Reformulating Pair-Merge, Inheritance and Valuation

by

Jun Omune

Submitted to Foreign Studies Research Division in partial fulfillment of the requirements for the degree of Doctor of English Linguistics

at

Kansai Gaidai University

Accepted in March 2018
(Submitted in November 2017)

Reformulating Pair-Merge, Inheritance and Valuation

Jun Omune

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Yukio Oba for the continuous support of my Ph.D. study and related research and for his patience, motivation and immense knowledge. His guidance helped me throughout the research and writing processes of this thesis. I could not imagine having a better advisor and mentor for my Ph.D. study.

Besides that, I would like to thank Prof. Nobuo Okada, for his insightful comments and encouragement as well as for the difficult questions that incented me to widen my research from various perspectives.

My sincere thanks also goes to Daisuke Hirai, who studied alongside me in Kansai Gaidai University and provided me an opportunity to join a team in Kindai University, and who provided me access to the laboratory and research facilities. Without his precious support, it would not have been possible to conduct this research.

I thank my friends for the stimulating discussions and for all the fun we have had in the last four years.

Last but not least, I would like to thank my family (my parents, my sister and my grandparents) for supporting me throughout the writing of this thesis and my life in general.

Chapter 2 in this thesis is based on my forthcoming paper in *English Linguistics* 34 (2). This forthcoming paper is a revised and extended version of the paper presented at the 153rd Conference of the Linguistic Society of Japan. I would like to express my gratitude to Yukio Oba and Daisuke Hirai for their advice at an early stage of this work. I also thank Hisatsugu Kitahara and the audience at the conference for their valuable comments and suggestions. In addition, I am grateful for the very valuable feedback from anonymous *EL* reviewers. Needless to say, all remaining errors are mine.

Chapter 4 is an amended and expanded version of the paper presented at the 2nd regular meeting of Fukuoka Linguistic Circle in 2017. I am grateful to the audience for the valuable comments and suggestions.

Section 5.3 (in chapter 5) stems from my forthcoming paper in *Kinki University (Kindai University), Center for Liberal Arts and Foreign Language Education Journal, Foreign Language Edition* 8 (2). I gratefully acknowledge the valuable comments and feedback from the two anonymous reviewers of this work. All remaining weaknesses and errors in this section are my own responsibility.

Contents

Chapter 1

Introduction

The main aim of this paper is to reformulate problematic operations that, to some extent, weaken minimalist theory, in terms of the SMT (Strong Minimalist Thesis), evolvavility, simplicity and/or computational efficiency.[1] The SMT holds that "[l]anguage is an optimal solution to legibility conditions" (Chomsky (2000: 96)). The recent minimalist model (cf. Chomsky (2013, 2015a, 2016a, b, c, 2017a, b) and Chomsky, Gallego and Ott (to appear)) eliminate many stipulations specific to human language. The model simplified by the elimination suggests that the *sole* operation specific to the language capacity is Merge, and other things indispensable for C_{HL} (computational system for human language), which fundamentally yields the *language of thought*, conform to third-factor principles. The third-factor principles are not specific to language but assumed to follow the laws of nature, which are the laws that science

[1] This dissertation is based on my works (i.e. Omune (2016, 2017, to appear-a, b)). As for further details, see Acknowledgements.

attempts to explicate. Thus, we assume that the operations obeying the third factor are not flaws but fit the SMT in minimalist syntax. To establish a derivational system involving Merge, the single operation specific to human language, however, we are required to reformulate operations such as pair-Merge, feature inheritance and feature valuation. These operations seem to be specific to human language against the minimalist spirit. This paper reformulates these operations for the conceptual reasons discussed above. Furthermore, it will be shown that the reformulation explains various linguistic phenomena such as (non-)extractability of the indirect object in the double object construction and the non-referential cognate object in the cognate object construction, a symmetric property in the *there* construction and among others.

1.1 Basic Assumptions and the Architecture of Grammar

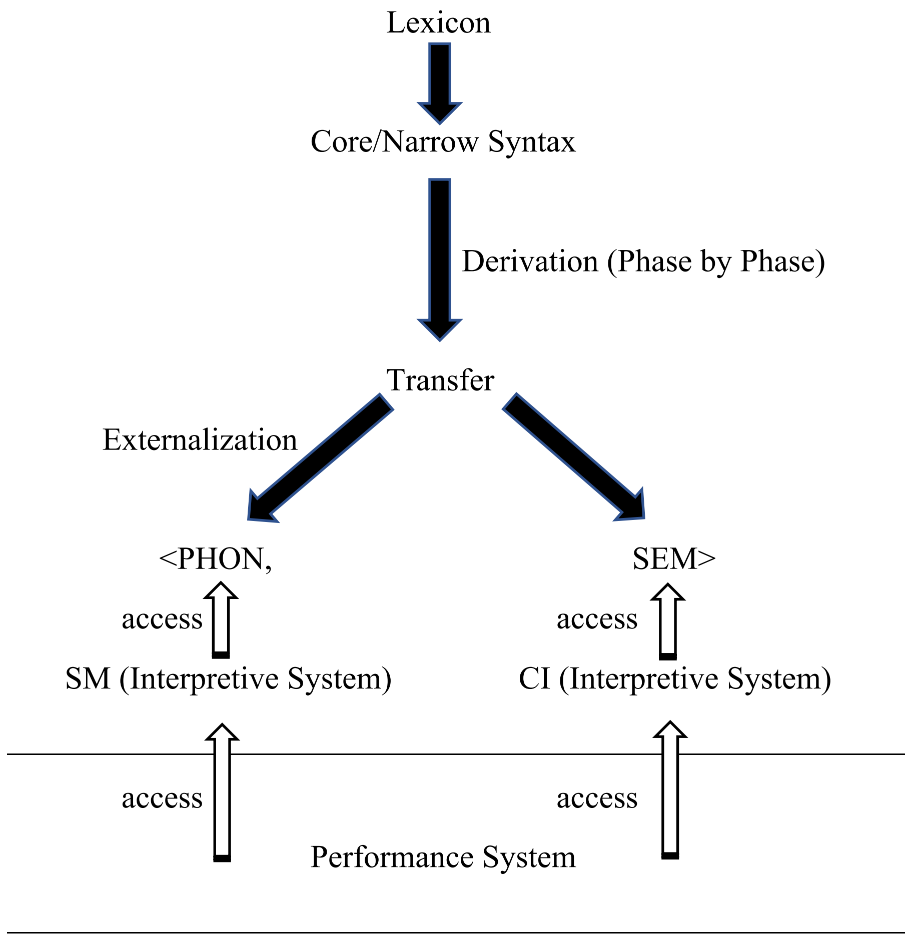
In the current minimalist model in Chomsky (2013, 2015a) and Chomsky, Gallego and Ott (to appear), we assume that structures built by Merge get transferred phase by phase after relevant operations have applied. The operation *Transfer*, however, arguably does not *Spell-Out* structures because transferred phases are not pronounced at the original positions but at other positions as discussed in Obata (2010) (see also Chomsky (2016c) and Chomsky, Gallego and Ott (to appear)). Spell-Out, therefore, does not exist (see Chomsky (2016c) and Chomsky, Gallego and Ott (to appear)), but Transfer at least yields the effects of the PIC (Phase Impenetrability Condition). Chomsky (2000: 108) defines the PIC as follows:

(1) Phase Impenetrability Condition:
In phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

Transfer is also generally assumed to map structures generated in core/narrow syntax onto <PHON, SEM>. PHON is accessed by the SM (sensorimotor) system, and SEM is accessed by the CI (conceptual-intentional) system. Thus, we can assume that Transfer sends *information* of objects constructed in core syntax to those two interpretive systems.

After that, computations do not need to look at or care about the objects left in core syntax. Accordingly, the original conception of *phase*, reducing a burden of memory, is still ensured. Notice that the representations accessed by the SM and CI systems are further accessed by the performance system. The architecture of grammar (i.e. Y-model) in the current minimalist model can be illustrated as follows:

(2) Architecture of Grammar (Y-model):



Externalization, which is mapping structures onto PHON, is complex and could be the locus of cross-linguistic variation.[2] In contrast, the course from the core syntax to SEM

[2] The process of externalization contains *linearization*, which alters hierarchically structured expressions (or set-theoretic objects) into linear order. For more on linearization, see Kayne (1994). Moreover, the process of externalization seems to involve *vocabulary insertion* in the framework of Distributed Morphology (see Halle and Marantz (1993), Embick and Marantz (2008), Embick and

is universal. The contents of the lexicon are debatable and seem to be mysterious in the current model. This paper mainly focuses on core syntax, derivation and Transfer. Particularly, I pursue the Merge-only derivation that third-factor principles appropriately affect because of the reasons briefly discussed earlier. I do not discuss the lexicon, externalization, PHON and the performance system.

1.2 Derivation

Merge has been simplified in the Minimalist Program since Chomsky (1995a, b). The bare phrase structure theory (Chomsky (1995a, b)) eliminated X′-level objects, but left projections and labels. Chomsky redefined the projections or labels as the equivalent to the features of lexical items (i.e. heads) and proposed the following definition:

(3) Merge (α, β) = {α, {α, β}}

The italicized α is the label of {α, β}, which is identical to the head or the feature α. There is no X′ projection such as α′ in the sense of X′ theory. In addition, there is no maximal projection like α_{max}. Chomsky (1995a, b) built the label α into the definition for the empirical necessity, although he notes that Merge = {α, β} is the simplest case. In other words, we do not have any conceptual necessity to assume the label in the definition of Merge. Many researchers (e.g. Collins (2002), Seely (2006) and Epstein, Kitahara and Seely (2014a)) explored the simplest case of Merge, and Chomsky (2013, 2015a) finally eliminated the label from the definition. It follows that the outer brackets in {α, {α, β}}, which denote a projection (see Seely (2006)), were also removed. Now that simplest Merge forms just an unordered set {α, β}, syntactic structures are formed without labels. However, it is obvious that unlabeled structures are not interpretable at the CI interface and the SM interface. For example, nominal phrases have nominal

Noyer (2007), Marantz (1997, 2013) and among others). Roughly, Vocabulary Insertion post-syntactically inserts phonological features.

interpretations and verb phrases have verbal interpretations. The remaining option to label unlabeled structures is arguably minimal search which conforms to a third-factor principle, Minimal Computation. Chomsky (2013) calls this labeling process a *labeling algorithm*, but it is important to note that the labeling algorithm is just minimal search, and, hence, not an arbitrary stipulation.[3]

As has been briefly reviewed above, we have simplified Merge and now assume that Merge is simply a set-formation operation that produces an unordered set {α, β}. We assume that Merge recursively applies, regardless of whether its application is external or internal. Hence, Merge ensures discrete infinity in the mind. In addition, if its application is internal to a set, internal Merge instantiates displacement in language. Because the basic properties of language, discrete infinity and displacement, are arguably exclusive to humans, Merge does not entirely fit third-factor principles. Merge, therefore, conforms to first-factor genetic endowments that are sometimes called *universal grammar* or UG.

There is another set-formation operation, pair-Merge, which yields an ordered pair $<\alpha, \beta>$. As in Chomsky (2000, 2004), pair-Merge is assumed to be another primitive operation. That is, we have two primitive operations, which form sets.

(4) a. Merge (α, β) = {α, β}
  b. Pair-Merge (α, β) = $<\alpha, \beta>$

While Merge takes two elements and forms an unordered set, pair-Merge takes two elements and forms an ordered set. Pair-Merge is used to explain the asymmetric properties of adjunction in the current framework of minimalist syntax (see Chomsky (2000, 2004, 2015a)). That is, this operation has the inherent asymmetricity (see Chomsky (2000: 133)). However, this inherent asymmetric property should not be part of genetic endowments with regard to simplicity and evolvavility. In the Minimalist

[3] For more on the recent analyses utilizing this algorithm, see Bošković (2016), Carstens, Hornstein and Seely (2016), Epstein Kitahara and Seely (2016), Rizzi (2016) and Saito (2016), among others. This paper, particularly in chapter 2, adapts the approach by Epstein Kitahara and Seely (2016).

Program, simplicity is required for methodological/procedural minimalism that is compatible to Ockham's razor. Evolvavility also needs to be addressed in order to explain the evolution of the language capacity. Simplest Merge is so simple that we can use it to explain the sudden emergence of the language capacity on the evolutionary timescale. In contrast, pair-Merge is somewhat complex due to the inherent asymmetricity or the inherent order. It is uncertain that pair-Merge emerged suddenly on the evolutionally timescale. Therefore, UG should be composed of the sole operation, simplest Merge, on conceptual grounds. To eliminate pair-Merge as a primitive operation, I will argue that ordered pairs can be derived from unordered sets formed by simplest Merge.[4]

In the current minimalist model, the structure and derivation of *John hit Mary* is as follows (in the fifth step, <Phi, Phi > and √P are labels, and see chapter 2 for more on these labels):[5]

---

[4] Note that the notion of *ordered pair* is necessary to interpret the paired expressions. For instance, we can understand pair-list readings and single-pair readings as follows:

(i) Who bought what?
  a. John bought a book, Mary bought a bag, Bill bought a cup. (pair-list reading)
  b. ??John bought a book. (single-pair reading)

(ii) Who just bought what?
  a. ??John bought a book, Mary bought a bag, Bill bought a cup. (pair-list reading)
  b. John bought a book. (single-pair reading)

These facts show that we undoubtedly have the notion of ordered pair in the mind. However, these facts do not suggest that the notion of ordered pair should be created by a *primitive* operation. Internal Merge forms unordered sets {X, {Y, t_X}} (*t* for a lower copy) but also creates a pair <X, X>, which is identified as two copies (see Chomsky (2007, 2008) and Epstein, Kitahara and Seely (2014a)). This pair is created *as a consequence* of the application of internal Merge but not directly created by Merge. Recall that Merge just forms unordered sets.

[5] External arguments might be *severed* from verbs and the head *voice* might appear in relevant argument structures (see Kratzer (1996), Marantz (1981, 2005, 2013) and Wood and Marantz (2015)). However, I adopt the standard assumption that external arguments externally merge to SPEC-v*.

(5) $\{\mathrm{C}, \{_{\beta}\ \mathrm{John}, \{_{\alpha}\ \mathrm{T}, \{t_{\mathrm{John}}, \{\langle\sqrt{}\mathrm{hit}, v^{*}\rangle, \{_{\delta}\ \mathrm{Mary}, \{_{\gamma}\ t_{\sqrt{}\mathrm{hit}}, t_{\mathrm{Mary}}\}\}\}\}\}\}\}$

- i. Merge externally forms $\{_{\gamma}\ \sqrt{}\mathrm{hit}, \mathrm{Mary}\}$.
- ii. Merge internally forms $\{_{\delta}\ \mathrm{Mary}, \{_{\gamma}\ \sqrt{}\mathrm{hit}, t_{\mathrm{Mary}}\}\}$.
- iii. Merge externally forms $\{\mathrm{John}, \{v^{*}, \{_{\delta}\ \mathrm{Mary}, \{_{\gamma}\ \sqrt{}\mathrm{hit}, t_{\mathrm{Mary}}\}\}\}\}$.
- iv. √hit inherits features from v*.
- v. Labeling takes place by minimal search: δ and γ are labeled as <Phi, Phi> and √P, respectively.
- vi. Pair-Merge internally forms <√hit, v*> with v* affixed: such a v* becomes invisible, and the phase-hood is activated on $t_{\sqrt{}\mathrm{hit}}$.
- vii. The complement of $t_{\sqrt{}\mathrm{hit}}$ gets transferred.
- viii. Merge externally forms $\{_{\alpha}\ \mathrm{T}, \{\mathrm{John}, \{\langle\sqrt{}\mathrm{hit}, v^{*}\rangle, \{_{\delta}\ \mathrm{Mary}, \{_{\gamma}\ t_{\sqrt{}\mathrm{hit}}, t_{\mathrm{Mary}}\}\}\}\}\}$.
- ix. Merge internally forms $\{_{\beta}\ \mathrm{John}, \{_{\alpha}\ \mathrm{T}, \{t_{\mathrm{John}}, \{\langle\sqrt{}\mathrm{hit}, v^{*}\rangle, \{_{\delta}\ \mathrm{Mary}, \{_{\gamma}\ t_{\sqrt{}\mathrm{hit}}, t_{\mathrm{Mary}}\}\}\}\}\}\}$.
- x. Merge externally forms $\{\mathrm{C}, \{_{\beta}\ \mathrm{John}, \{_{\alpha}\ \mathrm{T}, \{t_{\mathrm{John}}, \{\langle\sqrt{}\mathrm{hit}, v^{*}\rangle, \{_{\delta}\ \mathrm{Mary}, \{_{\gamma}\ t_{\sqrt{}\mathrm{hit}}, t_{\mathrm{Mary}}\}\}\}\}\}\}\}$.
- xi. T inherits features from C.
- xii. Labeling takes place by minimal search: β and α are labeled as <Phi, Phi> and TP.
- xiii. The complement of C gets transferred after labeling.

In this system, Merge applies freely, regardless of whether its application is external or internal.[6] It follows that Merge produces structures that are interpreted as *gibberish* by the CI system, contrary to *crash-proof syntax* (see Frampton and Gutmann (2002)). The <Phi, Phi> labels in (5v, xii) denote that agreement occurs and that Phi features are shared (see Chomsky (2013)). One might say that the <Phi, Phi> labels in (5v, xii) do not

[6] It is important to note that Gallego (2009) shows that Chomsky's (2000, 2001) *Activity Condition* no longer holds. The activity condition was introduced on behalf of the formulation of *Probe-Goal Agree*. Roughly, this condition states that "an XP is 'frozen in place' when it is inactive" Gallego (2009: 36, fn. 2).

provide any semantic meaning, though labels are required for interpretations. Despite the concern, however, <Phi, Phi> seems to be necessary. To argue for its necessity, we consider unvalued features that might be apparent imperfection.

In Chomsky (2008) and his subsequent work, unvalued features are markers of phases. Heads bearing uPhi (unvalued Phi) (i.e. C and v*) are therefore qualified to be phases. uPhi must be valued before being transferred because features without values violate Full Interpretation and cannot possibly receive phonological realizations. Notice that some heads inherit features from phase heads (see (5iv, xi)). This *feature inheritance* or *inheritance*, which was proposed by Chomsky (2008), is assumed to occur primarily because of theoretical reasons. As Richards (2007) (cf. Chomsky (2007, 2008) and Richards (2012)) argues, Transfer and feature valuation (i.e. agree) occur at the same time. This is because, once unvalued features are valued, the CI system cannot distinguish them from inherently/lexically valued features (see Richards (2007: 566)). Moreover, it is generally assumed that the PIC states that "The edge and nonedge (complement) of a phase are transferred separately" (Richards (2007: 568)). On these premises, Richards (2007) (see also Chomsky (2007)) argues that feature inheritance is required because uPhi is transferred as soon as Transfer-Valuation applies. It is important to note that this logic does not completely hold in the current system as in (5). Some valued uPhi features are not transferred but remain within the non-edge of the v*P phase (see (5vi, vii)). As Goto (2016) argues, it seems that feature inheritance, nevertheless, still needs to be postulated for labeling to take place correctly. This type of reasoning suggests that the <Phi, Phi> label actually provides an interpretation, which is probably a *derivational history* (cf. Munataka (2017)). If <Phi, Phi> labels can provide the derivational history, then we can solve the problem of indistinguishability between *derivationally* valued features and *lexically* valued features. <Phi, Phi> labels are sufficient to distinguish these features because they encode the information that unvalued features have been valued through derivations. Therefore, feature inheritance and <Phi, Phi> are (at least theoretically) essential.

As has argued above, we need to postulate feature inheritance in (5iv, xi), but it is still a problem on purely conceptual grounds. As discussed earlier, we should assume only one structure-building operation that conforms to genetic endowments. *Operation* feature inheritance should not be part of UG. However, how do heads inherit features without this operation? In the earlier framework, C-T or v*-√root relations were established by Merge, but there is no such establishment in the derivation above because Merge applies strictly cyclically. Additionally, it is obscure how agreement or feature valuation occurs in the current model. In chapter 4, I reformulate feature inheritance and feature valuation, based on the third-factor Minimal Computation.

1.3 The Structure of the Paper

This paper is organized as follows. Chapter 2 reformulates pair-Merge of heads, based on simplest Merge and a definition of ordered pairs in set theory. This reformulation further accounts for the double object construction, the cognate object construction, the small clause construction and the ECM (Exceptional Case Marking) construction. Chapter 3 reformulates pair-Merge of phrases, based on proposals in chapter 2. This expanded reformulation of pair-Merge explains phenomena concerning phrasal adjunction and the Specificity Effect. Chapter 4 reformulates feature inheritance and feature valuation to be minimal search. The reformulated version of inheritance and valuation can explicate the structure and the derivation of the *there* construction such as labeling failure and a long-distance agreement. Furthermore, the analyses in this chapter provide an explanation for the empirical facts of extractability of associates and scopes and anaphors in the *there* construction. Chapter 5 presents alternative analyses. It will be shown that one of the alternatives, in a different manner, potentially solves a technical problem of feature inheritance, which had already been solved in chapter 4. Furthermore, the other alternative explicates the double object construction under the maximality principle in Rizzi (2015a, b, 2016). However, the main purpose of this chapter is to show that those alternatives are not solutions but raise new research problems. Chapter 6 concludes the paper.

Chapter 2

Reformulating Pair-Merge of Heads

2.1 Introduction

As Chomsky (2013: 42) notes, (set-)Merge takes two objects α, β and forms the simplest unordered set {α, β} (see Collins (2002) and Seely (2006) for more on simplified Merge). Neither label nor node is created by this simplest operation, Merge, based on UG (Universal Grammar), contrary to Chomsky (1995a, b). Other indispensable mechanisms (e.g. labels for interpretation) are, thus, motivated by "principles not specific to the faculty of language, the third factor (Chomsky (2005: 6))." Accordingly, the SMT (Strong Minimalist Thesis) holds as follows, satisfying interface conditions:

(1) (SMT) Interfaces + Merge = Language

(Chomsky (2010: 52))

Ideally, the simplest account of UG reduces to the single, simple and primitive operation Merge in the current Minimalist Program (Chomsky (2013, 2014, 2015a, b, 2016a, b), etc.), conforming to third-factor principles such as minimal computation and computational efficiency. If not, the theory in the program readily faces the serious problem of the evolution of the language capacity. As for evolvability, Chomsky (2014: 11) states (see also Hauser, Chomsky and Fitch (2002), Hauser et al. (2014)):

> "… It appears that there has been no evolution of [the] language [capacity] (or virtually none; or of cognitive capacities generally) since our ancestors left Africa, perhaps about 50,000 years ago … If we go back roughly 50,000 years before that, there is little evidence that human language existed at all; archaeological evidence suggests that language, and with it complex cognition, emerged within this very narrow window, in what Jared Diamond called a 'great leap forward.' Doubling the numbers or more changes little; the window remains very narrow in evolutionary time, and any millions of years after separation from other surviving species. These facts suggest that at some point within this narrow range some slight rewiring of the brain occurred yielding the core property of language: Merge with its output linked to the CI [conceptual-intentional] interface. Mutations occur in an individual, not a group. The individual endowed with this rewiring would therefore have had a 'language of thought' LOT: a means to interpret, reflect, plan, etc., in principle unbounded."

Therefore, Merge, the sole operation for the basic properties of discrete infinity and displacement, should ideally be as simple as possible so that it emerges quite suddenly on the evolutionary timescale.

There is, however, another primitive operation pair-Merge empirically required to account for the asymmetric property of adjunction (see Chomsky (2004: 117–118)). Pair-Merge is not simple but complex by virtue of its definition. When pair-Merge applies, it takes two objects α, β and forms an ordered pair <α, β>. To form the ordered pair, however, pair-Merge must relegate one object (α or β) to a separate plane. Clearly, it is complex to introduce such an extra plane and to move an element to a separate plane.

Such a complex operation cannot possibly emerge suddenly on the evolutionary timescale. Thus, pair-Merge should not be part of UG for such purely conceptual reasons.

In the current Minimalist Program, Merge and pair-Merge are formulated as follows:

(2) a. Merge (α, β) = {α, β}
b. Pair-Merge (α, β) = <α, β>

Merge forms a simple unordered set while pair-Merge yields an ordered set (see Chomsky (2013, 2015a), Epstein, Kitahara and Seely (2014a, 2015, 2016) and works cited by them). Pair-Merge does so because pair-Merge (i.e. adjunction) inherently has an asymmetry as noted by Chomsky (2000: 133). This inherent asymmetry is not, however, welcome in the spirit of simplest Merge which we use to justify evolvability as mentioned above. Furthermore, pair-Merge is not only controversial on purely conceptual grounds but also contradictory to the invisibility of one element in an ordered pair. This contradiction will be discussed in detail in the next section. To resolve the conceptual problems above, I propose that there is no pair-Merge of heads as a primitive operation, but a *derivational* ordered pair derives from unordered sets.[1] Consequently, it will be shown that the proposal clearly fits the SMT in the best possible way.

The paper is organized as follows. Section 2.2 reviews Chomsky's (2015a) derivation of a typical transitive construction such as *John hit Mary* and Epstein, Kitahara and Seely's (2016) analysis of the bridge verb construction such as *John thinks that she will sing*. I then point out the conceptual contradiction of pair-Merge regarding the invisibility of one element in <α, β> formed by pair-Merge. Section 2.3 resolves this contradiction, proposing a new formulation of pair-Merge. This reformulation lets a derivational ordered pair derive from unordered sets, conforming to the simplest and ideal account of UG including Merge only. Section 2.4 shows that the proposal accounts for empirical facts concerning the double object construction. Section 2.5 explains

[1] As for the possible elimination of pair-Merge of phrases, see chapter 3. Also, see Oseki (2015). He eliminates the stipulation on pair-Merge by proposing that pair-Merge of phrases is Merge which forms the double peaked structure as in Epstein, Kitahara and Seely (2012, 2014a, 2015).

empirical facts of the cognate object construction. Section 2.6 further explains empirical facts with respect to the small clause construction and the ECM construction. Section 2.7 summarizes this chapter.

2.2 Theoretical Background, Analytical Assumptions and Problems

In addition to the conception of Merge in (2), Chomsky (2013, 2015a) argues that labels are required for interpretation at two interfaces: the SM (sensorimotor) system for externalization and the CI (conceptual-intentional) system for thought. Conforming to a third-factor principle of minimal computation, labeling takes place by minimal search as follows (where *t* is a copy and used only for expository purposes):

(3) Labeling by Minimal Search:

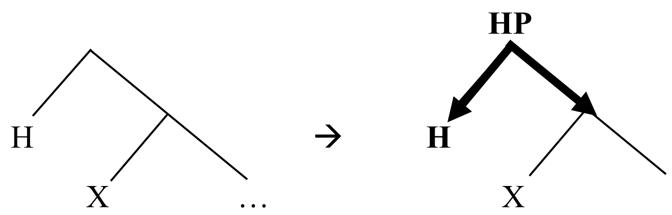
a. $\{H, XP\} \rightarrow \{_{HP}\ H, XP\}$

b. $\{XP, t_{YP}\} \rightarrow \{_{XP}\ XP, t_{YP}\}$

c. $\{XP, YP\}$ / X sharing features (e.g. Phi-set) with Y $\rightarrow \{_{<Phi, Phi>}\ XP, YP\}$

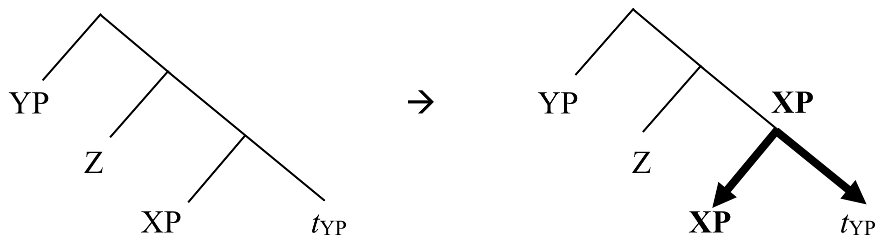
In the case of Head-Phrase SO (syntactic objects) such as (3a), minimal search unambiguously finds a head because it is the closest computational atom. The graph-theoretic notation of Head-Phrase labeling is illustrated below:

(4)



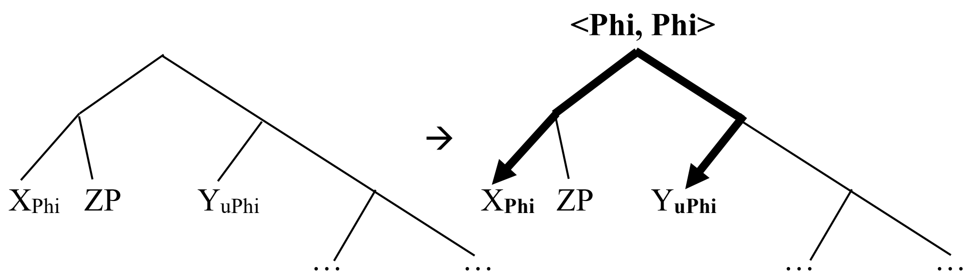
In the case of Phrase-Phrase SO, minimal search fails to find a head because neither XP nor YP is atomic. To label {XP, YP}, we need to modify it before labeling takes place. In (3b), the lower copy of YP (i.e. t_{YP}) remains after internal Merge has modified {XP, YP} following the No Tampering Condition (Chomsky (2008: 138)). Minimal search then successfully finds a head X in XP since a lower copy is, by definition, invisible to minimal search.[2]

(5)



There is another case for labeling {XP, YP} where X and Y share some feature(s), as shown in (3c). That is, agreement takes place when they share features. For example, Agree (X, Y) takes place by minimal search as suggested by Chomsky (2013, 2015a, b), and X and Y share Phi (phi-features) under the valuation of uPhi (unvalued phi-features). Such Phi becomes a label <Phi, Phi> for {XP, YP}.[3]

(6)



Based on the framework sketched out in Chomsky (2013), Chomsky (2015a) further assumes that Merge, which is strictly cyclic, applies freely. Let us consider the following typical transitive structure for the v*P phase of *John hit Mary* where √hit is the

---

[2] Chomsky (2013: 44) assumes "α to be 'in the domain D' if and only if every occurrence of α is a term of D." Epstein, Kitahara and Seely (2016: 91) define *occurrence* as follows: "an occurrence of [α] is a sister-category merged to [α] by set-Merge."

