP, like adjunct, so it is not a relevant structure we are investigating in this paper. (b) There might be a possibility that the elements in the <T,C> amalgam can be extracted. Thus the derivation would be like this.

b. (i) Merge <T,C>: {<T,C>, {…

(ii) Merge John: $\{_\alpha$ John, {<T,C,>, {…

(iii) Internal Merge C to the whole structure: {$C_{i,}$ $\{_\alpha$ John, {<T,C_i>, {…

(iv) once C is extracted, unvalued features on C are again activated and agree with John.

(v) the whole structure is labeled by C and α is labeled by <phi,phi>.

Although the reviewer points out that the canceling property of amalgams would lead to no for-trace effect, this also explains the comp-trace effect in for-to infinitival clause. The examples are provided by the reviewer.

(ii) *who did she want for to show up? (cf. *who did she want that show up?)

a. Merge *who* to the infinitival clause: {who, {<T,C>, …

b. internal Merge of C: {C(for), {who, {<T,C>, …

c. C is activated, and agrees with *who*

d. the valued feature on C is inherited by <T,C> and <T,C> acquires phasehood after

4.2 Merge-over-Move is a labeling problem

In this section, assuming that pair-Merge as well as set-Merge apply freely, I argue that the Merge-over-Move principle is just a stipulation and it can be deduced from the labeling theory. Consider from Chomsky (1995, 2000).

(11) [$_\gamma$ There, [is, [likely, [$_\beta$ to, [be, [$_\alpha$ a man, in the room]]]]]]

There are unlabeled structure problems in (11). The label of β is not determined by *to* assuming that "T is too weak to serve as a label." In the early minimalist program, the question at hand was how do we generate this case, but not (12) (if there is no S-structure

---

C(*for*) is deleted, otherwise the complement of *for* is transferred.

This derivation adopts Chomsky's (2015) that-trace effect approach, making the right prediction.

[16] As one of the reviewers correctly points out, we might extend our analysis to *tough*-movement. Thus, *though*-movement looks like a long-distance A-movement; the de-phasing by pair-Merge applies to the embedded clause twice, as shown in (ii).

(i) John is easy to please

(ii) {$_{TP}$ John$_i$, {T, {is-easy, {John$_i$, {<T,C>=to, {<R,v*>, John$_i$}}}}}

John moves from the embedded object position to the [spec, TP] in the matrix clause via [spec,<T,C>]. Since examining the details of *tough*-movement is beyond the scope of this paper, I will put it aside, but see Hartman (2009), Hickes (2009), Epstein and Obata (2012), Fleisher (2013), Keine and Poole (2017) Longenbauch (2017) for details.

and hence no Case Filter applied at S-structure).

(12) *There is likely [$_{TP}$ a man$_i$ to be t$_i$ in the room].

At the derivational stage of the embedded TP ({T, {be, {a man, in the room}}}), there are two possibilities to satisfy EPP; (i) *a man* moves to [spec, TP] by internal set-Merge or (ii) the expletive *there* is externally set-Merged. Chomsky (2000) argues that Merge is simpler than Move since the definition of Move (Merge and Agree) was more complex than (just) Merge. This is called "the Merge-over-Move principle." Chomsky (2000) solves these phenomena by stipulating phases along with lexical subarrays (SAs). Thus, a phase is defined by a lexical subarray, which is assembled from a lexical array (LA).

(13) : Lexical Subarray: Chomsky (2000: 106-107)

(i) Select LA from Lexicon, then Select SA from LA

(ii) "Take a *phase* of a derivation to be an SO derived by choice of SA."

This lexical subarray/phase-based approach correctly excludes (12), as desired, while allowing (11). In (12), at the derivational stage of embedded TP ({T, {be, {a man, in the room}}}), Merge of expletive *there* is preferred to Move of *a man* by the Merge-over-move Principle (see Chomsky (2000) for details).

Another case which seems to be violation of the Merge-over-Move is in (14) (cf. Chomsky 1995). However, the examples are subsumed in (sub)lexical phasal array analysis.

