10.15860/sigg.27.2.201705.327.

Labelability and Interpretability*

Manabu Mizuguchi
(Toyo University)

Manabu Mizuguchi. 2017. **Labelability and Interpretability**. *Studies in Generative Grammar,* 27-2, 327-365. Syntactic objects are constructed by Merge in its simplest form and are labeled by the Labeling Algorithm. This paper considers the labelability of heads and proposes a principled explanation by arguing that it is reducible to interpretability at the interfaces. This proposal removes unwanted stipulations on labeling assumed in Chomsky (2015) and deduces the labelability of T in rich agreement languages as well as that of T and V (i.e., R categorized by *v**/*v*) when they are pair-merged to phase heads; the proposal also suggests that there are other unlabelable heads and it is shown that *n* is an unlabelable head. The paper reconsiders EPP, arguing that it is not attributable to labeling but to externalization. It is concluded that the discussion upholds the Strong Minimalist Thesis, the basic hypothesis in minimalism.

Keywords: labelability, unlabelable heads, unvalued features, EPP, externalization

1. Introduction

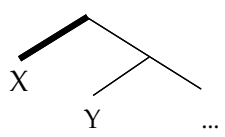
The Minimalist Program (Chomsky 1995, especially Chomsky 2000, 2004 and his subsequent works) assumes that language is an optimal system meeting interface conditions (i.e., Full Interpretation) in a way satisfying general principles of

* I would like to thank three anonymous *SIGG* reviewers for helpful comments and valuable suggestions on an earlier version of this paper, which have helped me improve the quality of the paper. This paper was presented in earlier form at the 7th International Conference on Formal Linguistics (ICFL-7), held at Nankai University, Tianjin, China, on December 2-4, 2016. I also thank the audience for comments and discussion. All remaining errors and inadequacies are my own. The research reported in this paper was supported by Japan Society for the Promotion of Science under Grant-in-Aid for Young Scientists (B) (#16K16835).

minimal computation or "third factor" principles. This assumption, which is intended to achieve principled explanation deeper than explanatory adequacy, is articulated by the Strong Minimalist Thesis (SMT) and is a cornerstone hypothesis of the current minimalist research. Given third factor principles, Merge, the basic operation in syntax, is formulated in its simplest form (hence Merge operates as "simplest Merge"). One important consequence of simplest Merge is that contrary to what has been standardly assumed, mother nodes and endocentric projections are not properties of phrase structure as Merge conforms to the Condition of Inclusiveness (Chomsky 1995: 228), one of the third factor principles, and expressions or syntactic objects (SOs) yielded by simplest Merge are without labels. At the Conceptual-Intentional (CI) interface as well as for the processes of externalization, however, SOs must be labeled; that is, they must be identified (i.e., for interpretation, CI and externalization processes must know what kind of object a derived SO is). Chomsky (2013, 2015) proposes that the identification or labeling is carried out independently by the Labeling Algorithm (LA), which is just one instance of minimal search. As illustrated in (1) and (2), when applied to the SO marked as α, LA locates through minimal search the closest head (in the case of a set of the form X-YP) or agreeing heads (in the case of an XP-YP set), determining the properties of the SO and providing necessary information about it for its CI interpretation and externalization:[1]

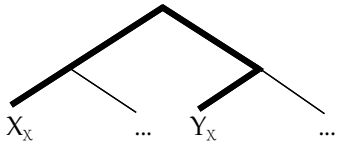
(1) a. $[_{\alpha}$ X [Y ...]] (α = X)

b.



(2) a. $[_{\alpha}$ $[X_{\chi}$ ...] $[Y_{\chi}$ ...]] (α = <χ, χ>)

b.



With this background in place, in this paper, I consider labelability and explore its principled explanation under the minimalist setting. In the current discussion on labeling, the labelability of heads is simply assumed or stipulated; a question remains why heads are labelable. Moreover, not all heads can label; for instance, Chomsky (2015) says that T and R by their nature cannot label. I

[1] Bold lines in (1)/(2) and elsewhere in this paper represent minimal search by LA. Remember that tree notations are quite misleading; Merge simply yields a set {α, β} (alternatively represented as [α, β] for convenience) or an ordered pair <α, β>. Only for expository purposes, tree notations are used here, under the assumption that sets and their familiar tree notations are equivalent. In this paper, angle brackets are used for ordered pairs (i.e., SOs formed by pair-Merge) as well as for labels of XP-YP structure identified by LA.

propose that the labelability of heads is reducible to interpretatibility, arguing that it receives a principled explanation in terms of the interfaces. I also submit that labeling does not explain EPP. The discussion in this paper demonstrates that the two proposals bring theoretically and empirically favorable consequences.

The structure of this paper is as follows. In section 2, as our starting point, I discuss Chomsky (2015) and consider two assumptions on labeling introduced in the literature. In section 3, I spell out my proposals, showing that a principled explanation can be given to labelability. In sections 4 and 5, I consider theoretical and empirical consequences of my proposals. In section 4, I discuss labelable T and in section 5, I argue that *n* is also an unlabelable head. In section 6, I summarize and conclude the paper.

2. Weakness of T as a Label

In the discussion of labeling, Chomsky (2015: 9) introduces the following two assumptions: (i) T is too weak to label; and (ii) T can label after strengthening by Spec-T: something must be in Spec-T for T to work as a label.[2] To see this, consider (3):

(3) [C [$_{\lambda}$ T_{ϕ} [$_{\alpha}$ *n*P [*v** [...]]]]] (λ = unlabeled)

Applied to the T-headed set, which is conveniently marked as λ in (3), minimal search LA can unambiguously locate T as the label (see (1)); however, since T is not qualified to label for its weakness as a label, the set will not be labeled. Chomsky says that if the *n*P is internally merged with λ (to put it informally, if it moves to the Spec of T) and Spec or XP-YP is created as in (4), then T will strengthen and can label:[3]

(4) [C [$_{\gamma}$ *n*P$_i$ [$_{\lambda}$ T_{ϕ} [$_{\alpha}$ t_i [*v** [...]]]]]] (λ = T)

Chomsky (2015: 9-10) claims that Spec-T must be visible when LA applies.

---

[2] The same argument applies to the root R (or V); if the *n*P moves to Spec-R, however, R strengthens and can work as a label. In this paper, I mainly discuss T, keeping in mind that the same argument applies to R. I consider R in section 3.3.

[3] t_i is intended as a copy (i.e., a lower occurrence created by movement or Internal Merge, IM) of the i-indexed SO. In this paper, t_i is used only for expository purposes since indexed traces cannot be produced by simplest Merge for inclusiveness.

The two assumptions on labeling, Chomsky argues, deduce from labeling subject EPP, a long-standing problem since its first introduction in the early 1980s, which stipulates that the clause must have a subject (Chomsky 1981). In (5), the λ-marked set will be left unlabeled (→ (3)) and cannot be interpreted at CI or externalized, with the result that Full Interpretation is violated and the sentences are ruled out as ill-formed at the interfaces:[4]

(5) a. *[$_{\lambda}$ Has [$_{\alpha}$ the student read the book]]
b. *[$_{\lambda}$ Was [$_{\alpha}$ read the book by the student]]

Though subject EPP may follow from (i) and (ii), *why*-questions remain on the relevant assumptions. One is: why is T, unlike other heads, too weak to label? Chomsky says that T is like roots R, which cannot serve as a label (see footnote 2). This does not answer the question, however, unless the inability of roots to label is given a principled explanation. Moreover, the labeling assumption (i) is challenged empirically. As evidenced by (6), raising T can stand without visible Spec-T:

(6) a. The student$_i$ seems [$_{\lambda}$ to [be likely [$_{\lambda}$ to [t$_i$ understand the theory]]]]
b. There is likely [$_{\lambda}$ to [be a student in the auditorium]]

This argues that raising T can serve as a label and that the λ-marked set is indeed labeled (λ = T) (see Epstein, Kitahara, and Seely 2014 for relevant discussion). If T in (3) (that is, finite T) cannot serve as a label, a further question will be raised: why can T in (6) or raising T serve as a label?

The second question is: why can Spec-T (or creation of XP-YP) strengthen, hence qualify, T as a label? If T cannot label and the λ-marked set cannot be labeled in (3), then it is more reasonable to assume that the set cannot be labeled in (4), either, for weakness of T as a label. Chomsky only states that T strengthens by Spec-T. It might be argued that T strengthens by Spec-T through labeling of γ, which is labeled <ϕ, ϕ> by minimal detection of agreeing *n* and T heads. However, this simply states in different terms what needs to be explained: why can such labeling qualify T as a label? Given the labeling assumption (i), it is more reasonable to assume that the γ-marked set cannot be labeled: if one of the agreeing heads (in this case, T) is too weak to serve as a label and cannot

[4] Note that in (5a), the α-marked set, which is XP-YP created by merge of the external argument (EA) with the *v**-headed set, will not be labeled, either, because the head of EA and *v** do not agree.

be a candidate for labeling, then it should be the case that ɣ cannot be labeled by minimal detection of *n* and T heads. In fact, it will be wrongly labeled (ɣ = *n*) by minimal search LA and will be ruled out at the interfaces: in [$_{ɣ}$ [*n* ...], [T ...]], since T cannot label for its weakness when located by LA, the only candidate for labeling will necessarily be *n* and LA will locate *n* and take it as the label of ɣ (see also Hiraiwa 2016, Takano 2015 for relevant discussion). It is a stipulation to say that T can label thanks to Spec-T, which simply redescribes subject EPP in terms of labeling, and is not a principled explanation.

3. Labelable and Unlabelable Heads

In the last section, I have argued that the labeling assumptions (i) and (ii) are simply stipulated and cannot be deduced from deeper principles of language; moreover, they are empirically problematic as well. In this section, I address *why*-questions regarding (i) and (ii). In fact, the answers to the questions rest, in the first place, on the answer to the question why heads can label. I argue that the labelability of heads is reducible to interpretability, deducing from the interfaces why some heads are labelable while others are not.

3.1. Labelability

I claim that label weakness or unlabelability is attributable to unvalued features borne by heads. Here I propose the following:

(7) Heads can label only when they are without unvalued features.

I submit that heads with unvalued features are unlabelable. Unvalued features cause weakness as labels because such features are uninterpretable: unvalued features are unspecified and hence, illegible to the interfaces, violating Full Interpretation. If labels are required to identify SOs for interpretation at CI as well as for the processes of externalization (hence for Full Interpretation at the interfaces), they must be interpretable and heads with unvalued features, as they are uninterpretable, cannot serve as labels. Label weakness is reducible to independently motivated and arguably ineliminable unvalued features borne by heads.