[3] The angle brackets of <Phi, Phi> have nothing to do with those of the ordered pair <α, β>. The notation <Phi, Phi> just indicates that Phi is shared.

root HIT, and v* is a phase-head, an affix and a verbalizer (I omit several irrelevant notations such as labels and curly brackets for nominals):[4]

(7) {John, {<√hit, v*>, $\{_{\delta}$ Mary, $\{_{\gamma}$ $t_{\surd \mathrm{hit}}$, t_{Mary}}}}}

i. Merge externally forms $\{_{\gamma}$ √hit, Mary}.

ii. Merge internally forms $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, t_{Mary}}}.

iii. Merge externally forms {John, {v*, $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, t_{Mary}}}}}.

iv. √hit inherits features from v*.

v. Labeling and Agree take place by minimal search: δ and γ are labeled as <Phi, Phi> and √P, respectively.

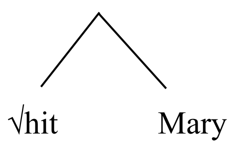
vi. Pair-Merge internally forms <√hit, v*> with v* affixed: such a v* becomes invisible, and the phase-hood is activated on $t_{\surd \mathrm{hit}}$.

vii. The complement of $t_{\surd \mathrm{hit}}$ gets transferred.

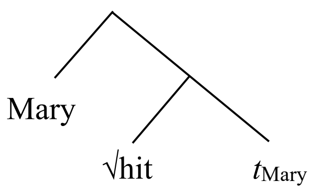
The graph-theoretic notation of this structure and derivation is provided below:

(8) {John, {<√hit, v*>, $\{_{\delta}$ Mary, $\{_{\gamma}$ $t_{\surd \mathrm{hit}}$, t_{Mary}}}}}

i.

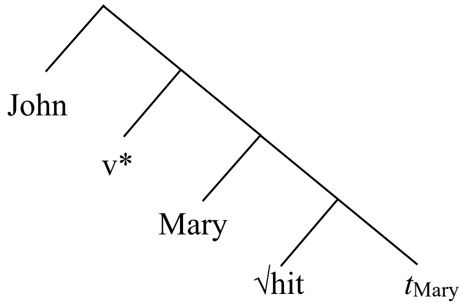


ii.

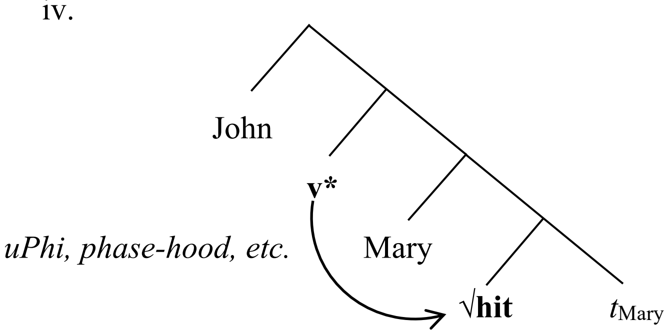


[4] Chomsky (2015a) does not refer to the timing of agreement, but he implies that feature valuation (i.e. Agree) is minimal search as well as labeling. Epstein, Kitahara and Seely (2014b) explicitly argue that Agree/Valuation is minimal search, and Kasai et al. (2016: 7) also conclude that "'[a]greement' is obtained as a by-product of labeling." On this account, I assume both labeling and Agree take place with the same timing. Also, see Nomura (2017) for an elaborated feature-valuation system under minimal search.

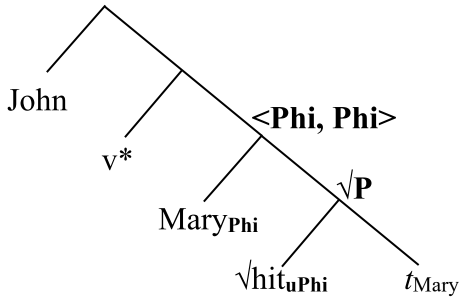
iii.


John
v*
Mary
√hit
t_{Mary}

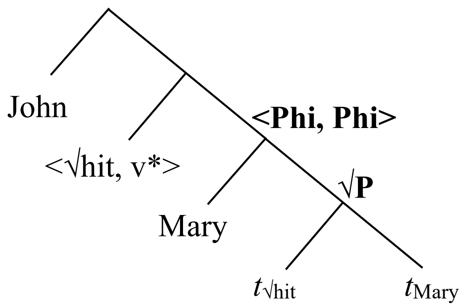
iv.


John
v*
uPhi, phase-hood, etc.
Mary
√hit
t_{Mary}

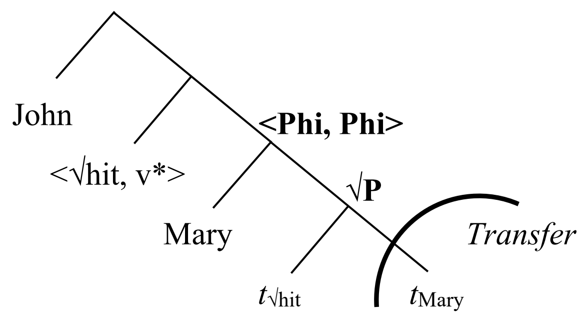
v.


John
v*
<Phi, Phi>
√P
$Mary_{Phi}$
$√hit_{uPhi}$
t_{Mary}

vi


John
<√hit, v*>
<Phi, Phi>
√P
Mary
$t_{√hit}$
t_{Mary}

vii.



Chomsky (2015a: 7–8) assumes that √root alone is universally too weak to serve as a label because it is unspecified as to categories. However, √root strengthens by SPEC-√root just as English-type T strengthens by SPEC-T under labeling/agreement. Thus, √hit strengthens by *Mary*, and minimal search labels γ as √P in (7v).[5] Internal pair-Merge (i.e. head-raising) then applies for categorizing √hit as *verb*. That is, √hit raises to v*, leaving the visible copy $t_{\sqrt{\text{hit}}}$ in (7vi).[6] Interestingly, the affix v* adjoins to √hit and becomes invisible against "conventional treatments of head-raising" (Chomsky (2015a: 12)). In other words, the traditional approach to head-raising has the raised element adjoin to its host, but Chomsky (2015a) assumes that the host adjoins to the raised element. Therefore, Chomsky states that internal pair-Merge of heads seems to be part of the process of externalization as reported in Nomura (2017: 399, fn. 3) (see also Chomsky, Gallego and Ott (to appear)). Internal pair-Merge (of heads) is, however, arguably a syntactic operation because it is just one mode of syntactic movement (see Richards (2009)). I would like to return later in this section to discuss the conceptual paradox between the traditional approach to head-raising and internal pair-Merge of heads by Chomsky (2015a).

Adopting the framework briefly reviewed above, Epstein, Kitahara and Seely (2016) point out that the system in Chomsky (2015a) has a conceptual problem, which is the

---

[5] Chomsky (2015a: 10) uses the notation RP instead of √P.

[6] This lower copy is visible because its phase-hood becomes active after internal pair-Merge has formed <√hit, v*> (see Chomsky (2015a) and Epstein, Kitahara and Seely (2016)). Accordingly, we should assume <√hit, v*> is a different SO to √hit. In other words, the lower copy √hit is in the domain D since every occurrence of √hit is a term of D (see also note 2).

failure of labeling, regarding the bridge verb construction. Let us consider the bridge verb construction for the matrix v*P phase of *John thinks that she will sing*. Provided that we follow the exact steps in (7), Merge externally forms {√think, {that, …}}, internally forms {{that, …}, {√think, $t_{\{\text{that}, \dots\}}$}}, and then externally forms {v*, {{that, …}, {√think, $t_{\{\text{that}, \dots\}}$}}}. This will clearly cause agreement failure because C (i.e. *that*) has no appropriate feature agreeing with √think (which has inherited uPhi from v*). Chomsky (2015a) tentatively solves this problem by no application of internal Merge to {that, …}. Epstein, Kitahara and Seely (2016), however, point out that this solution still has a problem with labeling. The γ label in {v*, $\{_\gamma$ √think, {that, …}}} is not determined since √think alone is too weak to serve as a label. Recall that √think needs to strengthen by the agreement relation of SPEC-√think.

To resolve the problems, Epstein, Kitahara and Seely (2016) propose an alternative analysis, assuming the null hypothesis where pair-Merge applies freely as well as Merge.

(9) {John, {<√think, v*>, {that, …}}}

i. Pair-Merge of heads externally forms <√think, v*>: v* becomes invisible with respect to both its uPhi and its phase-hood.

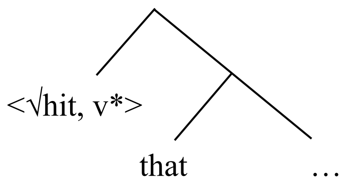
ii. Merge externally forms {John, {<√think, v*>, {that, …}}}.

The graph-theoretic notation of this structure and derivation is shown below:

(10) {John, {<√think, v*>, {that, …}}}

i. √hit v* → <√hit, v*>

ii.



In this analysis, pair-Merge *externally* applies and forms the amalgam <√think, v*>. The amalgamation, thereby, makes both the uPhi of v* and phase-hood of v* invisible.

Therefore, the failure of the agreement will not happen as there is no (visible) uPhi in <√think, v*>. In addition, because √think alone cannot label but <√think, v*> can (see Chomsky (2015a: 12)), labeling takes place unproblematically. Epstein, Kitahara and Seely (2016) expand this new type of rule application, external pair-Merge of heads, to other constructions. "[P]hase-cancellation by external pair-Merge of heads takes place in verbal phrases with passive, raising, unaccusative and bridge verbs" (Epstein, Kitahara and Seely (2016: 97)). Therefore, we do not need to stipulate the weak v (see Chomsky (2001)); the phase-head $v^{(*)}$ is always strong.

The analyses by Chomsky (2013, 2015a) and Epstein, Kitahara and Seely (2016) contribute to the theory of the Minimalist Program in that they apparently conform to the UG-based Merge and the third factor. Nevertheless, pair-Merge has enormous influence on syntactic computation, although it should not exist as a primitive operation on purely conceptual grounds. First, pair-Merge is not the UG-based-simplest Merge (α, β) (i.e. (2a)). Second, pair-Merge (α, β) = <α, β> (i.e. (2b)) is not simple as it has the ordered pair created by the adjunction on a separate plane. Pair-Merge is not, therefore, ideal in syntactic theory in terms of evolvability and simplicity.

The metaphorical notion *separate plane* is stipulated by Chomsky (2004: 117–118). According to him, SO on a separate plane becomes invisible because the operation sees SO only on a primary plane, which is the simple structure.[7] This mechanism, however, does not work well in the current framework developed by Chomsky (2013, 2015a) and Epstein, Kitahara and Seely (2016). Let us consider the typical transitive structure and the derivation of {John, {<√hit, v*>, {Mary, {$t_{\sqrt{\text{hit}}}$, t_{Mary}}}}} in (7) as if Chomsky's (2004) account of pair-Merge were tenable. When the verbal root √hit internally pair-merges to v*, √hit attaches on a separate plane. Hence, √hit becomes invisible, contrary to Chomsky's (2015a) account shown in (7vi). This contradiction does not entirely terminate the system in (7) because v* *empirically* adjoins to √hit as Chomsky (2015a:

[7] Chomsky (2004: 118) proposes an optional operation SIMPL (simplification) which converts an ordered pair to an unordered set when Transfer applies. Thus, the mate to the other SO becomes visible if SIMPL applies to the ordered pair containing them.

12) notes. It is, nevertheless, problematic on purely conceptual grounds.[8] Therefore, syntactic theory should dispense with the stipulation of *separate plane*. Alternatively, one might add the notion of direction to Merge for eliminating the notion of a separate plane. That is, in the case of (7), √hit *literally moves to v* and attaches to it*. This spatial movement is clearly not an ideal resolution because it makes Merge much more complex, going against the spirit of simplest Merge.

Another problem is that the system in Chomsky (2015a) largely depends on the invisibility of affixes without conceptually firm ground. In other words, there is no conceptually firm ground on which the claims of invisibility of affixes are based. Affixes, say v* and C, are trivially visible because Merge externally applies to these in the first place. That is, affixes do not have the special property which makes them invisible in ordered pairs. Then, what makes affixes invisible in ordered pairs?

(11) Problems of Pair-Merge of Heads in Chomsky (2015a):

a. Pair-Merge of heads necessarily involves a separate plane as far as it is a syntactic operation.

b. There is no affixal property which makes affixes invisible in ordered pairs.

In the next section, I reformulate pair-Merge by employing ZFC (Zermelo-Fraenkel Set-Theory with the Axiom of Choices).

2.3 Proposal: A Solution with ZFC

Syntactic structure has been formally defined by set-theoretic relations in the Minimalist Program since the introduction of the bare phrase structure theory by

---

[8] An anonymous reviewer points out that Chomsky (2015a) no longer assumes the separate plane. Because Chomsky (2015a) does not discuss pair-Merge of phrases, it seems that he discards this problematic plane. However, his later work (see Chomsky, Gallego and Ott (to appear) and Chomsky (2017a)) still adopts Chomsky's (2004) version of pair-Merge. Thus, as far as assuming pair-Merge is a syntactic operation, pair-Merge of heads necessarily involves the separate plane.

Chomsky (1995a, b) (cf. (14)). In Chomsky (1995a, b), Merge used to contain a label and a projection in its definition which was Merge = $\{\alpha_2, \{\alpha_1, \beta\}\}$ where α_2 is a label, and the outer brackets denote a projection (see Seely (2006)). Simplest Merge, however, does not contain a label and a projection as in (2a); Merge just takes two elements α, β and forms the simplest set-theoretic object {α, β}. Because syntactic theory has taken advantage of set theory, it is not impossible to reformulate pair-Merge by adopting a definition of an ordered pair in ZFC.

In ZFC or basic set theory, the widely accepted definition of an ordered pair is as follows (see Kuratowski (1921), Bagaria (2014), etc.):

(12) $<\alpha, \beta> = \{\{\alpha\}, \{\alpha, \beta\}\}$

This defines an ordered pair based on an unordered set in that the property of the ordered pair "$<a, b>=<c, d> \leftrightarrow a=c \land b=d$" also holds in "$\{\{a\}, \{a, b\}\}=\{\{c\}, \{c, d\}\} \leftrightarrow a=c \land b=d$." If we adopted this definition in syntactic theory, we could successfully reformulate pair-Merge based on Merge. That is because the ordered pair <α, β> is equal to the unordered sets {{α}, {α, β}} which Merge could form. However, a singleton set {α} typically does not exist in syntax; singleton sets are redundant.[9] Computational atoms are generally assumed to be LI (Lexical Items) (i.e. functional heads and roots in the current framework) in syntactic theory.[10] Assuming α and β are computational atoms, one might say {α} corresponds to a non-branching projection by LI in X′-theory, but bare phrase structure theory has abandoned such projection (see Chomsky (1995a, b)). Also, Merge equally applies to both LI and sets of SO. That is, both are SO, as Collins (2014: 3) describes in (13).

---

9 At least a noun might be the singleton set in terms of unvalued features, according to Kayne (2011). However, his analysis might not be relevant if we consider the view of anti-lexicalism (see Marantz (1997) and other related work).

10 In Distributed Morphology (Halle and Marantz (1993), Marantz (1997, 2005, 2013), Embick and Marantz (2008), etc.), abstract morphemes (i.e. categorizers/functional heads and roots) are assumed to be computational atoms in syntax. This view clearly affects the current framework of syntax. For related approaches, see Borer (2003), her subsequent work and Mateu (1997, 2002, 2010, 2012, 2014) as well.

(13) X is a syntactic object iff

i. X is a lexical item, or

ii. X is a set of syntactic objects.

The distinction between {α} and α is significant in set theory because {α} is not a member, but α is. In set theory, objects are either members of sets or not. In contrast, this is not very important in syntactic theory because both SO {α} and LI α are SO as in (13). In other words, both LI and SO can be terms in the technical sense. As for *term*, I adopt the simplified definition by Epstein, Kitahara and Seely (2012: 262, 2015: 162), which was originally proposed by Seely (2006: 201) (see also Chomsky (1995a, b), (2008: 158, fn. 16)).

(14) For any structure K,

i. K is a term of K, and

ii. if L is a term of K, then the members of L are terms of K.

Furthermore, computation clearly causes total chaos if it freely allows the formation of a singleton set without independent evidence.[11] Therefore, I adopt an alternative definition of an ordered pair in ZFC (see Tourlakis (2003: 182–183) and other related work).

(15) $<\alpha, \beta> = \{\alpha, \{\alpha, \beta\}\}$

Note that this definition also satisfies the characteristic property of ordered pairs: "$\{a, \{a, b\}\}=\{c, \{c, d\}\} \leftrightarrow a=c \land b=d$."

[11] (Simplest) Merge naturally forms a singleton set in arithmetic if there is only one element (see Chomsky (2012: 15, 263)). Merge yields the natural number system as follows: Merge takes 0 and recursively forms $\{0\} = 1$, $\{0, \{0\}\} = \{\{0\}\} = 2$, $\{0, \{0, \{0\}\}\} = \{\{\{0\}\}\} = 3, \ldots$ This case is irrelevant to the singleton formation of LI since Merge needs to apply to only one element in arithmetic.

Assuming that (15) also holds in syntactic theory, I propose to reformulate pair-Merge of heads as follows:

(16) Pair-Merge of heads formulated by Simplest Merge (PM by SM):

a. Pair-Merge (α, β) = Merge (α, (α, β)) = Merge (α, {α, β}) = {α, {α, β}} = <α, β>

b. Pair-Merge (α, β) = Merge (β, (α, β)) = Merge (β, {α, β}) = {β, {α, β}} = <β, α>

In this proposal, the original formulation of pair-Merge in (2b) is reformulated by simplest Merge in (2a) under the definition of an ordered pair in ZFC (see (15)).[12] Consequently, pair-Merge is to take two elements (e.g. α, β) and form a simple set (i.e. {α, β}), and then to take one element in the set (i.e. α in {α, β}) and the set (i.e. {α, β}) and form another simple set (i.e. {α, {α, β}}). In short, pair-Merge is Merge×2. I call this new kind of rule application *Pair-Merge of heads formulated by Simplest Merge* (PM by SM for expository purposes). Also, {α, {α, β}} (or {β, {β, α}}) in (16) is called a *derivational ordered pair* in this paper (cf. (17), note 13).

I assume α and β are atomic elements in this reformulation. As an anonymous reviewer and Yukio Oba (p.c.) point out, we could expand the double application of simplest Merge, forming a derivational ordered pair, to sets (i.e. XP-level objects). Yet, this has an empirical problem under the current minimalist system (Chomsky (2013, 2015a)). Reconsider the formation of $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, $t_{\text{Mary}}\}\}$ in (7). When Merge internally forms δ, it is the equivalent of {α, {α, β}} if we assume α may be an XP-level object. Namely, $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, $t_{\text{Mary}}\}\}$ is interpreted as <Mary, √hit>, though it is trivial that δ is not an adjunct-structure. It, nevertheless, seems worth exploring this

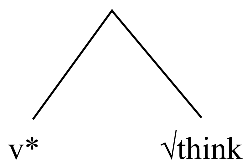
[12] The proposed definition of pair-Merge of heads is not an elementary operation, but it is a new type of application of set-Merge of heads. Accordingly, an anonymous reviewer indicates that it seems confusing to use the terminology "pair-Merge" of heads. To avoid such a confusing situation, I put the word *informally* when using pair-Merge as the informal cover term of PM by SM. See also note 14.

expansion for the quest of eliminating pair-Merge of phrases. I will explore the possibility of PM by SM applying to phrases in chapter 3.

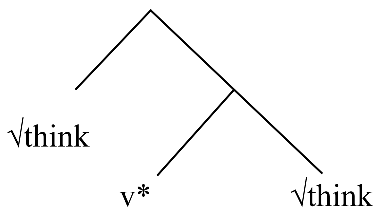
Under proposal (16), let us consider the case of <√think, v*> in (9), employing the graph-theoretic notation informally. Regarding the notation, I do not distinguish the higher copy from the lower copy. Namely, the lower copy √think in {√think, {√think, v*}} (see (17iii)) is not represented by using *t*.

(17) Derivational Steps of <√think, v*>:

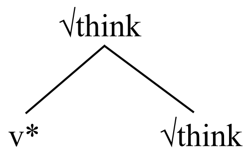
i. External Merge (√think, v*)



ii. Internal Merge (√think, {√think, v*})



iii. {√think {√think, v*}} = <√think, v*>



In the final step, {√think, {√think, v*}} is recognized as <√think, v*> in syntax. Namely, what the graph-theoretic notations in (17ii–iii) mean is that the ordered pair <√think, v*> is derivationally {√think, {√think, v*}} (i.e. (17ii)) but representationally <√think, v*> (i.e. (17iii)).[13] The notion of an ordered pair <α, β> is not abandoned since it is interpreted as adjunction at the CI interface and the process of externalization (see

[13] Hence, {α, {α, β}} (or {β, {β, α}}) in (16) is called a *derivational* ordered pair in this paper.

Chomsky (2004)). Thus, unless stipulated, the notion <α, β> is automatically available in syntax (under the definitions of (15) and (16)).[14] Note that √think on the so-called mother node of {√think, v*} is not a label. The graph-theoretic notation in (17iii) just indicates that it is the amalgam <√think, v*>, and v* is an adjunct.[15] Informally, <√think, v*> is not a *phrase* but a verbalized *word*.

PM by SM (16) is simpler than the original definition of pair-Merge in that it eliminates *a primitive operation*, namely pair-Merge, along with *a problematic separate plane*. In addition to the simplification achieved by eliminating an entire plane and a problematic one at that, postulating no operation is undoubtedly simpler than postulating an operation. Furthermore, Merge × 2 is not complicated. That is because the application of Merge is costless; hence the free application of Merge. We freely apply recursive Merge as many times as we want (the basic properties of discrete infinity and displacement). Notice that {α, {α, β}} = <α, β> in (16) is not an operation. As discussed above, it indicates that {α, {α, β}} can be representationally interpreted as <α, β> under (15); hence (17iii). Hereafter, I use <α, β> and {α, {α, β}} interchangeably as far as α and β are heads, but it does not mean there is a pure syntactic object like <α, β>. <α, β> is {α, {α, β}}, which is formed by PM by SM (i.e. Merge × 2).

As has been discussed above, one element in an ordered pair is invisible. This property is explained by the basic quality of an ordered pair. That is, in <α, β> (i.e. {α,

[14] An anonymous reviewer wonders why the proposed definition in (16) is essential if the notion of ordered sets is available in syntax. The proposed definition is significant because it eliminates an *operation*, pair-Merge of heads. In other words, the elementary operation, pair-Merge is discarded, but the notion of ordered sets is not discarded in (16).

[15] Interestingly, {α, {α, β}} used to be the definition of Merge in syntactic theory in Chomsky (1995a, b). As an anonymous reviewer mentions, this traditional definition Merge = {α, {α, β}} could be redefined in terms of ordered sets which include the information about the word order. As Chomsky (2017a) states, however, an ordered pair does not impose any linear order. Also, even if we adopt Merge = {α, {α, β}}, we still have two primitive operations, Merge = {α, {α, β}} and Merge = {α, β}. Again, the problem of evolvability emerges; how does Merge = {α, {α, β}} emerge suddenly on the evolutionary timescale? As far as evolvability is concerned, simplest Merge must be the sole operation for the basic property of recursive generation.

{α, β}}), α uniquely identifies β without seeing β as $f(\alpha) = \beta$ corresponds with <α, β>. In $f(\alpha) = \beta$, we only take care of α when determining β but not vice versa. Therefore, β in <α, β> is invisible. In the case of <√think, v*> in (17), v* is invisible since $f(\sqrt{}\text{think})$ = v*. Furthermore, as Chomsky (2015a) also notes that the ordered pair <α, β> is an amalgam, we can recognize it as a single object. Thus, v* is invisible because it has been part of the head √think in <√think, v*>. We regard <√think, v*> as the single verbalized head √think including v* which is invisible with regard to uPhi and phase-hood in syntax.[16]

The proposal, therefore, resolves the problems of the original formulation of pair-Merge (see (11)). The irreducible component of UG, simplest Merge eliminates the operation pair-Merge of heads, which involves the dubious separate plane and the questionable property of affixes discussed above.

This logic above immediately suggests the following interesting prediction. The reverse ordered pair <v*, √think>, which is {v*, {v*, √think}}, is freely formed by Merge. That is, there are two logically possible cases of an ordered pair unless it is arbitrarily stipulated. One case is (16a), and the other is (16b). Therefore, <v*, √think> entails that v* is visible, but √think is invisible; the invisibility does not relate to the dubious notions of *the separate plane* and the affixal property. This conceptually ideal assumption (in terms of evolvability, simplicity and the SMT) is supported by several empirical facts presented in the following sections (2.4, 2.5 and 2.6). Before moving to the next section, however, we will reconsider some simple structures and derivations.

First, reconsider the following structure and derivation for the matrix v*P phase of the bridge verb construction, *John thinks that she will sing* (see (9)) under proposal (16) (bold parts below are the entirely new analyses entailed by PM by SM):

(18) {John, {{√think, {√think, v*}}, {that, …}}}

---

[16] I assume the substantial property of v* is visible at least at the CI interface and at the process of externalization, following the implication by Chomsky (2015a) and Epstein, Kitahara and Seely (2016). If the entire property of v* in <√root, v*> is invisible there, the ordered pair cannot get the legitimate interpretation. See also note 7.

After the embedded CP phase {that, …} has been formed, …

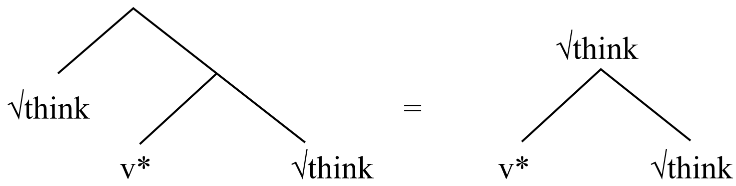
i. **Merge externally forms {√think, v*}, and then internally forms {√think, {√think, v*}}: v* becomes invisible thanks to the basic quality of an ordered pair**. (I.e., {√think, {√think, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√think, v*>**.)

ii. Merge externally forms {John, {{√think, {√think, v*}}, {that, …}}}.

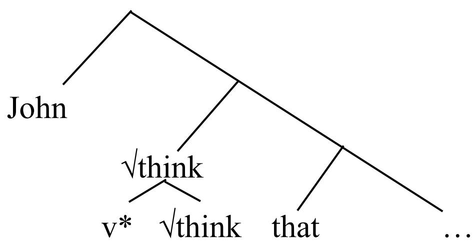
The graph-theoretic notation of this structure and derivation is as follows:

(19) {John, {{√think, {√think, v*}}, {that, …}}}

i.



ii.



The derivation of the bridge verb construction no longer involves external pair-Merge as a primitive operation. The v* in {√think, {√think, v*}} becomes invisible for the basic quality of an ordered pair. In (18i), Merge *externally* applies first and *internally* applies second. This type of application of PM by SM is called in this paper the *external* application of PM by SM for expository purposes.

Second, we reconsider the following typical transitive structure for the v*P phase of *John hit Mary* under proposal (16):

(20) {John, {{√hit, {√hit, v*}}, {$_\delta$ Mary, {$_\gamma$ $t_{\surd\text{hit}}$, t_{Mary}}}}}

i. Merge externally forms {$_\gamma$ √hit, Mary}.

ii. Merge internally forms {$_\delta$ Mary, {$_\gamma$ √hit, t_{Mary}}}.

iii. Merge externally forms {John, {v*, {$_\delta$ Mary, {$_\gamma$ √hit, t_{Mary}}}}}: √hit inherits features from v*.

iv. **Merge internally forms {√hit, v*}, and then {√hit, {√hit, v*}}: v* becomes invisible**. (I.e., {√hit, {√hit, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√hit, v*>**.)

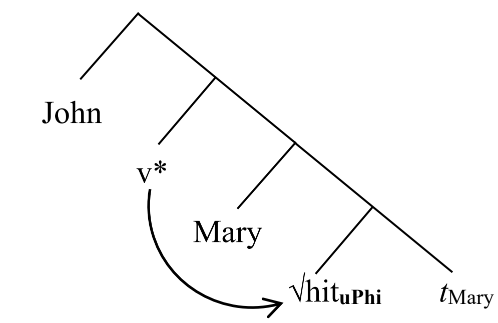
v. Labeling and Agree take place by minimal search: δ and γ are labeled as <Phi, Phi> and √P, respectively.

vi. The complement of $t_{\surd\text{hit}}$ gets transferred.

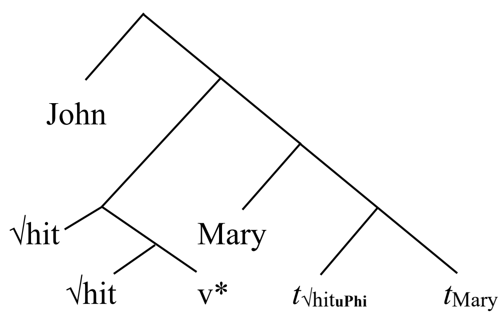
The graph-theoretic notation of this structure and derivation is presented below:

(21) {John, {{√hit, {√hit, v*}}, {$_\delta$ Mary, {$_\gamma$ $t_{\surd\text{hit}}$, t_{Mary}}}}}

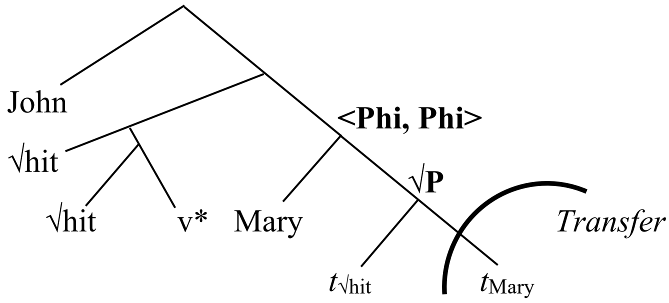
i–iii.



iv.



v–vi.



In (20iv), Merge internally forms {√hit, {√hit, v*}} under the *internal* application of PM by SM. Such a v* becomes invisible due to the basic property of an ordered pair. Also, it is necessary to change the order of the application for PM by SM since pair-Merge is reformulated to be Merge. There is an independent motivation for changing the order of PM by SM.[17] Because (internal) pair-Merge of heads is a special operation in Chomsky's (2015a) system (cf. Nomura (2017: 399, fn. 3)), it could be applied *just before Transfer* (see (7)). Given that labeling is part of Transfer (Chomsky (2015a: 6)), it is, nevertheless, dubious that pair-Merge applies *between labeling and actual Transfer* as long as we assume pair-Merge is a syntactic operation. Therefore, internal PM by SM should apply *just before labeling*. In other words, this timing is *after feature inheritance* in the canonical derivation. One might say that √hit in the derivational ordered pair {√hit, {√hit, v*}} bears uPhi, and it causes the derivation to crash with the violation of Full Interpretation. Adapting Epstein, Kitahara and Seely (2017) and Kitahara (2017), I assume that the valuation of uPhi on *any* copy satisfies Full Interpretation. This assumption is plausible because every copy is, in effect, the same element. The other aspects in (20) are the same as those of the normal transitive sentences such as (7).