(14) a. There is a possibility [$_{CP}$ that a man$_i$ will be t$_i$ in the room].

b. A possibility is [$_{CP}$ that there will be a man in the room].

(15) Subarray of embedded CP in (18a): {*that*, *will*, *be*, *a man*, *in*, *the*, *room*}

If a lexical subarray consists of *that*, *will*, *be*, *a man*, *in*, *the*, *room*, [spec, TP] is filled by Move of *a man*. There is no possibility of external set-Merge of *there* because there is no lexical item *there* in the lexical subarray. That is, "Merge over Move" allows MOVE when there is no *there* in the subarray to Merge. If a lexical subarray includes *there*, Merge of *there* is operated as in (17b). However, Epstein et al. (2014) argue that there is no support for the Merge-over-Move principle since Merge, including internal merge (which is called Move in Chomsky 2000), can apply freely (Chomsky 2004 *et seq*.). Thus, Merge is the simplest and unified single operation: Merge (α, β) = {α, β} (Chomsky 2013), hence there is no difference between Move and Merge, they are reduced to the single operation Merge rendering the Merge-over-Move principle unstatable. Furthermore, the effects of the 'lexical subarray' can be reduced to Labeling Theory. Consider again,

(16)

a. *There is likely [$_{\alpha}$ a man$_i$ to be t$_i$ in the room].

b. There$_i$ is likely [$_{\alpha}$ t$_i$ to be a man in the room].

At the derivational stage of the embedded TP, when *a man* moves to the embedded [spec, TP] in (19a), minimal search cannot find a unique feature shared by the head of *a man* (D) and the head of *to be in the room* (T), hence labeling by shared features is impossible since T is a non-finite head that does not have full φ-features. However, there is no problem in (19b). At the stage of α, *there* is a copy and minimal search cannot see it. The label of α is determined as T according to Epstein et al. (2014). They argue that "If *there* bears at least one φ-feature and undergoes further movement, then minimal search finds the only visible head T as the label of α, and a completely labeled, hence Full-Interpretation-compliant, CI representation is generated" (Epstein et al. 2014: 471). However, Chomsky (2015) suggests that T in English is too weak to serve as a label. If the proposal that the amalgam <T, C> can label (corresponding to the hypothesis that <R, *v**> can label as proposed in Chomsky 2015) is on the right track, the situation can be circumvented.

(17) There seems to be a man in the room.[17]

{There$_i$, {seems {$_{\alpha=<T, C>}$ t$_i$ **<T, C>**, {be, {a man, in the room}}}}}

⇒{$_\alpha$ H=<T, C>, XP}

(18) *There seems a man to be in the room.

{There, {seems {$_{\alpha=?}$ a man, {**<T, C>**, {be, {t$_i$, in the room}}}}}}

⇒*{$_\alpha$ XP, YP}

---

[17] As for the label of "the small clause" {a man, in the room}, Goto (2017) argues that there is no labeling problem. Based on Abe (2018), the derivation starts from merging *there* and *a man* in the small clause α, and *there* has unvalued φ-features, which may or may not be valued by its associate (*a man* in this case). After *there* and *a man* merge directly, both of them keep moving and the associate *a man* is valued at the embedded *v** phase. The expletive there moves successive cyclically to the [spec, TP] in the matrix clause.

(i) There is likely {to be {$_\alpha$ a man in the room}}.

a. {there$_i$, {<be(R)$_j$, *v**>, {{a, man}$_k$, {be$_j$ {$_\alpha${t$_i$, t$_k$,}, {in the room}}}}}

b. {likely, {there$_i$, {T(to), {t'$_i$, {<be(R)$_j$, *v**>, {{a, man}$_k$, {t$_j$ {$_\alpha${t$_i$, t$_k$,}, {in the room}}}}}

The structure of α becomes {t$_{\text{a man}}$, in the room}, so that the label of α is the label of {in the room} = P. As for cases like "I want a man in the room," they may be problematic, though see EKS (2014) for the suggestion that there is string vacuous movement of *a man* in such cases, leaving an invisible copy, in turn allowing labeling of {in the room} as P. See also Takita et al. (2016) for the labeling of small clauses.