The proposal gives a principled explanation to the labeling assumption that T is too weak to label. The answer to this question is straightforward: T inherits

unvalued ϕ-features from C and bears them as its features ([C_{ϕ} [T [...]]] → [C [T_{ϕ} [...]]]). Furthermore, the labelability of raising T noted in (6) also follows naturally from the proposal. Recall that in the raising construction, the λ-marked set does not cause labeling failure even though Spec-T is not visible, which implies that raising T, unlike finite T, can indeed label. In the raising construction, C, a phase head, is not merged in and T in the raising complement does not inherit unvalued ϕ-features.[5] Since it does not bear any unvalued features, it is interpretable and can serve as a label without any problems when located by minimal search LA. Given (7), the question does not arise why finite T is not labelable while raising T is.

The proposal can also explain labeling in under-inheritance. Legate (2011, 2014) argues that in subject-initial Verb Second and in subject *wh*-movement, feature inheritance fails to apply and C keeps ϕ, which leaves the head as the source of Case and agreement as well as the locus of EPP (see also Mizuguchi 2016b and Ouali 2008 for relevant discussion). The derived structure is illustrated as follows:

(8) [$_{\delta}$ Sub_i [C_{ϕ} [$_{\lambda}$ T [$_{\alpha}$ ... t_i ...]]]]

Under Chomsky's assumptions, λ will be left unlabeled as T is too weak to serve as a label and visible Spec is not created. Thus, if Legate's argument is on track and the subject moves directly to Spec-C in subject-initial Verb Second and in subject *wh*-movement, under-inheritance will pose another problem to Chomsky's labeling assumptions.[6]

The problem does not arise under the proposal in this paper. In (8) as well as in (6), T does not bear unvalued ϕ and it is interpretable in the absence of unvalued features. Hence, T can be a candidate for a label and λ can be labeled when minimal search LA locates the head.

---

[5] Alternatively, C is indeed merged in the derivation but is simply de-activated together with its ϕ-features due to its pair-merge to T (Mizuguchi 2016a). See section 4.2 for relevant discussion.

[6] Given simplest Merge, the derivation in (8) will be predicted for the subject in general. As discussed in Mizuguchi (2016b), however, with non-V2/non-*wh* subjects, it will be ruled out independently by the CI interface as it results in ill-formed interpretation. An anonymous reviewer notes that if (8) is not possible for non-V2/non-*wh* subjects, it will be necessary to delineate when feature inheritance takes place and when it does not. In fact, it is not necessary to do so since operations, including feature inheritance, apply freely (they are free to apply or not to apply in syntax) and there is no need to stipulate when certain operations apply and when they do not; the outcome of syntactic derivation is only ruled in or out at the interfaces.

3.2. Labeling and Spec

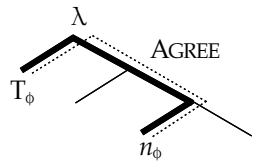
If unvalued features are the cause of weakness as labels, the proposal suggests that once unvalued features are valued by Agree, heads can label: thanks to valuation, no unvalued features will be present on heads. In the case of T, it can label once its unvalued ϕ-features are valued. This explains why the λ-marked set can be labeled in (4), which is repeated below: the ϕ-features of T have been valued by the time LA applies and there are no unvalued features present on T:

(4) [C [$_{\gamma}$ *n*P$_i$ [$_{\lambda}$ T$_{\phi}$ [$_{\alpha}$ t$_i$ [*v** [...]]]]]] (λ = T)

It also follows that the γ-marked set can be labeled: since T can label thanks to valuation of its unvalued ϕ-features via agreement, both *n* and T are labelable heads, and hence γ can be labeled <ϕ, ϕ> when minimal search LA detects the agreeing heads.

The argument that valuation makes unlabelable heads labelable might face a problem. Take T. The proposed analysis suggests that T can label without Spec-T; λ can be labeled in (3) as well as in (4). It has been standardly assumed in the literature that Agree can be executed in a long-distance manner and that valuation can take place in situ (Chomsky 2000). Operations apply freely as far as third factor principles are abided by and in the absence of a stipulation blocking it, this also holds true for Agree, and long-distance Agree follows as a straightforward consequence of the minimalist view of language design (see Carstens 2016 and references cited therein for relevant discussion); agreement or valuation is thus not restricted to Spec-head (cf. Kitahara 2017 for an alternative proposal). As illustrated in (9), applied to the set λ, Agree, which is nothing other than minimal search just like LA, locates the probe T$_{\phi}$ and the goal n_{ϕ}, and ϕ-feature agreement, hence valuation of T's ϕ, can take place in situ without Spec-head:

(9)



The λ-marked set will be labeled by (now labelable) T even if no *n*P is merged with the set. In examples like (5) as well as in (4), labeling is not problematic

and the second labeling assumption can be removed. Under the proposed analysis, contrary to Chomsky (2015), the creation of Spec-T or EPP is not reducible to labeling.

I maintain that labeling is irrelevant to EPP; instead, EPP is reducible to externalization, hence deducible from Full Interpretation on the side of the Sensory-Motor (SM) interface. Following Mizuguchi (2016b), I argue that Spec-T works as a syntactic instruction to the externalization component that agreement is spelled out canonically at T. Consider (10):

(10) $[_{\sigma}\ C\ [_{\chi}\ nP_{\phi}\ [_{\lambda}\ T_{\phi}\ [_{\alpha}\ \ldots\ (t)\ \ldots\ ]]]]$

↓

CANONICAL SPELL-OUT

This argument is empirically endorsed by examples in (11) and (12), which demonstrate that non-canonical (that is, partial/default) agreement, instead of full-fledged agreement, appears when subjects are post-verbal and do not move to Spec-T:

(11) a. ʔakal-at (/*ʔakal-na) ṭ-ṭaalibaat-u. *Standard Arabic*
eat.past-3.fem.sg eat.past-3.fem.pl the-students.fem.pl-nom
'The students ate.'
b. ṭ-ṭaalibaat-u ʔakal-na(/*ʔakal-at). (Benmamoun 1992)

(12) a. E' vegnú (/*L'è vegnuda) qualche putela. *Trentino*
is come they are come some girls
'Some girls have come.'
b. La Maria la parla.
the Maria she speaks
'Maria speaks.' (full agreement) (Brandi and Cordin 1989)

The argument is further supported by Guasti and Rizzi (2002: 175-176), who point out the following generalization, which they argue is a rather stable generalization: when the subject occupies a surface position in the higher parts of the inflectional system, typically higher than the inflected verb − presumably in the Spec of agreement or higher as in *[Subj* $Infl_{AGR}$ *...]*, the morphological expression of agreement is compulsory (provided that the language is equipped with the relevant morphology); on the other hand, if the subject stays in

VP-internal or in the lower part of the inflectional system, typically lower than the inflected verb – presumably, lower than the agreement layer as in *[... $Infl_{AGR}$ ... Subj ...]*, then agreement may or may not be morphologically expressed. Obligatory morphological expression of agreement can be taken as full-fledged or canonical agreement.

With (10) in place, suppose that in English, unlike in languages like Standard Arabic and Trentino, non-canonical agreement cannot be realized as the language is not equipped with the relevant morphology and that only canonical agreement is morphologically available. Then unless the *n*P is merged with λ and Spec-T is created, ϕ-features on T will not be externalized and Full Interpretation will be violated at the SM interface.[7] Syntactic agreement by Agree renders T labelable through valuation of unvalued ϕ-features, and labeling of λ is not problematic at all in (3). Externalization, however, will fail unless Spec-T is created to make canonical externalization possible. Given the theory of labeling, unless an SO agreeing with T moves to Spec-T, {{Spec}, {TP}}, an XP-YP set, will not be labeled in the absence of agreeing features even if Spec is produced. Hence, Spec-T is yielded by an SO that agrees with T in ϕ, not by any other SOs.

The proposed analysis also explains in a principled way languages with forced Spec-head agreement. For instance, consider the following examples from Kirundi (13) and Kilega (14):

(13) a. Ibitabo bi-á-somye Yohani.
books 3pl-past-read:perf John
'John (not Peter) has read (the) books.'
b. Yohani a-á-ra-somye ibitabo.
John 3sg-past-f-read:perf books
'John has read books.' (Ndayiragije 1999)

(14) Ku-Lúgushwá kú-kili ku-á-twag-a nzogu maswá.
17-Lúgushwá 17.sa-be.still 17.sa-a-stampede-fv 10elephant 6farm
'At Lugushwa elephants are still stampeding over the farms.' (Carstens 2011)

In (13a) and (14), morphological agreement is determined by preposed SOs, not by the subject. Kinyalolo (1991) and Ndayiragije (1999) argue that in Bantu languages, SOs controlling subject agreement are in Spec-T, not in any other

[7] For relevant discussion, see Richards (2016), who argues that subject EPP is attributable to morphological and phonological properties of languages (i.e., the properties imposed by the SM interface).

positions; in other words, Spec-head agreement is forced: T must agree with an SO in Spec-T. Given the theory of Agree, however, ϕ-feature agreement can occur in a long-distance manner even if an SO is not in Spec-T. This problem can be solved if morphological agreement in Bantu languages is canonical agreement and just as in English, only canonical agreement is morphologically available. As I have discussed, if a language cannot spell out non-canonical (partial/default) agreement for lack of its morphology, Spec-T must be created so that agreement can be externalized canonically (→ (10)). In syntax, ϕ-feature agreement is executed non-locally without Spec-head; however, unless an agreeing SO moves to Spec-T, externalization will fail and interface conditions imposed by the SM system cannot be satisfied. In Bantu, T can agree with non-subjects.[8] In (13a) and (14), if non-subjects agree with T in ϕ, they will move to Spec-T for canonical externalization, which makes it look as if Spec-head agreement takes place. Spec-head agreement in Bantu languages follows from (10) and labeling of XP-YP.

The argument that morphological realization of ϕ is not syntactic but a property of externalization is in line with the minimalist assumption that syntax (e.g., abstract agreement) is uniform across languages while externalization (e.g., morphological spell-out) is the locus of variations, expressing differences between and within languages (Chomsky 2001).

