⟨φ, φ⟩-less Labeling*

Jae-Young Shim†
Pukyong National University

ABSTRACT
Hypothesizing that every syntactic object reaching the interfaces must be labeled, Chomsky (2013, 2015) introduce an independent label-identifying operation, Labeling Algorithm (LA). This newly proposed LA, which is an instance of Minimal Search, is assumed to search for not only the categorial feature of heads but also any agreeing features shared by two heads to determine the label of a given syntactic object. This paper argues, however, that postulation of the latter type of search by LA (i.e., locating agreeing features) adds a computational burden to the grammar by forcing LA to perform two different types of search, namely, a "comparison search" in addition to the widely agreed-upon Minimal Search. This paper also addresses some conceptual and empirical problems that arise from the paired form of the non-categorial label, ⟨φ, φ⟩, whose introduction results from the problematic comparison search by LA. To solve these problems, I present an alternative analysis in which LA searches only for the categorial feature of heads. Consequently, neither comparison search nor the paired form of the label, ⟨φ, φ⟩, becomes necessary in the analysis.

Keywords: label, labeling algorithm, interpretability, categorial feature, minimal search, comparison search

1. Introduction

Throughout the history of generative grammar, the categorial information/feature of heads has been at the heart of the notion 'projection' (or 'label' in more recent terms). Thus, for example, the categorial information of a head was the (sole) obligatory element in the familiar Phrase Structure rules (e.g. NP → (Det) N (PP)) and in the X-bar template of the Government and Binding framework (Chomsky 1981), each phrase was understood to be 'projected' from the category of its corresponding head.

* I would like to thank the three anonymous reviewers for their insightful comments. All remaining errors and shortcomings are, of course, mine alone..
† Corresponding author: biolinguistics@gmail.com

https://doi.org/10.30961/lr.2018.54.1.23

Although the pivotal role of the categorial feature of heads in formation of phrases weakened to some degree in early minimalism (Chomsky 1993, 1995)[1], not only has it resurfaced in the more recent version of the minimalist framework (Chomsky 2013, 2015) but a novel form of label has also been introduced. One instance of this newly proposed label takes the form of a pair of agreeing φ-features between two heads, i.e. ⟨φ, φ⟩.[2]

As will be discussed in more detail in what follows, however, the label ⟨φ, φ⟩ that consists of agreeing features raises some conceptual as well as empirical problems, especially with respect to its interpretability at the Conceptual-Intentional (CI) interface. I thus argue against the proposed label ⟨φ, φ⟩ and present an alternative analysis where only the categorial feature of heads is relevant for labeling purposes.

The organization of the paper is as follows. Section 2 presents a brief overview of the labeling theory laid out in Chomsky (2013, 2015). Section 3 discusses theoretical and empirical problems with the label ⟨φ, φ⟩. An alternative analysis is presented in Section 4. Section 5 mentions some remaining problems and concludes the paper.

2. Labeling Theory in Chomsky (2013, 2015)

2.1. Overview

The operation Merge in minimalism has been taken, whether implicitly or explicitly, to incorporate two independent tasks: one is to 'combine' (two) syntactic objects (SOs) and the other to determine which one of the two combined SOs to 'project' or to become the 'label' of the resulting construction. Chomsky (2013: 43) thus writes "[u]nder PSG [Phrase Structure Grammar] and its offshoots, *labeling is a part of the process of forming a syntactic object SO* [italics added]."

Chomsky (2013, 2015; henceforth 'POP(E)'[3]), however, attempt to separate

1) This weakened role of the categorial feature was mainly due to the (then) assumption that what is projected is not just the categorial feature of a head but the entire feature set of it as well (Chomsky 1995).

2) The other type of the paired form of label proposed in (Chomsky 2013, 2015) is ⟨Q, Q⟩ where one member is (assumed to be) from the Q-feature of a wh-word and the other from the Q-feature of the interrogative C. Although I suspect that the analyses developed in later sections, especially those in section 4, can be applied to ⟨Q, Q⟩ without significant modifications and hence that ⟨Q, Q⟩ may also be eliminated, I will not discuss it here any further because the label ⟨Q, Q⟩ results from what has conventionally called A'-movement involving different properties from what I aim to focus in this paper, namely, the resulting label from A-movement.

labeling from Merge, reserving it for a new syntactic operation that he calls a Labeling Algorithm (LA). The operation LA, as he argues, seeks the structurally closest or the least embedded head (H) in a given SO, identifying such a head as the label of the SO. Consider the following two SOs in (1) to see in more detail how LA works in the POP(E) model.

(1) a. SO = {H, XP}
    b. SO = {XP, YP}

Identification of a unique label, i.e. the closest head, is trivial in (1a) since the structure contains a single head H that is least embedded so that LA can unambiguously identify it as the (unique) label of the structure. In (1b), however, identification of such a unique head is problematic because the structure contains two heads that are equally embedded, i.e. the head X of XP and the head Y of YP. Consequently, LA cannot unambiguously determine which of the two heads should become the label of the structure.

Chomsky (2013, 2015) discuss the following two scenarios in (2) where a unique label can nonetheless be identified in seemingly unlabelable {XP, YP} structures as in (1b).