I propose that we take infinitival marker *to* as <T, C> and this serves as a label, the same way <R, *v**> does since R is too weak to serve as a label by itself. Thus, we can label α to be <T, C> in (20). Note that <T, C> is a de-phased head (since C is invisible by pair-Merge) and there is no Transfer. If agreement is phase-bound (Bošković 2007 argues that it is not), there is no problem generating long-distance agreement between the matrix finite T and the associate *a man*, though the distance between them can be unbounded. Unbounded distance, but no phase boundary so no Transfer, as illustrated in (22).

(19) There **is(=T)** likely to seem to be likely … to be **a man** outside.

In this section, I argued that the Merge-over-Move stipulation can be eliminated from narrow syntax by assuming labeling theory. This is one of the consequences of simplest Merge (Merge (α, β) = {α, β}). Merge applies freely without any driving force (like EPP-feature or edge-feature or the 'need' to value features). Merge can form {H, XP} and {XP, YP}. However, if the labeling algorithm, reduced to third factor minimal search, cannot find a unique prominent lexical item or feature, the syntactic objects cannot be labeled. As we discussed above, the labeling problem arises in the infinitival clause. When the specifier of TP is occupied, the structure is of the form {XP, YP} and LA (minimal search) cannot label this set. In finite clauses, the labeling problem is solved by sharing features. This way, we can eliminate the stipulated hence unexplained EPP in terms of third factor labeling, coupled with the hypothesis that the interfaces require categorial information for proper interpretation.

5. Conclusion

In this paper, I put forward the proposal that external pair-Merge of a non-phase head to a phase head <non-Phase head, phase head> yields desirable results without any mechanism additional to those in Chomsky (2015). The main argument of this paper is as follows: (i) External Pair-Merge of a non-phase head to a phase head is allowed (in the absence of an unwanted stipulation on Merge preventing it) and derives the de-phasing property at the C-T level, (ii) External Pair-Merge of a phase head to a non-phase head makes uninterpretable/unvalued features on the phase head invisible for minimal search (Chomsky 2015, EKS 2016), (iii) non-phase heads are too weak to serve as labels alone, (iv) the amalgam <non-phase head, phase head> can serve as a label (Chomsky 2015 for internal Pair-Merge of R to *v**), (v) in the case of <T, C> = *to*, there is no EPP requirement because the amalgam <T, C> is sufficient to serve as the label. Moreover, the amalgam <T, C> is realized as the infinitive marker *to* in English.

As for EPP, both *v** phases and C phases require raising of subject/object in terms of labeling, but not for EPP. That is, infinitival clauses do not need anything to fill their subject position due to EPP effects (e.g. Epstein and Seely 2006). EPP effects are reduced to labeling theory, which in fact allows clauses without subjects in [spec, TP]. If this is on the right track, the timing of pair-Merge of a non-phase head and a phase head, in effect a morphological operation creating a relation between two lexical items, has important effects with respect to syntactic derivations (cf. Epstein et al. 2018). If the analyses presented here are on track, some of these effects are welcome. Further research is of course necessary to determine the broader range of empirical effects of a theory of human linguistic and third factor principles incorporating such formal mechanisms.

References

ABE, J. 2018. How to probe expletives. *Studia Linguistica* 72:1, 72–112.

ALEXIADOU, A & LOHNDAL, T. 2017. The structural configuration of root categorization. *Labels and Roots*. eds. L. Bauke & A. Blümel, 203–232. Mouton: De Gruyter.

BLÜMEL, A. 2017. Exocentric Root Declaratives: Evidence from V2. *Labels and Roots.* eds. L. Bauke & A. Blümel., 263–289. Mouton: De Gruyter.

BORER, H. 2013. *Structuring Sense vol Ⅲ: Taking Form*. Oxford: Oxford University Press.