I have argued above that Spec-T instructs to the externalization component that canonical agreement is realized on T. I claim that this, in turn, follows from labeling. In (10), the creation of Spec-T yields ɣ (i.e., a set of the form XP-YP), which, as I have argued, is labeled <ϕ, ϕ> by minimal detection of agreeing ϕ in *n* and T (=(2)). In the case of XP-YP or ɣ, agreeing heads provide the same label, which is taken as the label of ɣ: in terms of labeling, features shared by agreeing heads work as a unique head. Notice that mere matching of most prominent features does not suffice for labeling and that what is required is agreement, a stronger relation (Chomsky 2013). If so, values of *n*'s ϕ and those of T's ϕ must be identical in the label; otherwise, the two ϕ's would not provide the same label, with the result that ɣ would not be labeled by the agreeing

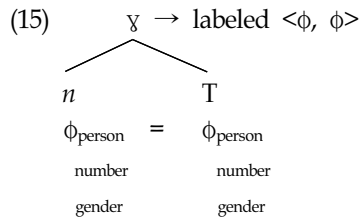
---

[8] Take (13) and consider how non-subject agreement comes out. In the course of the derivation, the object moves to Spec-*v** in order to move out of the *v**P phase. Since the object gets closer to T than the subject, it can agree with T:

(i) [C [T_{ϕ} [$_{v^*P}$ Obj_i [Subj [*v** [... t_i ...]]]]]]

In English, on the other hand, this derivation will be ruled out as it leaves unvalued the Case feature of the subject. In Bantu, Case is lacking altogether (Diercks 2012) and this problem does not arise. For relevant discussion, see section 5.

heads (for expository purposes, irrelevant details are omitted from (15)):

(15) 

With labeling of ɣ in mind, when the label information is processed by the externalization component to realize T's ϕ-features, the features will be spelled out in accordance with *n*'s ϕ-features for identical heads giving the label. Since the values of *n*'s ϕ-features are lexically determined and fully specified and the values of T's ϕ-features depend on those of the *n*P it agrees with, T's ϕ-features will be realized in a full-fledged manner, hence canonically spelled out when subject to externalization. Given that labels are required for externalization, labeling of ɣ or its label information <ϕ, ϕ> provides a principled explanation to (10).

Summarizing the discussions in sections 3.1 and 3.2, I have argued (i) that labelability is reducible to interpretability at the interfaces and (ii) that EPP follows from externalization, not from labeling. The proposals give principled explanations to the labelability of heads as well as to EPP, and remove the unwanted assumptions on labeling introduced by Chomsky (2015).

3.3. Unlabelable R

The proposed analysis can also explain the labelability of R (or V). As noted, R, like T, cannot label. Given my proposal here, weakness of R as a label is straightforwardly deduced: it follows from the fact that R inherits unvalued ϕ-features from *v**. R and T are alike in their labelability because they inherit unvalued features from phase heads (*v** and C, respectively).

Recall that raising T can label because it does not inherit ϕ-features from C in the absence of the phase head (or for de-activation of ϕ-features through pair-merge of C to T – see footnote 5): raising T is without unvalued ϕ-features. If so, it is predicted that R can label when it does not inherit ϕ-features. This prediction is, in fact, borne out. Consider (16):

(16) a. The book will be [$_{\lambda}$ read by the students]
b. The student is [$_{\lambda}$ believed to be intelligent]

The examples are well-formed, which argues that R (*read* and *believed*) can label, with λ labeled by LA. In the passive, *v* instead of *v** is merged in derivation and hence, no feature inheritance occurs in (16); R does not bear unvalued ϕ-features (or *v** is indeed merged but its ϕ-features are de-activated by its pair-merge to R – Epstein, Kitahara, and Seely 2016; see also section 4.2 for relevant discussion). It straightforwardly follows from the proposal in this paper that in (16), R is labelable.

Likewise, when *v** is merged in and R inherits ϕ-features, R will turn labelable thanks to valuation of its ϕ-features through Agree. Notice, however, that in languages like English, object agreement, unlike subject agreement, is not morphologically spelled out. If morphological object agreement is absent, the proposed analysis suggests that externalization of ϕ will be unproblematic at all even when the object does not move to Spec-R for canonical agreement; to put it differently, in the case of R, EPP is not forced by externalization and valuation of ϕ to make R labelable suffices for convergence at the interfaces. The proposed analysis thus predicts that the object EPP is optional, with (17a) possible along with (17b):

(17) a. $[R_j$-*v** $[t_j$ Obj]]
b. $[R_j$-*v** $[Obj_i$ $[t_j$ $t_i]]]$

This prediction is indeed borne out. It has been observed that unlike the subject, the object can stay in its first-merged position. Lasnik (2003, 2004) is one example that argues for the optionality of object raising or raising to Spec-R with ECM. For our discussion, first consider examples in (18) (Chomsky 1995: 327):

(18) a. (It seems that) everyone isn't there yet. [*every* > *Neg*; *Neg*> *every*]
b. Everyone seems not to be there yet. [*every* > *Neg*; **Neg*> *every*]

In (18a), the negation can have wide scope over the subject quantifier while it cannot in (18b). The contrast in (18) shows the following two things: (i) the subject in Spec-T can be interpreted inside the scope of clausal negation when they are in the same clause; and (ii) the *n*P is not interpreted in its reconstructed positions when it has A-moved (Chomsky 1995, Lasnik 2003, 2004). With these assumptions in place, consider (19) (all the examples below are cited from Lasnik 2003, 2004):

(19) a. I believe everyone not to have arrived yet.
b. I proved every Mersenne number not to be prime. [*every* > *Neg*; *Neg*> *every*]

In (19), the ECM subject can take scope under clausal negation, which suggests that it stays in the same clause with the negation and does not move out of the embedded clause to the Spec of the matrix R. This is also endorsed by (20), where A-movement out of the ECM complement into the matrix clause clearly takes place for passivization and wide scope reading of the negation is ruled out (also (18b)):

(20) a. Everyone is believed not to have arrived yet.
b. Every Mersenne number was proved not to be prime. [*every* > *Neg*; **Neg*> *every*]

The scope contrast between (19) and (20) demonstrates that object raising, unlike subject raising, is optional.

The optionality of object raising is also supported by the *make-out* ECM construction. Consider (21):

(21) a. Mary made John out to be a fool.
b. Mary made out John to be a fool.

In (21), the ECM subject can come to the left of the particle *out* as well as to the right. Since the particle is part of the matrix verb *make*, it can reasonably be considered that in (21a), *John* is in the Spec of the matrix R, with only the verb affixed to the matrix *v** through head movement and with *out* left behind in R. (21a) is thus analyzed as (22) (as usual, irrelevant details are omitted in (22)):

(22) [Mary T [*v**-made [John$_i$ [R-out [to be [t_i a fool]]]]]]

If (21a) involves overt raising to Spec-R, then it is reasonable to assume that there is no such raising in (21b). (23) endorses this argument: in (23), the ECM subject must take scope over the negation in the embedded clause when it comes to the left of *out* while it can take scope under the negation when it comes to the right:

(23) a. The mathematician made every even number out not to be the sum of two primes. [*every* > *Neg*; **Neg*> *every*]
b. The mathematician made out every even number not to be the sum of two primes. [*every* > *Neg*; *Neg*> *every*]

The scope facts in (23) argue that in (23b), the ECM subject does not raise out but stays in the embedded clause.

Licensing of negative polarity items (NPIs) as well as binding further argue for the same conclusion, supporting the optionality of raising to Spec-R. Consider the following data:

(24) a. The lawyer made no witnesses out to be idiots during any of the trials.
b. ?*The lawyer made out no witnesses to be idiots during any of the trials.
(25) a. The DA made the defendants out to be guilty during each other's trials.
b. ?*The DA made out the defendants to be guilty during each other's trials.

In (24b) and (25b), the ECM subject cannot c-command into the matrix clause adjunct and NPI licensing and binding fail. This straightforwardly follows if the subject does not raise out but stays in the embedded clause. As evidenced by (26), the subject, if it is embedded in the lower clause, can neither license the NPI nor bind the anaphor in the higher clause:

(26) a. ?*The DA proved [that no one was guilty] during any of the trials.
b. ?*The DA proved [that two men were at the scene of the crime] during each other's trials.

From the discussion thus far, it can be concluded that raising to Spec-R (object EPP), unlike raising to Spec-T (subject EPP), is not always forced but is optional, contrary to what has been assumed in Chomsky (2015). As I have demonstrated, the analysis I have proposed in this paper, which reduces EPP to externalization and not to labeling, can correctly predict the optionality of object raising as well as the obligatory nature of subject raising: externalization is unproblematic for R without its Spec. On the other hand, under Chomsky's analysis, the optionality of raising to Spec-R cannot be predicted, as object raising, just like subject raising, is always forced by labeling: unless visible Spec-R is created, R cannot label and labeling failure will arise. As I have discussed, this argument is not empirically supported.[9]

[9] If the ECM subject does not move to Spec-R in the higher clause, it will move to Spec-T in the ECM complement, which is forced by labeling: external merge of the ECM subject with the predicate phrase yields XP-YP, which will be unlabeled if neither XP nor YP moves out (Chomsky 2013, 2015). This movement, however, will raise the same problem in the higher position: internal merge of the ECM subject will produce {{*n*P}, {TP}} in the ECM complement but *n* and T do not agree in the absence of ϕ on T, which will cause a labeling problem. In this paper, I leave this problem to Mizuguchi (2017), where I discuss it and propose its solution, arguing that {{*n*P}, {TP}} yielded in the

4. Labelable T

In the discussion so far, I have proposed that label weakness is reducible to unvalued (hence, uninterpretable) features borne by heads, arguing that the labelability of T and R follows from this proposal. Notice that the proposal wipes out the assumption of "label strength" in labeling: if a head has unvalued features, it cannot label (i.e., it is too weak to label); if not, it can (i.e., it is strong enough to label). In this section and the next, I discuss theoretical and empirical consequences of the proposed analysis of labelability. In this section, I consider labelable T.

4.1. T in Null-Subject Languages

Chomsky (2015: 9) argues that unlike in English, in languages with rich agreement such as Italian, T can label: T in rich agreement languages is strong in labeling terms and can serve as a label. One argument for this claim is the fact that null subjects are possible in such languages (Rizzi 1982). Consider the following examples:

(27) [$_{\lambda}$ Parla italiano e inglese] *Italian*
speak-3.sg Italian and English
'He/she speaks Italian and English.'