(2) a. {XP ... {$_{\alpha}$ ⟨XP⟩, YP}}
    b. {$_{\beta}$ $XP_{[F]}$, $YP_{[F]}$}, where [F] of XP matches with that of YP.

Notice that unlike SO in (1b), ⟨XP⟩ in α of (2a) is a copy (of the moved XP). Chomsky claims that in this case the head of YP is unambiguously identified as the label of α on the assumption that copies such as ⟨XP⟩ are invisible to LA (i.e. copies are assumed to be irrelevant for labeling purposes). (2b) illustrates the second scenario where the seemingly unlabelable structure of {XP, YP} can nonetheless be unambiguously labeled. Unlike SO in (1b), the two phrases in (2b) share a feature indicated as [F]. Chomsky hypothesizes that in this case the pair of the shared feature F between X and Y (i.e. ⟨F, F⟩) becomes the label of β.[4]

3) For convenience, the abbreviation 'POP(E)' is used, when necessary, throughout the paper to collectively refer to the framework developed in Chomsky's 'Problems of Projection (POP, 2013)' and 'Problems of Projection: Extensions (POPE, 2015)'.

4) Chomsky (2013) adds the following with respect to the labelability of a feature shared by X and Y:

(i) "Searching {XP, YP}, LA finds the same most prominent element [i.e. a feature] [...] can take that to be the label of α [α = {XP, YP}]." (Chomsky 2013: 45)

2.2. ⟨φ, φ⟩ in Derivation

Consider (3) to see what type of SOs are labeled ⟨φ, φ⟩ in the course of derivation.

(3) a. The man bought a car.

b. $\{_{\alpha}$ NP(=the man), *v**P(=bought a car)$\}$ (α = unlabelable)

c. $\{_{\beta}$ NP, {T, $\{_{\alpha}$ ⟨NP⟩, *v**P}}} (α = *v**)

d. {C, $\{_{\beta}$ NP, {T, $\{_{\alpha}$ ⟨NP⟩, *v**P}}}} (β = ⟨φ, φ⟩)

[φ] [φ]

At some point in the derivation of (3a), the unlabelable structure in (3b) is generated. The structure is unlabelable because 1) it is of the form {XP(=NP), YP(=*v**P)} as discussed in section 2.1 and 2) there is no agreeing feature between the head of NP and the head of *v**P. Consequently, the label of α cannot be unambiguously identified by LA. However, if T merges with α and subsequently NP *the man* undergoes Internal Merge (IM) to Spec-T (from Spec-*v**) as illustrated in (3c),[5)] α turns into a labelable structure since the head of *v**P now becomes the unique head in the structure given the assumption that copies such as ⟨NP⟩ are not visible to LA.

Notice in (3c), however, that IM of NP *the man* to Spec-T results in generating another (seemingly) unlabelable SO, namely, SO=β. Since LA locates the two equally-embedded heads, i.e. N and T, it cannot unambiguously determine which of the two should be the label of β. As briefly discussed in section 2.1, however, the two heads in this case share the φ-features as indicated in (3d), one being inherent

---

(ii) "Mere matching of most prominent features does not suffice [...] What is required is not just matching but actual agreement [in the feature shared by X and Y]. (Chomsky 2013: 45)

According to (i), φ-features, for example, are found to be the most 'prominent' element when shared by X and Y. What is not clear here, however, is what exactly makes a feature 'prominent' to the eye of LA. In other words, is what makes the φ-features in our example prominent due to their 'inherent' prominence or simply to the fact that they are the ones 'shared' by two elements, or both, or neither? What about the categorial label *v** for {*v**, ZP}, for example? Is *v** identified as a label because it is (also) prominent by nature? If so, are the categorial features as prominent as the φ-features? Is there any hierarchy in prominence among features? I believe that all these (unanswered) questions make unclear, undefined and complicated how LA identifies a label.

5) As an anonymous reviewer pointed out, one may wonder how T can undergo merger with α in (3c) at the point when the label of α has not yet been determined. I assume with Chomsky (2013) that labeling is not a prerequisite for entering into computation such as Merge; hence no merging problem arises between T and α with no label in (3c).

to N and the other inherited by T from C.[6] Thanks to these agreeing φ-features between the two heads, the label of β can be unambiguously identified as $\langle\varphi, \varphi\rangle$.

Consider now the structures in (4) which illustrate another typical SO whose label is identified as $\langle\varphi, \varphi\rangle$ in the POP(E) framework.

(4) a. The man bought a car.
    b. $\{_\alpha$ R(=bought), NP(=a car)$\}$ (α = unlabelable)
    c. $\{_\beta$ NP, $\{_\alpha$ R, $\langle$NP$\rangle\}\}$
    d. $\{v^*$, $\{_\beta$ NP, $\{_\alpha$ R, $\langle$NP$\rangle\}\}\}$ (α = R, β = $\langle\varphi, \varphi\rangle$)
          [φ]   [φ]

Adopting ideas developed in Distributed Morphology (Halle and Marantz 1993, Borer 2005), Chomsky (2013, 2015) reanalyze the conventional lexical head V as the category-free root R, claiming that this R is (universally) too weak to serve as a label. He further assumes that though inherently unlabelable, R turns into a labelable head (or 'strengthens' in Chomsky's terms) if its Spec position is overtly filled by an element that bears an agreeing feature with it. With these two assumptions in mind, let's examine the structures in (4).