It is logically possible in (20) that freely applying Merge forms the reverse ordered pair {v*, {√hit, v*}}, but this leads to empirically unwelcome results. Direct objects are empirically extractable, but the direct object *Mary* is transferred at the v*P phase since the phase-hood of v* is activated on {v*, {√hit, v*}}. Also, {v*, {√hit, v*}} implies a

[17] As mentioned in section 2.2, it is the null hypothesis that pair-Merge applies freely in any order as well as Merge. See Epstein, Kitahara and Seely (2016), Mizuguchi (2016), Nomura (2017) and Sugimoto (2016) for this topic.

non-default semantic effect; √hit is interpreted as MANNER since it modifies v* as an adjunct. For MANNER adjunction, see Marantz (2005, 2013), Acedo-Matellán (2010), Acedo-Matellán and Mateu (2014), among others.

There remain two theory-internal concerns in (20). The first one is that the internal (pair-)Merge of heads in (20iv) apparently causes the problem of multidimensionality, which is a ternary relation (see Chomsky (2007: 6, 2013: 40, fn. 20, 2015b: 82)). Merge is binary but not ternary, quaternary or more in terms of simplicity and computational efficiency. That is, multidimensionality violates Merge's requirement for minimal search. As for the ternary relation, Chomsky (2015b: 82) notes that "[y]ou're finding one item in the workspace, you find something inside it – that's two – and then you're finding a third one which you attach this to." For example, in (20iv), PM by SM finds {John, {v*, {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}}}} first, it finds v* second and finds √hit last; this is the ternary relation. This problem is, however, not serious because the original internal pair-Merge of heads is also ternary in this sense. For instance, in (7vi), pair-Merge finds {John, {v*, {Mary, {√hit, t_{Mary}}}}} first, it finds v* second, and then it finds √hit as far as internal pair-Merge is one type of Merge. It is, therefore, reasonable to assume that morphological Merge(r) only sees atomic elements which are morphemes, heads or lexical items.[18] As discussed in (13), it is not incomprehensible that we assume Merge treats both a lexical item and a syntactic set similarly as SO. However, minimal search for Merge can distinguish them because minimal search for labeling recognizes lexical items to be simpler. I tentatively assume Merge takes advantage of this distinction when we apply morphological Merge(r) (16) which ultimately forms a derivational ordered pair.

(22) Merge only sees relevant heads iff it applies to lexical items. (tentative)

This kind of requirement is essential whether we assume PM by SM or not as far as internal pair-Merge of heads applies in syntax. It furthermore suggests that feature

[18] Here, morphological Merge(r) simply means that Merge applies to two atomic elements. Morphological Merger was first proposed by Marantz (1981).

inheritance activates condition (22) because PM by SM internally forming a derivational ordered pair such as (20iv) applies just after feature inheritance. Assumption (22) is, therefore, finally formulated as follows:

(23) Merge only sees relevant heads iff feature inheritance has established the relation of two lexical items.

The case of PM by SM externally forming a derivational ordered pair (informally, external pair-Merge) such as (18i) is, of course, independent of (23) (and (22)). Assumption (23) further suggests that feature inheritance is a special case of minimal search because minimal search may establish relations. Its special case is probably minimal search by uPhi for determining a receiver of uPhi.

As an alternative solution to the problem of multidimensionality, we can adopt the idea by T. Daniel Seely (p.c.) and Epstein, Kitahara and Seely (2014b). He or they assume that external Merge of heads establishes the relation of v*-√root (or C-T). In their assumption, Merge forms {IA, {{v*, √root}, …}} and then internally forms {v*, {IA, {{t_{v*}, √root}, …}}} in the normal course of a derivation (IA: internal argument). In v*P, internal pair-Merge finally forms {<√root, v*>, {IA, {{t_{v*}, $t_{\surd root}$}, …}}}. If this is tenable, the analogous effect to condition (23) holds since the v*-√root relation has been established by the first external Merge of v*-√root. I would like to return to the issues concerning (23) in chapter 4.

The second technical concern is a problem of replacement by counter-cyclic internal Merge (see Epstein, Kitahara and Seely (2012, 2013, 2014)). Epstein, Kitahara and Seely (2012) critically discuss that internal Merge cannot form {C, {EA, {T, {t_{EA}, {…}}}}} after {C, $\{_{\alpha}$ T, {EA, {…}}}} has been created as long as Merge forms the simplest sets.[19] If Merge first yields "{C, { T, {EA, {…}}}}" and then "$\{_{\beta}$ EA, {T, {t_{EA}, {…}}}}," that is all. Internal Merge cannot create the following set: "{C, $\{_{\beta}$ EA, {T, {t_{EA}, {…}}}}}" because the set α must be replaced with β. That is, the counter-cyclic

[19] EA: external argument.

internal Merge requires the additional operation *replacement*. Therefore, the application of Merge should be strictly cyclic in the current minimalist model. PM by SM potentially has this replacement problem. Let us consider the creation of {John, {$_{\alpha}$ v*, {Mary, {√hit, t_{Mary}}}}} in (20)–(21). In this derivation, specifically (20iii–iv), Merge first yields {John, {v*, {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}}}} and then {{√hit, {√hit, v*}}, {$_{\delta}$ Mary, {$_{\gamma}$ $t_{\sqrt{\text{hit}}}$, t_{Mary}}}}, and that is all. The problem is that we might need to assume that the counter-cyclic internal Merge has replacement to yield {John, {{√hit, {√hit, v*}}, {$_{\delta}$ Mary, {$_{\gamma}$ $t_{\sqrt{\text{hit}}}$, t_{Mary}}}}}. This assumption is not possible because it makes Merge complex. However, we could avoid this problem without assuming replacement.

It seems that v* and {√hit, {√hit, v*}} (or <√hit, v*>) are, in a sense, the same object. Recall that v* and {√hit, {√hit, v*}} are both heads. In other words, v* and the *head* {√hit, {√hit, v*}} should be the object bearing identical properties if we consider the No Tampering Condition. The property of syntactic objects should not be tampered by the application of Merge. I assume that the feature inheritance of √hit and v* ensures the uniformity of these heads. The properties of these heads are identical because {√hit, {√hit, v*}} is a head, and √hit inherits all features from v*. This reasoning seems to hold unless it is proved to be wrong. In the next section, it will be shown that PM by SM explains some empirical facts about the double object construction.

2.4 Case 1: The Double Object Construction

Given the proposal PM by SM, neither the separate plane nor the affixal property contributes to the invisibility of adjuncts. This view is supported by a widely known phenomenon: the (non-)extractability of grammatical objects in the double object construction.

(24) a. What did Mary give John t_{what}?

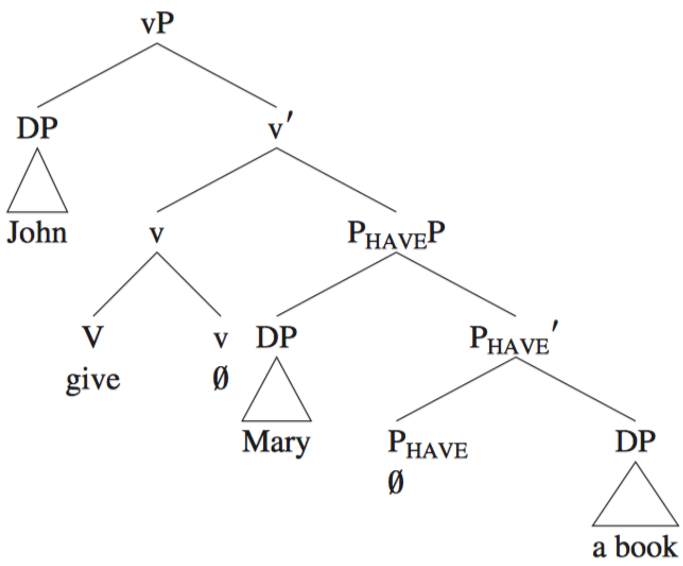
b. *Who did Mary give t_{who} the book?

Direct objects are extractable as in (24a), but indirect objects are not as in (24b). To explain the facts under PM by SM, I adopt Harley and Jung's (2015: 716) argument structure of the double object construction illustrated below (see also Harley (1995, 2002)):

(25) Mary gave John the book:
[[D Mary] [√give-v [[D John] [P_{HAVE} [the book]]]]]

For expository purposes, the graph-theoretic notation of their structure is shown below:

(26) John gave Mary a book:



(Harley and Jung (2015: 716))

In their structure, √give directly adjoins to the transitive v, and P_{HAVE} is a preposition denoting prospective possession. √give like other double object verbs is an adjunct which modifies the causative/transitive event under the version of Distributed Morphology that Harley and Jung (2015) adopt. Such modifiers specify the manner of the event by virtue of the adjunct position (MANNER adjunction).[20] Also, adopting the

[20] They do not explicitly state how √give semantically functions as MANNER. However, the core meaning of double object verbs (i.e. X causes Y to have Z) is clearly expressed by [Subj [v_{cause} [Obj_1 [P_{HAVE} Obj_2]]]] (v_{cause} for transitive/causative event) without a verbal root. See Marantz (2005, 2013),

basic derivational steps under freely applying Merge and labeling by minimal search in Chomsky (2015a) and Epstein, Kitahara and Seely (2016) as discussed in section 2.2, PM by SM entails the following structure and derivation for the v*P phase of *Mary gave John the book* (bold parts below are the entirely new analyses entailed by PM by SM):

(27) {Mary, {{v*, {v*, √give}}, {$_{\delta}$ John, {$_{\gamma}$ PHAVE, the book}}}}

i. Merge externally forms {John, {PHAVE, the book}}.

ii. **Merge externally forms {v*, √give} and internally forms {v*, {v*, √give}}: √give becomes invisible**. (I.e., {v* {v*, √give}} formed by PM by SM is **representationally interpreted as the ordered pair <v*, √give** >.)

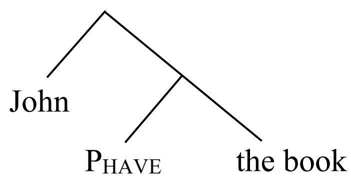
iii. Merge externally forms {Mary, {{v*, {v*, √give}}, {John, {PHAVE, the book}}}}, and PHAVE inherits features from v*.

iv. Labeling and Agree take place by minimal search, and **the complement of {v*, {v*, √give}}** (= <v*, √give >) **gets transferred**.

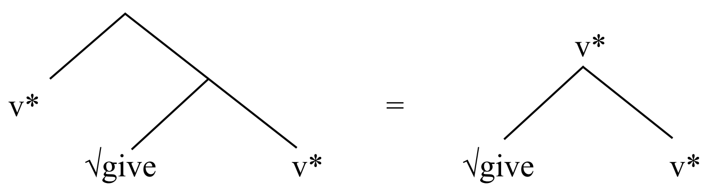
The graph-theoretic notation of this structure and derivation is illustrated below:

(28) {Mary, {{v*, {v*, √give}}, {$_{\delta}$ John, {$_{\gamma}$ PHAVE, the book}}}}

i.

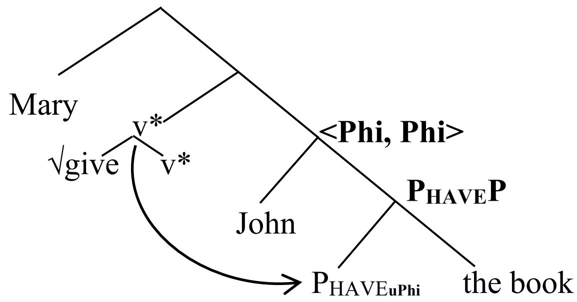


ii.

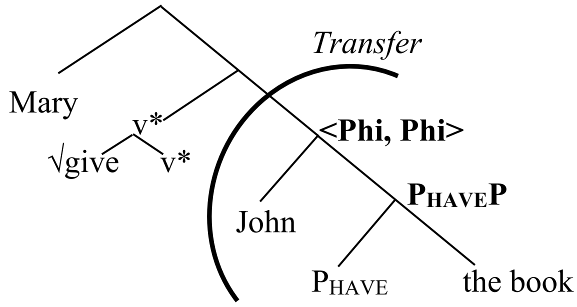


Acedo-Matellán (2010), Acedo-Matellán and Mateu (2014), Mateu (2002) and other related work on this topic.

iii.



iv.



In (27ii), instead of applying external pair-Merge of heads by Epstein, Kitahara and Seely (2016), Merge simply applies twice for forming the derivational ordered pair on account of proposal (16). Additionally, if P_{HAVE} inherits features from v* in the way that T inherits features from C, P_{HAVE} and its SPEC (i.e. DP *John*) share features (or agree). Thus, δ gets labeled as <Phi, Phi>, and γ gets labeled as P_{HAVE}P. Consequently, *de-phasing* as in (7vi) does not occur. Instead, the complement of the original phase head v* (i.e. the complement of {v*, {v*, √give}}) gets transferred in (27iv) because the v* is visible for the reasons discussed in section 2.3.

Crucially, the standard external pair-Merge of heads as in Epstein, Kitahara and Seely (2016) cannot explain the derivation of (25). That is because the standard external pair-Merge relies on the invisibility of affixes. If we adopt the standard pair-Merge, the derivation proceeds as follows. First, external pair-Merge yields <√root, v*> before external (set-)Merge applies to v*. Second, the v* including Phi becomes invisible since Epstein, Kitahara and Seely (2016) assume the invisibility of affixes which was originally suggested by Chomsky (2015a). (Recall that this invisibility is dubious because we do not have any conceptual reason why it becomes invisible in an ordered pair but is visible in other situations.) Finally, labeling and Agree fail because there is no candidate for the indirect object to agree with; the derivation crashes (see also (33)). On the other hand, PM by SM can explain the derivation of (25) as shown in (27). PM by SM thus has an empirical advantage over the standard external pair-Merge of heads in that PM by SM freely allows two patterns of a derivational ordered pair (i.e. {v* {v*, √give} and {√give, {√give, v*}}}.

The analysis in (27) explains the empirical facts in (24). Note that the shaded parts below indicate that they have been transferred in the v*P phase.

(29) a. What did Mary t_{what} give John t_{what}?

b. *Who did Mary t_{who} give t_{who} the book?

c. *___ did Mary give who the book?

First, in the double object construction, the preposition P_{HAVE} licenses Case on the direct object in the same way that the typical prepositions do.[21] Therefore, the direct object can internally merge from the original position to the edge of v*P, and then to SPEC-C after the introduction of C in (29a). Notice that if *what* remains at the original position, the sentence is interpreted as not a wh-question but a yes-no question which is gibberish (see Chomsky (2015a: 8)). In other words, the Q feature on the wh-phrase cannot get the correct operator-variable interpretation. Second, in (29b), the computation (i.e. minimal search) fails to value both uCase (unvalued Case) on the indirect object and uPhi on P_{HAVE} inheriting it from v* (see (27iii)). Once the indirect object *who* internally merges to the v*P, the lower copy t_{who} becomes invisible. That is, minimal search (i.e. Agree) cannot value unvalued features, but both t_{who} bearing uCase and P_{HAVE} bearing uPhi get transferred. It leads the derivation to crash at the interface(s), violating Full Interpretation. Finally, let us consider the other case in (29c); the indirect object would

[21] An anonymous reviewer wonders why Case on the direct object is not valued as a by-product of Phi-agreement, although uCase on the indirect object is valued in this way. This is because the former is not structural uCase, but inherent Case which is licensed by prepositions. Chomsky, Gallego and Ott (to appear) also note that one of the objects bears inherent Case in the double object construction.

try to internally merge after agreement had been completed successfully. In this case, it violates the PIC (Phase Impenetrability Condition) because *who* gets transferred immediately after labeling and Agree (i.e. minimal search) have taken place at the v*P phase (see (27iv)).

The analysis above expands to a variety of movement of the objects in the double object construction such as relativization, clefting, *tough* movement and topicalization.[22]

(30) a. This is the computer which Mary gave/sent the friend *t*.
b. *This is the friend who Mary gave/sent *t* the computer.
c. It is that computer that Mary gave/sent John *t*.
d. *It is John that Mary gave/sent *t* that computer.
e. That computer is impossible to give/send John *t*.
f. *John is impossible to give/send *t* that computer.
g. That computer, Mary gave/send John *t*.
h. *John, Mary gave/send *t* that computer.

(cf. Oba (2005: 61))

That is simply because movement is the instantiation of the application of internal Merge. Suppose that internal Merge applies to the indirect objects in all the deviant cases in (30). The uCase on the indirect object and uPhi on P_{HAVE} get transferred since there is no visible indirect object at SPEC-P_{HAVE}P. These transferred unvalued features violate Full Interpretation. Hence, the derivation crashes at the interface(s). Alternatively, even if agreement is completed prior to the internal merger of the indirect objects at the v*P phase, that internal merger violates the PIC. The indirect objects have been transferred because they are part of the complement of the phase head v* which does not de-phase. Thus, the empirical facts in (30) support PM by SM and the analysis in (27).

---

[22] Oba (2005) shows that the non-extractability of indirect object is due to Th/Ex in Chomsky (2001). Also, his subsequent work such as Oba (2016) explains the same phenomenon under Rizzi's (2010) criterial freezing.

As is well known, the indirect object is extractable in passives, such as *John was given the book*. This case is not a counterexample against the analysis above. The structure and derivation for the CP-v*P phase of *John was given the book* is shown below:

(31) {C, {John, {T, {{√give, {√give, v*}}, {t_{John}, {P_{HAVE}, the book}}}}}}

i. Merge externally forms {John, {P_{HAVE}, the book}}.

ii. **Merge externally forms {√give, v*} and internally forms {√give, {√give, v*}}: v* becomes invisible**. (I.e., {√give {√give, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√give, v*>**.)

iii. Merge externally forms {{√give, {√give, v*}}, {John, {P_{HAVE}, the book}}}: **nothing gets transferred because the phase-hood of v* has been canceled**.

iv. Merge externally forms {T, {{√give, {√give, v*}}, {John, {P_{HAVE}, the book}}}}.

v. Merge internally forms {John, {T, {{√give, {√give, v*}}, {t_{John}, {P_{HAVE}, the book}}}}}.

vi. Merge externally forms {C, {John, {T, {{√give, {√give, v*}}, {t_{John}, {P_{HAVE}, the book}}}}}}: T inherits features from C.

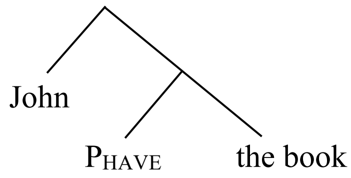
vii. Labeling and Agree take place at the timing of Transfer.

The graph-theoretic notation of this structure and derivation is illustrated below:[23]

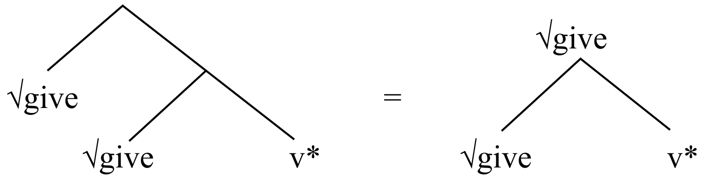
(32) {C, {John, {T, {{√give, {√give, v*}}, {t_{John}, {P_{HAVE}, the book}}}}}}

[23] The label of {{√give, {√give, v*}}, {t_{John}, {P_{HAVE}, the book}}} is represented as vP in the graph-theoretic notation, although its more accurate representation is {√give, {√give, v*}}P or <√give, v*>P.

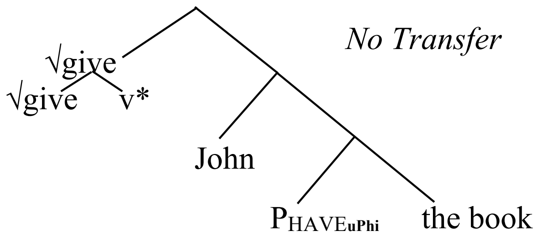
i.



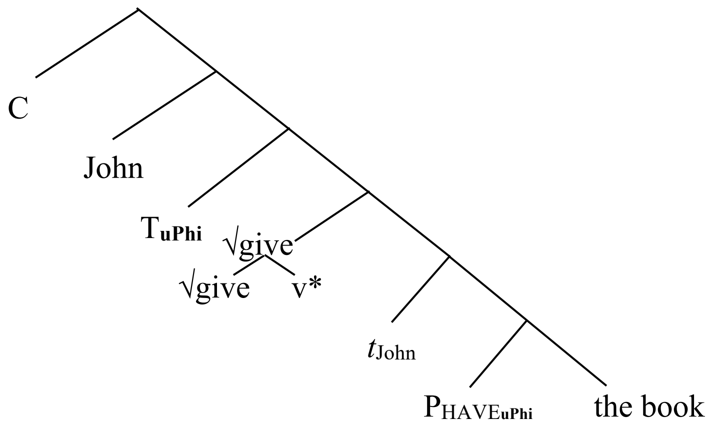
ii.



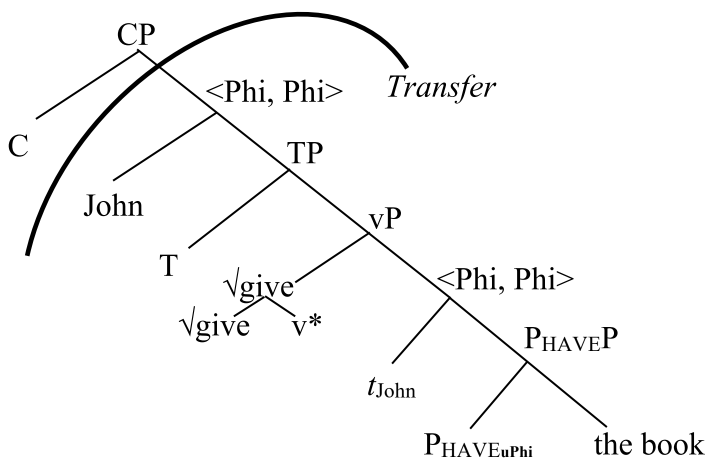
iii.



iv–vi.



iv–vi.



Given that Epstein, Kitahara and Seely (2016) note that pair-Merge externally forms <√root, v*> in passives (see section 2.2), the phase-hood of v* is canceled. Thus, the indirect object is extractable in such a case since Merge forms {√give, {√give, v*}} in (31(31ii), and v* becomes invisible prior to feature inheritance from v* to √give.[24]

Let us further consider the logically possible case of (27) by virtue of freely applying Merge. In this case, not the verbal root but the affix v* becomes invisible in the double object construction. This follows the standard analyses of the invisibility of affixes in ordered pairs (see Chomsky (2015a) and Epstein, Kitahara and Seely (2016)). The alternative structure and derivation for the v*P phase of *Mary gave John the book* is as follows:

(33) {Mary, {{√give, {√give, v*}}, {$_\delta$ John, {$_\gamma$ $\mathrm{P_{HAVE}}$, the book}}}}
- i. Merge externally forms {$_\delta$ John, {$_\gamma$ $\mathrm{P_{HAVE}}$, the book}}.
- ii. **Merge externally forms {v*, √give} and internally forms {√give, {√give, v*}}: v* becomes invisible**. (I.e., {√give {√give, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√give, v*>**.)
- iii. Merge externally forms {Mary, {{√give, {√give, v*}}, {$_\delta$ John, {$_\gamma$ $\mathrm{P_{HAVE}}$, the book}}}}.
- iv. Labeling and Agree take place by minimal search. **However, minimal search fails to label δ and fails to value uCase on *John***.

[24] An anonymous reviewer points out that the system in (31) cannot explain the grammaticality of *the book was given him* since uCase on *him* is not valued. Many speakers, however, accept such sentences when the indirect object is cliticized.

(i) The book was given'im.

(ii) ?*The book was given him.

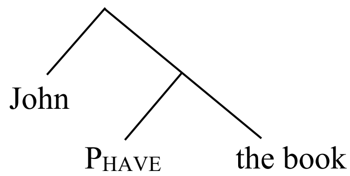
(iii) ?*The book was given John.

I do not explore how this cliticization works in the current framework, but the fact in (i) should be treated accordingly. Rather, the facts in (ii–iii) support the analysis in (31).

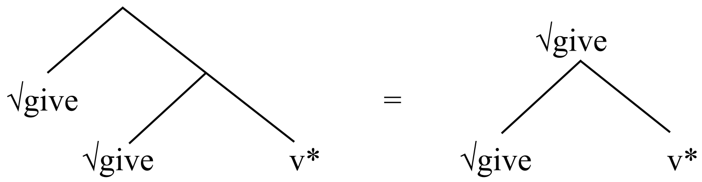
The graph-theoretic notation of this alternative structure and derivation is illustrated below:

(34) {Mary, {{√give, {√give, v*}}, {$_{\delta}$ John, {$_{\gamma}$ P$_{\text{HAVE}}$, the book}}}}

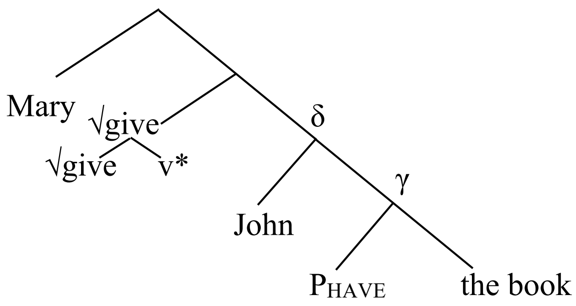
i.



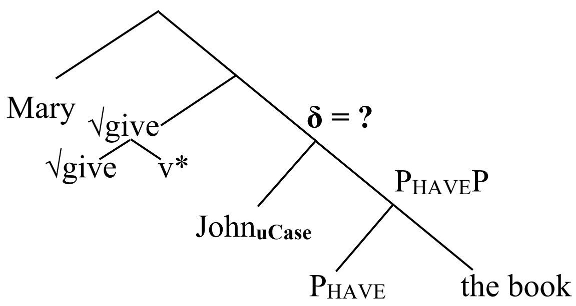
ii.



iii.



iv.



In (33ii), Merge applies twice and forms the derivational ordered pair {√give, {√give, v*}} (i.e. <√give, v*>). Then, v* becomes invisible. Thus, the feature inheritance from v* to P$_{\text{HAVE}}$ does not occur (cf. (27iii)); this will cause the failure of labeling and Agree in (33iv). *John* and P$_{\text{HAVE}}$ cannot share any feature since P$_{\text{HAVE}}$ does not have its own uPhi. The v* in the double object construction (in active voice), therefore, should be visible after Merge has formed the derivational ordered pair. In conclusion, the typical analysis where pair-Merge makes the affix v* invisible is no longer tenable in (27). Notice that I assume the null hypothesis where $\mathrm{P_{HAVE}}$ does not have its original uPhi, following the general property of prepositions. Generally, prepositions are not assumed to have uPhi. If prepositions had their original uPhi, they should be phase heads because only phase heads can have uPhi as noted by Chomsky (2008).[25]

2.5 Case 2: The Cognate Object Construction

The cognate object construction is defined as "involving verb-noun pairs which are either zero-related or which share a root morpheme and are not derived by means of affixation" (Macfarland (1995: 48)). Typical cognate object constructions are shown below.

(35) a. Chris danced a traditional dance.
b. Rose laughed a bitter laugh.
c. Bill smiled a happy smile.

There are at least two types of cognate objects in the cognate object construction (see Oba (2013)). One is referential, and the other is non-referential.

(36) a. Chris danced a traditional dance, and Rose danced it, too.
b. Chris danced a traditional dance, and Rose danced one, too.

(37) a. *Chris danced a staggering dance, and Rose danced it, too.
b. Chris danced a staggering dance, and Rose danced one, too.

25 Some prepositions (e.g. certain instances of French *à*) might have their own uPhi in terms of Probe-Goal Agree if we follow Kayne (2004). However, the original Probe-Goal Agree by Chomsky (2000, 2001) does not entirely hold in the present framework developed by Chomsky (2013, 2015a). See note 4 for Agree/Valuation in the present framework. See also Chomsky (2007) and Richards (2007) for topics related to phase and uPhi.

((36)–(37): Oba (2013: 73))

Both *it* and *one* can make up for *a traditional dance*, but not *a staggering dance*. The facts show that we cannot interpret the cognate object *a staggering dance* as the referential cognate object. Thus, *a staggering dance* is the non-referential cognate object. Furthermore, the non-referential cognate object is interpreted as the MANNER-adverbial. The examples below are from Horita (1996: 239):

(38) What (sort of dance) did the girls dance?
  a. They danced a traditional dance. (referential)
  b. *They danced a staggering dance. (non-referential)

(39) How did the girls dance?
  a. *They danced a traditional dance. (referential)
  b. They danced a staggering dance. (non-referential)

In (39b), *a staggering dance* (non-referential) cannot be the answer of *what*-question, but can be that of *how*-question because *staggering* is interpreted as the MANNER-adverbial.

Oba (2013: 72–73) reaches the conclusion that the (non-)referential cognate object is not the adjunct but the internal argument due to several empirical facts such as *though* movement, VP preposing and *do so* substitution (see also Macfarland (1995)).[26] *Though* movement can prepose the verb and argument, but not the whole v*P. Informally, *though* movement is V′ movement but not VP movement in the X′-theoretic terms.

(40) *though* movement:
  a. I read that book$_{\text{ARGUMENT}}$.
  b. Read that book though I did, (I didn't understand it).

(41) *though* movement:
  a. I read that day$_{\text{ADJUNCT}}$.

[26] For the argument-hood on the cognate object, see also Hale and Keyser (2002).

d. *Read that day though I did, (I didn't understand anything).

(Macfarland (1995: 103))

Moreover, VP presposing shows the similar behavior. Informally, it can prepose the V′-level object but not the whole VP.

(42) VP preposing:

a. I wanted Chris to read that book$_{\text{ARGUMENT}}$ on vacation, and read that book she did on vacation.

b. *I wanted Chris to read that day$_{\text{ADJUNCT}}$ on vacation, and read that day she did on vacation.

(Macfarland (1995: 103))

Oba (2013: 72) observes that both *though* movement and VP preposing can apply to the cognate object construction, regardless of whether the object is referential or non-referential.