(28) [$_{\lambda}$ Hemos trabajado todo el día] *Spanish*
have worked all the day
'We have worked all day.' (Perlmutter 1971)

(29) [$_{\lambda}$ Mila ellinika] *Greek*
speaks Greek
'He/she speaks Greek.' (Roberts and Holmberg 2010)

The examples above argue that labeling failure does not arise even if an overt *n*P is not merged with the λ-marked set to form Spec-T, which suggests that T alone can label. Note that merge of small *pro* does not help toward labeling: just like copies, it does not create a visible Spec. Thus, strong T in rich-agreement languages cannot be explained by the assumption of small *pro*. As far as my

ECM complement can indeed be labeled.

proposal is correct, T in null-subject languages, as it can label without agreement with the *n*P, does not bear any unvalued features. I argue that the relevant T can serve as a label because rich agreement has historically developed from subject/demonstrative pronouns (van Gelderen 2011, Givón 1976, Roberts 2010; see also Simpson and Wu 2002 for relevant discussion). For instance, van Gelderen (2011) argues that agreement comes from a series of changes, which she calls the "subject agreement cycle" or "subject cycle," demonstrating the process of this change as (30):

(30) a. demonstrative > third person pronoun > clitic > agreement > zero
b. noun/oblique/emphatic > first/second person pronoun > clitic > agreement > zero

In rich agreement (null-subject) languages, subject/demonstrative pronouns have been reanalyzed as rich agreement. We can think of the process by which this happened as follows: subject/demonstrative pronouns, which were originally merged as Spec-T, were confused with T as they were like heads and were structurally adjacent to T, and became part of the verbal morphology, hence T. Thanks to the process, rich agreement has the categorial status of pronouns (Alexiadou and Anagnostopoulou 1998, Baker and Hale 1990, Fassi-Fehri 1993, Platzack 2003, Taraldsen 1992 among others). In other words, verbal inflections in rich agreement languages like Italian are incorporated pronouns, functioning exactly in the same way as the corresponding pronouns in English:

(31)

	English		*Italian*	
	Singular	Plural	Singular	Plural
1st person:	**I** speak	**we** speak	parl**o**	parl**iamo**
2nd person:	**you** speak	**you** speak	parl**i**	parl**ate**
3rd person:	**he/she/it** speaks	**they** speak	parl**a**	parl**ano**

Based on these arguments, the conclusion can be reached that verbal inflections in null-subject languages are valued ϕ-features and hence that T has valued ϕ qua rich agreement just as *n*Ps do.[10] Since there are no unvalued

[10] Taking *parla* as an example, the verbal inflection *-a* is merged as T, to which R or *v**-R *parl* moves, given that the verb head-moves to T in null subject languages; the resulting SO *[[[R parl]*-v**]-a]* is morphologically spelled out as *parla* (cf. Alexiadou and Anagnostopoulou 1998):

features on T, T can label even without ϕ-feature agreement with the *n*P. "Strong" or labelable T in null-subject languages is deduced from my proposal.

In non-null-subject languages like English, on the other hand, valued ϕ has decayed through historical loss of rich agreement, losing specific values and changing into valueless (hence, unvalued) ϕ, which requires valuation through syntactic agreement with the *n*P. T in English bears unvalued ϕ, which is why the *n*P must be merged in derivation for agreement (i.e., null subjects are impossible) and hence T is "weak" as a label.[11]

The discussion in this paper also suggests that null subjects are possible when rich agreement or valued ϕ-features are totally missing. We have already observed one illustration of this in section 3.1, where we considered the labelability of raising T or more generally, T which does not inherit ϕ from C (as in under-inheritance): in such cases, the overt subject is not merged in the Spec of T and a null subject is possible. As I have argued there, such T does not bear unvalued ϕ-features and can label without agreement, hence without an overt *n*P. Moreover, externalization is not at all problematic without Spec-T for infinitival T or T without ϕ as there is no need to morphologically spell out ϕ. In fact, if Spec-T is created, labeling failure will result in the absence of agreement. The analysis I have proposed in this paper can correctly predict null subjects in languages which do not have rich agreement.

Another illustration is found in East Asian languages such as Chinese, Japanese and Korean, where null subjects are widely observed. Take Japanese as our example. As (32) illustrates, null subjects are observed in the language:

(32) a. Taroo-o tazune-ta.
Taroo-acc visit-past
'pro visited Taroo.'

b. Speaker A: Kinoo soko-ni itta no. Speaker B: Itta yo.
yesterday that place-to went q went excl
'Did you go there yesterday?' 'Yes, I did!'

Fukui (1986, 1995) argues that the functional head T is missing in Japanese and

(i) [[[[R parl]-*v**]-a] [*v** [R ...]]]

[11] Given the theory of feature inheritance, it is suggested that once valued ϕ-features have lost their values and decayed into unvalued ϕ, they are lexically reassigned to C. This change follows if the loci of unvalued features are phase heads, which can reasonably be defined as heads having unvalued features. For relevant discussion on this, see Chomsky (2016). With this argument in mind, it is no wonder that the locus of ϕ-features has changed from T (a non-phase head) to C (a phase head) in languages like English, with the result that ϕ-features are inherited onto T.

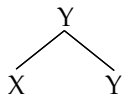
that phrases are not licensed by agreement/feature checking. In terms of the present discussion, Fukui's proposal indicates that agreement has been completely lost, there being no unvalued ϕ-features on T (or on C, from which it inherits ϕ) (see Saito 1982 for relevant discussion). This outcome is expected given subject (agreement) cycle in (30). We get the following typology in language thanks to the process of change in agreement:

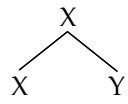
(33) a. valued ϕ on T: rich agreement languages (e.g., Italian)
b. unvalued ϕ on T: English
c. zero ϕ on T: Japanese

In (33c), since there is no unvalued ϕ on T, it can label without agreement as in the raising complement as well as in rich agreement languages and consequently, there is no need to merge the *n*P (= the subject) in the derivation in order for T to label. The null-subject phenomenon follows as a natural consequence of the proposed deduction of labelability. The same argument can apply to Chinese and Korean, which, like Japanese, do not at all spell out subject-verb agreement morphologically. This can be taken to suggest that ϕ-features are missing in these languages as well (=(33c)), and that T can label without agreement with the *n*P as in Japanese.

4.2. Labelability by Pair-Merge

Chomsky (2015) discusses another case of labelable heads. In considering R to *v** raising, he proposes a fresh treatment of head movement (which is reformulated as internal pair-merge of heads under current minimalist assumptions) by which the host head X is affixed or adjoined to the raised head Y. The derived end result is (34a), instead of traditional (34b):[12]

(34) a. 

b. 

[12] Again, tree notations are used for expository purposes only; recall that given simplest Merge, Merge will only produce label-free/unidentified SOs ({α, β}, <α, β>), with no mother nodes or projection labels.

Notice that as discussed in Bobaljik and Brown (1997), internal pair-Merge (head movement) can apply in a way abiding by the No-Tampering Condition and does not pose a problem to cyclic derivation.

What is interesting about Chomsky's proposal is not just that X is adjoined to Y (=(34a)), which is the opposite of the standardly accepted formulation; it also has a consequence for labeling. He argues that although R alone cannot label, it can label when *v** is pair-merged to it and the pair-merged SO <R, *v**> is yielded: β cannot be labeled in (35a) while it can in (35b):

(35) a. [$_{\alpha}$ *v** [$_{\beta}$ R [...]]] (β = ?)
b. [$_{\beta}$ <R, *v**> [...]] (β labeled by <R, *v**>)

Likewise, it is argued in Mizuguchi (2016a), who develops Chomsky (2015), that T, which is unlabelable, can indeed label if C is pair-merged to it and T is <T, C>:[13]

(36) a. [$_{\gamma}$ C [$_{\lambda}$ T [$_{\alpha}$ *v* [$_{\beta}$ R ...]]]] (λ = ?)
b. [$_{\lambda}$ <T, C> [$_{\alpha}$ *v* [$_{\beta}$ R ...]]] (λ labeled by <T, C>)

Chomsky and Mizuguchi claim that R and T, which alone are too weak to label, can label once *v** and C are pair-merged to R and T, with the former adjoined to the latter, and <R, *v**> and <T, C> are yielded. Notice, however, that the labelability is simply stipulated. The question is left unanswered why R and T are rendered labelable if *v** and C are pair-merged to R and T; they just say that R and T can label when pair-merge of the phase heads applies, which does not explain at all. I argue that the analysis proposed in this paper can give a principled answer to the relevant question and that the labelability of the pair-merged SOs follows. To begin with, consider pair-Merge. Pair-Merge takes two SOs and forms an ordered pair out of them. Suppose that Y is pair-merged to X. When Y is pair-merged, it will get syntactically de-activated, becoming invisible to syntax (Chomsky 2015). De-activation or invisibility follows from the assumption that pair-Merge, unlike set-Merge, asymmetrically merges two SOs by affixing or adjoining one to the other: the operation puts Y on a "separate" plane; set-Merge, on the other hand, symmetrically merges two SOs and puts them on the same plane (Chomsky 2004). When Y is pair-merged to X, X and Y are asymmetrical in structure, with X retaining all its properties on the primary plane and with Y, as it is merged on a different plane, behaving as if it were not there. For expository convenience, this is graphically illustrated in (37):

[13] As Epstein, Kitahara, and Seely (2016) and Mizuguchi (2016a) argue, given simplest Merge, pair-Merge can apply externally as well as internally to yield (34a). Thus, <R, *v**> and <T, C> can be produced by both external and internal pair-Merge.

(37) a. <X, Y> b. {X, Y}

Thus, those SOs affixed or adjoined to other SOs by pair-Merge are syntactically invisible, being de-activated. This argument is not unreasonable given simplest Merge: to the extent that Merge applies freely, it will apply symmetrically as well as asymmetrically, and two types of Merge can reasonably be assumed.[14] Moreover, there is much evidence in language for substitution (set-Merge) and adjunction (pair-Merge). If set-Merge and pair-Merge are available in language as consequences of simplest Merge, then the structures yielded by one type of Merge will be different from those produced by the other; otherwise, there would be only one type of Merge.

With this assumption in place, when *v** and C are pair-merged to R and T, only the latter will be visible to operations as the former gets de-activated, being invisible. Thus, when applied to (38), LA will unambiguously locate R and T through minimal search (SOs de-activated by pair-Merge are shaded in (38)):

(38) a. χ (χ = R) b. ψ (ψ = T)

Since unvalued ϕ-features are borne by *v** and C, and there are no unvalued ϕ-features on R and T, χ and ψ can be labeled by R and T. Notice that ϕ-feature inheritance will not take place in (38) since *v** and C become syntactically inert by pair-Merge and ϕ-features cannot be subject to operations. R and T are without any unvalued features thanks to de-activation of *v** and C by pair-merge of the heads to R and T, and hence the pair-merged SOs <R, *v**> and <T, C> can label. Under my proposal, the labelability of the pair-merged SOs is not at all stipulated but can be deduced in a principled fashion.