At the derivational stage of (4a), the unlabelable SO=α in (4b) is generated[7]; unlabelable since the least embedded unique head in the structure, namely, R, is (assumed to be) too weak to serve as a label.[8] However, if the object NP *a car* undergoes IM to SPEC-R and subsequently R inherits φ-features from v^* as illustrated in (4c) and (4d), respectively, both α and the newly generated β turn into a labelable structure. More specifically, the label of α is identified as R since R has now been strengthened thanks to NP in its Spec position, and the label of β can be identified by the agreeing φ-features between the head of NP and R.

I have shown that the two typical SOs whose label is identified as $\langle\varphi, \varphi\rangle$ in the POP(E) model, one being SO=$\{$NP, TP$\}$ as discussed in (3) and the other $\{$NP, RP$\}$

6) Chomsky (2013) claims that φ-feature inheritance from C to T leaves [φ] on C intact as he takes the mechanism of inheritance as copying. For different approaches to the status of [φ] on C after inheritance, see Ouali (2008), among others.

7) Technically speaking, however, labeling in Chomsky's (2013, 2015) framework does not take place immediately after the structure in (4b) is constructed but it is rather (assumed to be) executed at the phase level (i.e. once a phase head such as v^* is introduced into the workspace). For ease of exposition, however, we continue to discuss labeling of a given structure as if LA operated at the point when the structure is generated unless the timing of the application of LA plays a crucial role. See also footnote 14.

8) It is not clear to me, though, what prevents LA from seeking the head of NP instead.

as shown in (4). In the following section, I discuss some conceptual and empirical problems that arise from the proposed paired form of label, ⟨φ, φ⟩.

3. Problems

3.1. Conceptual Issues

The first conceptual problem with the label ⟨φ, φ⟩ pertains to its interpretive contribution to the Conceptual-Intentional (CI) interface. If a label is indeed required for an SO to be interpreted at the CI interface as claimed in the POP(E) model, the identified label of an SO is expected to make some interpretive contribution to the very interpretation of the SO. This indeed seems to be the case with the conventional category-based labels such as V and N given that the CI interface should have a means to distinguish whether the given SO is 'verbal' (e.g. *grow tomatoes*) or 'nominal' (e.g. *growth of tomatoes*).[9] It is not entirely clear, however, what type of such a discriminative interpretive contribution the non-categorial label ⟨φ, φ⟩ makes for the interpretation of SO.

The second problem, which is related to the first one, concerns the very interpretability of the label ⟨φ, φ⟩. Consider (5).

(5) {$_α$ NP, {T, {⟨NP⟩, ... }}}   α = ⟨φ, φ⟩
  [*i*φ] [*u*φ]   (*i* = interpretable, *u* = uninterpretable)

The label of α in (5) is identified as ⟨φ, φ⟩ by the agreeing φ-features between the two heads (i.e. between N and T). Notice, however, that the two φ-feature sets comprising the label ⟨φ, φ⟩ are not identical with respect to their interpretability at the CI interface. To be more specific, one member of the pair is from N and it is therefore (inherently) interpretable at the CI interface, whereas the other member is not since it is that of the head T (from C via feature inheritance). Given the assumption that uninterpretable features such as the φ-features of T must delete before reaching the CI interface (Chomsky 2008)[10], it is unclear how the uninter-

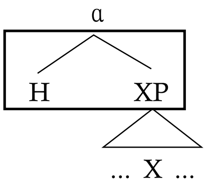
---

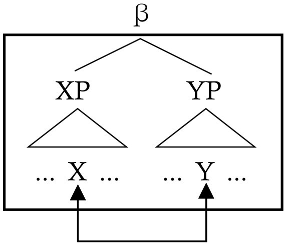
9) Chomsky (1995: 243) thus writes that "verbal and nominal elements are interpreted differently at LF and behave differently in the phonological component. K [i.e. the outcome of Merge] must therefore at least [...] be of the form [K=] {γ, {α, β}}, where γ identifies the type to which K belongs [...]." For similar discussions on the interpretive roles of labels, see also Narita and Fukui (2016).

10) An anonymous reviewer suggested that the problem of uninterpretability of T's φ-feature set in ⟨φ,

pretable part in ⟨φ, φ⟩ can survive to the CI interface, making an interpretive contribution (see also Stockwell 2014, 2016 for similar discussion).

The third problem has to do with computational efficiency of the operation LA. Consider the structures in (6).

(6) a.


b.