(43) *though* movement:

a. Dance a traditional dance though Chris did, no one praised her ability. (referential)

b. Dance a staggering dance though Chris did, no one applauded her afterward. (non-referential)

(44) VP preposing:

a. ?I wanted Chris to dance a traditional dance at the ball, and dance a traditional dance she did at the ball. (referential)

b. ?I wanted Chris to dance a traditional dance at the ball, and dance a traditional dance she did at the ball. (non-referential)

((43)–(44): Oba (2013: 72))

These facts show that both referential objects and non-referential cognate objects are not adjuncts but arguments. According to Oba (2013), the argument-hood is also supported by *do so* substitution. As is well known, *do so* substitutes for the verb and argument, which is informally V′-level.

(45) *do so* substitution:
  a. I saw Chris that day, and John did so, too.
  b. *I saw Chris that day, and John did so Susan, too.
(Oba (2013: 72))

In (45b), we cannot leave the object *Susan*. Oba (2013) observes the same pattern in the both types of cognate object constructions.

(46) *do so* substitution:
  a. Chris danced a traditional dance, and Mary did so, too. (referential)
  b. *Chris danced a traditional dance, and Mary did so a popular dance, too. (non-referential)
  c. Chris danced a staggering dance, and Mary did so, too. (referential)
  d. *Chris danced a staggering dance, and Mary did so a nervous dance, too. (non-referential)
(Oba (2013: 73))

In conclusion, cognate objects are internal arguments, regardless of whether they are referential or non-referential.

Oba's (2013) observations and (38b) imply that we cannot extract the non-referential cognate object by wh-movement though it is the internal argument. In contrast, his observations and (38a) show that we can extract the referential object because it is the internal argument. The following facts support this analysis:

(47) a. What traditional dance did they dance? (referential)

b. *? How staggering a dance did they dance? (non-referential)

The (non-)extractability is also borne out by the following facts on topicalization:

(48) a. A traditional song, they sang. (referential)

b. *A silly smile, John smiled. (non-referential)

The facts in (48) also show that the referential object is extractable, but the non-referential object is not extractable.

The proposal in (16) accounts for the (non-)extractability of the cognate object. I assume that v* becomes invisible in the referential cognate object construction while a verbal root becomes invisible in the non-referential cognate object construction. First, let us consider the following structure and derivation for the v*P phase of *They danced a traditional dance* (referential):[27]

(49) {they, {{√dance, {√dance, v*}}, {$_{\delta}$ a traditional dance, {$_{\gamma}$ $t_{\surd\text{dance}}$, $t_{\text{a traditional dance}}$}}}}

i. Merge externally forms {$_{\gamma}$ √dance, a traditional dance}.

ii. Merge internally forms {$_{\delta}$ a traditional dance, {$_{\gamma}$ √dance, $t_{\text{a traditional dance}}$}}.

iii. Merge externally forms {they, {v*, {$_{\delta}$ a traditional dance, {$_{\gamma}$ √dance, $t_{\text{a traditional dance}}$}}}}: √dance inherits features from v*.

iv. **Merge internally forms {√dance, v*}, and then {√dance, {√dance, v*}}: <u>v* becomes invisible</u>**. (I.e., {√dance {√dance, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√dance, v*>**.)

[27] The lower copy √dance (i.e. $t_{\surd\text{dance}}$) is visible because the higher copies become part of the derivational ordered pair {√dance, {√dance, v*}} (i.e. <√dance, v*>). The same is true of (7), (51) and other derivations. See Chomsky (2015a) and Epstein, Kitahara and Seely (2016) for this special property of the lower copy created by internal pair-Merge. See also note 6.

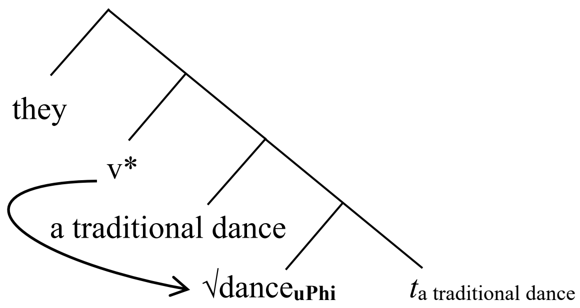
v. Labeling and Agree take place by minimal search: δ and γ are labeled <Phi, Phi> and √P, respectively.

vi. **The complement of $\boldsymbol{t}_{\surd\textbf{dance}}$ gets transferred** because the phase-hood is activated on $t_{\surd\text{dance}}$.

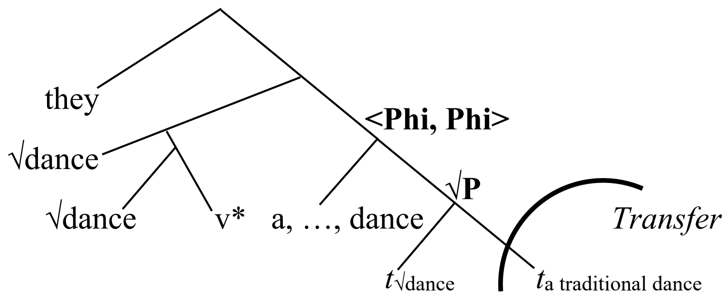
The graph-theoretic notation of this structure and derivation is illustrated below:

(50) {they, {{√dance, {√dance, v*}}, $\{_{\delta}$ a traditional dance, $\{_{\gamma}$ $t_{\surd\text{dance}}$, $t_{\text{a traditional dance}}$}}}}

i–iii.



v–vi.



This structure and derivation is essentially same as that of the simple transitive construction in (20). An anonymous reviewer mentions why the internal PM by SM follows external Merge of the EA *they*. It is logically possible that it might apply before the formation of {*they*, {v*, {a staggering dance, {√dance, $t_{\text{a staggering dance}}$}}}} as long as Merge applies freely. If feature inheritance takes place after the completion of a phase (see (7)), the internal PM by SM, nevertheless, applies after the Merge of EA by virtue of condition (23).

Then, let us consider the following structure and derivation for the v*P phase of *They danced a staggering dance* (non-referential):

(51) {they, {{v*, {v*, √dance}}, {$_\delta$ a staggering dance, {$_\gamma$ $t_{\sqrt{\text{dance}}}$, $t_{\text{a staggering dance}}$}}}}

i. Merge externally forms {$_\gamma$ √dance, a staggering dance}.

ii. Merge internally forms {$_\delta$ a staggering dance, {$_\gamma$ √dance, $t_{\text{a staggering dance}}$}}.

iii. Merge externally forms {they, {v*, {$_\delta$ a staggering dance, {$_\gamma$ √dance, $t_{\text{a staggering dance}}$}}}}: √dance inherits features from v*.

iv. **Merge internally forms {v*, √dance}, and then {v*, {v*, √dance}}:** **<u>√dance becomes invisible</u>**. (I.e., {v* {v*, √dance}} formed by PM by SM is **representationally interpreted as the ordered pair <v*, √dance>**.)

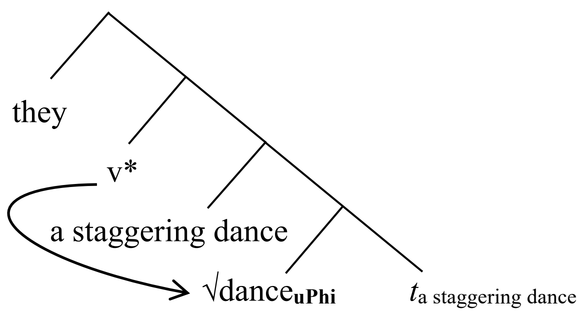
v. Labeling and Agree take place by minimal search: δ and γ are labeled <Phi, Phi> and √P, respectively.

vi. **<Phi, Phi> (i.e. δ) gets transferred**.

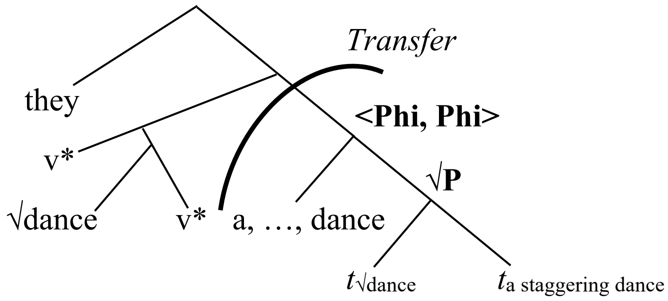
The graph-theoretic notation of this structure and derivation is illustrated below:

(52) {they, {{v*, {v*, √dance}}, {$_\delta$ a staggering dance, {$_\gamma$ $t_{\sqrt{\text{dance}}}$, $t_{\text{a staggering dance}}$}}}}

i–iii.



v–vi.



In (51iv), √dance becomes invisible as opposed to several transitive/unergative cases in (18), (20) and (49). This point is similar to the case of the double object construction in (27) except that (pair-)Merge *internally* applies to the root √dance in (51iv). In (51iv), the phase-hood of v* is active, and the complement of {v*, {v*, √dance}} (i.e. <Phi, Phi>) gets transferred.

It is also crucial that a verbal root (e.g. √dance) specifies the manner of the verbal event (i.e. v*) as an adjunct (cf. (25) and note 20). The non-referential cognate object construction is clearly related to the manner of the event as shown in (39); (39b) can be the answer to the *how*-question. The verbal root in (51iv), therefore, needs to be an adjunct of v* so that the non-referential cognate object construction conveys the correct interpretation of MANNER at the CI interface. However, MANNER does not relate to the verbal event of the referential cognate object construction as shown in (38a) and (39a). The verbal root in (49iv), therefore, should not be the adjunct of v* in terms of MANNER adjunction. In sum, the CI interface requires the reverse ordered pair <v*, √root> in the non-referential cognate object construction, but it requires the canonical ordered pair <√root, v*> in the referential cognate object construction although both ordered pairs are logically possible in syntax in each construction as long as Merge applies freely.

The analyses in (49)–(51) explain why the referential cognate object is extractable, but the non-referential cognate object is not.

(53) a. What traditional dance did they dance $t_{\text{what traditional dance}}$ $t_{\sqrt{\text{dance}}}$ $t_{\text{what traditional dance}}$? (referential)

b. *___ did they dance how staggering a dance $t_{\sqrt{\text{dance}}}$ $t_{\text{how staggering a dance}}$?

(non-referential)

c. *? How staggering a dance did they $t_{\text{how staggering a dance}}$ dance $t_{\text{how staggering a dance}}$ $t_{\sqrt{\text{dance}}}$ $t_{\text{how staggering a dance}}$? (non-referential)

Because the derivation of (53a) is the same as the standard transitive sentence, *what traditional dance* is extractable. In contrast, the non-referential cognate object is not extractable since the derivations of (53b, c) are essentially based on (51). In (53b), *how staggering a dance* is frozen at SPEC-√dance under PIC because it has been transferred immediately after labeling and Agree apply at the v*P phase. This sentence is interpreted as a yes-no question but not as a wh-question due to the matrix QP label ([$_{\text{QP}}$ did they dance, …]) (see Chomsky (2015a: 8)). Hence, this is gibberish and crashes at the CI interface. In (53c), *how staggering a dance* internally merges to the edge of v*P for escaping from Transfer. By doing so, Agree (i.e. minimal search) values neither uCase of *how staggering a dance* nor uPhi of √dance which has inherited it from v*. These unvalued features cause the violation of Full Interpretation at the CI interface. It is, therefore, impossible to extract non-referential cognate objects.

The proposal in (16) and the analyses in (49)–(51) account for the same effects in the case of topicalization shown in (48).

(54) a. A traditional song, they sang $t_{\text{a traditional song}}$ $t_{\sqrt{\text{sing}}}$ $t_{\text{a traditional song}}$. (referential)

b. *___, John smiled a silly smile $t_{\sqrt{\text{smile}}}$ $t_{\text{a silly smile}}$. (non-referential)

c. *A silly smile, John $t_{\text{a silly smile}}$ smiled $t_{\text{a silly smile}}$ $t_{\sqrt{\text{smile}}}$ $t_{\text{a silly smile}}$. (non-referential)

As discussed above in the case of wh-movement, we violate the PIC by extracting *a silly smile* in (54b). Even if it remains at the edge of v*P, the sentence is not going to be a topicalized interpretation. In (54c), we violate Full Interpretation due to transferred unvalued features.

PM by SM, recall, can representationally form the reverse ordered pair <v*, √dance>. In contrast, the standard approach to pair-Merge of heads cannot form <v*, √dance> because it involves the invisibility of affixes or a separate plane. Namely, the standard pair-Merge cannot explain the non-extractability of the non-referential cognate object as far as we assume simple transitive structures like (49) and (51) in the cognate object construction. This is because objects are generally extractable in terms of the ECP (Empty Category Principle) which is reanalyzed by labeling, freely applying Merge and internal pair-Merge of heads in the recent framework of minimalist syntax (see Chomsky (2015a)). In other words, PM by SM, freely forming {√root, {√root, v*}}, is thus supported by the exceptions to the ECP (i.e. the indirect object in the double object construction and the non-referential cognate object).

2.6 Case 3: The Small Clause Construction and ECM

In sections 2.4 and 2.5, we primarily discussed two cases. One is the case of PM by SM externally forming a derivational ordered pair (i.e. the double object construction). The other is the case of PM by SM internally forming a derivational ordered pair (i.e. the cognate object construction). In this section, I will show the last case of PM by SM which externally and internally forms a derivational ordered pair in one construction, the small clause construction. Furthermore, I will argue that the analyses of the small clause construction can be expanded to the ECM construction.

In Standard English, there are small clauses taken by certain verbs, such as *consider* and *think*, which are exemplified below:

(55) a. Mary considers [$_{\gamma}$ John foolish].

b. Mary thinks [$_{\gamma}$ John foolish].

c. John is considered [$_{\delta}$ t_{John} foolish].

d. John is thought [$_{\delta}$ t_{John} foolish].

Rizzi (2015a) assumes that the label γ in the small clause in (55a, b) is different from the label δ in (55c, d). His assumption is supported by the following facts in some varieties of English (the examples are taken from Rizzi (2015a: 41)):

(56) Some Varieties of English:

a. *I think [$_{\gamma}$ John intelligent].

b. John is thought [$_{\delta}$ t_{John} intelligent].

c. A man who I think [$_{\delta}$ t_{who} intelligent]

That is, *think* does not select the label γ, but it selects the label δ in (56). In contrast, *consider* (or *think* in Standard English) selects both γ and δ as in (55). Importantly, γ allows the subject of the small clause to remain in situ, but δ does not. Rizzi (2015a) argues that this distribution is explained by his proposal of maximality under Chomsky's (2013) labeling algorithm.

(57) Maximality: Phrasal movement can only involve maximal objects with a given label.

(Rizzi (2015a: 22))

Accordingly, *John* in (56a) cannot move out of γ since γ is maximal, but the label of *John* is not maximal. That is, γ is the maximal Subj_{def} (defective subject head), but the label of John is the non-maximal Subj_{def}.

(58) a. *I think [$_{\gamma}$ John [Subj_{def} [$_{\text{AP}}$ intelligent]]].

b. John is thought [$_{\delta}$ t_{John} [$_{\text{AP}}$ intelligent]].

c. A man who I think [$_{\delta}$ t_{who} [$_{\text{AP}}$ intelligent]]

In his analysis, the subject of small clause γ must remain in the small clause since Subj_{def} occurs in the small clause γ. The subject *John* at SPEC-Subj_{def}P and the Subj_{def} share a criterial feature (i.e. +Subj), and then it causes the subject to freeze in the place (i.e.

Criterial Freezing) in terms of (57). In other words, *John* cannot move because the label of *John* is not maximal, but γ (i.e. $\mathrm{Subj}_{\mathrm{def}}\mathrm{P}$) is maximal. On the other hand, the label δ is assumed to be AP (or a functional head about Phi). The subject of δ must move from SPEC-AP for labeling δ since they do not share any criterial feature. In sum, the subject in γ must remain in γ, but the subject in δ must move out of δ.

Rizzi (2015a) adopts Chomsky's (2013) labeling algorithm shown in (3). Rizzi, however, does not adopt the timing of labeling explicitly shown by Chomsky (2015a). In Rizzi's analysis, labels are determined immediately after Merge applies in order to obtain the unmovable property of maximal objects (57). Thus, labeling is not part of Transfer in his system, contrary to Chomsky (2013, 2015a).

It is problematic to postulate the features not inherent to lexical items such as discourse-related criterial features (e.g. Subj, Topic, Focus) in terms of the Inclusiveness Condition (see Chomsky, Gallego and Ott (to appear)). In addition, the maximality principle in (57) is unnecessary if the PIC can replace it. PM by SM (16) and selectional properties of verbs explain the interesting facts in (55)–(56) without postulating Rizzi's (2015) $\mathrm{Subj}_{\mathrm{def}}$ and his maximality principle. The non-extractability of the small clause subject is ultimately explained by the PIC and Full Interpretation under PM by SM. First of all, I borrow Larson's (1988: 349) structure of the small clause construction for the type γ.[28]

(59) Mary [√consider-v* [$_{\gamma}$ John [$t_{\surd\mathrm{consider}}$ intelligent]]].

Adopting this structure and PM by SM, we will consider the following structure and derivation for the v*P phase of *Mary considers John intelligent*:

(60) {Mary, {{v*, {v*, √consider}}, {$_{\gamma}$ John, {$_{\iota}$ $t_{\surd\mathrm{consider}}$, intelligent}}}}

[28] In Larson's original notation, v* and √consider are represented as a double V. This Larsonian VP-shell structure corresponds to v* and a verbal root in the current minimalist framework.

i. Merge externally forms {Mary, {v*, {$_\gamma$ John, {$_\iota$ √consider, intelligent}}}}: √consider inherits features from v*.

ii. **Merge internally forms {v*, √consider}, and then {v*, {v*, √consider}}: √consider becomes invisible**. (I.e., {v* {v*, √consider}} formed by PM by SM is **representationally interpreted as the ordered pair <v*, √consider>**.)

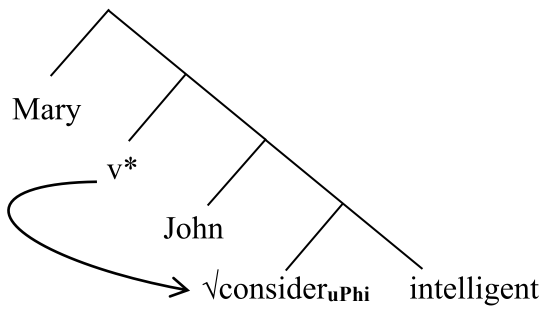
iii. Labeling and Agree take place by minimal search: **γ and ι are labeled <Phi, Phi> and √P, respectively**.

iv. **<Phi, Phi> (i.e. γ) gets transferred**.

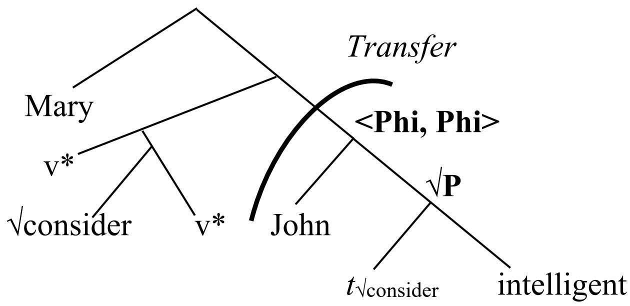
The graph-theoretic notation of this structure and derivation is illustrated below:

(61) {Mary, {{v*, {v*, √consider}}, {$_\gamma$ John, {$_\iota$ $t_{\surd\text{consider}}$, intelligent}}}}

i–ii.



iii–vi.



In (60ii), Merge *internally* forms the set {v*, √consider}, and then forms the derivational ordered pair {v*, {v*, √consider}} by PM by SM. This pattern is the same as (51iv) in the non-referential cognate object construction. That is, *John* must remain in its original position. If not, it causes a violation of the PIC or Full Interpretation. As for the violation of the PIC, *John* is forbidden to internally merge after Transfer in (60iv). Concerning the violation of Full Interpretation, *John* must remain in SPEC-√consider for the valuation of uCase or uPhi inherited to √root. Therefore, the subject of the small clause must remain in its original place for the type γ = <Phi, Phi>.

Suppose that (60) is on the right track, and let us consider the examples in (58) again. In some varieties of English, I assume *think* has the same structure and derivation as (60). Thus, *think* cannot select the label <Phi, Phi>.[29]

(62) a. *I think $[_{\gamma = \langle \mathrm{Phi, Phi} \rangle}$ John $[_{\iota = \sqrt{}\mathrm{P}}$ $t_{\sqrt{}\mathrm{think}}$, intelligent]].

b. John is thought $[_{\delta}$ t_{John} intelligent].

c. A man who I think $[_{\delta}$ t_{who} intelligent]

Instead, *think* can select the label δ which is assumed to be AP by Rizzi (2015a). Let us consider this type of small clause below.

As Chomsky (2013: 43–44) asserts, some small clauses have the symmetric structure [XP, YP] where XP or YP will raise in terms of labeling. He shows that two cases (a copular structure and a split-topic construction) have this symmetric structure. I argue that the small clause type δ also has the symmetric structure [XP, YP]. Let us consider the structure and derivation for CP-v*P phase of (62b): *John is thought intelligent* under this assumption and PM by SM.

(63) {C, {John, {T, {{√think, {√think, v*}}, $\{_{\delta}$ t_{John}, intelligent}}}}}

i. Merge externally forms {John, intelligent}.

ii. **Merge externally forms {√think, v*}, and then internally forms {√think, {√think, v*}}: v* becomes invisible** (phase cancellation). (I.e., {√think, {√think, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√think, v*>**.)

[29] Selectional properties are (if any) evaluated when Transfer applies or at the CI interface in the current framework since a syntactic set is not labeled when it is formed by Merge.

iii. Merge externally forms {T, {{√think, {√think, v*}}, {$_{\delta}$ John, intelligent}}}.

iv. Merge internally forms {John, {T, {{√think, {√think, v*}}, {$_{\delta}$ t_{John}, intelligent}}}}.

v. Merge externally forms {C, {John, {T, {{√think, {√think, v*}}, {$_{\delta}$ t_{John}, intelligent}}}}}: T inherits features from C.

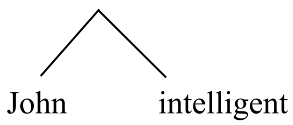
vi. Labeling and Agree take place by minimal search: **δ is labeled AP**.

vii. The complement of C gets transferred.

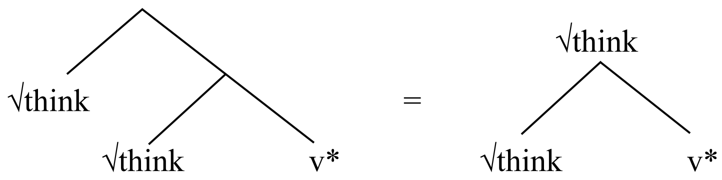
The graph-theoretic notation of this structure and derivation is illustrated below:

(64) {C, {John, {T, {{√think, {√think, v*}}, {$_{\delta}$ t_{John}, intelligent}}}}}

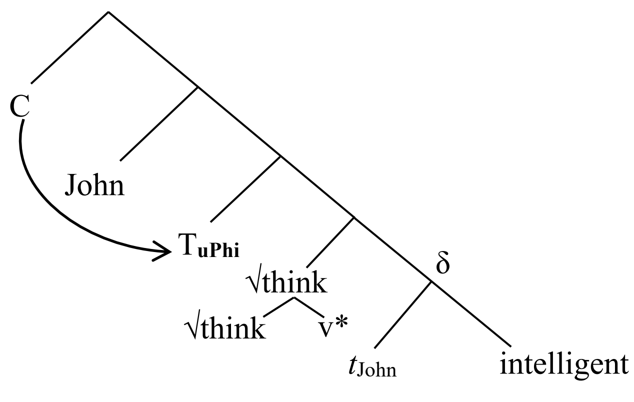
i.



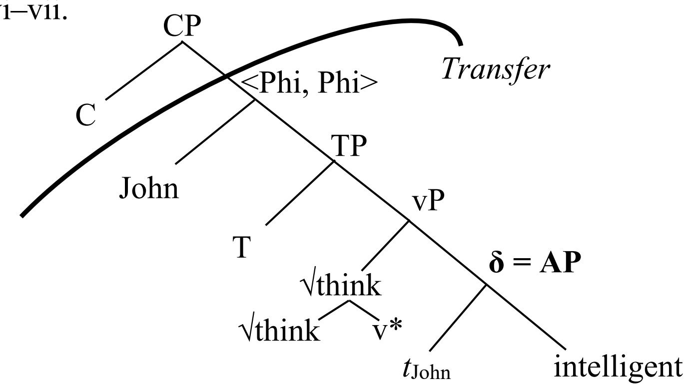
ii.



iii–v.



vi–vii.



Following Epstein, Kitahara and Seely (2016), v* becomes invisible in passives when PM by SM forms the derivational pair in (63ii). This reminds us of the structure and derivation regarding the passive version of the double object construction in (31). As with (31), the v*P phase is canceled, and the derivation continues. Thus, the label δ, which is not labeled at the v*P phase, gets labeled as AP at the CP phase thanks to the internal merger of *John* in (63iv). Furthermore, we can assume that the construction in (62c) also has the structure and derivation analogous to that of (63).

In conclusion, the structures and derivations in (60) and (63) entail that *think* only selects AP in some varieties of English. Put differently, the reverse ordered pair <v*, think> is not allowed because the CI interface requires that *think* in these dialects selects only AP.

(65) a. *I think [$_{\gamma = \langle \text{Phi, Phi} \rangle}$ John [$_{\iota = \sqrt{}\text{P}}$ $t_{\sqrt{}\text{think}}$ intelligent]].

b. John is thought [$_{\delta = \text{AP}}$ t_{John} intelligent].

c. A man who I think [$_{\delta = \text{AP}}$ t_{who} intelligent]

In Standard English, however, *think* selects both <Phi, Phi> and AP; the subject of the small clause remains in situ in <Phi, Phi> but not in AP.

(66) a. Mary thinks [$_{\langle \text{Phi, Phi} \rangle}$ John [$_{\sqrt{}\text{P}}$ $t_{\sqrt{}\text{think}}$ foolish]].

b. John is thought [$_{\text{AP}}$ t_{John} foolish].

c. A man who I think [$_{\text{AP}}$ t_{who} intelligent]

The verb *consider* also selects both <Phi, Phi> and AP.

(67) a. Mary considers $[_{\langle \mathrm{Phi,\ Phi}\rangle}$ John $[_{\surd \mathrm{P}}$ $t_{\surd \mathrm{consider}}$ foolish]].

b. John is considered $[_{\mathrm{AP}}$ t_{John} foolish].

As Rizzi (2015a) notes, this selectional property is also observed in the ECM verbs originally shown by Postal (1974). The following facts are from Rizzi (2015a: 42):[30]

(68) a. Mary believed/*alleged $[_{\langle \mathrm{Phi,\ Phi}\rangle}$ John $[_{\surd \mathrm{P}}$ $t_{\surd \mathrm{believe}/\surd \mathrm{allege}}$ $[_{\mathrm{TP}}$ t_{John} to be an idiot]]].

b. John was believed/alleged $[_{\mathrm{TP}}$ t_{John} to be an idiot].

c. Who did John believe/allege $[_{\mathrm{TP}}$ t_{who} to be an idiot]?

The facts suggest that *allege* allows the ordered pair <√allege, v*> but not the reverse order <v*, √allege> because the CI interface requires that *allege*-type verb selects TP but not <Phi, Phi>. Thus, PM by SM and the analyses about small clauses entail the following structure and derivation for the matrix CP-v*P phase of *John was alleged to be an idiot* in (68b):[31]

(69) {C, {John, {T, {{√allege, {√allege, v*}}, $\{_{\delta}$ t_{John}, {to, …}}}}}}

i. **Merge externally forms {√allege, v*}, and then internally forms {√allege, {√allege, v*}}** (phase cancellation). (I.e., {√allege, {√allege, v*}} formed by PM by SM is **representationally interpreted as the ordered pair <√allege, v*>**.)

---

[30] Rizzi (2015a) does not propose the structure shown in (68). This paper adopts Chomsky's (2013: 47) structure for ECM in (68a).

[31] The infinitival *to* alone is too weak to serve as a label if it is an instantiation of the single head T. It is plausible that we assume that *to* is the ordered pair <T, C> which is the other *weak* phase. See Mizuguchi (2016) and Sugimoto (2016) for relevant discussions.

ii. Merge externally forms {T, {{√allege, {√allege, v*}}, {$_{\delta}$ John, {to, …}}}}.

iii. Merge internally/externally forms {C, {John, {T, {{√allege, {√allege, v*}}, {$_{\delta}$ t_{John}, {to, …}}}}}}: T inherits features from C.

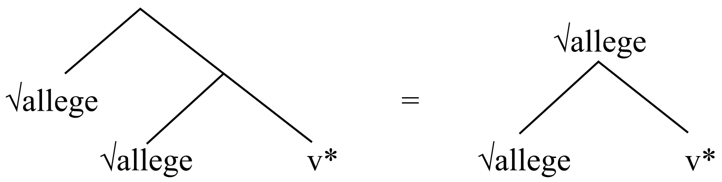
iv. Labeling and Agree take place by minimal search: **δ is labeled TP**.

v. The complement of C gets transferred.

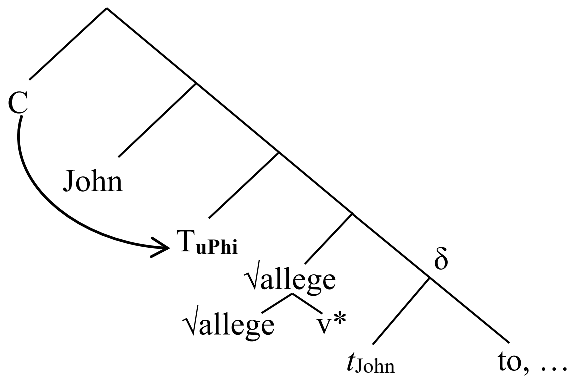
The graph-theoretic notation of this structure and derivation is shown below:

(70) {C, {John, {T, {{√allege, {√allege, v*}}, {$_{\delta}$ t_{John}, {to, …}}}}}}

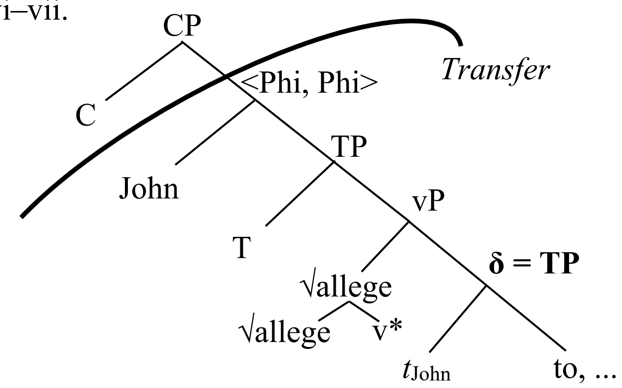
i.



ii–iii.



vi–vii.