To the extent that the discussion in this section is correct, the labelability of <R, *v**> assumed in Chomsky contradicts his assumption that R cannot label. Recall once again that pair-merged elements are syntactically inert due to

[14] An anonymous reviewer says that the assumption of two types of Merge seems to go against the minimalist spirit. In fact, it does not. As I have argued, given simplest Merge, Merge can apply both symmetrically and asymmetrically; pair-Merge (the Merge that applies asymmetrically) is straightforwardly predicted and is empirically well-founded as well.

de-activation by pair-Merge. Then the SO <R, *v**>, in which *v** is de-activated, is syntactically on par with R; only R, as I have argued, is visible to LA. If R cannot label for its weakness as a label as Chomsky assumes, then it should be the case that the pair-merged SO <R, *v**> cannot label, either. This further strengthens the argument that the labelability of <R, *v**> is just stipulated. The same argument applies to the labelability of T as well as that of the pair-merged object <T, C>. LA can see and locate only T in <T, C> for labeling. Even if *v** and C are pair-merged to R and T, respectively, the derived SOs (<R, *v**> and <T, C>), in terms of labeling by LA, are no different from R and T. This poses problems to Chomsky but is favorable to us in that they do not bear unvalued features and hence can label.

Notice that pair-Merge can explain null subjects in raising and ECM infinitives. Consider (39). In (39b), the ECM subject moves to Spec-R in the higher clause:

(39) a. The student seems [to be happy]
b. I believe-*v** [the student [R [to be [intelligent]]]]

Suppose that in raising and ECM infinitives, C, a phase head, is merged in the derivation just as in their finite counterparts (Epstein and Seely 2006, Mizuguchi 2016a, 2017, Ormazabal 1995, Rooryck 1997); however, imagine that C is not set-merged with a T-headed set (that is, [C [T [...]]]) but is pair-merged to T (i.e., [<T, C> [...]]), which is freely available given simplest Merge. If so, the labelability of infinitival T or the fact that it can label without an overt *n*P is deduced from the pair-merge, which de-activates C and hence its unvalued ϕ (=(38b)); agreement with an overt *n*P will not be necessary for labeling. In languages like English, unlike in rich agreement languages, null subjects are made derivationally possible by pair-Merge, coupled with the proposed analysis of labelability.

The same argument can also explain the labelability of λ in the passive construction. Consider once again (16), which are repeated below:

(16) a. The book will be [$_{\lambda}$ read by the students]
b. The student is [$_{\lambda}$ believed to be intelligent]

Suppose that the grammar is not equipped with two types of *v* (i.e., *v* with ϕ (*v**) and *v* without (*v*)) but that only *v** is available (Epstein, Kitahara, and Seely 2016). *v** is thus merged in (16). However, if pair-merged to R, the head, hence

its unvalued ϕ, will be de-activated, and λ can be labeled by <R, *v**> in the absence of unvalued ϕ on the pair-merged SO (=(38a)). Pair-Merge, together with the proposed analysis of labelability, can also derivationally explain the fact that R can label without an overt *n*P in the passive construction.[15]

5. The Labelability of Other Heads

In this paper, I have proposed that label weakness is reducible to unvalued features borne by heads. As far as this proposal is correct, it suggests as its consequence that unlabelable heads are not limited to T and R as conventionally assumed but include other heads as well. For instance, consider interrogative C and *n*. Chomsky (2000) argues that interrogative C bears an unvalued Q-feature; likewise, it has been assumed that *n* bears an unvalued Case feature. If unvalued features are the cause of label weakness and heads bearing such features cannot label, then interrogative C and *n*, just like T and R, cannot serve as labels for their unvalued features. They will turn labelable (and hence, C-headed and *n*-headed sets can be labeled by C and *n*) when their unvalued features are valued by agreement.

With this extension in mind, take *n* and consider what happens when the *n*-headed set is internally merged.[16] If *n* cannot label unless its Case feature is valued, the feature must be valued before *n* or the *n*-headed set {*n*, NP} is internally merged; otherwise, unlabelable (hence uninterpretable) copies will be created in the course of the derivation by its movement, which satisfies the

[15] SOs adjoined by pair-Merge will be subject to interpretation; otherwise, adjuncts would not be interpreted. If so, unvalued features, when SOs bearing them are pair-merged and put on a different plane, will be invisible to the interfaces as well. I suggest that unvalued features on pair-merged SOs are rendered interpretively invisible by pair-Merge and can be ignored at the interfaces. This assumption is empirically supported by (i), where lexical items such as *yesterday* and *last year*, which are *n*Ps and bear unvalued Case features as shown in (ii), do not cause crash or are not left unlabeled at the interfaces when adjoined and used as adverbials, even if their Case features are not valued:

(i) John went to the party yesterday / last year.

(ii) a. Yesterday was a wonderful day for both of us.
b. Last year saw a big change in the political system.

This argues that unvalued Case features become invisible to the interfaces when *n*Ps are pair-merged.

[16] The argument here also applies to interrogative C (as well as T and R), which can head-move.

No-Tampering Condition (NTC) under the simplest-Merge hypothesis and as illustrated in (40), produces identical occurrences of {*n*, NP}:[17]

(40) ... $\{\{n, NP\}_3$ ... $\{\{n, NP\}_2$ ... { ... $\{n, NP\}_1$ ... $\}\}\}$

Unlabeled copies will not be interpreted by CI and the processes of externalization, violating Full Interpretation. On the assumption that SOs, once transferred and crashed, cannot be de-crashed (i.e., saved) by subsequent derivation, which is reasonable in that SOs are cyclically interpreted and externalized through Transfer and transferred SOs cannot be subject to further computation (Chomsky 2008 and Epstein, Kitahara, and Seely 2012 for this argument; but cf. Epstein and Seely 2006 for an alternative proposal), the relevant derivation will be ruled out as ill-formed at the interfaces. Given my proposal of labelability, it follows that the *n*P will be internally merged once its Case feature is valued and *n* can label; unvalued features must be valued before movement in order for labeling to be successful.

If this conclusion is correct, it has implications for both A- and A´-movement. Let us start with the discussion of A-movement.

5.1. A-Movement as One-Fell Swoop

As for A-movement, the proposed analysis implies that it is not successive cyclic but proceeds in one fell swoop to the target Spec-T. Consider (41):

(41) The student seems to be likely to be in the library.

(41) will be analyzed not as (42a) but as (42b):

(42) a. [$_{CP}$ [$_{TP}$ The student [seems [$_{TP}$ t_3 [to [be likely [$_{TP}$ t_2 [to [be t_1 in the library]]]]]]]]]
b. [$_{CP}$ [$_{TP}$ The student [seems [$_{TP}$ to [be likely [$_{TP}$ to [be t in the library]]]]]]]

In (41), the Case feature of *the student* is valued when C is merged in the root phase and T inherits ϕ-features to agree the *n*P. If so, successive-cyclic movement

[17] Recall that subscripted numbers are used only for expository purposes, intended to show that multiple occurrences of {*n*, NP} are yielded by IM.

as in (42a), as it starts before the merge of the matrix C, will yield copies with unvalued Case (t_2 and t_1), which will be left unlabeled for the feature and cannot be interpreted. On the other hand, in (42b), the Case of the *n*P is valued in situ by ϕ-feature agreement thanks to Agree, which, applied when the matrix T inherits ϕ from C, locates the probe and the goal. Thus, the copy in (42b) (marked as t) can be labeled *n* and hence can be interpreted without any problems.[18]

This implication is theoretically and empirically supported. Theoretically, successive-cyclic A-movement will incur interpretive deviance at the CI interface. As I have argued, IM, which satisfies the Inclusiveness Condition, will yield multiple occurrences of an SO when it moves (=(40)). For proper interpretation at the interfaces, such occurrences must be interpreted as one, hence must be identified as copies, which will yield a single or chain interpretation. Take (43). In (43a), unlike in (43b), the four occurrences of the *n*P *the student* created in the derivation are interpreted as one, forming a discontinuous element:

(43) a. [The student$_4$ [seems [the student$_3$ [to be believed [the student$_2$ [to be [the student$_1$ intelligent]]]]]]]
b. [The student$_4$ [believes that [the student$_3$ [said that [the student$_2$ was likely to visit the country where [the student$_1$ [was born]]]]]]]

Multiple occurrences of an SO will be identified as copies at the interfaces if IM applies at the phase level in simultaneity with Transfer, and the derivational information is locally available to the CI interface and to the processes of externalization that the SO is moved (Chomsky 2007, 2008); otherwise, such occurrences will be identified as repetitions (distinct occurrences of the same SO) as in (43b) because the interfaces cannot tell whether the occurrences created are by IM or by External Merge (EM) as indices cannot be added to SOs for the distinction under simplest computation and both IM and EM can produce multiple occurrences of the same SO. Since raising TP is not selected by a C phase head (or C is de-activated, hence de-phased by its pair-merge to T − see section 4.2) and is not cyclically transferred, if A-movement is successive cyclic,

[18] An anonymous reviewer suggests that if all the copies created by successive-cyclic A-movement can agree with T in the matrix clause, their Case features can be valued before Transfer, in which case unlabeled copies will not be sent to the interfaces. This sort of multiple agreement, however, is not possible in languages like English. In such languages, there is a one-to-one relation between the probe and the goal, which means that when the matrix T agrees with the copy marked as t_3 in (42a), Agree will be completed and halt: there are no more ϕ-features left unvalued on T which make multiple agreement by the probe possible.

the IM and Transfer do not co-occur and the occurrences of the *n*P created in Specs of intermediate TP will be identified as distinct (repetitions) at the interfaces and each will be independently interpreted. This will result in distinct/non-chain interpretations of such occurrences, not a chain interpretation, which will violate Full Interpretation and cause interpretive deviance at the CI interface:

(44) *No single/chain interpretation*
[$_{CP}$ [$_{TP}$ The student$_3$ [seems [$_{TP}$ the student$_3$ [to [be likely [$_{TP}$ the student$_2$ [to [be [the student$_1$ in the library]]]]]]]]]
(the student$_3$ ≠ the student$_2$ ≠ the student$_1$)

Furthermore, successive-cyclic A-movement will result in anomalous externalization: the occurrences *the student*$_2$ and *the student*$_1$ will be spelled out along with *the student*$_3$, as they are not identified as forming a discontinuous element with *the student*$_3$ and are recognized as distinct by the externalization component.