For expository convenience, let us suppose that LA scans down into a given SO to locate a label, i.e. to find the least embedded head. In (6a), scanning stops when LA locates H because H is the unique head that is least embedded in the structure. Turn to (6b): LA this time scans down into both XP and YP until it finds the two heads, X and Y. Although the identification of a unique label has already become ambiguous at this point due to the equal embeddedness of X and Y, the final decision cannot be made at this stage of search since LA does not know yet whether there could be any agreeing features between X and Y. In other words, LA needs to execute what we may call a 'comparison' search to see if there is any agreeing feature between X and Y. If LA locates any such agreeing features via this comparison search, the pair of those features will be identified as the label of β. If not, the label of β will then be finally determined as unidentifiable. Put simply, a comparison search for potential agreeing features complicates the operation LA in such a way that LA is required to perform an additional search besides the conventionally agreed-upon Minimal Search.

3.2. Empirical Issues

The empirical issues we discuss here pertain to Chomsky's (2013, 2015) claim that

---

φ⟩ may be circumvented if we adopt the 'feature-sharing' analysis proposed in Pesetsky and Torrego (2007) where the biconditional relation is rejected between the interpretability of a feature and its valuation (e.g. an 'unvalued' feature can nonetheless be 'interpretable'). Though intriguing, I suspect that the interpretability problem we raise above may still persist even in the feature-sharing analysis since the question of what 'semantic' contribution T's φ-feature set makes still remains unclear for the interpretation of T(ense) itself as well as of ⟨φ, φ⟩, regardless of whether one assumes it to be interpretable or not.

both the root R and T (in English) are too weak to serve as a label. Although these issues do not directly concern the very validity of the label $\langle\varphi, \varphi\rangle$ but they nonetheless play a conducive role in its creation. Thus, let us first consider the weakness of the root R.

(7) a. $\{_{\alpha}$ R, NP$\}$
b. $\{v^*$, $\{_{\beta}$ NP, $\{_{\alpha}$ R, $\langle$NP$\rangle\}\}\}$ (α = R, β = $\langle\varphi, \varphi\rangle$)
c. $\{$[R-v*], $\{_{\alpha}$ $\langle$R$\rangle$, NP$\}\}$

In the POP(E) model where R is assumed to be universally too weak to serve as a label, the label of α in (7a) cannot be identified. Once Spec-R is (overtly) filled via IM of NP[11], however, R gets strengthened and thus turns into a labelable head so that it can now serve as the label of α. Consequently, α in (7b) is identified as R, the label of β being identified as $\langle\varphi, \varphi\rangle$ by the agreeing φ-features between N and R.

Note, however, that (7b) is not the only conceivable legitimate derivation. There is, in fact, an alternative derivation which does not lead to creation of $\langle\varphi, \varphi\rangle$ but nonetheless poses no problem in terms of labeling. This alternative legitimate derivation proceeds as follows. Given the widely-accepted assumption that in the course of the derivation the root R raises to *v**[12], R in α of (7b) becomes a copy invisible to LA as illustrated in (7c). What this suggests is that the label of α in (7c) can be unambiguously identified (as the head of NP), rendering IM of NP (to Spec-R) unnecessary.[13] What this alternative derivation further suggests is that the same SO, i.e. *α* in (7), can be given a legitimate interpretation, for at least labeling purposes,

---

11) Although we will not discuss in this paper, it should be pointed out that IM of NP to Spec-R can raise problems in terms of anti-locality which bans movement that is too local. See, among others, Grohmann (2000, 2003, 2011) and Bošković (2013) for the Anti-locality Hypothesis and related discussions.

12) 'Widely-accepted' as we view R to *v** raising as an instance of what has traditionally been termed 'Head Movement'. For detailed discussion of head movement including R(=V) to *v** raising and related issues, see Funakoshi (2014) and references therein. As will be discussed in section 4.1, however, R-to-*v** raising is not assumed in my own analysis, *contra* Chomsky (2013, 2015) discussed in 3.2 above. For related issues, see also footnote 18 and 20 below.

13) It is thus crucial in Chomsky's (2015) framework that labeling of α in (7c) precedes raising of R to *v**. Furthermore, if we consider the following scenario, the ordering between raising of R and labeling becomes even more critical in the POP(E) model.

(i) $\{$[R-v^*], $\{_{\beta}$ NP, $\{_{\alpha}$ $\langle$R$\rangle$, $\langle$NP$\rangle\}\}\}$

If labeling were to take place after both IM of NP and raising of R, there would remain no element in α visible to LA, which would in turn result in a labeling failure. For related discussion of these issues, see Narita (2015). An anonymous reviewer pointed out that labeling of α should 'follow' (rather than 'precede') raising of R so as for α to be able to be labeled N(P). Note, however, α in (7) is argued to be labeled R(P), not N(P), in Chomsky's framework.

despite absence of the paired label ⟨φ, φ⟩.

Let us now turn to the claim of the weak T (in English). Consider (8).

(8) $\{_\beta$ T, $\{_\alpha$ NP, v*P}} (α, β = unlabelable)

Although neither α nor β in (8) is labelable in the POP(E) framework, the source of the unlabelability is different in each case. That is, the unlabelability of α is due to its own structural properties, i.e. α takes the form of the (unlabelable) {XP, YP} with no agreeing feature between the two. On the other hand, the unlabelability of β results not from such a structural anomaly but rather from Chomsky's (2015: 9) assumption that 'T in English is too weak to serve as a label.' In other words, T in (8) is 'somehow' defective when it comes to labeling, which prevents the label of β from being properly identified (unless its Spec position, i.e. Spec-T, is overtly filled). However, typical raising constructions as in (9) seem to pose a serious challenge to Chomsky's claim that T (in English) is inherently weak.