The δ can be TP in (69iv) since the CI interface allows *allege* to select it.[32] As for the unacceptable case of *allege* in (68a), PM by SM and the analyses about small clauses above entail the following structure and derivation for the matrix v*P phase of **Mary alleged John to be an idiot*:

(71) {Mary, {{v*, {√allege, v*}}, $\{_{\gamma}$ John, $\{_{\iota}$ $t_{\sqrt{\text{allege}}}$, $\{t_{\text{John}}$, {to, …}}}}}}

i. Merge forms {Mary, {v*, $\{_{\gamma}$ John, $\{_{\iota}$ √allege, $\{t_{\text{John}}$, {to, …}}}}}}: √allege inherits features from v*.

ii. **Merge internally forms {√allege, v*}, and then internally forms {v*, {√allege, v*}}**. (I.e., {v* {v*, √allege}} formed by PM by SM is **representationally interpreted as the ordered pair <v*, √allege >**.)

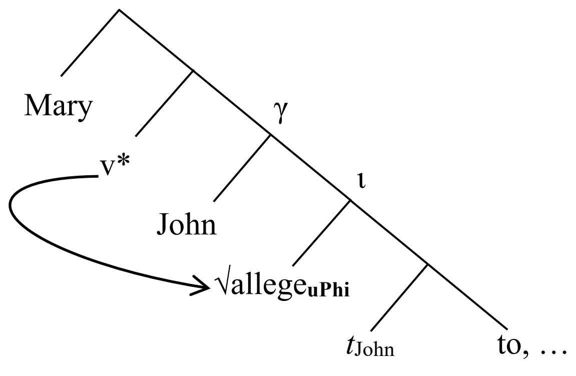
iii. Labeling and Agree take place by minimal search: **γ and ι are labeled <Phi, Phi> and √P, respectively**.

iv. <Phi, Phi> (i.e. γ) gets transferred.

The graph-theoretic notation of this structure and derivation is shown below:

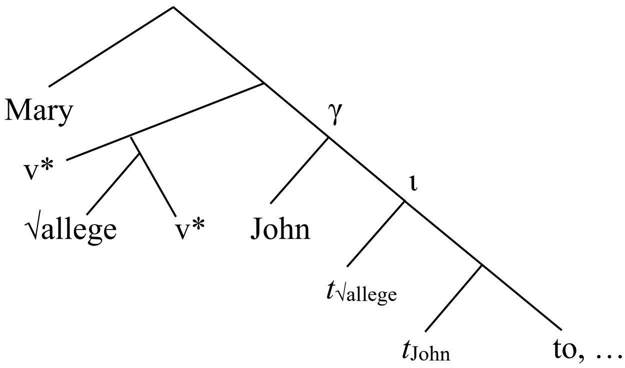
(72) {Mary, {{v*, {√allege, v*}}, $\{_{\gamma}$ John, $\{_{\iota}$ $t_{\sqrt{\text{allege}}}$, $\{t_{\text{John}}$, {to, …}}}}}}

i.

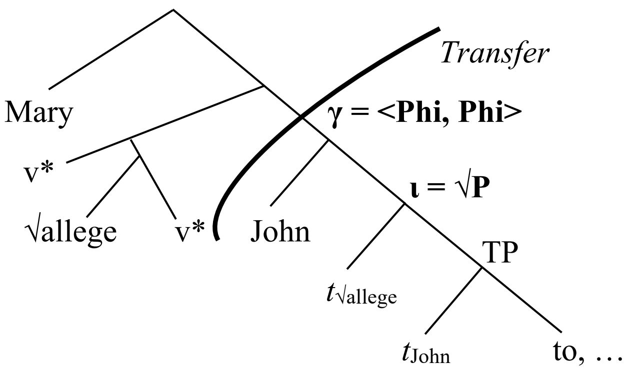


[32] In (69iv), the notation TP is used as the cover term for the label of <T, C>. Also, note that <T, C> is in effect {T, {T, C}} formed by the external application of PM by SM (cf. note 31).

ii.



iii–iv.



The γ cannot be <Phi, Phi> since the CI interface does not allow *allege* to select it. In addition, the standard pair-Merge, recall, cannot yield the reverse ordered pair <v*, √allege > because it only forms <√allege, v*>.

In principle, both orders for an ordered pair are available as far as Merge applies freely, but *allege*-type verbs and *think* in a dialect only have the default order <√root, v*>. That is, PM by SM forms the reverse order <v*, √root> only when it is necessary in terms of the requirement(s) by the interface(s). Accordingly, one might say the process of externalization for the SM interface affects this phenomenon because of exceptional cases of *allege*-type verbs. The following examples are borrowed from Bošković (1997: 58–59):[33]

(73) a. He alleged there to be stolen documents in the drawer.

[33] Bošković (1997) owes the grammatical judgment to Howard Lasnik, Noam Chomsky, Postal (1993) and Ura (1993).

b. *He alleged stolen documents to be in the drawer.
c. He acknowledged it to be impossible to square circles.
d. John wagered there to have been a stranger in that haunted house.
e. *John wagered a stranger to have been in that haunted house.
f. Mary alleged him to have kissed Jane.
g. *Mary alleged that man to have kissed Jane.
h. Mary never alleged him to be crazy.
i. *Mary never alleged the students to be crazy.

The notable elements *it*, *him* and *there* are unstressed in (73). Why these unstressed elements can remain in the position is unclear because the *allege*-type verbs seem to select <Phi, Phi> in (73a, c, d, f, h). Bošković (1997), however, argues that cliticization of these elements affects the acceptability in (73a, c, d, f, h) because the following sentences are not acceptable:

(74) a. *Mary alleged him and her to have kissed Jane.
b. *Mary never alleged him and her to be crazy.

(Bošković (1997: 59))

The empirical facts in (73)–(74) suggest that *it*, *him* or *there* incorporates with <√root, v*> at syntax, and this incorporation allows the *allege*-type verbs to select <Phi, Phi>. Individuals may analyze the incorporation or cliticization differently, but the process of externalization may be irrelevant to the phenomenon in (73) if this analysis is tenable. To pursue this possibility, we need to await the further research. Whatever the correct analysis is, interface conditions filter out the unacceptable structures in (73)–(74), conforming to the SMT.

In this section, I have shown that {v*, {v*, √root}} formed by PM by SM and its representational counterpart <v*, √root>, which the standard pair-Merge cannot yield, explain interesting facts concerning the small clause construction and the ECM construction under interface conditions. PM by SM is thus supported by the selectional properties in those constructions. The analyses by PM by SM have also shown that we do not need to postulate Subj head and the maximality principle. The former violates the Inclusiveness Condition, and the latter is just redundant because PIC makes up for it under PM by SM and general requirements of the CI system.

2.7 Summary

In this chapter, I have proposed that simplest Merge yields the derivational ordered pair {α, {α, β}} which is identical to the representational ordered pair <α, β> in terms of ZFC. Through forming this ordered pair, it is logically possible to create the canonical order <α, β> and the reverse order <β, α>. This possibility eventually contributes to the invisibility of their head. We do not have to look at β in <α, β> because <α, β> is equal to $f(\alpha) = \beta$. For $f(\alpha)$, α uniquely identifies β without looking at β. Thus, β becomes invisible in <α, β>, and α becomes invisible in <β, α>.

It also has been shown that this formalization of ordered pairs, PM by SM, explains various empirical facts: the double object construction, the cognate object construction, the small clause construction and the ECM construction. Particularly, it explains the non-extractability of an indirect object, a non-referential cognate object and a grammatical object (i.e. logical subject) in a certain small clause. Grammatical objects are typically extractable in terms of ECP, but these objects are not. Hence, the proposal in this chapter offers a principled explanation of these exceptional phenomena. For the small clause construction, I have argued that the proposal of ordered pairs also accounts for two types of small clauses observed by Rizzi (2015a) without depending on his analysis. It has been shown that there are at least two kinds of labels for small clauses: <Phi, Phi> and AP. The derivational ordered pair {v*, {v*, √root}} (i.e. <v*, √root>) is formed if the label is <Phi, Phi>. In contrast, the derivational ordered pair {√root, {√root, v*}} (i.e. <√root, v*>) is formed if the label is AP. Therefore, we have two kinds of structures involved in the small clause construction. This analysis further suggests the ECM construction also has the same two kinds of structures.

In conclusion, this chapter has shown that PM by SM conceptually and empirically surpasses the original definition of pair-Merge as a primitive operation; it reduces the syntactic operation to the simpler system, strictly conforming to the SMT. The "single computational system C_{HL} for human language" (Chomsky (1995b: 8) is, thus, Merge which ultimately explains the evolution of the language capacity within "[t]he limited evidence [available] from the evolutionary record" (Chomsky (2016a: 1)).

Chapter 3

Reformulating Pair-Merge of Phrases

3.1 Introduction

I reformulated pair-Merge of *heads* in the previous chapter but did not propose a reformulation of pair-Merge of *phrases*. In this chapter, I would like to show that the proposed PM by SM also applies to phrasal cases. Thus, the true reformulation of pair-Merge will finally be completed in this chapter.

3.1.1 Towards the True Reformulation of Pair-Merge

I have argued that PM by SM applies only to heads/atomic elements in syntax. *But, why*? As we discussed in the previous chapter, it is logically possible for PM by SM to apply to both X^0-level heads and XP-level phrases. The possibility, nevertheless, clearly raises an empirical problem. We consider the definition of PM by SM again:

(1) Pair-Merge of heads formulated by Simplest Merge (PM by SM):

a. Pair-Merge (α, β) = Merge (α, (α, β)) = Merge (α, {α, β}) = {α, {α, β}} = <α, β>

b. Pair-Merge (α, β) = Merge (β, (α, β)) = Merge (β, {α, β}) = {β, {α, β}} = <β, α>

If β and α could be anything, we would have the following structure and derivation of a simple sentence, *John hit Mary*:[1]

(2) {John, {{√hit, {√hit, v*}}, $\{_{\delta}$ Mary, $\{_{\gamma}$ $t_{\sqrt{\text{hit}}}$, t_{Mary}}}}}

i. Merge externally forms $\{_{\gamma}$ √hit, Mary}.

ii. Merge internally forms $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, t_{Mary}}}. (I.e., **$\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, t_{Mary}}} formed by PM by SM is representationally interpreted as the ordered pair <Mary, √hit>**.)

iii. The derivation continues, **but it will crash because of the derivational ordered pair $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, t_{Mary}}}**.

We discussed that step (ii) is the problem in chapter 2. Given the definition of PM by SM, $\{_{\delta}$ Mary, $\{_{\gamma}$ √hit, t_{Mary}}} is representationally interpreted as the ordered pair <Mary, √hit> though it should not be so. In this chapter, I obviate this unwelcome result without adding new mechanisms that make C_{HL} more complex in terms of methodological minimalism and the SMT.

The chapter is organized as follows. In section 3.2, we discuss problems of PM by SM proposed in the previous chapter. Additionally, we review the approach by Oseki (2015), who eliminates pair-Merge of phrases as a primitive operation, and I point out a problem in his approach. Section 3.3 is the main proposal, which extends PM by SM of heads to PM by SM of either two heads or two phrases. In section 3.4, it will be shown that the proposal explains the (anti-)adjunction condition effects. In Section 3.5, the

[1] Notice that *Mary* is assumed to be a set which is informally called a phrase. See section 3.6 for structures of nominals.

proposal further accounts for the Condition C effects. In section 3.6, we will discuss the internal structures of nominals under the proposal. In this section, the proposal also explains the Specificity Effects. Section 3.7 summarizes this chapter.

3.2 Problems

3.2.1 A Problem of PM by SM

As briefly mentioned above, PM by SM has an empirical problem if it applies to a phrase. Namely, the phrase $\{_{\delta}$ Mary, $\{_{\gamma}$ $t_{\surd\text{hit}}$, $t_{\text{Mary}}\}\}$ formed by PM by SM is representationally interpreted as the ordered pair <Mary, √hit> in (2ii). The δ should not be an ordered pair because both uCase on *Mary* and uPhi on √hit need to be valued to satisfy Full Interpretation. Even more problematic, the lower copy of √hit, which is $t_{\surd\text{hit}}$, must be visible, if we consider the basic derivational system in Chomsky (2015a). Accordingly, we have discussed that PM by SM applies only to heads, which are computational atoms in the narrow syntax.

(3) The Stipulation of PM by SM
PM by SM applies to computational atoms but not to other objects.

We need to eliminate this stipulation in the spirit of the Minimalist Program, methodological minimalism and, more crucially, the SMT. Given the basic idea of recursion, it is the null hypothesis that any object is accessible to further operations.[2] Accordingly, the stipulation in (3) is conceptually not welcome at all. Even empirically, this stipulation prevents us from explaining ubiquitous facts concerning XP-adjunction. In other words, PM by SM, by definition, cannot explain linguistic phenomena concerning the phrasal adjunction as long as it applies only to heads. It thus weakens the explanatory power of PM by SM. Technically speaking, either β or α in the definition of PM by SM cannot be an XP-level object.

[2] See also Chomsky (2017b) for an important discussion of recursion. Recursion is one of seven desiderata in Chomsky's (2017b).

(4) A Problem of PM by SM:

In the definition of PM by SM (Pair-Merge (α, β) = Merge (α, (α, β)) = Merge (α, {α, β}) = {α, {α, β}} = <α, β>), either β or α cannot be an XP-level object.

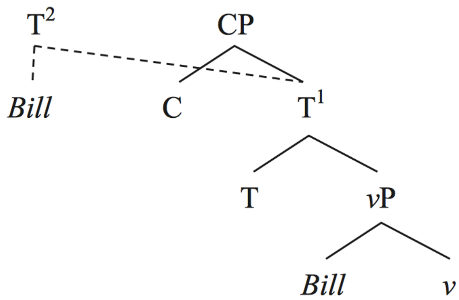
Tentatively putting this problem aside, we will review an alternative approach towards elimination of pair-Merge of phrases in the next section.

3.2.2 No Recourse to the Label Accessibility Condition

According to Oseki's (2015) proposal and analysis, it is possible to explain the phenomena concerning the phrasal adjunction without postulating an elementary operation, pair-Merge. Oseki (2015) eliminates Chomsky's (2004) original pair-Merge of phrases by proposing an alternative mechanism based on *the two-peaked structure*. We will first review the two-peaked structure and then consider Oseki's (2015) approach.

The two-peaked structure was originally presented by Epstein, Kitahara and Seely (2012, 2014). Their approach deduces the cyclically transferred domains of the phases (i.e. the complement of C/v*) without introducing the lexical array and the lexical sub-array. In other words, it is the counter-cyclic internal Merge that triggers Transfer in order to repair violations of the Extension Condition and the No-Tampering Condition.

(5) a. $[_{CP}$ [C $[_{TP}$ [T $[_{vP}$ Bill [v $[_{VP}$ ate rice]]]]]]] → $[_{CP}$ [C $[_{TP}$ Bill [T $[_{vP}$ Bill [v $[_{VP}$ ate rice]]]]]]]

b. 

(a: adapted from Epstein, Kitahara and Seely (2012: 255))

(b: Epstein, Kitahara and Seely (2012: 256))

In (5), Merge does not apply freely because the derivational system adopted above is essentially based on the earlier framework in Chomsky (2000, 2001, 2008) and among others. For example, the original Probe-Goal Agree applies to this system; Probe-Goal Agree triggers internal Merge/Move. After Agree (T, *Bill*) takes place, *Bill* moves to T^1 and forms the two-peaked structure as in (5b), following the LAC (Label Accessibility Condition).

(6) The Label Accessibility Condition (LAC):
Only the label of an entire syntactic object, the root, is accessible to the narrow syntax.
(adapted from Epstein, Kitahara and Seely (2012: 156))

Now that the two peaked-structure needs to be repaired, T^2 (i.e. the higher *Bill* and the complement of the *Bill* (T^1)) gets transferred. A similar analysis applies to v*P phases because the derivational steps of v*P are almost identical to those of CP phases under the analyses proposed by Epstein, Kitahara and Seely (2012) and Chomsky (2000, 2001, 2008). The two-peaked structure deduced by the counter-cyclic internal Merge, therefore, explains the timing and the nature of cyclic Transfer.

Oseki (2015) extends their analysis to adjunct structures and eliminates pair-Merge, which traces back to Chomsky (2000, 2004). Oseki (2015) adopts the two-peaked structure, which is a theorem deduced from the following three axioms:[3]

(7) Three Axioms of Phrase Structure:
a. *{XP, YP}*: The defining geometry of adjunction is Merge (XP, YP).
b. *Labeling Algorithm*: Outputs of Merge are labeled algorithmically.

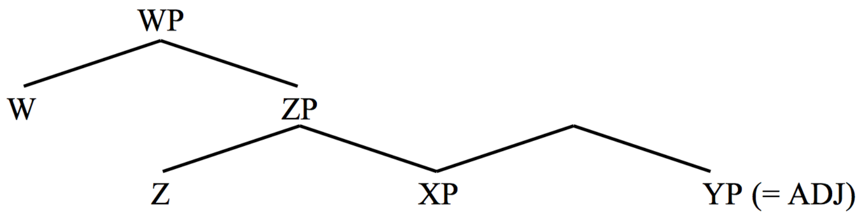
[3] For the first axiom, see May (1985), Chomsky (1986), Kayne (1994), Narita (2011), Ott (2011) and among others. As for the second, see Chomsky (2007, 2008, 2013, 2015a), Cecchetto and Donati (2015) and among others. For the final axiom, see Chomsky (2000), Collins (2002), Hornstein (2009) and among others.

c. *Label Accessibility Condition*: Unlabeled SOs are inaccessible to Merge.

(Oseki (2015: 307))

Note that the version of the LAC adopted by Oseki (2015) is not completely analogous to the one adopted by Epstein, Kitahara and Seely (2012). Note also that, in Chomsky (2008), labels are essential for derivations to proceed in the narrow syntax, contrary to Chomsky (2013, 2015a) and recent work on freely applying Merge. Following the axioms in (7), Merge yields the following adjunct structure:

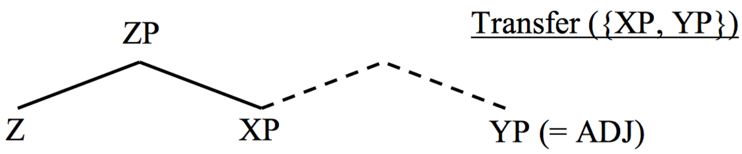
(8) The Two-Peaked Structure for Adjuncts:



(Oseki (2015: 307))

In Oseki's (2015) system, Merge (XP, YP) = {XP, YP} applies first. Second, Label (XP, YP) applies, but it fails since XP and YP are not in an agreement relation. Third, the subsequent external Merge of Z does not apply to {XP, YP} but applies to either XP or YP, following the LAC.[4] Fourth, if Merge of Z applies to XP, {Z, XP} is labeled as ZP, and {XP, YP} is transferred, following the proposal by Epstein, Kitahara and Seely (2012).

(9)



(Oseki (2015: 308))

[4] Oseki (2015) adopts Horstein's (2009) version of the LAC: only the label of a syntactic object is accessible to Merge.

Finally, W, following the LAC, externally merges with label ZP, with YP being inaccessible to further operations in the narrow syntax; hence, it yields (8).

Oseki's (2015) proposal is fairly appealing in that we do not need to assume a problematic operation, pair-Merge (α, β) = <α, β>. Furthermore, it explains the phenomena of the adjunct condition and others related to adjuncts.

(10) The Adjunct Condition:

a. *Who did Mary cry [$_{\text{ADJ}}$ after John hit t_{wh}]?

b. *Which paper did you read Don Quixote [$_{\text{ADJ}}$ before filing t_{wh}]?

c. *Who did an advocate speak of Betsy [$_{\text{ADJ}}$ before a discussion of t_{wh}]?

(a: Huang (1982: 503))

(b: Nunes and Uriagereka (2000: 21))

(c: Johnson (2003: 188))

In Oseki's (2015) analysis, ADJ, which denotes an adjunct phrase, has been transferred when wh-phrases internally merge. All wh-phrases in (10), therefore, cannot be extracted under the PIC.

Oseki (2015) also argues that his analysis may also explain the phenomena of the anti-adjunct condition. Adjunct phrases generally yield island effects, but the effects loosen in certain cases.

(11) The Anti-Adjunct Condition:

a. What did John arrive [$_{\text{ADJ}}$ whistling t_{wh}]?

b. What did John drive Mary crazy [$_{\text{ADJ}}$ trying to fix t_{wh}]?

c. Kinél szívott [$_{\text{ADJ}}$ nagyobbat $t_{\text{kinél}}$]?
Who-to smoke-Past large-Cpr-Acc
'He smoked more than who?'

d. Yamada-sensei-ga shinsatsu-shita-yori(-mo) Tanaka-sensei-ga
Dr. Yamada-Nom examination-did-than(-also) Dr.Tanaka-Nom

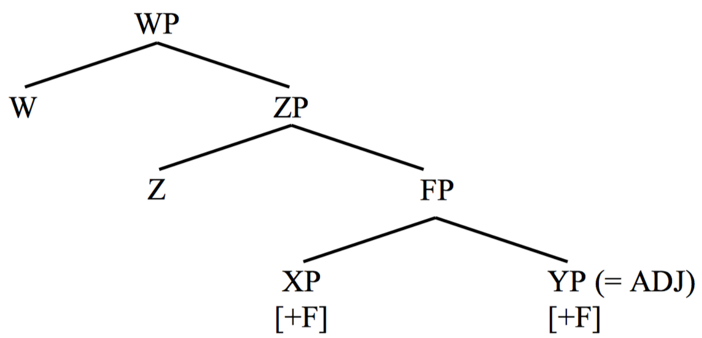
kanja-o [$_{\text{ADJ}}$ $t_{\text{Yamada, …, yori(-mo)}}$ oozei] shinsatsu-shita.
patient-Acc many examination-did
'Dr. Tanaka examined more patients than Dr. Yamada examined.'

(a: Borgonovo and Neeleman (2000: 200))
(b: Trueswell (2007: 1356))
(c: den Dikken (2012: 11))
(d: Miyamoto (2012: 344))

According to Oseki (2015), YP in (8) shares the prominent feature [+F] with XP in the cases of (11). Therefore, the following generalization holds:

(12) The adjunct-island effect loosens when an adjunct phrase is in a special-agreement relation.

Notice that the feature sharing in this sense is not the typical/normal agreement such as Phi-agreement of EA-T and IA-√root because such agreements must involve some unvalued Phi-feature which only phase heads arguably bear. Thus, the feature sharing of adjuncts is a *special* agreement. The abstract graph-theoretic notation of the cases of (11) is shown below:

(13)


(Oseki (2015: 309))

{XP, YP} is, therefore, labeled as FP by the special agreement relation. In this structure, YP is not an adjunct but just one of typical complements (see (7a)). Consequently, the phrases in ADJ can be extractable in (11).

Despite its explanatory averages, the analysis has a serious problem under the current minimalist framework. Let us reconsider three axioms that Oseki (2015) adopts:

(14) Three Axioms of Phrase Structure:

a. *{XP, YP}*: The defining geometry of adjunction is Merge (XP, YP).

b. *Labeling Algorithm*: Outputs of Merge are labeled algorithmically.

c. *Label Accessibility Condition*: Unlabeled SOs are inaccessible to Merge. (=(7))

The first axiom seems to be inconsistent with PM by SM, but I will argue that some special cases of adjunct phrases, in effect, follow this symmetric geometry. The second axiom undoubtedly holds in the current minimalist model though the timing of labeling is different from the assumption in Oseki (2015). He implicitly assumes that labels are determined per Merge, but minimal search identifies labels per phase under the current model employing freely applying Merge (see Chomsky (2013, 2015a)). That is, there is almost no labeled set/phrase when Merge applies because labeling, repeated in (15), takes place at the timing of Transfer:

(15) Labeling by Minimal Search:

a. $\{H, XP\} \rightarrow \{_{HP}\ H, XP\}$

b. $\{XP, t_{YP}\} \rightarrow \{_{XP}\ XP, t_{YP}\}$

c. $\{XP, YP\}$ / X sharing features (e.g. Phi-set) with Y $\rightarrow \{_{<Phi, Phi>}\ XP, YP\}$

Freely applying Merge is summarized below:

(16) Freely Applying Merge:

Merge applies freely in any order, regardless of whether internal or external.

What the current mechanism requires is, therefore, no recourse to (any version of) the LAC. Formally, syntactic structures do not even have *nodes* although they are illustrated in informal graph-theoretic notation. Simplest Merge merely yields unordered sets without labels and projections as shown below:

(17) Merge (α, β) = {α, β}

Merge is, thus, accessible to any set-theoretic object. In other words, Merge is inaccessible to labels. More crucially, if we ensure recursion, which is a universal property of human language, any generated object should be accessible to further operations (see Chomsky (2017b)). Limiting its accessibility is, then, adding the stipulation. Hence, something like the LAC puts the extra condition on recursion unless it is motivated on conceptual grounds.

(18) No Recourse to the LAC:
Recursion must be free.
(cf. (Chomsky (2017b))

The last axiom in (14), therefore, raises the serious problem with respect to recursion. Consequently, we cannot adopt Oseki's (2015) approach in the current minimalist framework. In the next section, I will present an alternative approach to eliminate pair-Merge of phrases as a primitive operation. Concretely, PM by SM replaces both pair-Merge of phrases and the two-peaked structure for adjunction with simplest Merge without postulating any stipulation and redundant mechanism.

3.3 A Proposal: Reformulating Pair-Merge of Phrases

I not only argued in chapter 2 that PM by SM can resolve the problems of pair-Merge of heads but also argued that PM by SM can explain a variety of empirical facts. The original definition of PM by SM is repeated below:

(19) Pair-Merge of heads formulated by Simplest Merge (PM by SM):

a. Pair-Merge (α, β) = Merge $(\alpha, (\alpha, \beta))$ = Merge $(\alpha, \{\alpha, \beta\}) = \{\alpha, \{\alpha, \beta\}\}$ $= <\alpha, \beta>$

b. Pair-Merge (α, β) = Merge $(\beta, (\alpha, \beta))$ = Merge $(\beta, \{\alpha, \beta\}) = \{\beta, \{\alpha, \beta\}\}$ $= <\beta, \alpha>$

(=(1))

As we have already seen, α and β in this definition are limited to computational atoms, but this limitation is just a stipulation.

(20) The Stipulation on PM by SM

PM by SM applies to computational atoms but not to other objects.

(=(3))

This stipulation must be reduced if we ensure recursion in the narrow syntax. Recursion, a universal property of language, must be free.

(21) A Problem of PM by SM:

In the definition of PM by SM (Pair-Merge (α, β) = Merge $(\alpha, (\alpha, \beta))$ = Merge $(\alpha, \{\alpha, \beta\}) = \{\alpha, \{\alpha, \beta\}\} = <\alpha, \beta>$), either β or α cannot be an XP-level object.

(=(4))

To eliminate the stipulation and the problem above, I argue that α and β in the formulation of PM by SM can be XP-level objects and X^0-level objects, but the two elements targeted by PM by SM must be of the same level.

(22) The Condition on PM by SM:

In the definition of PM by SM (Pair-Merge (α, β) = Merge $(\alpha, (\alpha, \beta))$ = Merge $(\alpha, \{\alpha, \beta\}) = \{\alpha, \{\alpha, \beta\}\} = <\alpha, \beta>$), both β and α must be same-level objects.

Given the definition of an ordered pair in ZFC (Zermelo-Fraenkel Set-Theory with the Axiom of Choices) and the empirical data, the condition is naturally deduced. In chapter 2, I adopted the refined definition of ordered pairs (see Tourlakis (2003: 182–183), among others).

(23) $<\alpha, \beta> = \{\alpha, \{\alpha, \beta\}\}$

In this definition, let us suppose that both α and β are not the singleton sets $\{\alpha\}$ and $\{\beta\}$ but members. In set theory, it is trivial that singleton sets are not equal to members that are not sets, but it does not mean that sets cannot be members. In other words, sets can be members even if their *levels* are different from atomic elements like α and β. Thus, the following definition of the *levelness* of members trivially holds in set theory:

(24) A and α are *different-level* members where A is a set, but α is not.

Note that (24) is not a new definition but only a naturally holding generalization in set theory. I assume that this levelness affects the formation of ordered pairs in syntactic theory because the following formulation emprically holds where α and β are *atomic(-level) members*, and A and B are *set(-level) members*:

(25) a. $<\alpha, B> \neq \{\alpha, \{\alpha, B\}\}$

b. $<A, \beta> \neq \{A, \{A, \beta\}\}$

Recall that in the structure of *John hit Mary* in (2), $\{_{\delta}$ Mary, $\{_{\gamma}\ t_{\sqrt{\text{hit}}}, t_{\text{Mary}}\}\}$ is not an ordered pair but a unordered set. In linguistic literature, the head-head adjunction and the phrase-phrase adjunction are generally observed, but the head-phrase adjuction is not.

That is, the empirical observations motivate (25), which says if we assume $A = \{\alpha\}$ and $B = \{\beta\}$, neither $<\alpha, \{\beta\}> = \{\alpha, \{\alpha, \{\beta\}\}\}$ nor $<\{\alpha\}, \beta> = \{\{\alpha\}, \{\{\alpha\}, \beta\}\}$ holds. Therefore, given the levelness of members in (24), definitions of ordered pairs can be restated as follows:

(26) a. $<\alpha, \beta> = \{\alpha, \{\alpha, \beta\}\}$

b. $<A, B> = \{A, \{A, B\}\}$

In this definition, the members in sets must be same-level objects in order to form ordered pairs. Using familiar symbols in syntactic theory, the definition can be expressed as follows:[5]

(27) a. $<X^0, Y^0> = \{X^0, \{X^0, Y^0\}\}$

b. $<XP, YP> = \{XP, \{XP, YP\}\}$

On the contrary, the following formulation holds since X^0 and Y^0 are not the same-level members as XP and YP (see (25)):

(28) a. $<X^0, YP> \neq \{X^0, \{X^0, YP\}\}$

b. $<XP, Y^0> \neq \{XP, \{XP, Y^0\}\}$

Accordingly, the formulation of PM by SM is finally completed as follows.[6]

(29) PM by SM (the final version):

<u>*head-head adjunction*</u>:

---

[5] X^0/Y^0 and XP/YP denote computational atoms (i.e. heads) and sets (i.e. phrases), respectively. Also, notice that X^0/Y^0 and XP/YP do not denote labels.