A labeling problem will also arise if A-movement is successive cyclic. As I have argued above, the lower occurrences of the *n*P (*the student*$_2$ and *the student*$_1$) will be identified as distinct, not being part of a discontinuous element. Since the two occurrences are not identified as copies, they will be visible to LA (Chomsky 2013, 2015). The visibility, however, has the effect that α and β in (45), forming XP-YP structure (XP is *{the, student}* and YP is *{to, RP}*/*{in,* n*P}*), will not be labeled as the heads of the two SOs do not agree:

(45) [$_{CP}$ C [$_{TP}$ The student$_3$ [seems [$_{TP}$ t$_3$ [to be likely [$_{\alpha}$ the student$_2$ [to be [$_{\beta}$ the student$_1$ in the library]]]]]]]]

Successive-cyclic A-movement will leave the α-marked and β-marked sets unlabeled, incurring labeling failure. This will also violate Full Interpretation at the CI interface and cause externalization problems on the side of the SM interface.

On the empirical side, there has been a debate in the literature on whether long-distance A-movement is successive cyclic. Arguments have been provided to argue for "traces" in the Spec of raising T in long-distance A-movement, hence supporting its successive cyclicity. Such arguments include reflexive binding, quantifier float and extraction from subjects (e.g., Bošković 2002, Chomsky 2008 and Lasnik 2003). As Bobaljik (1995), Castillo, Dury, and Grohmann (2009),

Epstein and Seely (2006), Grohmann, Dury, and Castillo (2000) among others persuasively argue, however, alternative explanations are available for the alleged evidence which do not assume occurrences or copies in the Spec of raising T; moreover, the studies demonstrate that such evidence is also weak in firmly endorsing successive-cyclic A-movement when a wider range of examples is considered. For instance, consider reflexive binding in (46):

(46) a. *Mary seems to *John* to appear to *himself* to be happy.
b. … [seems to *John* [〈 Mary 〉 [to appear to *himself* to be happy]]]

Proponents for successive-cyclic A-movement say that the ill-formedness of (46a) is attributable to a copy of *Mary* in the Spec of raising T (indicated as "〈 Mary 〉"), which is closer to the reflexive and blocks binding between *John* and *himself*. This argument, however, does not go through. Suppose that Transfer defines a binding domain for reflexives, which is reasonable given that interpretation applies cyclically by Transfer; transferred TP thus constitutes a binding domain. With this assumption in place, consider (47). In the example, although *Mary* is closer to *himself* than *John*, the binding relation between *John* and *himself* is NOT blocked, which argues that the copy in (46) is irrelevant to the ill-formedness:

(47) *John* told Mary about *himself*.

Epstein and Seely (2006) show that the binding failure comes from the fact that *John*, which is embedded in PP, cannot c-command out. They provide (48) as evidence for this claim; in (48), Condition B is not violated, indicating that *John* does not c-command *him*:

(48) It seems to *John* to appear to *him* that the earth is flat. (Condition B)

Thus, it can be concluded that (46) does not constitute evidence for successive-cyclic A-movement.[19]

---

[19] A potential problem is with Condition C cases, which suggest that the experiencer c-commands into the lower clause:

(i) *It seems to him to appear to John that the earth is flat.

As Epstein and Seely (2006: 135) note, however, in other cases, the experiencer seems not to c-command out. I leave for a future occasion the discussion of how Condition C effects can be captured.

Next, consider (49), which might suggest that c-commanding out of PP is possible:

(49) a. It seems to *John* to appear to *himself* that the earth is flat.
b. It appears to *Mary* to seem to *herself* that Bill is ill.

However, if we consider the fact that the reflexives in (49), as shown in (50), can be replaced with pronouns, the well-formedness of (49) cannot be attributed to binding out of PP:

(50) a. It seems to *John* to appear to *him* that the earth is flat.
b. It appears to *Mary* to seem to *her* that Bill is ill.

In fact, Castillo, Dury, and Grohmann (2009) argue that the raising construction with the expletive *it* as the subject assigns logophoric properties to reflexives, which is endorsed by (51):[20]

(51) a. *John* kissed Mary since it appeared to *himself* that this would be the last chance.
b. Mary kissed *John* since it appeared to *himself* that this would be the last chance.

Given Castillo, Dury, and Grohmann's argument, the well-formedness of (49) is due to logophoricity, not to binding, and (49) does not argue for c-commanding out of PP.[21]

Quantifier float does not constitute evidence for successive cyclicity of A-movement, either. It has been assumed in the literature that floating quantifiers indicate positions of intermediate traces of A-movement (Sportiche 1988). Bobaljik

---

[20] The same observation is found in NPI licensing (=(i)) and bound variable interpretation (=(ii)) (Epstein and Seely 2006 and Castillo, Dury, and Grohmann 2009):

(i) a. ?*Bill seems to *no* linguist to like *any* recent theory.
b. It seems to *no* man to appear to *any* woman that the earth is flat.

(ii) a. ?*Mary seems to *every* man to like *his* theory.
b. It seems to *every* boy to appear to *his* mother that the earth is flat.

[21] Unclarities and complications are involved in examples related to binding. The reader is referred to Epstein and Seely (2006) for how they are dealt with in one-step A-movement. They show that such unclarities and complications do not provide evidence for successive-cyclic A-movement.

(1995) argues, however, that floating quantifiers are adverbs and modify the predicate in a predictable manner with respect to some nominal phrase. As evidence, he demonstrates that floating quantifiers can appear in positions where copies of A-movement are not expected. Consider (52):

(52) a. The magicians (all) should (all) have (all) arrived before the show begins.
b. The votes (all) will (all) have (all) been counted by midnight.

Then floating quantifiers in (53) can be adjoined as adverbs, which does not endorse the argument that they indicate traces or copies of successive-cyclic A-movement.

(53) The boys seem (all) to appear (all) to (all) like ice cream.

Finally, Chomsky (2008: 153) discusses (54) and argues that (54a) provides evidence for successive-cyclic A-movement:

(54) a. Of which car is the driver/picture likely to cause a scandal?
b. *Of which car did the driver/picture cause a scandal?

His argument is based on the following assumption: movement out of an SO (or sub-extraction) is prohibited when the SO is at the phase edge while such movement is possible when it is not at the edge of a phase. With this assumption in place, the well-formedness of (54a) follows if the subject moves to the Spec of raising T (a non-phase edge) and *wh*-movement (sub-extraction) is executed from that Spec; in (54b), on the other hand, the sub-extraction is forced from Spec-*v** (a phase edge where the subject is externally merged) and the movement is ruled out.

Chomsky's evidence (54), however, does not argue for successive-cyclic A-movement. Chomsky crucially assumes that the phase edge is a barrier to sub-extraction while the non-phase edge is not. This assumption is not empirically supported. As discussed, among many others, in Diesing (1992), Gallego and Uriagereka (2007), Stepanov (2007) as well as references cited therein, sub-extraction from the subject is indeed possible when it stays in Spec-*v**. Consider the following German (55) and Spanish (56) examples:

(55) [$_{CP}$ Was$_i$ haben [$_{TP}$ denn [$_{v*P}$ [$_{vP}$ t$_i$ für Ameisen] einen Postbeamten
what have indeed for ants a postman

gebissen]]]?
bitten
'What kind of ants have bitten a postman?' (Diesing 1992)

(56) (?)[$_{CP}$ [De qué conferenciantes]$_i$ te parece [$_{CP}$ que me van
of what speakers cl.to.you seem.3.sg that cl.to.me go.3.pl
a impresionar [$_{nP}$ las propuestas t_i]]]?
to impress.inf the proposals
'Which speakers does it seem to you that the proposals by will impress me?' (Gallego and Uriagereka 2007)

These data demonstrate that sub-extraction from the phase edge is in fact possible.

Furthermore, sub-extraction from the non-phase edge is impossible (Chomsky 1973, Kayne 1984). Consider the following examples:

(57) ?*Who$_i$ did Mary consider [friends of t_i] to be fools?

(58) a. ?*Who$_i$ did Mary make [friends of t_i] out to be fools?
b. ?*Who$_i$ did Mary make out [friends of t_i] to be fools? (Lasnik 2001)

Whether the ECM subject moves to Spec-T in the ECM complement or to Spec-R in the higher clause, it moves to a non-phase edge position in the ECM construction. As evidenced by (57) and (58), sub-extraction causes ill-formedness, which argues that the non-phase edge forms a barrier to sub-extraction. Both (55)/(56) and (57)/(58) show that Chomsky's assumption is empirically wrong. A correct descriptive generalization is that sub-extraction is impossible if the subject (more generally, an SO) moves (Wexler and Culicover 1980, Stepanov 2007). To the extent that sub-extraction takes place in (54), ill-formedness will result in (54a) just as in (54b) and (57)/(58) since the subject moves to Spec-T in English. A reasonable conclusion is that (54a) does not argue for successive-cyclic A-movement.[22]

---

[22] Well-formed (54a) can be captured independently by an analysis which does not employ sub-extraction: adnominal PPs can often be extraposed from derived subjects (e.g., Broekhuis 2015, van Craenenbroeck and den Dikken 2006, Mizuguchi 2014; see Rochemont and Culicover 1990 for PP extraposition as base generation). As shown in (i), extraposition, unlike *wh*-movement, is possible from derived subjects:

(i) a. A biography appeared [of the Marx brothers].

Notice that Chomsky's analysis will wrongly predict that (59) is well-formed, since just as in (54a), the *wh*-phrase can be sub-extracted from the non-phase edge (intermediate Spec-T) thanks to successive-cyclic A-movement of the subject to the Spec:

(59) *Which car is the driver/picture of likely to cause a scandal?

Unlike (54a), (59) is ill-formed, suggesting that the non-phase edge does not explain the well-formedness of sub-extraction.[23]

In sum, the theoretical and empirical arguments considered in this section can reasonably lead us to conclude that A-movement is not successive cyclic but proceeds in one fell swoop to Spec-T. The conclusion endorses the proposed analysis of labelability. It should be noted that the problem of EPP does not arise in one fell swoop A-movement: there is no EPP (or its equivalents like an edge feature) under the assumption of simplest Merge.[24]

---

b. *[Which Marx brother] did a biography of appear?