(9) a. John seems to like Mary.
 b. {C $\{_\delta$ John $\{_\gamma$ T_2-seems $\{_\beta$ ⟨John⟩ $\{_\alpha$ T_1-to {v*P ⟨John⟩ like Mary}}}}}}

As illustrated in (9b), the subject *John* undergoes IM from its base-generated Spec-v* position to the matrix Spec-T_2 via Spec-T_1. We now have (at least) four SOs that need to be labeled, namely, α, β, γ, and δ. Identification of a label for δ and γ is unproblematic under Chomsky's weak T analysis: the label of δ is identified as ⟨φ, φ⟩ by means of the agreeing φ-features between T_2 and the head of NP *John*, and the label of γ is identified as T by the strengthened T_2. The label of α and β, however, cannot be identified in this analysis[14]: since neither Spec-T_1 is overtly filled (i.e. what occupies Spec-T_1 is a 'copy' of *John*) nor there is any agreeing feature between T_1 and the head of NP *John*, T_1 must remain weak throughout the derivation. Consequently, the label of neither α nor β can be identified under Chomsky's weak T analysis.

14) One may wonder, as an anonymous pointed out, what prevents α from being labeled at the point when *John* moves to Spec-T_1. There are (at least) two reasons for this. One is that as briefly noted in footnote 7, labeling in Chomsky's (2013, 2015) framework takes place at the phase level. That is, LA does not operate until C is introduced into the workspace when *John* has already moved to Spec-T_2. The other is that in Chomsky's framework actual agreement between two elements is required for their shared feature(s) to serve as a label. Since there is no agreement between *John* and T_1, α cannot be labeled even if we assume that LA applies when *John* is (overtly) present in Spec-T_1.

4. Proposal: Copy Invisibility and Elimination of ⟨φ, φ⟩

4.1. Assumptions

In the Government and Binding (GB) model (Chomsky 1981), a moved element such as *John* in (10) was assumed to leave behind a coindexed phonetically-null trace.

(10) John$_i$ ..... { t$_i$ ... }

In the context of the Minimalist Program (Chomsky 1993 *et. seq.*), however, elements such as traces and indices that are not part of the lexicon are banned by the following condition in (11) proposed in Chomsky (1995).

(11) **Inclusiveness Condition** (Chomsky 1995: 225)
Outputs consist of nothing beyond properties of items of the lexicon.

Complying with the Inclusiveness Condition in (11) while keeping the attested ontological advantages of trace[15], Chomsky (1995) proposes a 'copy theory of movement' according to which, as illustrated in (12), a moved element does not leave behind a coindexed trace, but a copy of itself instead.

(12) John ... { ⟨John⟩ ... }

It is important to note, however, that contrary to the widely-held misconception, each of the two occurrences of *John* in (12) is a 'copy' of the other. In other words, it is not the case that the lower ⟨John⟩ is indeed a copy, while *John* in the higher position is something else other than a copy (although the former is often notationally distinguished by the angle bracket for expository convenience). Chomsky (2008: 140) thus writes that "IM yields *two copies* of Y in {X, Y}, one external to X, one within X" (emphasis on *two* is added).

Notice, however, that there arises some inconsistency between Chomsky's claim about the status of *John* and ⟨John⟩ in (12) and his claim of the invisibility of copy

15) The necessity that a moved element should leave behind an element identical to it can follow from the No Tampering Condition:

No Tampering Condition (from Narita 2011, originally proposed in Chomsky 2008)
No elements introduced by syntax are deleted or modified in the course of linguistic derivation.

to LA. According to the former, *John* is as much a copy as ⟨John⟩ is, but the latter distinguishes between the two, claiming only ⟨John⟩ is invisible to LA. To resolve this inconsistency, I propose (13).

(13) **Copy Invisibility** (to LA)
All copies are invisible to LA.

According to (13), copies are *all* visible to LA regardless of their being the head or the (intermediate) trace.[16] To highlight invisibility of (all the) copies to LA in our system, I will indicate copies by way of outline font as in (14).

(14) John ... { John ...}

Our second assumption concerns the target of LA, i.e. what LA searches for. In the POP(E) model LA seeks two types of feature, i.e. the categorial features and any agreeing features.[17] I have shown in section 3.1 that the latter search, which I called a 'comparison' search, complicates LA in that it forces LA to perform an additional search besides Minimal Search. I thus propose (15) which reduces computational burden on LA by rending comparison search unnecessary.

(15) LA seeks only the categorial feature of LIs.[18],[19]

---

16) An anonymous reviewer pointed out that the proposal (13) may be as much of a stipulation as Chomsky's (strong vs.) weak T given that only the closest copy to the probe typically becomes the target for movement, i.e. the highest copy (or the 'head' (of a chain) in traditional terms) is considered (somehow) special. Though I acknowledge the potential problem with (13), I, unfortunately, leave further justification for (13) for future research, pointing out only that there is some case where syntactic operations such as movement seem to ignore the closest head (of a chain) to them:

(i) C ... T ... [$_{v^*P}$ what [John ... <what>]] 'what did John buy?'