[6] I continue to call the final version of proposal *PM by SM*, but its full name, *Pair-Merge of heads formulated by Simplest Merge*, should be changed to *Pair-Merge formulated by Simplest Merge*.

a. Pair-Merge (X^0, Y^0) = Merge $(X^0, (X^0, Y^0))$ = Merge $(X^0, \{X^0, Y^0\})$ = $\{X^0, \{X^0, Y^0\}\} = \langle X^0, Y^0 \rangle$

b. Pair-Merge (Y^0, X^0) = Merge $(Y^0, (Y^0, X^0))$ = Merge $(Y^0, \{Y^0, X^0\})$ = $\{Y^0, \{Y^0, X^0\}\} = \langle Y^0, X^0 \rangle$

phrase-phrase adjunction:

c. Pair-Merge (XP, YP) = Merge (XP, (XP, YP)) = Merge (XP, {XP, YP}) = {XP, {XP, YP}} = <XP, YP>

d. Pair-Merge (YP, XP) = Merge (YP, (YP, XP)) = Merge (YP, {YP, XP}) = {YP, {YP, XP}} = <YP, XP>

Note that notation such as X^0 and XP is used only for expository purposes to indicate that X^0 and XP are a computational atom and a syntactic set, respectively. There are no real X^0 and XP objects as X´ theoretic terms.

The proposed formulation readily resolves the empirical problem shown in (2). The derivation (2) now proceeds successfully under the final version of PM by SM.

(30) {John, {{√hit, {√hit, v*}}, $\{_\delta$ Mary, $\{_\gamma$ $t_{\sqrt{hit}}$, t_{Mary}}}}}

i. Merge externally forms $\{_\gamma$ √hit, Mary}.

ii. Merge internally forms $\{_\delta$ Mary, $\{_\gamma$ √hit, t_{Mary}}}. (I.e., $\{_\delta$ Mary, $\{_\gamma$ √hit, t_{Mary}}} formed by PM by SM **is not** representationally interpreted as the ordered pair <Mary, √hit>.)

iii. Merge externally forms {John, {v*, $\{_\delta$ Mary, $\{_\gamma$ √hit, t_{Mary}}}}}: √hit inherits features from v*

iv. Merge internally forms {√hit, v*}, and then {√hit, {√hit, v*}}: v* becomes invisible. (I.e., {√hit, {√hit, v*}} formed by PM by SM **is** representationally interpreted as the ordered pair <√hit, v*>.)

v. Labeling and Agree take place by minimal search: δ and γ are labeled <Phi, Phi> and √P, respectively.

vi. The complement of $t_{\sqrt{hit}}$ gets transferred.

In this derivation, {$_δ$ Mary, {$_γ$ $t_{\sqrt{\text{hit}}}$, t_{Mary}}} is not a derivational ordered pair, but {√hit, {√hit, v*}} is. Because *Mary* and √hit are XP and X^0, respectively, {$_δ$ Mary, {$_γ$ $t_{\sqrt{\text{hit}}}$, t_{Mary}}} is not interpreted as <Mary, √hit> due to (28). In contrast, because both √hit and v* are X^0, {√hit, {√hit, v*}} is interpreted as <√hit, v*> due to (27). This correctly captures the basic derivation of *John hit Mary*. Accordingly, the proposed formulation resolves the empirical problem of the extension of PM by SM of *heads* to that of *phrases*. The reformulation of pair-Merge is thus finally completed in a true sense. In the following sections, it will be shown that PM by SM (the final version) explains several phenomena of phrasal adjuncts and nominals (hereafter, I use the term PM by SM in the sense of (29), unless otherwise noted).

We know that an entire adjunct phrase is extractable:

(31) How many times did you upload the file to the webpage t_{how}?

This linguistic phenomenon might suggest that YP (i.e. {how many times}) in {XP, {XP, YP}}, which is formed by PM by SM, is invisible and non-extractable under the basic quality of ordered pairs. However, it seems that what this phenomenon actually suggests is that *{XP, YP} (i.e. {{$_{v*P}$ ...}, {How many times}}) in {XP, {XP, YP}} is visible and extractable*. For example, the set-theoretic notation of the sentence in (31) can be illustrated as follows (irrelevant parts are omitted):[7]

(32) {{$_{\text{v*P}}$ …}, {How many times}} did you {{$_{\text{v*P}}$ upload the file to the webpage}, $t_{\{\{\text{v*P} \ldots\}, \{\text{how many times}\}\}}$}?

In {XP, {XP, YP}}, YP is invisible, but {XP, YP} is visible because {XP, YP} is merely a set that contains YP and a lower copy of XP. Therefore, the following descriptive generalization holds:

[7] {how many times} may externally merge to √P.

(33) In {XP, {XP, YP}} = <XP, YP>, YP is invisible but XP and {XP, YP} are visible.

I will use this generalization and PM by SM to explain the Condition C anti-reconstruction effects later. Yet, before explaining the effects, I will show that PM by SM can explain the adjunct island effects.

3.4 Explaining the (Anti-)Adjunct Condition

Oseki (2015) accounted for the phenomena of both the adjunct condition and the anti-adjunction condition under his proposal, the two-peaked structure for adjunct phrases. Its empirical coverage is significant because Merge, under the LAC, explains not only the non-extractable domain of adjunct phrases but also their extractable domain. His proposal, nevertheless, has a serious conceptual problem, which has already been mentioned. The LAC and the two-peaked structure are no longer tenable on purely conceptual grounds (see section 3.2.2). We thus need to reanalyze the phenomena in more ideal ways, one of which capitalizes on PM by SM.

3.4.1 Facts

Adjunct phrases are generally assumed to be islands since they are opaque to extraction (see Huang (1982), Chomsky (1986), among others).

(34) The Adjunct Condition:

a. *Who did Mary cry [$_{ADJ}$ after John hit t_{wh}]?

b. *Which paper did you read Don Quixote [$_{ADJ}$ before filing t_{wh}]?

c. *Who did an advocate speak of Betsy [$_{ADJ}$ before a discussion of t_{wh}]?

(=(10))

In all examples above, wh-phrases cannot be extractable since they externally merge within the adjunct islands or phrases. As shown earlier, it has been reported in linguistic literature that this condition is sometimes relaxed.

(35) The Anti-Adjunct Condition:

a. What did John arrive [$_{\text{ADJ}}$ whistling t_{wh}]?

b. What did John drive Mary crazy [$_{\text{ADJ}}$ trying to fix t_{wh}]?

c. Kinél szívott [$_{\text{ADJ}}$ nagyobbat $t_{\text{kinél}}$]?
Who-to smoke-Past large-Cpr-Acc
'He smoked more than who?'

d. Yamada-sensei-ga shinsatsu-shita-yori(-mo) Tanaka-sensei-ga
Dr. Yamada-Nom examination-did-than(-also) Dr.Tanaka-Nom
kanja-o [$_{\text{ADJ}}$ $t_{\text{Yamada, …, yori(-mo)}}$ oozei] shinsatsu-shita.
patient-Acc many examination-did
'Dr. Tanaka examined more patients than Dr. Yamada examined.'

(=(11))

Contrary to the cases of the adjunct condition in (34), we can extract all wh-phrases from the adjunct phrases in (35). In the following section, it will be shown that PM by SM explains the adjunct condition, and Merge with one axiom explains the anti-adjunction condition.

3.4.2 Analyzing the Adjunction Condition

Given the refined version of PM by SM in (29), the sentences in (34) have the following underlying structures (note that the notation t is not used as a lower copy in the derivational ordered pairs, and irrelevant parts (e.g. functional heads C, T and v*) are excluded):

(36) a. *Who did you cry after John hit t_{wh}?
{{you cry}, {{you cry}, {$_{\text{ADJ}}$ after John hit who}}}
=<{you cry}, {$_{\text{ADJ}}$ after John hit who}>

b. *Which paper did you read Don Quixote before filing t_{wh}?

{{you read Don Quixote}, {{you read Don Quixote}, {$_{ADJ}$ before filing which paper}}}

=<{you read Don Quixote},{ $_{ADJ}$ before filing which paper}>

c. *Who did an advocate speak of Betsy before a discussion of t_{wh}?

{{an advocate speak of Betsy}, {{an advocate speak of Betsy}, {$_{ADJ}$ before a discussion of who}}}

=<{an advocate speak of Betsy},{$_{ADJ}$ before a discussion of who}>

In the first example of (36), (i) Merge externally forms {{you cry}, {$_{ADJ}$ after John hit who}} at some stages of the derivation. The label of {you cry} is C_QP, which means that the closest head of the set is C bearing the Q feature. Given PM by SM, (ii) Merge internally forms {{you cry}, {{you cry}, {$_{ADJ}$ after John hit who}}}.[8] Under the definition of ordered pairs in ZFC, (iii) the resultant sets are representationally interpreted as <{you cry}, { $_{ADJ}$ after John hit who}>. As a result, (iv) the wh-phrase becomes invisible since it is one part of YP in the ordered pair <XP, YP>. Recall that XP is visible, but YP is, by definition, invisible in <XP, YP>. The wh-phrase *who* is, therefore, non-extractable in (36a). The identical derivational steps occur in the other examples, and explain the non-extractability of elements in adjunct phrases. PM by SM thus accounts for the adjunction condition without any arbitrary stipulation and redundant mechanisms.

3.4.3 Analyzing the Anti-Adjunction Condition

The anti-adjunction condition is also explained by the combination of generalization (12) and axiom (14a) repeated here as (37) and (38), respectively.

---

[8] It is logically possible for PM by SM to yield {{after John hit who}, {{you cry}, {after John hit who}}}, the *reverse* derivational ordered pair, which is representationally interpreted as <{after John hit who}, {you cry}>. After the resultant pair has been transferred, the CI system interprets {you cry} as an adjunct. I am not confident with regard to what type of interpretation this is, but I am confident that it is intuitively gibberish.

(37) The adjunct-island effect loosens when an adjunct phrase is in a special-agreement relation.

(38) The defining geometry of adjunction is Merge (XP, YP) = {XP, YP}.

Given PM by SM, the latter does not hold in the true sense because the typical-defining geometry of adjunction is not {XP, YP} but {XP, {XP, YP}}. Nevertheless, it holds if it is in a special-agreement relation. That is, the ADJ-phrases in (35) are no longer adjuncts but the normal sets due to the special relation. I will not elaborate on what the special-agreement relation is for the reasons listed in the following sentences. It apparently does not meet the usual criteria for Agree or feature valuation. One of the criteria involves (the set of) unvalued features, where feature valuation takes place as an instantiation of agreement. Only phase heads, recall, can have such features in terms of the phase theory.[9] We certainly need to comprehend the true nature of this kind of agreement, but it seems to me that we should refine the theory in advance for that purpose under the SMT. Putting many questions aside, I follow the descriptive generalization as follows:

(39) The so-called adjunct phrases may or may not be visible. If they are invisible, they are YP in {XP, {XP, YP}}. If they are visible, they are YP in {XP, YP}, where X and Y are in a special relation.

In this generalization, X and Y in {XP, YP} should be in a special relation for labeling purposes.

---

[9] Kawashima and Kitahara (2015), Epstein, Kitahara and Seely (2017) and Kitahara (2017) argue that wh-phrases bear unvalued Q features which agree with valued C_Q. C_Q and a wh-phrase at SPEC-C_Q, therefore, agree or share a Q feature. Lower copies or a lower copy of such wh-phrases is interpreted as a variable for the operator-variable interpretation at the CI interface. Hence, the remaining lower copies of the phrases, therefore, do not violate Full Interpretation. This suggests that all agreement phenomena are instantiations of the SPEC-head relation, at least in English.

Given the considerations above, the sentences in (35) have the following underlying and resulting structures:[10]

(40) a. What did John arrive whistling t_{wh}?
[John arrive {$_{\text{ADJ}}$ whistling what}]
→ [what [John arrive {$_{\text{ADJ}}$ whistling t_{wh}}]]
b. What did John drive Mary crazy trying to fix t_{wh}?
[John drive Mary crazy {$_{\text{ADJ}}$ trying to fix what}]
→ [what [John drive Mary crazy {$_{\text{ADJ}}$ trying to fix t_{wh}}]]
c. Kinél szívott nagyobbat $t_{\text{kinél}}$?
[szívott {$_{\text{ADJ}}$ nagyobbat Kinél}]
→ [Kinél [szívott {$_{\text{ADJ}}$ nagyobbat $t_{\text{Kinél}}$}]]
d. Yamada-sensei-ga shinsatsu-shita-yori(-mo) Tanaka-sensei-ga kanja-o $t_{\text{Yamada, …, yori(-mo)}}$ oozei shinsatsu-shita.
[Tanaka-sensei-ga kanja-o {$_{\text{ADJ}}$ Yamada-sensei-ga shinsatsu-shita-yori(-mo) oozei} shinsatsu-shita]
→ [{Yamada-sensei-ga shinsatsu-shita-yori(-mo)} [Tanaka-sensei-ga kanja-o {$_{\text{ADJ}}$ $t_{\text{Yamada, …, yori(-mo)}}$ oozei} shinsatsu-shita]]

In the derivations above, all ADJ-phrases are part of unordered sets. Hence, those phrases are visible, and the lower copies of internally merged elements may be in a special relation in terms of (39). Importantly, Merge does not yield derivational ordered pairs or representational ordered pairs as in (36). Therefore, the anti-adjunct condition is explained by freely applying simplest Merge and the descriptive generalization of visible adjunct phrases in (39). PM by SM does not apply in these cases because its application is simply superfluous. In other words, PM by SM applies as long as it is necessary.

[10] I omit several irrelevant parts such as functional heads, curly brackets for wh-phrases and, if any, the intermediate copies of internally merged elements. Additionally, square brackets denote sets (or structures), regardless of whether they are ordered or unordered

3.5 Explaining Condition C (Anti-)Reconstruction

3.5.1 Condition C Anti-Reconstruction Effects

Condition C was formulated in the government and binding model as follows:

(41) Condition C
An R-expression cannot have an antecedent that c-commands it.

Oseki (2015) argues that his approach to pair-Merge can correctly explain Condition C effects as well. In what follows, this section shows that the analyses capitalizing on PM by SM also explain the Condition C effects without depending on the standard analysis of pair-Merge by Chomsky (2004) and the two-peaked structure for adjuncts by Oseki (2015).

It is crucial that the fronted adjuncts do not give rise to a violation of Condition C: Condition C Anti-Reconstruction (see Lebeaux (2000), among others). In other words, the fronted adjuncts generally do not observe the reconstruction effects as shown below:

(42) Condition C Anti-Reconstruction
a. *He_i likes the story [$_{\text{ADJ}}$ that John_i wrote].
b. Which story [$_{\text{ADJ}}$ that John_i wrote] did he_i like?
c. [$_{\text{ADJ}}$ In Ben_i's office], he_i is an absolute director.
d. [$_{\text{ADJ}}$ With John_i's novel finished], he_i began to write a book of poetry.
e. [$_{\text{ADJ}}$ To Ben_i's surprise], he_i noticed that the others had left.
f. [$_{\text{ADJ}}$ For Mary_i's valor], I heard she_i was given a medal.
(a, b: Lebeaux (2000: 103))
(c–f: Speas (1991: 250))

In (42a), Condition C is violated. The violation implies that the representational ordered pair <XP, YP> is interpreted as the derivational ordered pair {XP, {XP, YP}} at *SEM* (SEM is a semantic representation in the sematic component accessed by the CI system).

The interpretation reminds us of an optional operation *SIMPL* (simplification), which converts <α, β> to {α, β} (see Chomsky (2004: 118)). Namely, the antecedent *he*, in (42a), c-commands *John* in {the story, {the story, that John wrote}}, which had been mapped onto SEM by Transfer.

In the examples on wh-movement and topicalization above, the fronted adjuncts avoid reconstruction effects.[11] It follows that the antecedents do not c-command the lower copies of the R-expressions in the fronted adjuncts as long as we assume that there is no covert movement and that SEM capitalizes on the structural configurations formed in core syntax.[12] In the standard minimalist model, an adjunct-phrase yields the anti-reconstruction because the phrase is included in an ordered pair <XP, YP> where the adjunct is attached on a separate plane. If an optional operation SIMPL does not apply to the pair, it is still invisible even at SEM or the CI interface (see Chomsky (2004)). In contrast, if SIMPL applies, it becomes visible. As far as SIMPL applies optionally, there is apparently no reason why it *must* apply in certain cases (e.g. (42a)). However, it seems that SIMPL still applies freely *with only certain choices converging* (although Chomsky (2004) does not explicitly state so). In the case of (42a), SIMPL must apply though it causes the derivation to diverge because it is the only option that SIMPL can choose. If SIMPL does not apply in this case, the CI system simply cannot interpret [$_{\text{ADJ}}$ that John$_i$ wrote] since it is invisible or does not exist there. To interpret the phrase,

---

[11] In (42a), the wh-phrases containing the adjunct phrase moves, but in (42c–f), adjunct phrases are fronted. The latter seems to be debatable because it is obscure how adjunct phrases can be moved (or internally merged) once it becomes invisible after the application of pair-Merge. That is, adjunct phrases in (42c–f) is fronted, although they are already invisible in the base positions due to the application of pair-Merge. This problem arises in Oseki's (2015) analysis and my analysis in the next section too. One might assume that those adjunct phrases merge with a null head K, and this K phrase is actually fronted in (42c–f). If it is on the right track, (42c) could be as follows: [$_{\text{KP}}$ K [$_{\text{ADJ}}$ In Ben$_i$'s office]], he$_i$ is an absolute director. However, I rather propose an alternative analysis (see (31)–(33) and (44)).

[12] See Chomsky (1995b, 2007), Epstein, Kitahara and Seely (2013, 2015), among others for the non-existence of covert movement. In addition, see chapter 4 for or evidence revealing no support of covert movement in the current model of the Minimalist Program.

SIMPL must apply even when the resulting interpretation is complete gibberish. Recall that derivations converge even if they get gibberish interpretations at the interfaces. In contrast, when the ADJ-phrases are fronted as shown in (42), SIMPL does not apply to the lower copies of those phrases. If SIMPL applies to the fronted ADJ-phrases, the operation does not need to apply further to the lower copies. It is redundant to convert lower ordered pairs of ADJ-phrases into unordered sets if ADJ-phrases are interpreted at the highest position by the CI system.[13] When adjuncts are fronted, the reconstruction effects are, therefore, not observed. However, SIMPL does not seem to exist in the current model because Chomsky (2004) proposed that it applied when *Spell-Out* took place, although Chomsky (2016c) argues that *Spell-Out* is no longer tenable in the current model (see also Obata (2010)). To eliminate SIMPL and explain Condition C anti-reconstruction Effects, I capitalize on PM by SM and axiom (38).

3.5.2 Analyzing Condition C Anti-Reconstruction Effects

Given the consideration of SIMPL in the earlier section, PM by SM can explain Condition C anti-reconstruction effects without postulating SIMPL. In this section, it will be shown that PM by SM can account for the Condition C anti-reconstruction effects in a theoretically desirable manner.

First, I explain the case in (42a). Recall that PM by SM essentially forms unordered sets. These unordered sets are called derivational ordered pairs, which are representationally interpreted as ordered pairs (i.e. representational ordered pairs). There is no ordered pair $<\alpha, \beta>$ as a pure-syntactic object. Hence, it is unnecessary to convert $<\alpha, \beta>$ into $\{\alpha, \beta\}$. Instead, if the CI system treats {XP, {XP, YP}} as unordered sets, the system interprets it accordingly. In contrast, if the CI system judges {XP, {XP, YP}} as a representational ordered pair, the system representationally interprets the

[13] Phase-level memory is enough for the information that syntactic objects have internally merged. That is, the phase-level memory suffices to determine whether objects are copies formed by internal Merge or repetitions formed by external Merge when Transfer applies (see Chomsky (2007, 2012b)). Accordingly, the fronted ADJ-phrases can be interpreted as one of copies even if there is no application of SIMPL to the lower copies.

unordered sets {XP, {XP, YP}} to be the ordered pair <XP, YP>. For example, in (42a), [the story [$_{\text{ADJ}}$ that John$_i$ wrote]] is virtually {{the story}, {{the story}, {that John wrote}}} = <{the story}, {that John wrote}> formed by PM by SM. The set {that John wrote} is invisible in the narrow syntax because it is representationally interpreted as <{the story}, {that John wrote}> which is, by definition, *f*(the story) = <{the story}, {that John wrote}> (see relevant discussions in chapter 2). At SEM, <{the story}, {that John wrote}> can be interpreted as either {{the story}, {{the story}, {that John wrote}}} or <{the story}, {that John wrote}>. It is the null hypothesis that both options are freely available unless stipulated. Given the structure of *He$_i$ likes the story that John$_i$ wrote*, the former option {{the story}, {{the story}, {that John wrote}}} is chosen. That is because if the latter <{the story}, {that John wrote}> is chosen, {that John wrote} is not interpreted by the CI system. However, when <{the story}, {that John wrote}> is interpreted as {{the story}, {{the story}, {that John wrote}}} at SEM, it gives rise to a violation of Condition C since *he* c-commands *John*. The only non-gibberish interpretation is not to co-index *he* and *John*. Consequently, there is no need for SIMPL to make computation complex in terms of the <α, β>-to-{α, β}conversion in which an element (i.e. α or β) at a separate plane comes back to a primary plane.

Second, I explain remaining cases in (42). (42b) is a simple case: {the story, {the story, {that John$_i$ wrote}}} internally merges. After this internal merger, the CI system interprets the higher copy of {the story, {the story, {that John$_i$ wrote}}} as a derivational ordered pair but not its lower copy. The CI system interprets the lower copy as a representational ordered pair <{the story}, {that John wrote}>. Therefore, we do not observe the reconstruction effect. The cases in (42c–f) are a little complicated. I argue that {XP, YP} in {XP, {XP, YP}} internally merges in (42c–f). Recall that generalization (33), repeated here as (43), holds because an entire adjunct can be extracted.

(43) In {XP, {XP, YP}} = <XP, YP>, YP is invisible but XP and {XP, YP} are visible.

Following this generalization, say in (42c), $\{_{\text{v*P}}$ …, {in Ben$_i$'s office}} in $\{_{\text{v*P}}$ …, $\{_{\text{v*P}}$ …, {in Ben$_i$'s office}}} internally merges. The CI system does not need to see the lower copy of $\{_{\text{v*P}}$ …, {in Ben$_i$'s office}} because the same object exists in the higher position. Consequently, in this case as well, we do not observe the reconstruction effect. The set theoretic notation of the sentences in (42) is shown below:[14]

(44) Condition C Anti-Reconstruction

a. *He$_i$ likes {the story, {the story, {that John$_i$ wrote}}}.

b. {Which story, {which story, {that John$_i$ wrote}}} did he$_i$ like $\underline{t_{\text{\{which story, \{which story, \{that Johni wrote\}\}\}}}}$?

c. $\{_{\text{v*P}}$ …, {In Ben$_i$'s office}, he$_i$ is $\underline{\{\{_{\text{v*P}}\text{ an absolute director}\},\ t_{\{\text{v*P …, \{in Beni's office\}\}}}\}}$.

The CI interprets the underline parts as not {XP, {XP, YP}} but <XP, YP>. YP in <XP, YP> is invisible at SEM with respect to Condition C. The other cases in (42d–f) can be explained in the same way as (42c) (but see note 11).

3.5.3 Condition C Reconstruction Effects

As Speas (1991) points out, Condition C anti-reconstruction does not hold in some cases. In such cases, the reconstruction is obligatory as shown below:

(45) Condition C Reconstruction

a. *He$_i$ believes the claim that John$_i$ is nice.

b. *Whose claim that John$_i$ is nice did he$_i$ believe?

c. *[$_{\text{ADJ}}$ In Ben$_i$'s office], he$_i$ lay on his desk.

d. *[$_{\text{ADJ}}$ With John$_i$'s computer], he$_i$ began to write a book.

e. *[$_{\text{ADJ}}$ To Ben$_i$'s office], he$_i$ takes the bus.

---

[14] In a copular sentence, such as (42c), v* could be v. Or, neither v* nor v could occur. However, in chapter 4, I will argue that v* can occur at least in some copular sentences, such as *there* constructions.

f. *[ADJ For Mary$_i$'s brother], I heard she$_i$ was given some clothes.

(a, b: Lebeaux (2000: 103))

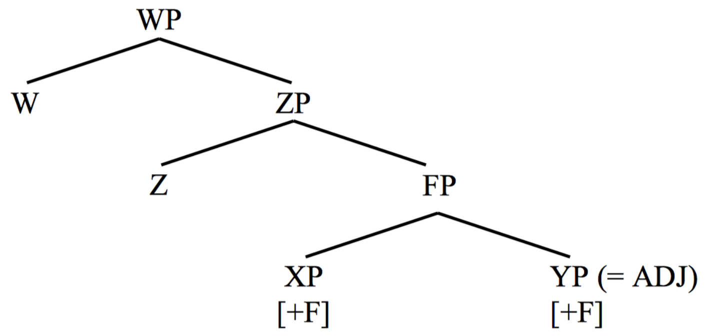
(c–f: Speas (1991: 250))

These peripheral phenomena are explained by the approach of the two-peaked structure for adjuncts proposed by Oseki (2015). That is because Oseki (2015), following den Dikken (2012) and Miyamoto (2012), assumes that agreement (i.e. feature sharing) occurs in these phenomena. If the assumption is on the right track, the ADJ-phrases in (45) are no longer adjunct phrases but just set-Merged phrases (i.e. unordered sets) in a special relation. The peripheral phenomena are, thus, not empirically problematic. Adopting Oseki's (2015) analysis, I will show in the next section that his analysis is compatible with PM by SM.

3.5.4 Analyzing Condition C Reconstruction Effects

As shown in section 3.2.2, Oseki (2015) argues that adjuncts can be in a special relation in some circumstances. One of those is the anti-adjunction condition, and another is the Condition C reconstruction. The notation in (13), repeated here as (46), indicates that ADJ is in a special relation and shares a prominent feature [+F] with XP.

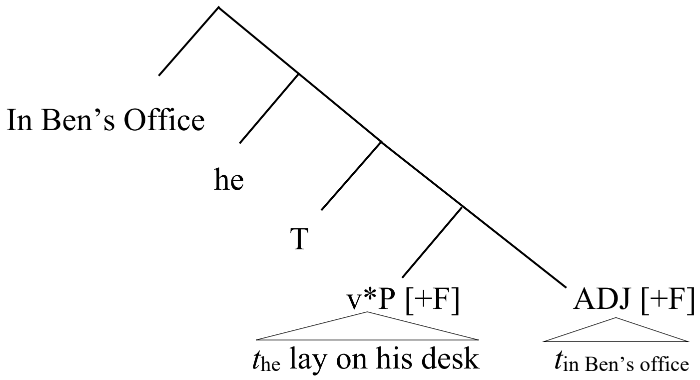
(46)



(=(13))

If this analysis is tenable, we do not need to present a new alternative. Let us consider (45c) under (46). In (45c), YP is [ADJ in Ben$_i$'s office], and XP is probably v*P [he lay on his desk] in the graph-theoretic notation above.[15] The detailed notation of (45c) is illustrated below:

(47)



If we follow Oseki (2015), ADJ and v*P in (45c) somehow share a prominent feature due to a special relation. The reconstruction affects the lower copy of ADJ because the ADJ sharing [+F] is in the sister relation with its merged mate and *must be visible and interpreted at SEM* (see Oseki (2015)). Accordingly, the deviancy of (45c) and the other examples is explained. This analysis is compatible with PM by SM since the analysis assumes no operation that creates ordered pairs and two-peaked structures. However, a question arises: how does a special relation allow feature sharing to occur? Further research is needed to resolve the problem.

3.6 Nominals

3.6.1 Essential Properties of Phase

While phases are generally assumed to be v*P and CP, smaller units such as DP, NP and PP also have been said to be phases in linguistic literature (e.g. Bošković (2014), Marantz (2007), Oba (1999, 2003), Uriagereka (1999)). It is technically possible to postulate these units as phases. Nevertheless, postulation of extra phases is conceptually undesirable. If these units are phases, the increasing number of phase heads creates

---

15 The external argument *he* internally merges to SPEC-T after the introduction of T. In addition, XP could be √P.

complexity and makes the notion of *phase* meaningless. Phases are, in their original definition, v*P and CP where we observe the phenomena concerning successive cyclic movement. On the contrary, the phenomena seem to be absent in smaller units. In addition, it is generally assumed that v*P and CP yield the *duality of interpretation*. The former, in essence, corresponds with D-structure, and the latter corresponds with S-structure in earlier theories. Hale and Keyser (1993, 2002), among others, generalize the v*P phase as argument structures formed by external Merge. In contrast, Chomsky (2000, 2001, 2004, 2007), among others, argues that internal Merge creates the chain regarding scope/discourse properties that are interpreted by the CI system. The verbal-v*P and clausal-CP phases are thus propositional. Other smaller units are therefore not phases in that they do not satisfy these properties of the two phases. In sum, phases should, at least, satisfy the following two properties to avoid rendering the notion of *phase* meaningless.

(48) a. Phases yield successive cyclic movement.
b. Phases are propositional.

Smaller units (e.g. DP, NP, PP) do not seem to satisfy the essential properties stated above.[16] Furthermore, a phase head is assumed to bear unvalued Phi-features, which trigger Transfer (see Richards (2007), Chomsky (2007, 2016b), among others).[17] Smaller units do not seem to have this property either.

---

[16] Yukio Oba (p.c.) points out that DP might yield successive cyclic movement. For example, if we assume that the following structure in (i) exists, D yields successive cyclic movement.

(i) Do [$_{DP}$ [$_{NP/DP}$ whatever] D T you want to do $t_{whatever}$].
(ii) Do [$_{DP}$ [$_{NP/DP}$ whatever] D $t_{whatever}$ C you want to do $t_{whatever}$].

However, it seems to me that C is not optional but obligatory in this kind of a structure, as illustrated in (ii), and this C yields successive cyclic movement under the PIC.

[17] As discussed in chapter 1, uPhi is arguably essential for labeling to take place successfully in the current model.

However, if smaller units are not phases, how do we explain facts about smaller units such as the *Specificity Effect* discussed in Oba (1999, 2003)? Section 3.6 focuses on nominals and attempts to explain extraction from NP without postulating extra phases but with PM by SM.