With the extraposition analysis in mind, in (54a), the *wh*-phrase can be base-generated in a position external to the subject and from there, undergoes *wh*-movement to Spec-C. The derivation, as it does not involve sub-extraction from the subject, does not violate the Subject Condition if the subject moves.

As for ill-formed (54b), Mizuguchi (2014) argues that an extraposed PP and the *n*P must c-command each other in the same transfer domain so that the former can be interpreted as part of the latter. He shows that mutual c-command fails in (54b) for cyclic transfer of VP while it succeeds in (54a) thanks to the fact that the raising predicate provides VP to which *of which car* can be extraposed, which enables *the driver/picture* and the extraposed *wh*-phrase to c-command each other in the same transfer domain (i.e., in the matrix TP). See Mizuguchi (2014) for more detailed discussion.

[23] The extraposition analysis, on the other hand, can correctly predict (59) because the *n*P, unlike the PP, cannot be extraposed. In (59), the *wh*-phrase will be forced to be sub-extracted from the subject, which violates the Subject Condition as the subject moves to Spec-T.

Note that given Chomsky's assumption, (59) argues for the opposite conclusion: that is, A-movement is not successive cyclic. If the subject does not undergo successive-cyclic A-movement, the *wh*-phrase will be forced to be sub-extracted from the phase edge (Spec-*v**), which violates the Subject Condition (see (54b)); if it does, sub-extraction will be possible as the *wh*-phrase is sub-extracted from the non-phase edge (the Spec of raising T).

[24] An anonymous reviewer has pointed out that some languages like Greek show full-fledged agreement in raising and ECM complements, which, given the argument in this paper, seems to suggest that A-movement should be successive cyclic for the sake of canonical spell-out of ϕ in the embedded clause. If so, this will leave the problem of how multiple occurrences created by IM can be identified as copies. I leave detailed discussion for future but suggest that in languages such as Greek, the embedded clause is cyclically transferred thanks to a C phase head, the merge of which is indicated by ϕ in the embedded clause. Since successive-cyclic A-movement can occur in tandem with cyclic Transfer, multiple occurrences created by the IM can be identified as copies; at the same time,

5.2. Agreement in A'-Movement

As for A´-movement, it is suggested that SOs undergoing A´-movement do not bear unvalued features and do not agree with A´-heads to the Spec of which they move. As an example of A´-movement, consider *wh*-movement in (60):

(60) a. Which book did John say Bill will buy?
b. Which student do you think will read the book?

In the derivation of *wh*-movement, due to cyclic transfer of phase complements, IM applies in such a way that the *wh*-phrase moves successive cyclically through the phase edge for trans-phasal movement. In (60), the Case feature of the *wh*-phrase is valued in its first-merged position before movement, and *wh*-movement, unlike successive-cyclic A-movement, does not yield copies with unvalued Case features in the course of the derivation. Successive-cyclic A´-movement does not cause a problem of unlabelable (hence, uninterpretable) copies.

Now suppose that the *wh*-phrase bears an unvalued feature in addition to unvalued Case and agrees with C. For our purpose here, suppose that the unvalued feature is an unvalued wh-feature, which is valued thanks to agreement of the *wh*-phrase with C in Q-feature (valued for the *wh*-phrase but unvalued for C − Chomsky 2000, 2013). With this in mind, let us go back to (60). In these examples, the C with which the *wh*-phrase agrees is in the matrix clause and the unvalued feature cannot be valued in situ but in the Spec of the matrix *v**. Thus, *wh*-movement, just like successive-cyclic A-movement, will produce copies bearing an unvalued feature in the derivational process (marked as $\boxed{\langle\, WhP_{u,wh} \,\rangle}$ in (61)) (see also Kitahara 2017 and Kawashima and Kitahara 2015 for relevant discussion):

(61) $[_{CP}\ WhP_{wh}\ [C_{Q}\ \ldots\ [_{v^*P}\ \langle\, WhP_{wh} \,\rangle\ [v^*\ \ldots\ [_{CP}\ \boxed{\langle\, WhP_{u,wh} \,\rangle}\ [C\ \ldots\ [_{v^*P}\ \boxed{\langle\, WhP_{u,wh} \,\rangle}\ [v^*\ [\ \ldots\ \boxed{\langle\, WhP_{u,wh} \,\rangle}\ ]]]]]]]]]$

The derivation will violate Full Interpretation as unlabeled copies are transferred to the interfaces. Notice that unlike in long-distance A-movement, there are phase boundaries in (61) and long-distance agreement cannot be executed; consequently, if the *wh*-phrase does not move successive cyclically, the derivation will crash at

---

canonical spell-out of ϕ will be possible thanks to the creation of Spec. The same argument can apply to Bantu hyperraising (e.g., (66) in section 5.2).

the lower phase level as it bears an unvalued wh-feature and hence is unlabelable. Given that labelability, as I have claimed, is reducible to interpretability at the interfaces, it follows that agreement is irrelevant to A′-movement (one instance of which is *wh*-movement), with no unvalued features (except for Case, which can be valued in situ before movement) present in SOs undergoing A′-movement. This is in line with Chomsky (2007, 2008), where unvalued features (unvalued Q, unvalued wh) are removed from the derivation of A′-movement.

This conclusion, however, may face problems, both theoretical and empirical. First, consider a theoretical problem. Recall that labeling of a set of the form XP-YP requires agreement between X and Y; prominent feature sharing between the two heads for labeling of {{XP}, {YP}} is warranted by Agree (Chomsky 2013, 2015). Considering (62), Chomsky (2013: 45) says that "the most prominent feature of α and of β is shared, namely the interrogative feature Q, a feature of C and the head of α if we adopt a plausible analysis of interrogative *wh*-phrases that takes Q to be the most prominent element ⋯ Searching {XP, YP}, then, LA finds the same most prominent element − Q − in both terms, and can take that to be the label of α":

(62) They wondered [$_{\alpha}$ in which Texas city [$_{\beta}$ C [JFK was assassinated]]]

If C does not agree with the *wh*-phrase in Q, then α, which forms XP-YP structure, will not be labeled by LA:

(63) [$_{\alpha}$ *Wh*P [$_{\beta}$ C [$_{\lambda}$ T [...]]]] (α = ?)

On the empirical side, Japanese *wh*-sentences suggest that *wh*-phrases agree with C. Saito (2017) claims that Japanese *wh*-expressions are operators without any specified quantificational force of their own; instead, their quantificational force is determined by quantificational particles associated with them such as *ka* and *mo*. To see this, consider (64) from Saito (2017):

(64) a. Taroo-wa [[dare-ga sore-o tabeta] ka] sitteiru.
Taroo-top who-nom it-acc ate q know
'Taroo knows who ate it.'
b. [[[dare-ga kaita] hon] mo] omosiroi.
who-nom wrote book also interesting
'For every x, x a person, the book that x wrote is interesting.'

In (64), the same *wh*-expression *dare* is interpreted differently: it is construed as an interrogative *wh*-phrase in (64a) but as a universal quantifier in (64b). If the quantificational force of a *wh*-expression is determined by associated particles, this implies that probe-goal or Agree is involved to establish a relation between C and the *wh*-expression. In fact, Saito proposes that the mechanism of feature valuation, hence probe-goal or Agree, plays a key role in the determination of the quantificational force of a Japanese *wh*-expression.

Theoretical and empirical discussions above show conflicting results for agreement in A′-movement. To solve the problem, I argue that in A′-constructions, C does agree with the *n*P (say, the *wh*-phrase) but that it does so without assuming in the *n*P any unvalued features for agreement; I submit that the implementation of agreement does not require any unvalued features (i.e., "activation") on the side of goals. Consider (65):

(65) $[P_{u.F}\ [\ \ldots\ [\ \ldots\ G_{F}\ \ldots\ ]]]$
AGREE

Taking the *wh*-construction as our example, with the assumption that interrogative C has an unvalued Q-feature, this proposal allows the head to agree with the *wh*-phrase without postulating any unvalued features like wh on the side of the *n*P. Consequently, even if the *wh*-phrase moves successive cyclically through phase edges, it will not produce copies bearing unvalued features along the way, with the result that no unlabeled copies will be sent to the interfaces. Moreover, since agreement *is* executed in A′-constructions, α can be labeled in (63) and Japanese *wh*-expressions can be correctly interpreted in (64). (65) can solve conflicting requirements posed by theoretical and empirical consideration.[25]

The proposal that activation of goals is not necessary for Agree to be implemented is supported by hyperraising in Bantu languages. Consider (66):[26]

---

[25] Saito (2017), following Bošković (2007), proposes that *wh*-expressions covertly move to Spec-C to probe particles in C in (64). However, probe-goal or Agree can be implemented as shown in (65) even in Saito's examples. Take the particle *ka*. I argue that *ka* has an unvalued quantification feature and must agree with a *wh*-expression, which bears a valued counterpart. The agreement allows the unvalued feature to be valued and enables the *wh*-expression to be interpreted at CI as an interrogative *wh*-phrase (that is, its quantificational force is specified as interrogative) thanks to the probe-goal with now CI interface compliant *ka*.

[26] See Carstens and Diercks (2013) as well as references cited therein for arguments that (66a) and (66b) are examples of hyperraising (A-movement across CP) and not those of so-called "copy raising" as in English (i):

(66) a. Chisaang'i_i chi-lolekhana [$_{CP}$ t_i [mbo [t_i chi-kona]]] *Lubukusu*
10animal 10sa-seem that 10sa-sleep.pres
'The animals seem to be sleeping.'
b. Efula yi-bonekhana i-na-kwa muchiri. *Lussamia*
9rain 9sa-appear 9sa-fut-fall tomorrow
'It seems that it will rain tomorrow.' (Carstens and Diercks 2013)

As persuasively argued by Carstens and Diercks (2013), in (66), the embedded clause out of which the *n*P is raised is a finite clause exhibiting a full range of tense and agreement. This is to say that a C phase head is merged in and its ϕ-features are inherited onto the embedded T. If so, Agree applies in the embedded clause and the Case feature of the *n*P will get valued at the point when it agrees with the embedded T. As we can see, however, the hyperraised *n*P also agrees in the higher clause, undergoing multiple agreement. This argues that activation of goals is not a prerequisite for agreement. The proposal that activation is not required on the side of goals is also endorsed by Diercks (2012), who argues that Case is lacking altogether in Bantu languages. This suggests that the *n*P can agree with T even though it does not bear an unvalued Case feature (i.e., even when it is not activated by the feature).