BOŠKOVIĆ, Ž. 1996. Selection and the categorical status of infinitival complements. *Natural language and linguistic theory* 14, 269–304.

BOŠKOVIĆ, Ž. 2002. A-movement and EPP. *Syntax* 5:3, 167–218.

BOŠKOVIĆ, Ž. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:4, 589–644.

BOŠKOVIĆ, Ž. 2016. Contextual phasehood and the ban on extraction from complements of lexical heads: when does X become a phase? *Phase Theory and its consequences: The initial and recursive symbol S*. eds. M. Yasui & M. Mizuguchi., 5–39. Tokyo, Kaitakusha.

CECCHETT, C & DONATI, C. 2015. *(Re)labeling*. Cambridge: MIT Press.

CHOMSKY, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

CHOMSKY, N. 2000. Minimalist inquiries: The framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. eds. R. Martin, D. Michaels & J. Uriagereka, 89–156. Cambridge, MA: MIT Press.

CHOMSKY, N. 2001. Derivation by phase. *Ken Hale: A life in Language*. ed. M. Kenstowicz, 1–52. Cambridge, MA: MIT Press.

CHOMSKY, N. 2004. Beyond explanatory adequacy. *The cartography of syntactic structures: Vol. 3. Structures and beyond*. ed. A. Belletti, 104–131. Oxford: Oxford University. Press.

CHOMSKY, N. 2005. Three factors in the language design. *Linguistic Inquiry*, 36:1, 1–22.

CHOMSKY, N. 2007. Approaching UG from below. *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics* ed. U. Sauerland & H.-M, Gärtner, 1–30. Berlin Mouton de Gruyter.

CHOMSKY, N. 2008. On phases. *Foundational issues in linguistic theory*. eds. R. Freidin, C. P. Otero & M.-L. Zubizarreta, 133-166. Cambridge, MA: MIT Press.

CHOMSKY, N. 2013. Problems of projection. *Lingua* 130, 33–49.

CHOMSKY, N. 2015. Problems of projection Extensions. *Structures, Strategies and Beyond*. eds. E. Di Domecinco, C. Hamann & S. Matteini, 3–16. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

CHOMSKY, N. 2017. Puzzles about phases. *Linguistic Variation: Structure and Interpretation*. to appear in eds. L. Franco & P. Lorusso. De Gruyter Mouton.

COLLINS, C. 2002. Eliminating labels. *Derivation and Explanation in the minimalist program.* eds. S. D. Epstein and T. D. Seely., 42–64. Blackwell Publishers.

DAVIS, L. 1981. Argument-binding and control. *Journal of Linguistic Research* 2, 89–103.

DEN DIKKEN, M. 2006. Phase extension: contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1, 1–44.

EPSTEIN, S. D. 1986. The local binding condition and LF chains. *Linguistic Inquiry* 17:2, 187–205

EPSTEIN, S. D. 1998. Overt scope marking and covert verb-second. *Linguistic Inquiry* 29:2, 181–227.

EPSTEIN, S. D, KITAHARA, H. & SEELY, T. D . 2014. Labeling by minimal search. *Linguistic inquiry* 45:3, 463–481.

EPSTEIN, S. D, KITAHARA, H. & SEELY. T. D. 2016. Phase-Cancellation by external pair-merge of heads. *The Linguistic Review* 33:1, 87–102.

EPSTEIN, S. D & OBATA, M. 2012. Feature-splitting Internal Merge: The Case of *Tough*-constructions. *Ways of Structure Building*. eds. Uribe-Etxebarria, M. & Valmala, M. 366-384. Oxford: Oxford University Press.

EPSTEIN, S. D & SEELY, T. D. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.

EPSTEIN, S. D, OBATA, M. & SEELY, T. D. 2018. Is Linguistic Variation Entirely Linguistic? *Linguistic Analysis* vol. 41, 3–4, 481–516.

FLEISHER, N. 2013. On the absence of scope reconstruction in *tough* subject A-chain. *Linguistic Inquiry* 44:2, 321-332.