3.6.2 Nominal = Phase?

If nominals are not specific, the extraction from them is possible. In contrast, if nominals are specific, the extraction from them is impossible. This is called the Specificity Effect in linguistic literature.

(49) a. Who did John see a picture of t_{wh}?
b. *Who did John see Mary's picture of t_{wh}?

To explain a fact such as (49) under a minimalist model, Oba (1999, 2003) assumes that D becomes a phase head, so DP becomes a phase if D bears a [+specific] feature. Adapting Jakendoff (1977), Bowers (1987) and Karimi (1999), Oba (1999, 2003) proposes the following structures of nominals:

(50) a. $[_{DP}$ X $[_{D'}$ D (+specific) NP]]
b. X: Bill's/the/a/this/that/these/those/each/all/no/any/every/most/few
(51) a. $[_{NP}$ Y $[_{N'}$ N …]]
b. Y: a/little/many/several/a few/some

(Oba (2003: 205))

Following the PIC, the complement of D (+specific) becomes non-extractable because D (+specific) is a phase head.[18] The shaded part below denotes that it has been inaccessible under the PIC (the intermediate copies of *who*, if any, are omitted):

[18] Note that X plays the important role of completing the *proposition* serving as one of the phasal properties. See Oba (1999, 2003).

(52) a. Who did he see a picture of t_{wh}?

b. *Who did he see Bill's picture of t_{wh}?

Given the proposals above, *a picture of who* and *Bill's picture of who* are NP and DP (+specific), respectively. In NP, the wh-phrase *who* is extractable since it has not been inaccessible before it moves out of the original position. In DP, the wh-phrase *who* is, in contrast, non-extractable since it has been inaccessible in [$_{DP}$ Bill's picture of who].

In this approach, it is crucial that D is a phase head. Chomsky (2007), in contrast, suggests that a phase head of nominals is not D but n*, assuming the parallelism of v*P and n*P. That is, a verbal root and D inherit features from and raise to v* and n*, respectively. "Therefore, the structure [of n*P] is a nominal phrase headed by n*, not a determiner phrase headed by D, which is what we intuitively always wanted to say" (Chomsky (2007: 26)). In sum, his analysis suggests the following derivation of *the book*:[19]

(53) A Derivation of *the book*:

i. Merge externally forms {D_{the}, √book}.

ii. Merge externally forms {n*, {D_{the}, √book}}.

iii. Feature inheritance from n* to D_{the} takes place.

iv. Merge internally forms {n*, {√book, {D_{the}, $t_{\surd book}$}}}.

v. Pair-Merge internally forms {<n*, D_{the}>, {√book, {D_{the} $t_{\surd book}$}}}.

Although Chomsky (2007) does not show analyses of the Subjacency Effect, this derivation can explain it. That is simply because the complement of <D_{the}, n*> gets transferred, and it becomes inaccessible under the PIC.[20]

[19] In Chomsky (2007), internal Merge does not apply freely but applies as a triggered operation under the Probe-Goal-Agree system.

[20] De-phasing from n* to a nominal root does not occur since this derivational system is based on Chomsky (2000, 2001, 2004, 2007). De-phasing was first proposed by Chomsky (2015a).

(54) A Derivation of *Bill's picture of who*:

i. Merge externally forms {√picture, (of) who}.

ii. Merge externally forms {n*, {D_s, {√picture, (of) who}}}.

iii. Feature inheritance from n* to D_s takes place.

iv. Merge internally forms {n*, {{√picture, (of) who}, {D_s, $t_{\{\surd\text{picture, (of) who}\}}$}}}.

v. Merge externally forms {Bill, {n*, {{√picture, (of) who}, {D_s, $t_{\{\surd\text{picture, (of) who}\}}$}}}}.

vi. Pair-Merge internally forms {Bill, {<n*, D_s>, {{√picture, (of) who}, {$t_{D_{the}}$, $t_{\{\surd\text{picture, (of) who}\}}$}}}}.

vii. The complement of <n*, D_s> gets transferred.

The analyses by Oba (1999, 2003) and Chomsky (2007) are significant in that these proposed structures can explain (49) with the PIC but without the *Subjacency Condition* which was originally proposed by Chomsky (1973). The number of phases, nevertheless, should be reduced on conceptual grounds as discussed above. It will be shown later that an alternative approach can explain the Specificity Effect without postulating the DP/n*P phase. In addition, I will also show that the analyses based on Oba (1999, 2003) and Chomsky (2007) do not hold in the current framework of minimalist syntax (cf. Chomsky (2013, 2015a)).

In the next section, we will review the recent research translating Chomsky's (2007) derivations in (53)–(54) into the current system under labeling by minimal search and freely applying Merge.

3.6.3 Elaborated Internal Structures of Nominals

Oishi (2015), following the suggestions by Chomsky (2007), elaborates internal structures of nominals under the system of labeling by minimal search and freely applying Merge (Chomsky (2013, 2015a)). As Chomsky (2007) suggests, Oishi (2015) assumes that the internal structures of nominals essentially correspond with those of verbal phrases, as shown in (53)–(54). Namely, the parallelism holds between nominal phrases and verbal phrases. Before discussing the internal structures of nominals, let us briefly review the structure and derivation of v*P under the current minimalist model (see Chomsky (2015a)):

(55) {EA, {<√root, v*>, {IA, {$t_{\sqrt{root}}$, t_{IA}}}}}

- i. Merge externally forms {√root, IA}.
- ii. Merge internally forms {IA, {√root, t_{IA}}}.
- iii. Merge externally forms {EA, {v*, {IA, {√root, t_{IA}}}}}.
- iv. Feature Inheritance from v* to √root takes place.
- v. Labeling by minimal search takes place.
- vi. Pair-Merge internally forms {EA, {<√root, v*>, {IA, {$t_{\sqrt{root}}$, t_{IA}}}}}.
- vii. The phase-hood gets active at $t_{\sqrt{root}}$ because $t_{\sqrt{root}}$ is, by definition, visible, and v*, being part of <√root, v*>, becomes invisible.
- viii. The complement of $t_{\sqrt{root}}$ gets transferred.

Adopting Chomsky's (2007) suggestion, Oishi (2015) unifies two functional heads: n* for definite nominals and n for indefinite nominals. He argues that n* or n is, in effect, not for definiteness; rather, what is involved is specificity, although n* or n itself lacks a feature concerning specificity. Accordingly, Oishi (2015: 326) states, "[s]pecificity will be characterized as a part of the features that the nominal heads (n and n*) do not bear along with a categorial specification." Under these assumptions, Oishi (2015) proposes the following derivations of nominal phrases in abstraction away from phase-hood of n/n*:

(56) Indefinite (or non-specific) nominals, pattern 1 (e.g. *authors*):

- i. Merge externally forms {n, √author}.
- ii. Pair-Merge internally forms {<√author, n>, $t_{\sqrt{author}}$}.

(57) Indefinite (or non-specific) nominals, pattern 2 (e.g. *authors*):

- i. Pair-Merge externally forms <√author, n>.

(58) Definite (or specific) nominals, pattern 1 (e.g. *the book*, *that book*, *John's book*):

i. Merge externally forms {n, √book}.

ii. Pair-Merge internally forms {<√book, n>, $t_{\sqrt{\text{book}}}$}.

iii. Pair-Merge externally forms <DP, {<√book, n>, $t_{\sqrt{\text{book}}}$}>.

In Oishi's (2015) original notation, the order in the canonical ordered pairs is opposite. That is, <√author, n> is expressed as <n, √author>. This paper, nevertheless, keeps adopting the canonical notation <√root, affix/categorizer> unless otherwise noted. In addition, Oishi (2015) does not mention the possibility of (59) below, but it is logically possible unless stipulated.

(59) Definite (or specific) nominals, pattern 2 (e.g. *the book*, *that book*, *John's book*):

i. Pair-Merge externally forms <√book, n>.

iii. Pair-Merge externally forms <DP, <√book, n>>.

Furthermore, notice that Oishi (2015) assumes that articles, demonstratives and Saxon genitive noun phrases are all instances of XP. DP in the notation above is, thus, for *the*, *that* and *John's*. This is quite similar to Oba's (1999, 2003) assumption. Recall that Oba (1999, 2003) assumes that D may bear a feature [+specific] in certain cases, and that X for actual articles, such as demonstratives and the Saxon genitive noun phrases, generates in SPEC-D. This similarity suggests that determiners can be decomposed into at least two elements, d and √root. It is plausible that *the* is an instantiation of {d, √the} or <d, √the>. Putting aside this possibility until a later section, we will review how Oishi (2015)'s analysis can explain nominalization.

Oishi (2015) proposes the following structure and derivation for the nominalization of √destroy under the derivations in (56) and (58):

(60) (*the enemy's*/*the*) *destruction of the city*:

<DP$_{\text{the enemy's/the}}$, {<√destroy, n>, {$t_{\sqrt{\text{destroy}}}$, (of) the city}}>

i. Merge externally forms {√destroy, (of) the city}.

ii. Merge externally forms {n, {√destroy, (of) the city}} (see (56), (58)).

iii. Pair-Merge internally forms {<√destroy, n>, {$t_{\sqrt{\text{destroy}}}$, (of) the city}} (see (56), (58)).

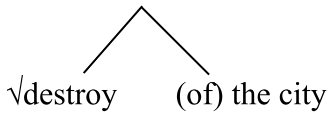
iv. Pair-Merge externally forms <DP$_{\text{the enemy's/the}}$, {<√destroy, n>, {$t_{\sqrt{\text{destroy}}}$, (of) the city}}> (see (58)).

(adapted from Oishi (2015: 332))

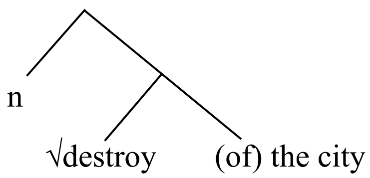
The graph-theoretic notation of (60) is illustrated below:

(61) <DP$_{\text{the enemy's/the}}$, {<√destroy, n>, {$t_{\sqrt{\text{destroy}}}$, (of) the city}}>:

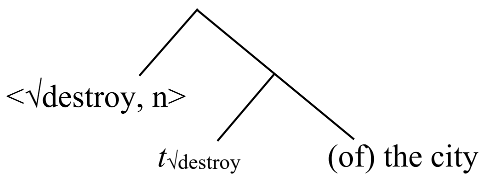
i.



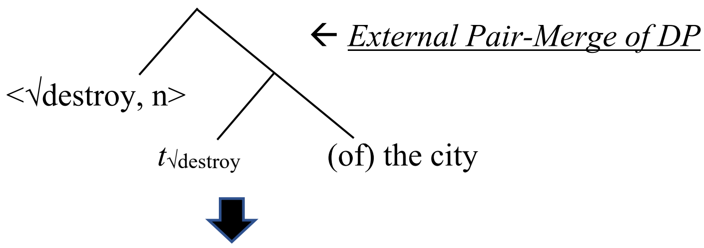
ii.



iii.



iv.



<DP$_{\text{the enemy's/the}}$, {<√destroy, n>, {$t_{\sqrt{\text{destroy}}}$, (of) the city}}>

In this structure, $DP_{\text{the enemy's/the}}$ is optional because external arguments are not essential in nominal phrases, contrary to v*P verbal phrases.[21] The following examples further show that $DP_{\text{the enemy's/the}}$ in (60)–(61) is optional:

(62) a. John dedicated his life to Mary's creation of the computer.
b. John dedicated his life to the creation of the computer.
c. John dedicated his life to creation of the computer.

(63) a. … to <$DP_{\text{Mary's}}$, {<√create, n>, {$_{\alpha}$ $t_{\surd\text{create}}$, (of) the computer}}> …
b. … to <DP_{the}, {<√create, n>, {$_{\alpha}$ $t_{\surd\text{create}}$, (of) the computer}}> …
c. … to {<√create, n>, {$_{\alpha}$ $t_{\surd\text{create}}$, (of) the computer}} …

((62): Oishi (2015: 332))
((63): adapted from Oishi (2015: 332))

(63a–c) are structures of (62a–c), respectively. In (63c), external arguments for nominal structures do not appear. Oishi (2015) also notes that even PRO does not appear as external arguments for nominals.

3.6.4 Problems of Nominal Structures

Oishi's (2015) approach is appealing in that we do not have to assume the extra phase head, D/n*, as argued in Oba (1999, 2003) and Chomsky (2007), among others. In addition, explaining the Specificity Effect in terms of the invisibility of one element in an ordered pair seems promising, although Oishi (2015) does not explicitly demonstrate how his approach accounts for it. His analyses in (60) and (63), however, feature the technical problem of labeling. Recall that the lower copy left by internal pair-Merge is, by definition, visible (see Chomsky (2105a) and Epstein, Kitahara and Seely (2016)). If so, α in (63) cannot be labeled because a root alone is too weak to serve as a label, and no agreement exists between the lower copy of √create and *(of) the computer*. Thus, the

[21] Notice that Oishi (2015) assumes that $DP_{\text{the enemy's/the}}$ is semantically analogous to the familiar external argument of v*P, which is a set-Merged element, although $DP_{\text{the enemy's/the}}$ is, in effect, a pair-Merged element.

analyses in (60) and (63) should be revised under the other logical possibility shown in (59).

(64) A Revised Version of (60):
<the enemy's/the, <<n, √destroy>, (of) the city>>

(65) A Revised Version of (63):

a. … to <$\mathrm{DP}_{\mathrm{Mary's}}$, $\{_{\alpha}$ <√create, n>, (of) the computer}> …

b. … to <$\mathrm{DP}_{\mathrm{the}}$, $\{_{\alpha}$ <√create, n>, (of) the computer}> …

c. … to $\{_{\alpha}$ <√create, n>, (of) the computer} …

In this alternative, the labeling problem does not appear. This is because <√create, n> can be the label of α in the same way that <√root, v*> is labeled as <√root, v*> (i.e. v*P). If it is on the right track, for instance, {<√create, n>, (of) the computer} is labeled as <√create, n> (i.e. nP). However, the alternative still has another problem, which Oishi (2015)'s original approach also features: *pair-Merge applies as a primitive operation*. Pair-Merge should not be the primitive operation for the reasons discussed in chapter 2. In the next section, I will demonstrate that PM by SM solves the problem and provides a new analysis of the internal structures of nominals and the Specificity Effect.

3.6.5 PM by SM for Nominals

As seen in (64)–(65), the internal structures of nominal expressions are essentially based on ordered pairs. If this analysis is on the right track, PM by SM entails the following structure and derivation of a specific nominal, *Mary's creation of the computer*:

(66) {Mary's, {Mary's, {{√create, {√create, n}}, (of) the computer}}} = <Mary's, {<√create, n>, (of) the computer}>

i. PM by SM (i.e. Merge) forms {√create, {√create, n}}, which is representationally interpreted as <√create, n>.

ii. Merge externally forms {{√create, {√create, n}}, (of) the computer}.

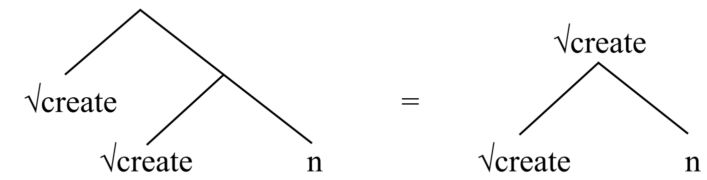
iii. Merge externally forms {Mary’s, {{√create, {√create, n}}, (of) the computer}}.

iv. **PM by SM (i.e. Merge) internally forms {Mary’s, {Mary’s, {{√create, {√create, n}}, (of) the computer}}}, which is representationally interpreted as <Mary’s, {<√create, n>, (of) the computer}>**.

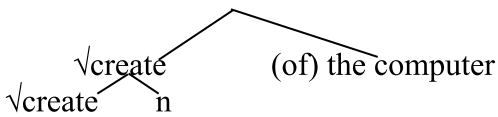
The graph-theoretic notation of this structure and derivation is illustrated below:

(67) {Mary’s, {Mary’s, {{√create, {√create, n}}, (of) the computer}}} = <Mary’s, {<√create, n>, (of) the computer}>

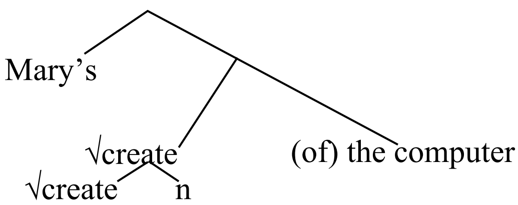
i.



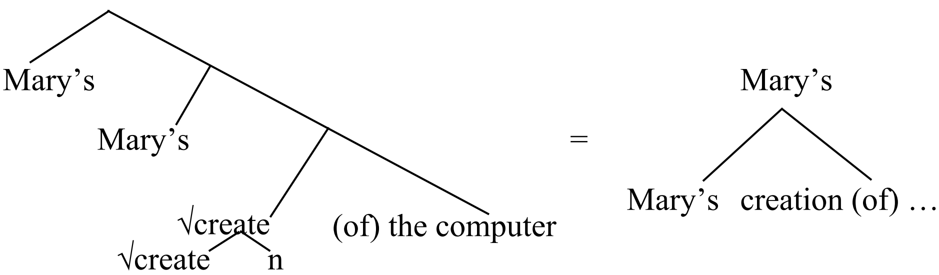
ii.



iii.



iv.



Notice that PM by SM of *heads* (i.e. (29a, b)) applies in the first step. In contrast, PM by SM of *phrases* (i.e. (29c, d)) applies in the final step. It is also crucial that the resulting structure (i.e. <Mary’s, {<√create, n>, (of) the computer}>) can be labeled. Recall the discussion in the previous chapter. In <Mary’s, {<√create, n>, (of) the computer}>, {<√create, n>, (of) the computer} is invisible due to the basic quality of ordered pairs, but an entire ordered pair provides a label if a categorizer is a part of that entire ordered pair or the amalgam.

Now, let us consider the structure and derivation of a non-specific nominal, *creation of the computer*:

(68) {{√create, {√create, n}}, (of) the computer}} = {<√create, n>, (of) the computer}

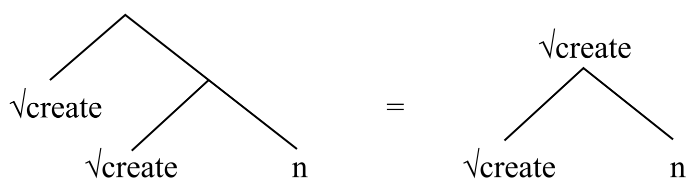
i. PM by SM (i.e. Merge) forms {√create, {√create, n}}, which is representationally interpreted as <√create, n>.

ii. Merge externally forms {{√create, {√create, n}}, (of) the computer}.

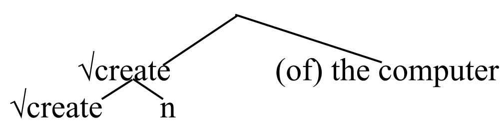
The graph-theoretic notation of the structure and derivation is illustrated below:

(69) {{√create, {√create, n}}, (of) the computer}} = {<√create, n>, (of) the computer}

i.



ii.



In this derivation, PM by SM of *heads* (i.e. (29a, b)) applies in the first step, but PM by SM of *phrases* (i.e. (29c, d)) does not apply.

These new analyses, solving the problems in Oishi’s (2015) approach, can account for the Specificity Effect without postulating the extra phase.

(70) a. *Who did you see John’s picture of t_{wh}?

b. Who did you see a picture of t_{wh}?

Given the analysis in (66), the internal structure of the specific nP is {John’s, {John’s, {{√picture, {√picture, n}}, (of) who}}} = <John’s, {<√picture, n>, (of) who}>. It entails that {{√picture, {√picture, n}}, (of) who} is invisible under the basic quality of ordered pairs.

(71) {John’s, {John’s, {{√picture, {√picture, n}}, (of) who}}} = <John’s, {<√picture, n>, (of) who}>

i. PM by SM (i.e. Merge) forms {√picture, {√picture, n}}, which is representationally interpreted as <√picture, n>.

ii. Merge externally forms {{√picture, {√picture, n}}, (of) who}.

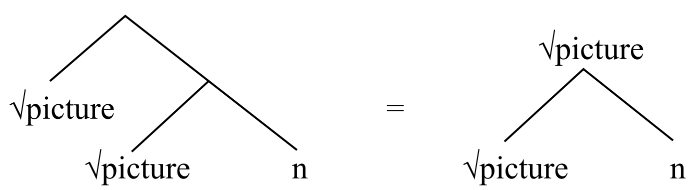
iii. Merge externally forms {John’s, {{√picture, {√picture, n}}, (of) who}}.

iv. **PM by SM (i.e. Merge) internally forms {John’s, {John’s, {{√picture, {√picture, n}}, (of) who}}}, which is representationally interpreted as <John’s, {<√picture, n>, (of) who}>**.

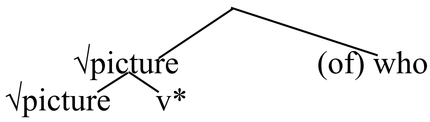
The graph-theoretic notation of this structure and derivation is illustrated below:

(72) {John’s, {John’s, {{√picture, {√picture, n}}, (of) who}}} = <John’s, {<√picture, n>, (of) who}>

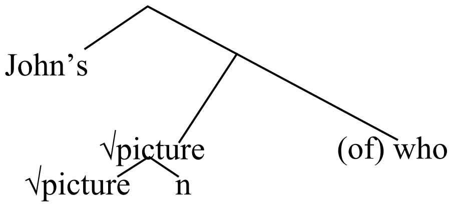
i.



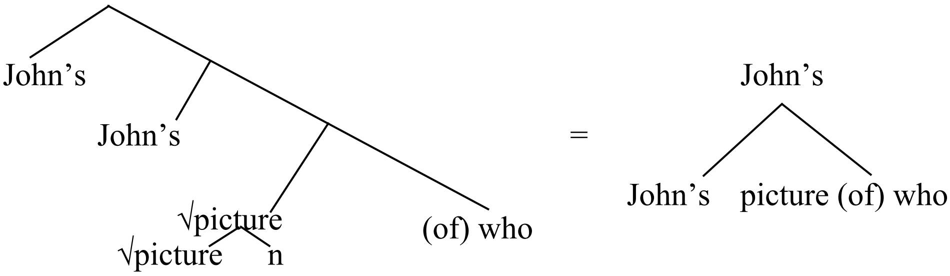
ii.



iii.



iv.



Technically speaking, {{√picture, {√picture, n}}, (of) who} is an invisible term of {John's, {John's, {{√picture, {√picture, n}}, (of) who}}}. Therefore, *who* is non-extractable from the complement of *John's picture of* in (70a). In contrast, given the analysis in (68), the internal structure of the non-specific nP is {{√picture, {√picture, n_a}}, (of) who}}} = {<√picture, n_a>, (of) who}.

(73) {{√picture, {√picture, n_a}}, (of) who}}} = {<√picture, n_a>, (of) who}

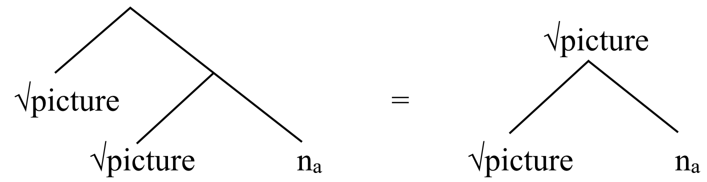
i. PM by SM (i.e. Merge) forms {√picture, {√picture, n_a}}, which is representationally interpreted as <√picture, n_a>.

ii. Merge externally forms {{√picture, {√picture, n_a}}, (of) who}.

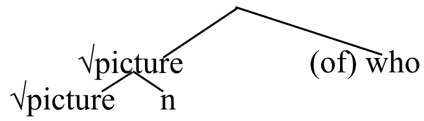
The graph-theoretic notation of the structure and derivation is illustrated below:

(74) {{√picture, {√picture, n_a}}, (of) who}}} = {<√picture, n_a>, (of) who}

i.



ii.



I assume that the root for article *a* in *a picture of who* directly merges with n and creates the amalgam <√a, n> in which √a denotes the root. See the detailed discussion of this assumption in the following section. In this structure, *who* is a typical/standard complement of the amalgam <√picture, n_a>. More precisely, *who* is a visible term of {<√picture, n_a>, (of) who}; hence, it is extractable from the complement of *a picture of* in (70b).

The other case remains to be explained. It is the case of a specific and non-specific nominal expression with *the* as shown below:

(75) $^{(*)}$ Who did you see the picture of t_{wh}?

As Chomsky (1973) points out, the acceptability of extraction from this type of nominal is better than the specific-nominals as in (70a) but worse than the non-specific ones as in (70b). In addition, linguistic literature (e.g. Rodman (1977) and Fiengo and Higginbotham (1981), among others) has shown that *who* is non-extractable when the nominal including this wh-phrase is non-specific, but it is extractable when the nominal including this wh-phrase is specific. Given the consideration, the indeterminacy of acceptance is due to two kinds of the internal structures as in (66)–(68). Thus, we can explain the remaining case too.

Suppose that *the picture of who* is specific. Then, the structure and derivation of the nominal is such as (66) translated into (76) here.

(76) {the, {the, {{√picture, {√picture, n}}, (of) who}}} = <the, {<√picture, n>, (of) who}>

i. PM by SM (i.e. Merge) forms {√picture, {√picture, n}}, which is representationally interpreted as <√picture, n>.

ii. Merge externally forms {{√picture, {√picture, n}}, (of) who}.

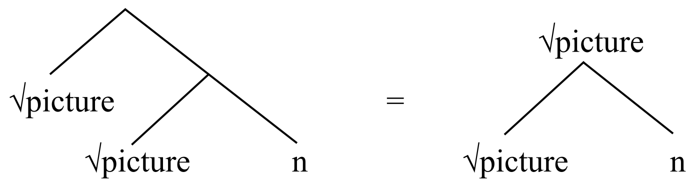
iii. Merge externally forms {the, {the, {{√picture, {√picture, n}}, (of) who}}.

iv. **PM by SM (i.e. Merge) internally forms {the, {the, {{√picture, {√picture, n}}, (of) who}}}, which is representationally interpreted as <the, {<√picture, n>, (of) who}>**.

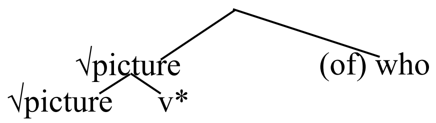
The graph-theoretic notation of this structure and derivation is illustrated below:

(77) {the, {the, {{√picture, {√picture, n}}, (of) who}}} = <the, {<√picture, n>, (of) who}>

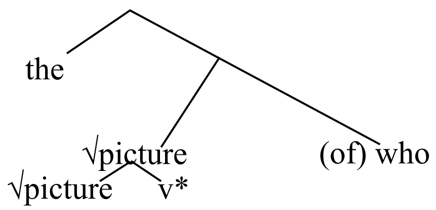
i.



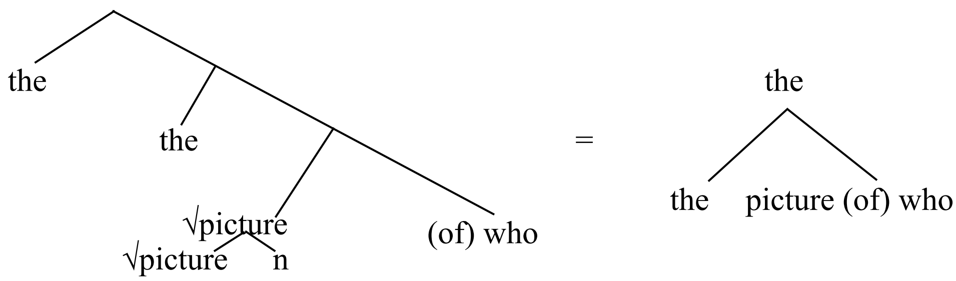
ii.



iii.



iv.



On the contrary, if *the picture of who* is non-specific, the structure and derivation of the nominal is as (68) translated into (78) here.

(78) {{√picture, {√picture, n_{the}}}, (of) who} = {<√picture, n_{the}>, (of) who}

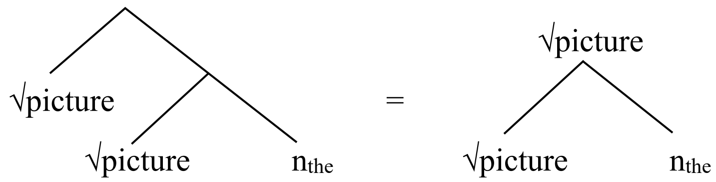
i. PM by SM (i.e. Merge) forms {√picture, {√picture, n_{the}}}, which is representationally interpreted as <√picture, n>.

ii. Merge externally forms {{√picture, {√picture, n_{the}}}, (of) who}.

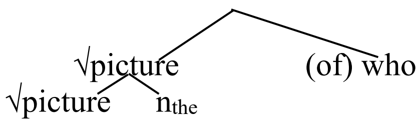
The graph-theoretic notation of the structure and derivation is illustrated below:

(79) {{√picture, {√picture, n_{the}}}, (of) who} = {<√picture, n_{the}>, (of) who}

i.



ii.



who is impossible to extract in (76) because it is invisible as a part of the ordered pair <the, {<√picture, n>, (of) who}>. Notice that *the* in (78) is the instantiation of n_{the}, and n_{the} is the instantiation of <√the, n>. I will discuss the internal structures of determiners in the next section.

3.6.6 Internal Structures of Determiners

As mentioned earlier, we discuss the internal structures of determiners in this section. Recall that both Oba (1999, 2003) and Oishi (2015) assume that determiners are not the simple D but more complex objects. Translating their assumption into the current minimalist model, a determiner consists of a categorizer d and a √root, such as √the, √that and √John. As Oishi (2105) argues, *the* may be a weaker version of *that*. If this is true, √root for *the* and *that* is the same element, and categorizer d makes these elements different. As generally assumed, categorizers or functional heads, especially phase heads C/v*, and roots are sets of linguistic features. The feature specification of d is too abstract and too hard to capture. However, if d can bear something such as [+specific] as in Oba (1999, 2003), it is plausible that d can also bear another feature such as [+definiteness] and [+universal (∀)]. If this line of reasoning is on the right track, we speculate that √root for articles and demonstratives is the sole element, and that its instantiation depends on the contents of d. In contrast, other elements, for example, prenominal genitives (Saxon genitives), numerals and negatives, seem to have substantive roots. Whatever the correct analysis is, I assume that a determiner, at least, consists of √root and d.[22]

Following the assumption, the structures in (76) and (78) are further elaborated as follows:

(80) {**{d, √the}**, {**{d, √the}**, {{√picture, {√picture, n}}, (of) who}}} = <**{d, the}**, {<√picture, n>, (of) who}>

i. PM by SM (i.e. Merge) forms {√picture, {√picture, n}}, which is representationally interpreted as <√picture, n>.