As discussed in Chomsky (2008), T in (1) targets John and attracts to its Spec, despite the presence of the (intervening) closest head what in the outer Spec of v^*.

17) In fact, it is not explicit in Chomsky's work what type of features LA is sensitive to. Instead, it is stated as follows:

LA seeks features [...] perhaps seeks only features [...] [O]nly certain features can serve as labels. (Chomsky 2013: 45)
LABEL [i.e. Labeling Algorithm] locates a *feature* of H. (Chomsky *et al.* 2017: 22, italics in the original)

18) An anonymous reviewer pointed out that computational burden alone does not seem enough to justify the proposal (15). Note, however, that what I claim by (15) is that it reduces computational

A conclusion that naturally follows from (15) is that a head can serve as a label in our system only if it bears a categorial feature. This conclusion in turn has two implications for the labelability of heads: First, recall that in the POP(E) model the root R is merely 'stipulated' to be unlabelable. Our system, however, can give an account of why that is or should be the case, i.e. R is unlabelable because it is a head that lacks a categorial feature. I follow Embick and Noyer (2007), Borer (2009) and Embick (2012) in assuming that this category-lacking R is assigned a categorial feature when it combines with a category-defining functional head such as *v**.[20] The second implication that follows from (15) is that *contra* Chomsky (2013, 2015), T is invariably labelable regardless of its finiteness.

To sum up this section, I list in (16) the two assumptions discussed and what follows from them.

(16) a. **Copy Invisibility**

All copies are invisible to LA.

b. **LA seeks only the categorial feature.**

- R is unlabelable unless it combines with a category-defining head.
- T is invariably labelable.

---

burden on LA 'as a result', not that considerations of computational burden can be a 'justification' for the proposal.

19) One may wonder, as an anonymous reviewer pointed out, what would be behind the uniqueness of categorial features when it comes to labeling. I believe that some feature needs to be visible to LA if labeling of SO will ever to be possible. In Chomsky's (2013, 2015) framework, the relevant features are the (undefined) 'prominent' features in addition to the categorial features. What I propose instead is that only the categorial features are relevant, which is not only what Chomsky (2013: 37) claims ("identification of the *category* of a phrase (projection, more recently called "labeling" [italics added]) but also what has been conventionally assumed (see also discussion in Introduction). The proposal in (15) may be viewed as our working hypothesis at the least.

20) Embick and Noyer (2007: 296) propose the following:

Categorization Assumption

Roots cannot appear without being categorized; Roots are categorized by combining with category-defining functional heads [such as *v* and *n*].

Embick (2012: 74) also states, "The Roots are by definition acategorial, [...] there is a subtype that *categorizes* Roots: these are called *category-defining* heads; by definition, varieties of *v*, *n*, and *a*." (emphasis on 'categorizes' is added; others in the original). Also, Borer (2009) writes, "Roots [...] are categorized contextually by their merger environment. Consequently, *bare roots, without a category label, are not a syntactic option.*" (Italics added). Although, as an anonymous reviewer correctly pointed out, other researchers such as Borer (2014) claim otherwise, I follow with Embick (and Noyer) in assuming that R is (derivationally) categorized by the phase head *v**.

In the next section I present an alternative labeling process in which all the issues discussed in section 3 are resolved.

4.2. Labeling without ⟨φ, φ⟩

To see how labeling can proceed without creation of ⟨φ, φ⟩, let us first consider (17).

(17) a. John bought a car. (Compare with (3))

b. $\{_{\alpha}$ $[_{NP}$ John], $[_{v^*P}$ v^* ...]}

c. $\{_{\gamma}$ John $\{_{\beta}$ T $\{_{\alpha}$ John $\{v^*$ ... }}}}

At some point in the derivation of (17a), the unlabelable SO=α in (17b) is generated. (17c) illustrates the structure where the NP *John* undergoes IM to Spec-T, which results in creating the two 'copies' of *John* in the structure. What this IM of *John* means in our system is that both the higher and the lower *John* become irrelevant to labeling process as we assume that copies are 'all' invisible to LA (see (13) and (16a)).[21] With this in mind, let us examine how the labeling process of the three SOs (i.e. α, β, and γ) in (17c) proceeds. First, the label of α is identified as v^* since the copy *John* in Spec-v^* is irrelevant in terms of labeling. Second, β is labeled T as T with its categorial feature is a legitimate, inherently labelable head in our analysis. Finally, the label of γ is (also) identified as T because *John* in Spec-T, just like *John* in Spec-v^*, is a copy and hence invisible to LA. Notice that the cause of the EPP-effect of T (i.e. the obligatory IM of *John* to Spec-T) remains unchanged; IM of *John* is driven, as assumed in the POP(E) model, by labeling failures in α.[22] The (only) major difference between the POP(E) model and our alternative analysis

21) An anonymous reviewer pointed out that the operation Move (or IM in more recent terms) would be unconstrained if it can serve as a mechanism to make SO immune to labeling as I suggest. In fact, I do believe that it is indeed the case, i.e. IM is unconstrained or '*free*' in the sense of what Chomsky (2015) and Chomsky et al. (2017) claim:

> "the lingering idea [...] that each operation has to be motivated by satisfying some demand. But there is no reason to retain this condition. Operations can be free, with the outcome evaluated at the phase level for transfer and interpretation at the interfaces." (Chomsky 2015: 14 [italics added])
>
> "MERGE thus applies *freely*, generating expressions that receive whatever interpretation they are assigned by interfacing systems." (Chomsky et al. 2017: 11 [italics added])

22) Our analysis suggests that the function of the operation Internal Merge could be understood to render SO invisible to LA. I will not pursue this interesting suggestion any further, however. For similar roles of suffixal Case in rendering SO invisible to LA, see Saito (2016).

is that the former stipulates only a certain copy (e.g. the lower *John*) is invisible to LA, while all copies are so in our analysis.

Let us now consider the labeling process within the *v**P domain in our analysis.

(18) a. John bought a car. (Compare with (4))

b. $\{_\alpha$ R(=bought), a car$\}$

c. $\{v^*$, $\{_\alpha$ R, a car$\}\}$ (α = R)

d. $\{v^*$, $\{_\beta$ a car, $\{_\alpha$ R, a car$\}\}\}$ (α, β = R)

At the derivational stage of (18a), the unlabelable SO=α in (18b) is generated. The reason for the unlabelability of α is straightforward in our analysis; the closest head R bears no categorial feature at this stage of the derivation and thus cannot serve as a label. However, once this category-lacking R is combined with the category-defining functional head *v** as shown in (18c), it turns into a labelable head and can serve as a label.[23] Consequently, α is labeled R.

Note that IM of NP *a car* to Spec-R is neither required nor prohibited in our system to assure the labeling of SO=$\{$R, NP$\}$. In other words, $\{$R, NP$\}$ can be labeled R with or without IM of NP, as illustrated in (18c) and (18d), respectively.[24]

Finally, let us consider the raising constructions which pose a serious challenge to Chomsky's weak T analysis (see (9)).

(19) a. John seems to like Mary.

b. $\{$C $\{_\delta$ John $\{_\gamma$ T_2-seems $\{_\beta$ ⟨John⟩ $\{_\alpha$ T_1-to $\{_{v^*P}$ ⟨John⟩ like Mary$\}\}\}\}\}\}$

c. $\{$C $\{_\delta$ John $\{_\gamma$ T_2-seems $\{_\beta$ John $\{_\alpha$ T_1-to $\{_{v^*P}$ John like Mary$\}\}\}\}\}\}$

As discussed in section 3.2, α and β in (19b) are the SOs causing problems for labeling in the POP(E) model since no agreement between T_1 and (the head of) *John* leaves T_1 weak to serve as a label. Our system, however, can readily account for those labeling issues: since T_1 with its own categorial feature can serve as a label and John in its Spec is invisible to LA, the label of α and β is both identified as T_1.

Notice that (19c), where *John* undergoes IM 'successive-cyclically' through every Spec-T, is not the only conceivable derivation for (19a) in our analysis. Consider (20).

---

23) It may be the case that a process analogous to 'feature inheritance' (Chomsky 2008) is involved in assignment of a categorial feature to R. However, I put aside this issue for future research.

24) Chomsky (p.c.) also suggests that 'IM of NP to Spec-R does hold for Exceptional Case Marking constructions but is not clear for R-NP constructions.'

(20) {C $\{_{\delta}$ John $\{_{\gamma}$ T_2-seems $\{_{\alpha}$ T_1-to $\{_{v^*P}$ John like Mary}}}}}}

Since T is an inherently labelable head in our system, no labeling problems arise even if *John* moves from Spec-*v** to the matrix Spec-T_2 in one fell swoop as illustrated in (20), an alternative derivation suggested and proposed in Chomsky (2001) and Mizuguchi (2017), respectively.

5. Remaining Problems and Conclusion

This paper focused on the paired form of label, ⟨φ, φ⟩, proposed in Chomsky (2013, 2015) and addressed some theoretical and empirical issues regarding its validity. For the theoretical issues, I showed that 1) ⟨φ, φ⟩ raises problems in terms of its interpretability and that 2) creation of it complicates the grammar by forcing LA to perform an additional comparison search. We discussed the empirical issues by showing that the claim of the weak T (in English) in the POP(E) model is untenable especially when we consider raising constructions. I finally presented an alternative analysis in which LA is reinterpreted as being sensitive only to the categorial feature of heads. Our analysis not only reduces computational burden on the operation LA but also provides a more principled account of the claim of the root R being universally weak as a label.

Many other related issues remain to be resolved, of course, but I conclude this study by briefly pointing out one of the issues which I think should be addressed in future research. The issue is that, although I showed in detail that the interpretive contribution of the label ⟨φ, φ⟩ to the CI interface is questionable, the proposed role the label plays in the narrow syntax was not addressed in our discussion. That is, in addition to its role at the interface, the label ⟨φ, φ⟩ in the POP(E) model is assumed to play an important role in indicating what Rizzi (2016) calls the 'criterial positions' in syntactic computation. According to this assumption, the NP that undergoes IM is 'frozen' and becomes unavailable for further movement operations once it reaches a (specifier) position of SO whose label is identified as ⟨φ, φ⟩. If this turns out to be indeed the case, our future research should provide an alternative solution.