GALLEGO, Á. 2007. Phase theory and parametric variation. Ph.D. dissertation. Universitat Autònoma de Barcelona.

GOTO, N. 2017. How to label *there*-constructions. *English Literature, Regional Branches Combined Issue*, vol. 9, 33–43.

HARTMAN, J. 2009. Intervention in *tough* constructions. In Proceedings of *NELS39*, 387-397.

HICKS, G. 2009. *Tough*-constructions and their derivations. *Linguistic Inquiry* 40:4, 535-566.

HORNSTEIN, N. 1999. Movement and control. *Linguistic inquiry* 30:1, 69–96.

IATRIDOU, S. 1988. On nominative case assignment and a few related things. Ms.

KEINE, S & POOLE, E. 2017. Intervention in *tough*-constructions revisited. *The Linguistic Review* 34:2, 295-329.

LANDOU, I. 2003. Movement out of control. *Linguistic Inquiry 34:3*, 471-498.

LASNI, H & SAITO, M. 1992. *Move α: conditions on its application and output*. Cambridge, MA: MIT Press.

LONGENBAUGH, N. 2017. Composite A/A'-movement: Evidence from English *tough*-movement. Ms. MIT.

MARANTZ, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon's. eds. Dimitriadis, A., L. Siegel, C. Surek-Clark, and A. Williams. *UPenn Working Papers in Linguistics* 4.2, 201-225.

MARTIN, R. 2001. Null case and the Distribution of PRO. *Linguistic Inquiry* 32:1, 141–166.

NARITA, H. 2014. *Endocentric Structuring of Projection - free Syntax*. John Benjamins Publishing Company.

NOMURA, M. 2014. Case and Agreement Revisited: Why do Unvalued Features Exist? Presented at 32nd Conference of the English Linguistic Society of Japan.

NOMURA, M. 2015. Pair-Merge and its consequences. Presented at the 87th General Meeting of The English Literary Society of Japan.

OBATA, M. & EPSTEIN, S. D. 2012. Feature-splitting internal merge: The case of *tough*-constructions. *Ways of structure building.* eds. M. Uribe-Etxebarria and V. Valmala, 366–384. Oxford: Oxford University Press.

O'NEIL, J. 1995. Out of control. In proceedings of *Proceedings of the 25th Annual Meeting of the North East Linguistic Society (NELS* 25), 361–371. GLSA.

ORMAZABAL, J. 1995. The syntax of complementation: on the connection between syntactic structure and selection. Ph.D. dissertation, University of Connecticut, Storrs.

OSEKI, Y. 2015. Eliminating Pair-Merge. *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, eds. Ulrike Steindl et al., 303312. Momerville, MA. Cascadilla Proceedings Project.

PESETSKY, D. 1992. Zero Syntax. vol. 2. Ms. MIT.

PESETSKY, D. 2016. Exfoliation: towards a derivational theory of clause size. Paper presented at Workshop on shrinking trees, October 10, 2016.

RICHARDS, M. 2009. Internal pair-Merge: the missing mode of movement. *Catalan Journal of Linguistics 8*, 55-73.

SEELY, D. T. 2006. Merge, derivational c-command, and subcategorization in a label-free syntax. *Minimalist Essays*. ed. C. Boeckx., 182–217. Amsterdam/Philadelphia. John Benjamins Publishing Company.

SHIM, J-Y. and EPSTEIN, S. D. 2016. Two Notes on Possible Approaches to the Unification of Theta Relations. *Linguistic Analysis*, 40:1-2, 3–20.

TAKITA, K., GOTO, N. & SHIBATA, Y. 2016. Labeling through Spell-Out. *The*

Linguistic Review 33:1, 177–198.

Received June 19, 2019

Accepted October 26, 2020

Yushi Sugimoto

Department of Linguistics

University of Michigan

440D Lorch Hall

611 Tappan Street

Ann Arbor, MI

48109-1220

The United States

yushis@umich.edu