Summarizing the discussion in this section, the proposed deduction of labelability suggests that T and R are not the only unlabelable heads. I have argued that interrogative C and *n* are also unlabelable heads as they bear unvalued features. I have considered what this extension has to say about movement, taking up *n*. The two implications discussed in this section are theoretically and empirically supported, which in turn endorses the argument in this section that *n* is also an unlabelable head for its unvalued Case.

6. Conclusion

Expressions or SOs created by simplest Merge are without labels and must be identified or labeled by LA for CI interpretation and SM externalization. In this paper, I have addressed the labelability of heads and proposed that it is reducible to interpretability at the interfaces: heads are unlabelable because of unvalued features; they can label when they are without such features. As I have discussed, the proposal removes unwanted stipulations on labeling introduced by

(i) John seems like he is sick.

Chomsky (2015) and hence, we do not need to stipulate "label strength." Furthermore, I have argued that EPP is not attributable to labeling but to externalization. The proposed analysis deduces the labelability of pair-merged SOs as well as explains the null-subject phenomenon (i.e., T can label without the *n*P). It also follows that T and R are not the only unlabelable heads.

Since interpretability is a notion that is rooted in the interfaces, if labelability is reducible to interpretability, then the discussion in this paper provides another argument for the Strong Minimalist Thesis, the basic hypothesis in the Minimalist Program, which states that language keeps to the simplest recursive operation, Merge and the interfaces (i.e., Full Interpretation).

References

Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parameterizing AGR: Word order, V-movement, and EPP-checking. *Natural Language and Linguistic Theory* 16:491–539.

Baker, Mark and Ken Hale. 1990. Relativized minimality and pronoun incorporation. *Linguistic Inquiry* 21:289–298.

Benmamoun, Elabbas. 1992. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

Bobaljik, Jonathan David. 1995. *Morphosyntax: The syntax of verbal inflection*. Doctoral dissertation, MIT, Cambridge, MA.

Bobaljik, Jonathan David and Samuel Brown. 1997. Head movement and the extension requirement. *Linguistic Inquiry* 28:345–356.

Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5:167–218.

Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.

Brandi, Luciana and Patrizia Cordin. 1989. Two Italian dialects and the null subject parameter. In *The null subject parameter*, ed. by Osvaldo Jaeggli and Ken Safir, 111–162. Dordrecht: Kluwer.

Broekhuis, Hans. 2015. Feature inheritance versus extended projections. Ms., Meertens Institute.

Carstens, Vicki. 2011. Hyperactivity and hyperagreement in Bantu. *Lingua* 121:721–741.

Carstens, Vicki. 2016. Delayed valuation: A reanalysis of goal features, "upward" complementizer agreement, and the mechanics of case. *Syntax* 19:1–42.

Carstens, Vicki and Michael Diecks. 2013. Parameterizing case and activity:

Hyper-raising in Bantu. In *NELS 40*, ed. by Seda Kan, Claire Moore-Cantwell, and Robert Staubs, 99–118. Amherst: University of Massachusetts, Graduate Student Association.

Castillo, Juan Carlos, John Dury, and Kleanthes K. Grohmann. 2009. Merge over move and the extended projection principle: MOM and the EPP revisited. *Iberia: An International Journal of Theoretical Linguistics* 1:53–114.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. by Stephen R. Anderson and Paul Kiparsky, 232–286. New York: Holt, Rinehart, and Winston.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

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

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *The cartography of syntactic structures*. Vol. 3, *Structures and beyond*, ed. by Adriana Belletti, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies, and beyond: Studies in honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 1–16. Amsterdam: John Benjamins.

Chomsky, Noam. 2016. Puzzles about phases. Ms., MIT.

Craenenbroeck, Jeroen van and Marcel den Dikken. 2006. Ellipsis and EPP repair. *Linguistic Inquiry* 37:653–664.

Diecks, Michael. 2012. Parameterizing case: Evidence from Bantu. *Syntax* 15:253–286.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Epstein, Samuel David and T. Daniel Seely. 2006. *Derivations in minimalism*. Cambridge: Cambridge University Press.

Epstein, Samuel David, Hisatsugu Kitahara, and T. Daniel Seely. 2012. Exploring

phase based implications regarding clausal architecture. A case study: Why structural Case cannot precede theta. In *Phases: Developing the framework*, ed. by Ángel J. Gallego, 103–123. Berlin: Mouton de Gruyter.

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

Epstein, Samuel David, Hisatsugu Kitahara, and T. Daniel Seely. 2016. Phase cancellation by external pair-merge of heads. *The Linguistic Review* 33:87–102.

Fassi-Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer.

Fukui, Naoki. 1986. *A theory of category projection and its applications*. Doctoral dissertation, MIT, Cambridge, MA.

Fukui, Naoki. 1995. The principles-and-parameters approach: A comparative syntax of English and Japanese. In *Approaches to language typology*, ed. by Masayoshi Shibatani and Theodora Bynon, 327–372. Oxford: Oxford University Press.

Gallego, Ángel J. and Juan Uriagereka. 2007. Conditions on sub-extraction. In *Coreference, modality, and focus*, ed. by Luis Eguren and Olga Fernández Soriano, 45–70. Amsterdam: John Benjamins.

Gelderen, Elly van. 2011. *The linguistic cycle: Language change and the language faculty*. Oxford: Oxford University Press.

Givón Talmy. 1976. Topic, pronoun, and grammatical agreement. In *Subject and topic*, ed. by Charles Li, 149–188. New York: Academic Press.

Grohmann, Kleanthes K., John Dury, and Juan Carlos Castillo. 2000. No more EPP. In *WCCFL 19*, ed. by Roger Billerey and Brook Danielle Lillehaugen, 153–166. Somerville, MA: Cascadilla Press.

Guasti, Maria Teresa and Luigi Rizzi. 2002. Agreement and tense as distinct syntactic positions: Evidence from acquisition. In *The cartography of syntactic structures*. Vol. 1, *Functional structure in DP and IP*, ed. by Guglielmo Cinque, 167–194. Oxford: Oxford University Press.

Hiraiwa, Ken. 2016. Labeling roots and pronouns. Paper presented at North East Linguistic Society (NELS) 47, University of Massachusetts, Amherst, 14 October.

Kawashima, Ruriko and Hisatsugu Kitahara. 2015. On the core ideas of generative grammar – Then and now. In *In untiring pursuit of better alternatives*, ed. by Hiroki Egashira et al., 198–208. Tokyo: Kaitakusha.

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

Kinyalolo, Kasangati K. W. 1991. *Syntactic dependencies and the spec-head agreement*

hypothesis in KiLega. Doctoral dissertation, UCLA, Los Angeles, CA.
Kitahara, Hisatsugu. 2017. Maximizing syntactic minimization: Five preliminary studies. *Reports of the Keio Institute of Cultural and Linguistic Studies* 48:239–265.
Lasnik, Howard. 2003. *Minimalist investigations in linguistic theory*. London: Routledge.
Lasnik, Howard. 2004. The position of the accusative subject in the accusative-infinitive construction. In *Non-nominative subjects*. Volume 1, ed. by Peri Bhaskararao and Karumuri Venkata Subbarao, 269–281. Amsterdam: John Benjamins.
Legate, Julie Anne. 2011. Under-inheritance. Paper presented at North East Linguistic Society (NELS) 42, University of Toronto, 12 November.
Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Cambridge, MA: MIT Press.
Mizuguchi, Manabu. 2014. Phases and counter-cyclicity of A-movement. In *Proceedings of the 16th Seoul international conference on generative grammar (SICOGG)*, ed. by Jong Un Park and Il-Jae Lee, 257–277. The Korean Generative Grammar Circle.
Mizuguchi, Manabu. 2016a. De-activation of ϕ through pair-Merge. *English Linguistics* 33:108–118.
Mizuguchi, Manabu. 2016b. Simplest Merge, labeling, and A′-movement of the subject. In *Phase theory and its consequences: The initial and recursive symbol* S, ed. by Miyoko Yasui and Manabu Mizuguchi, 41–83. Tokyo: Kaitakusha.
Mizuguchi, Manabu. 2017. Labeling in ECM and its implications for phase theory. Paper presented at Manchester Forum in Linguistics (mFiL) 2017, University of Manchester, 29 April.
Ndayiragije, Juvénal. 1999. Checking economy. *Linguistic Inquiry* 30:399–444.
Ormazabal, Javier. 1995. The syntax of complementation: On the connection between syntactic structure and selection. Doctoral dissertation, University of Connecticut, Storrs.
Ouali, Hamid. 2008. On C-to-T ϕ-feature transfer: The nature of agreement and anti-agreement in Berber. In *Agreement restrictions*, ed. by Roberta D'Alessandro, Susann Fischer, and Gunnar Hrafn Hrafnbjargarson, 159–180. Berlin: Mouton de Gruyter.
Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart, and Winston.
Platzack, Christer. 2003. Agreement and null subjects. *Nordlyd* 31:326–355.
Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.
Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
Roberts, Ian. 2010. A deletion analysis of null subjects. In *Parametric variation: Null*

subjects in minimalist theory, ed. by Theresa Biberauer et al., 58–87. Cambridge: Cambridge University Press.
Roberts, Ian and Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In *Parametric variation: Null subjects in minimalist theory*, ed. by Theresa Biberauer et al., 1–57. Cambridge: Cambridge University Press.
Rochemont, Michael and Peter Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
Rooryck, Johan. 1997. On the interaction between raising and focus in sentential complementation. *Studia Linguistica* 51:1–49.
Saito, Mamoru. 1982. Case marking in Japanese: A preliminary study. Ms., MIT.
Saito, Mamoru. 2017. Japanese wh-phrases as operators with unspecified quantificational force. *Language and Linguistics* 18:1–25.
Simpson, Andrew and Zoe Wu. 2002. Agreement, shells, and focus. *Language* 79:287–313.
Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425–450.
Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10:80–126.
Takano, Yuji. 2015. Surprising constituents as unlabeled syntactic objects. *Nanzan Linguistics* 10:55–73.
Taraldsen, Knut Tarald. 1992. Agreement as pronoun incorporation. *Glow Newsletter* 28:50–51.
Wexler, Kenneth and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Manabu Mizuguchi
Toyo University
Department of Social Psychology
Faculty of Sociology
5-28-20 Hakusan Bunkyo-ku
Tokyo 112-8606, Japan

mizuguchi@toyo.jp

Received: 2017. 4. 11
Revised: 2017. 5. 4
Accepted: 2017. 5. 16