References

Borer, Hagit. (2005). *In Name Only: Structuring Sense vol I*. Oxford: Oxford University Press.

Borer, Hagit. (2009). Roots and categories. Handout of the 19th Colloquium on Generative Grammar, University of the Basque Country, April, 1-3.

Borer, Hagit. (2014). The category of roots. In A. Alexiadou, H. Borer and F. Schäfer, eds., *The Roots of Syntax, the Syntax of Roots*, 91-121. Oxford: Oxford University Press.

Bošković, Željko. (2013). Phases beyond clauses. In Lilia Schürcks, Anastasia Giannakidou, and Urtzi Etxeberria, eds., *The Nominal Structure in Slavic and Beyond*, 75-128. Berlin: de Gruyter.

Chomsky, Noam. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. (1993). A minimalist program for linguistics theory. In K. Hale and S. J. Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sivain Bromberger*, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. (1995). *The Minimalist program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. (2000). Minimalist Inquiries: The Framework. In R. Martin, D. Michaels and J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-156. Cambridge, Mass.: MIT Press.

Chomsky, Noam. (2001). Derivation by Phase. In Michael Kenstowicz, ed., *Ken Hale: A Life in Language*, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. (2008). On phases. In R. Freidi, C. P. Otero and M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133-166. Cambridge, Mass.: MIT Press.

Chomsky, Noam. (2013). Problems of projection. *Lingua* 130, 33-49.

Chomsky, Noam. (2015). Problems of projection: Extensions. In Elisa Di Domenico, Comelia Hamann and Simona Matteini, eds., *Structures, Strategies and Beyond - Studies in Honor of Adriana Belletti*, 3-16. Amsterdam & Philadelphia: John Benjamins.

Chomsky, Noam, Ángel J. Gallego and Dennis Ott. (2017). Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. Retrieved from http://ling.auf.net/lingbuzz/003507. Revised version to appear in *Catalan Journal of Linguistics*.

Embick, David. 2012. Roots and features (an acategorial postscript). *Theoretical Linguistics*, 38(1-2), 73-89.

Funakoshi, Kenshi. (2014). *Syntactic Head Movement and Its Consequences*. Doctoral dissertation. University of Maryland, College Park.

Grohmann, Kleanthes K. (2000). *Prolific Peripheries: A Radical View from the Left*. Doctoral dissertation. University of Maryland, College Park.

Grohmann, Kleanthes K. (2003). *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam: John Benjamins.

Grohmann, Kleanthes K. (2011). Anti-locality: Too-close relations in grammar. In Cedric

Boeckx, ed., *The Oxford Handbook of Linguistic Minimalism*, 260-290. Oxford: Oxford University Press.

Halle, Morris and Alec Marantz. (1993). Distributed Morphology and the pieces of inflection. In Hale, Kenneth and Samuel Jay Keyser, eds., *The View from Building 20*, 111-176. Cambridge, Mass.: MIT Press.

Mizuguchi, Manabu. (2017). Labelability and interpretability. *Studies in Generative Grammar*, 27-2, 327-365.

Narita, Hiroki. (2011). *Phasing in Full Interpretation*. Doctoral dissertation. MIT.

Narita, Hiroki. (2015). *{*t*, *t*}. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo G. Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara, and Iris C. Ouyang, eds., *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, 286-295.

Naria, Hiroki and Naoki Fukui. (2016). Feature-equilibria in syntax. In Fujita, Koji and Cedric Boeckx, eds., *Advances in Biolinguistics: The Human Language Faculty and Its Biological Basis*, 9-28. New York: Routledge.

Ouali, Hamid. (2008). On C-to-T feature transfer: The nature of agreement and anti-agreement in Berber. In Roberta D'Alessandro, Gunnar H. Hrafnbjargarson and Susann Fischer, eds., *Agreement Restrictions*, 159-180. Berlin: Mouton de Gruyter.

Pesetsky, David and Esther Torrego. (2007). The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian, and Wendy Wilkins, eds., *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, 262-294. Amsterdam: Benjamins.

Rizzi, Luigi. (2016). Labeling, maximality and the head-phrase distinction. *The Linguistic Review*, 33(1), 103-127.

Saito, Mamoru. (2016). (A) Case for labeling: labeling in languages without φ-feature agreement. *The Linguistic Review* 33(1), 129-175.

Stockwell, Richard. (2014). *The End of Endocentricity*? Ms., University of Cambridge.

Stockwell, Richard. (2016). Labeling in Syntax. *Cambridge Occasional Papers in Linguistics* 9, 130-155.

Jae-Young Shim
Department of English Language and Literature
Pukyong National University
45, Yongso-ro, Nam-gu, Busan 48513, Korea
E-mail: biolinguistics@gmail.com

Received: January 30, 2018
Revised version received: March 23, 2018
Accepted: April 17, 2018