[22] Wood and Marantz (2015) argue that prepositional phrases consist of √root and p in their highly elaborated version of Distributed Morphology. If pP and dP both have their categorizers and roots, all functional categories (e.g. p, d, n and a) except for the core functional categories (C, T and v) would have their roots directly merge with their functional heads by external Merge. If it is on the right track, P_{HAVE} (see chapter 2) could be decomposed into p and √have.

ii. Merge externally forms {{√picture, {√picture, n}}, (of) who}.

iii. Merge externally forms {**{d, √the}**, {{√picture, {√picture, n}}, (of) who}}.

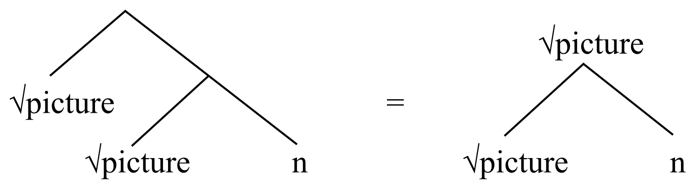
iv. PM by SM (i.e. Merge) internally forms {**{d, √the}**, {**{d, √the}**, {{√picture, {√picture, n}}, (of) who}}}, which is representationally interpreted as <**{d, √the}**, {<√picture, n>, (of) who}>.

(adapted from (76))

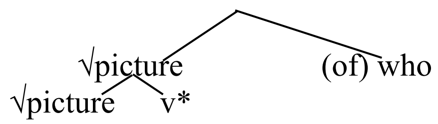
The graph-theoretic notation of this structure and derivation is illustrated below:

(81) {**{d, √the}**, {**{d, √the}**, {{√picture, {√picture, n}}, (of) who}}} = <**{d, √the}**, {<√picture, n>, (of) who}>

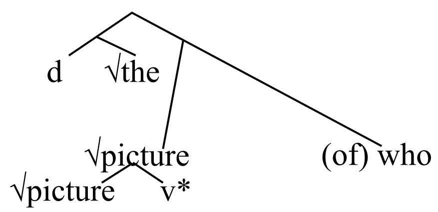
i.



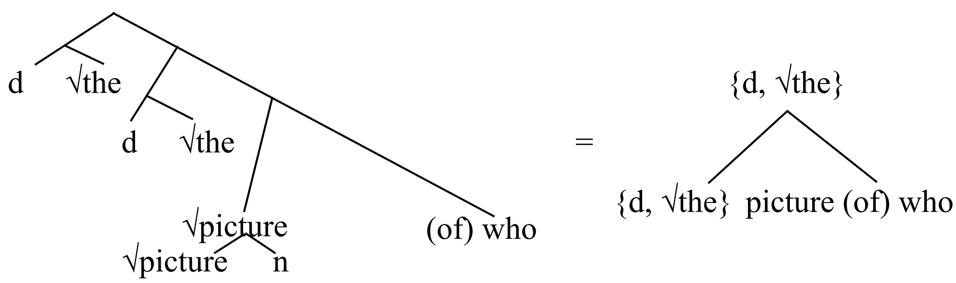
ii.



iii.



iv.



(82) {{√picture, {√picture, **<√the, n>**}}, (of) who} = {<√picture, **<√the, n>**>, (of) who}

  i. PM by SM (i.e. Merge) forms {√picture, {√picture, **<√the, n>**}}, which is representationally interpreted as <√picture, n>.

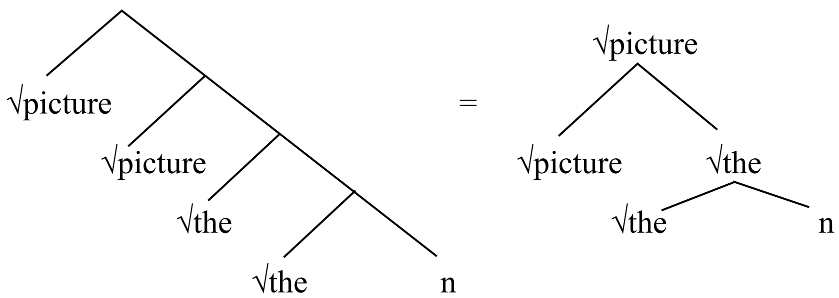
  ii. Merge externally forms {{√picture, {√picture, **<√the, n>**}}, (of) who}.

(adapted from (78))

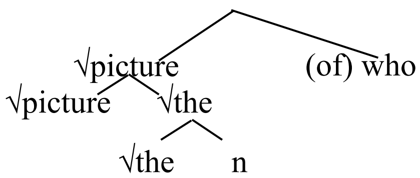
The graph-theoretic notation of the structure and derivation is illustrated below:

(83) {{√picture, {√picture, **<√the, n>**}}, (of) who} = {<√picture, **<√the, n>**>, (of) who}

  i.



  ii.



Crucially, {d, √the} should not be the logical alternative <√the, d> in (80). As argued in Chomsky (2007) and Oishi (2015), among others, noun phrases are not determiner phrases because both phrases are intuitively *nominals*. Following the intuition, it is naturally deduced that the label of the so-called NP and DP turns into *n*, which is the n head bearing the set of interpretive phi-features and plausibly other features. This prediction is technically welcome. Suppose that the label of nominals is not n but d. Then, we have at least two problems. The first problem is that PM by SM cannot correctly apply and cannot yield the desired result. The second one is that minimal search either labels nominals as dP or fails to label. Let us consider this case below:

(84) {$_\alpha$ **{√the, {d, √the}}**, {**{√the, {d, √the}}**, {$_\beta$ {√picture, {√picture, n}}, (of) who}}} ≠ <$_\alpha$ **<√the, d>**, {$_\beta$ <√picture, n>, (of) who}>

i. PM by SM (i.e. Merge) forms {$_\beta$ √picture, {√picture, n}}, which is representationally interpreted as <√picture, n>.

ii. Merge externally forms {{$_\beta$ √picture, {√picture, n}}, (of) who}.

iii. PM by SM forms **{√the, {d, √the}}**, which is representationally interpreted as **<√the, d>**.

iv. Merge externally forms {**{√the, {d, √the}}**, {$_\beta$ {√picture, {√picture, n}}, (of) who}}.

v. PM by SM (i.e. Merge) internally forms {$_\alpha$ **{√the, {d, √the}}**, {**{√the, {d, √the}}**, {$_\beta$ {√picture, {√picture, n}}, (of) who}}}, which **cannot be** representationally interpreted as <$_\alpha$ <√the, d>, {$_\beta$ <√picture, n>, (of) who}>.

First, PM by SM cannot form the derivational ordered pair or the representational ordered pair as in (84). Recall (29), repeated here as (85):

(85) PM by SM (the final version):

head-head adjunction

a. Pair-Merge (X^0, Y^0) = Merge (X^0, (X^0, Y^0)) = Merge (X^0, {X^0, Y^0}) = {X^0, {X^0, Y^0}} = $\langle X^0, Y^0 \rangle$

b. Pair-Merge (Y^0, X^0) = Merge (Y^0, (Y^0, X^0)) = Merge (Y^0, {Y^0, X^0}) = {Y^0, {Y^0, X^0}} = $\langle Y^0, X^0 \rangle$

phrase-phrase adjunction

c. Pair-Merge (XP, YP) = Merge (XP, (XP, YP)) = Merge (XP, {XP, YP}) = {XP, {XP, YP}} = <XP, YP>

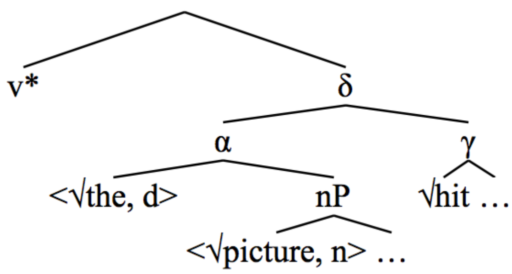
d. Pair-Merge (YP, XP) = Merge (YP, (YP, XP)) = Merge (YP, {YP, XP}) = {YP, {YP, XP}} = <YP, XP>

Given the formulation of PM by SM, Pair-Merge (X^0, YP) = Merge (X^0, (X^0, YP)) cannot form a derivational/representational ordered pair. In other words, phrase-head Merge cannot yield a derivational/representational ordered pair. In (84), {√the, {d, √the}} (i.e. representationally <√the, d>) is an X^0-level object (i.e. head), but {$_\beta$ <√picture, n>, (of) who} is an XP-level object (i.e. phrase). Recall that PM by SM applies only to same-level objects. The resulting structure in (84) is thus only a gibberish set at least because internally merging the determiner {√the, {d, √the}} is redundant.

Second, even if we form the ordered pair <$_\alpha$ **<√the, d>**, {$_\beta$ <√picture, n>, (of) who}> somehow, minimal search labels α as <√the, d>P, which is dP informally, whereas minimal search labels β as nP. This leads the derivation to crash in terms of the failure of agreement/feature sharing. For example, the problematic argument structure of *John hit the picture of who* is shown as follows (in the graph-theoretic notation, several irrelevant parts are omitted for expository purposes):

(86) a. {John, {v*, {$_\delta$ <$_\alpha$ <√the, d>, {$_{nP}$ <√picture, n>, (of) who}>, {$_\gamma$ √hit, $t_{\text{the picture of who}}$}}}}}

b.



As shown in the graph-theoretic notation, the verbal root √hit and the amalgam <√the, d> are at the same depth. At the relevant stage of the derivation, minimal search thus locates them, but the labeling of δ and γ fails in terms of no feature sharing (i.e. no agreement/no strong relation). Notice that the nP is also at the same depth as the amalgam and the root, but it is not an atomic element/head but rather a set with the given label nP. Accordingly, the nP is not a possible candidate for labeling α, δ and γ. In addition, it is crucial that labels do not affect the operations in the narrow syntax (e.g. no recourse to the LAC). Thus, while the nP is, in effect, n under the bare phrase structure theory (see Chomsky (1995a, b)), it cannot be a candidate for the labeler. The detailed steps of the derivation for the structure are as follows:

(87) i. Upon both the formation of the argument structure by Merge and feature inheritance from v* to √hit, labeling by minimal search takes place at the timing of Transfer.

ii. Minimal search labels {<√picture, n>, (of) who} as nP.

iii. Minimal search locates the nP and <√the, d> in α.

iv. Minimal search labels α as dP because <√the, d> is a head.

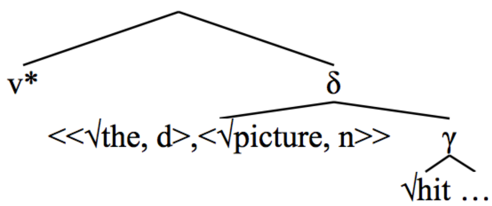
v. Minimal search locates the <√the, d> and √hit in δ.

vi. **Minimal search fails to label δ because <√the, d> and √hit are not in the strong relation (i.e. agreement)**.

One might say that labeling may still logically take place correctly as long as object nP does not contain a complement such as *(of) who*. This possibility is illustrated below:

(88) a. {John, {v*, {$_{\delta}$ <<√the, d>, <√picture, n>>, {$_{\gamma}$ √hit, $t_{\text{the picture}}$}}}}}

b.



In this case, minimal search labels δ as <<√the, d>, <√picture, n>> because it is a head. Then, γ remains unlabeled, and such a structure, by definition, cannot obtain the correct interpretation at the CI, thus violating Full Interpretation. What is worse, uCase on *the picture* also remains unvalued, and again, the structure crashes at the CI in terms of a violation of Full Interpretation. The steps of this derivation are delineated below:

(89) i. Upon both the formation of the argument structure by Merge and feature inheritance from v* to √hit, labeling by minimal search takes place at the timing of Transfer.

ii. Minimal search locates <<√the, d>, <√picture, n>> and $\{_{\gamma}$ √hit, $t_{\text{the picture}}\}$ in δ.

iii. Minimal search labels δ as <<√the, d>, <√picture, n>> **whereas fails to label γ because √hit alone is too weak to serve as a label**. **Furthermore, uCase on the nominal remains unvalued, and it causes the derivation to crash at the CI interface**.

3.7 Summary

In this chapter, I have extended PM by SM of heads to PM by SM of heads/phrases. PM by SM of heads was originally proposed as the operation applying only to heads/X^0-objects. The targets of Merge, however, conceptually should not be restricted, considering the nature of the faculty of language which assures recursion. Put differently, Merge should access any object in the narrow syntax unless stipulated. This paradoxical situation is resolved via an extended version of PM by SM. In this final version of PM by SM, the targets of the application are loosely restricted in terms of a definition of ordered pairs in ZFC. Considering the definition of ZFC, the targets of the application of PM by SM need to be the same levels, which are either set-members (i.e. XP-level) or atomic-members (i.e. X^0-level). If the targets are set-members, this application of PM by SM is equal to the phrasal adjunction. Consequently, it follows that the final reformulation of PM by SM can explain various empirical facts concerning the phrasal adjunction such as the (anti-)adjunction condition effects, the Condition C effects and the Specificity Effects. Furthermore, the proposed reformulation makes us deeply consider new possibilities of the internal structures of nominals. By doing so, we move toward the true nature of the faculty of language under the spirit of the Minimalist Program.

Chapter 4

Reformulating Inheritance and Valuation

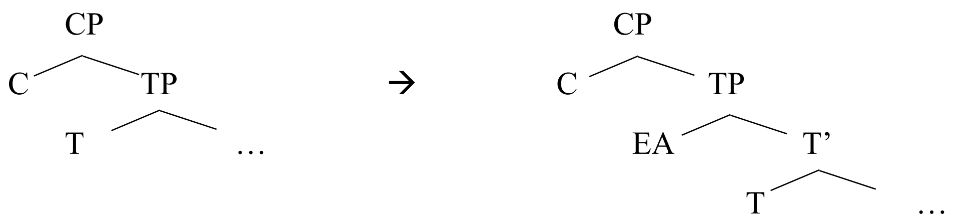
4.1 Introduction

In the previous chapters, we reformulated pair-Merge in terms of evolvability, simplicity and the basic properties of the faculty of language. This chapter mainly discusses and reformulates another problematic operation, feature inheritance, as it is conceptually ideal that we eliminate feature inheritance as an elemental operation (see Chomsky, Gallego and Ott (to appear), Epstein, Kitahara and Seely (2014b) and Kitahara (2017)). In this chapter, feature inheritance is not an elemental operation but rather a special case of minimal search regarding a set of unvalued Phi-features (uPhi). In addition, I reformulate feature valuation to be the same minimal search as labeling, adapting Chomsky's (2015a) suggestion. I also show that reformulations explain various empirical facts especially concerning various types of *there* constructions.

4.1.1 An Earlier Version of the Minimalist Framework

In an earlier version of the minimalist framework (e.g. Chomsky (2008)), a basic derivation proceeds as follows (the graph-theoretic notation is used only for expository purposes below):

(1)



i. External Merge forms the set-theoretic object [$_{CP}$ C [$_{TP}$ T [$_{v*P}$ EA v*…]]].

ii. Feature inheritance: T inherits features from C because external Merge has established the direct relation of C-T.

iii. Internal Merge forms [$_{CP}$ C [$_{TP}$ EA T [$_{v*P}$ t_{EA} v*…]]] because uPhi on T has probed into EA.

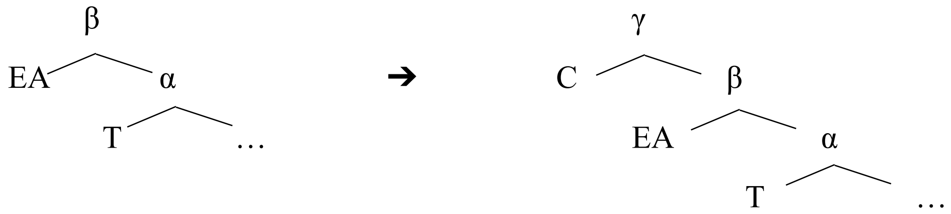
In this derivation, step (iii) is counter-cyclic because internal Merge inserts EA into C-TP. In other words, internal Merge *replaces* [$_{CP}$ C [$_{TP}$ T [$_{v*P}$ EA v*…]]] with [$_{CP}$ C [$_{TP}$ EA T [$_{v*P}$ t_{EA} v*…]]]. This kind of replacement is barred in the current minimalist model because the counter-cyclic internal Merge violates two conditions. The first violated condition is the Extension Condition which is not an independent principle but is deduced by the minimal search of Merge itself (see Chomsky (2015b: 82)). Recall that minimal search is part of minimal computation, a third factor principle. The second violated condition is the NTC (No-Tampering Condition). Note that the NTC, as well as the Inclusiveness Condition, conforms to the SMT (Strong Minimalist Thesis). In the derivation above, we cannot alter [$_{CP}$ C [$_{TP}$ T [$_{v*P}$ EA v*…]]] into the different syntactic object [$_{CP}$ C [$_{TP}$ EA T [$_{v*P}$ t_{EA} v*…]]]. In contrast, Epstein, Kitahara and Seely (2012, 2014) motivate this counter-cyclic internal Merge by elaborating the system of derivations. Their approach capitalizes on the two-peaked structures which deduce Transfer per phase without depending on the lexical array or sub-array (see chapter 3).

Chomsky (2013, 2015a), nevertheless, resolves the problem of the counter-cyclic internal Merge in a different way. We review this different way in the following section.

4.1.2 The Current Framework and a Problem

In the current framework (e.g. Chomsky (2013, 2015a)), Merge applies *freely* in any order, regardless of whether external or internal. The basic derivation proceeds as follows:

(2)



i. External Merge forms the set-theoretic object [$_{\alpha}$ T [EA v*…]].

ii. Internal Merge forms [$_{\beta}$ EA [$_{\alpha}$ T [$_{v*P}$ t_{EA} v*…]]].

iii. External Merge forms [$_{\gamma}$ C [$_{\beta}$ EA [$_{\alpha}$ T [$_{v*P}$ t_{EA} v*…]]]].

iv. Feature inheritance: T inherits features from C, **but how does it take place?**

The derivation based on freely applying Merge resolves the problem of the counter cyclic internal Merge, strictly following the Extension Condition and the NTC. This is because the application of (external and) internal Merge is strictly cyclic. This new system, however, leads to another problem concerning feature inheritance. Namely, feature inheritance takes place *problematically* because no appropriate relation exists between C and T (i.e., EA intervenes in the relation C-T). Put differently, it is obscure how feature inheritance applies in the current minimalist framework.

In this chapter, I resolve the problem by proposing the reformulation of feature inheritance (and also feature valuation). The main proposal of this chapter is to argue that feature inheritance is simply minimal search. More concretely, I propose an equation, Feature Inheritance = Matching = Minimal Search, which is defined as follows:

(3) Feature inheritance is minimal search by uPhi-set, which identifies the most appropriate goal; hence matching.

I also propose that feature valuation is the same minimal search as minimal search for labeling.

(4) Labeling is minimal search, which identifies unlabeled syntactic objects and unvalued features; hence valuation.

The two proposals explain the derivation of the *there* construction which has an empirical problem under the current minimalist model. I will point out the problem in the next section.

This chapter is organized as follows. Section 4.2 reviews analytic assumptions and points out that the current minimalist framework has an empirical problem in the *there* construction, which is called an XP-YP problem. In section 4.3, I reformulate feature inheritance and feature valuation. In this reformulation, both operations are just minimal search. I also demonstrate that the reformulated feature inheritance deduces the condition on internal PM by SM of heads. In section 4.4, it is demonstrated that the reformulation of feature inheritance resolves the XP-YP problem in the *there be* construction and the unaccusative *there* construction. In section 4.5, the proposals will further explain "unaccusativized" *there* constructions. In section 4.6, it is illustrated that the proposals and analyses can explain the (non-)extractablity and (non-)subextractablity of associate and scopal/binding phenomena in *there* constructions. Section 4.7 summarizes the chapter.

4.2 Analytic Assumptions and a Conceptual Problem

4.2.1 Simplest Merge and Labeling

In the earlier models of the Minimalist Program, Merge automatically produces a label (see Chomsky (1995a, b)). In contrast, Chomsky (2013, 2015a) has formulated Merge as simplest Merge, which does not yield the label automatically.

(5) Chomsky (1995a, b):

Merge $(\alpha, \beta) = \{\alpha, \{\alpha, \beta\}\}/\alpha$ = label

(6) Chomsky (2013, 2015a):

Simplest Merge $(\alpha, \beta) = \{\alpha, \beta\}$

Simplest Merge eliminates labels; labeling takes place by minimal search as has been elsewhere in this paper. This labeling process is repeated below:

(7) Labeling by Minimal Search

a. $\{H, XP\} \rightarrow \{_{HP}\ H, XP\}$

b. $\{XP, t_{YP}\} \rightarrow \{_{XP}\ XP, t_{YP}\}$

c. $\{XP, YP\}/X$ sharing features (e.g. Phi-set) with Y

$\rightarrow \{_{\langle Phi, Phi\rangle}\ XP, YP\}$

4.2.2 Freely Applying Merge

At the third step in (1), the probing of uPh triggers internal Merge, but Merge freely applies in the current minimalist framework.

(8) *John hit Mary*:

$\{C, \{_{\gamma}\ John, \{_{\delta}\ T, \{t_{John}, \{_{v^*P} \langle \surd hit, v^* \rangle, \ldots\}\}\}\}\}$

After all computations for v*P phase have taken place, (i) Merge externally forms $\{T, \{John, \{_{v^*P}\ \surd hit\text{-}v^* \ldots\}\}\}$. (ii) Merge internally forms $\{John\ \{T, \{t_{John}, \{_{v^*P}\ \surd hit\text{-}v^* \ldots\}\}\}\}$. (iii) Merge externally forms $\{C, \{John, \{T, \{t_{John}, \{_{v^*P}\ \surd hit\text{-}v^* \ldots\}\}\}\}\}$. (iv) T inherits features from C (feature inheritance). (v) Labeling by minimal search takes place: γ and δ are labeled as <Phi, Phi> and TP, respectively, as *John* and T agree. (vi) Pair-Merge internally forms <√hit, v*>. (vii) The complement of C gets transferred.

This basic system of derivations was further refined in earlier chapters, especially in chapters 2 and 3. That is, pair-Merge has been reformulated by the double application of Merge, PM by SM (Pair-Merge formulated by Simplest Merge). Adopting the reformulation, the derivational steps in (8) are refined as follows:

(9) *John hit Mary*:
{C, {$_{\gamma}$ John, {$_{\delta}$ T, {t$_{\text{John}}$, {$_{\text{v*P}}$ {√hit, {√hit, v*}}, …}}}}}

After the completion of v*P phase, (i) Merge externally forms {T, {John, {$_{\text{v*P}}$ √hit-v*…}}}. (ii) Merge internally forms {John {T, {*t*$_{\text{John}}$, {$_{\text{v*P}}$ √hit-v*…}}}}. (iii) Merge externally forms {C, {John, {T, {*t*$_{\text{John}}$, {$_{\text{v*P}}$ √hit-v*…}}}}}. (iv) T inherits features from C (feature inheritance). (v) PM by SM internally forms {√hit, {√hit, v*}} which is representationally interpreted as <√hit, v*>. (vi) Labeling by minimal search takes place: γ and δ are labeled as <Phi, Phi> and TP, respectively, as *John* and T agree. (vii) The complement of C gets transferred.

4.2.3 A Conceptual Problem

As we have seen, a C-T relation has never been established in (8i-vii) and (9i-vii), as *John* intervenes in the C-T relation. In short, the problem is how feature inheritance takes place within the derivational system shown in (8i-vii) and (9i-vii). Note that T. Daniel Seely (p.c.) and Epstein, Kitahara and Seely (2014b) are aware of the same problem and propose the alternative steps of Merge, but I do not adopt their idea, which I will discuss in the next chapter.

(10) A Conceptual Problem:
Under the typical syntactic derivation in the current minimalist framework (Chomsky (2013, 2015a)), freely and cyclically applying simplest Merge cannot establish an appropriate relation for feature inheritance.

In what follows, I will solve this problem and explain the structures and derivations of *there* constructions.

4.3 Reformulating Feature Inheritance and Feature Valuation

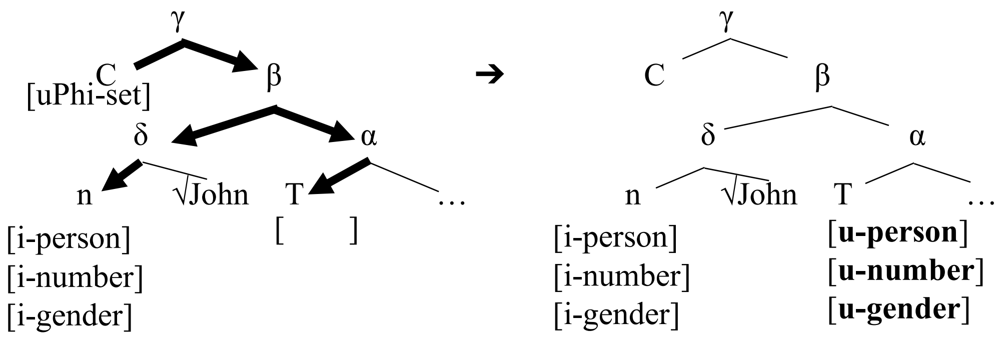
4.3.1 Reformulating Feature Inheritance

To resolve the conceptual problem above, I propose the following hypothesis:

(11) Feature Inheritance = Matching = Minimal Search:
Feature inheritance is minimal search by uPhi-set, which identifies the most appropriate goal; hence matching.

Given the hypothesis, the uPhi-set (perhaps along with other features, such as Tense) of a phase head probes into its c-commanding domain in a similar way to Probe-Goal Agree (see Chomsky (2000, 2001)). Let us consider how hypothesis (11) resolves the problem in (10). Feature Inheritance = Matching = Minimal Search works as illustrated below:

(12) $[_{\gamma}$ C $[_{\beta}$ $[_{\delta}$ n √John] $[_{\alpha}$ T …]]]:



The uPhi-set on C simultaneously finds both a weak head (e.g. T) and a nominal interpretable Phi-set on a head n in *John*. Crucially, I assume that this tells the uPhi-set which uPhi the weak head should inherit (i.e. matching). For instance, when the uPhi-set simultaneously finds T and the n head of *John* [3rd-person, singular, masculine], the matched [u-person, u-singular, u-gender] is transmitted to T.

(13) A weak head inherits features in a matching relation established by minimal search by uPhi-set.

On this account, the proposal (11) solves the problem (10) because minimal search establishes the appropriate relation for feature inheritance.

4.3.2 Reformulating Valuation

After the matching effect is maximized (cf. Chomsky (2001: 15)), the next step is feature valuation. Adapting Chomsky's (2013, 2015a) suggestion, I assume that feature valuation is a by-product of labeling by minimal search.

(14) Labeling = Feature Valuation = Minimal Search:
Labeling is minimal search, which identifies unlabeled syntactic objects and unvalued features; hence valuation.

Following this assumption, feature valuation takes place exactly when labeling does. The reformulation is reasonable. Labeling involves putting the names of simple heads on unspecified syntactic sets. Valuation also involves putting the names of matched features on unvalued/unspecified features. The parallelism should hold as long as both labeling and valuation are minimal search.

For the unification of labeling and feature valuation (or Agree), see Epstein, Kitahara and Seely (2014b) and Kato et al. (2016), among others. See also Epstein, Obata and Seely (to appear) and Chomsky (2016b) for an analysis still holding something as Probe-Goal Agree in the current minimalist framework.

4.3.3 Deducing the Condition on Internal PM by SM of Heads

As discussed in chapter 2, Chomsky (2007: 6, 2013: 40, fn. 20, 2015b: 82) argues that the concept of multidimensionality or multidominance, which late Merge, parallel Merge (see Nunes (1995) and Citko (2005)) and other *subspecies* of Merge involve, violates "the minimal search requirement for Merge itself" (Chomsky (2015b: 82)). Namely, Merge is binary but not ternary, quaternary or more in terms of simplicity and computational efficiency. As for the ternary relation, recall that Chomsky (2015b: 82) states that "[y]ou're finding one item in the workspace, you find something inside it – that's two – and then you're finding a third one which you attach this to."

(15) The Binary Requirement of Minimal Search for Merge:
$n = 2$/Merge $(X_1, \ldots, X_n) = \{X_1, \ldots, X_n\}$

This requirement must be applied to pair-Merge as long as it is one type of Merge.

(16) The Binary Requirement of Minimal Search for Pair-Merge:
$n = 2$/pair-Merge $(X_1, \ldots, X_n) = \langle X_1, \ldots, X_n \rangle$

If (16) holds, when pair-Merge internally applies to elements in a given workspace, head-movement cannot occur. Let us review the following typical structure and derivation for the v*P phase under the system of Chomsky (2015a).

(17) {John, {<√hit, v*>, $\{_\delta$ Mary, $\{_\gamma$ $t_{\sqrt{\text{hit}}}$, $t_{\text{Mary}}\}\}\}\}$

i. Merge externally forms $\{_\gamma$ √hit, Mary}.
ii. Merge internally forms $\{_\delta$ Mary, $\{_\gamma$ √hit, $t_{\text{Mary}}\}\}$.
iii. Merge externally forms {John, {v*, $\{_\delta$ Mary, $\{_\gamma$ √hit, $t_{\text{Mary}}\}\}\}\}$.
iv. √hit inherits features from v*.
v. Labeling and Agree take place by minimal search: δ and γ are labeled as <Phi, Phi> and √P, respectively.
vi. Pair-Merge internally forms <√hit, v*> with v* affixed: such a v* becomes invisible, and the phase-hood is activated on $t_{\sqrt{\text{hit}}}$.
vii. The complement of $t_{\sqrt{\text{hit}}}$ gets transferred.

(cf. (2i-iii), (8))

When internal pair-Merge applies in (17), you find {John, {v*, $\{_\delta$ Mary, $\{_\gamma$ √hit, $t_{\text{Mary}}\}\}\}\}$ in the workspace, you find √hit inside it – that is two – and then you find the third one, v* which you attach this to. This violates (16) because $n \neq 2$ but $n = 3$.

The proposals and analyses in the sections 4.3 and 4.4 resolve this problem. This is because minimal search for feature inheritance has already established the relation between two lexical items through minimal search for feature inheritance. Minimal search for pair-Merge, therefore, does not need to look at the SO (syntactic object) {John, {v*, {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}}}} in the workspace.

(18) Minimal search for pair-Merge does not apply iff minimal search for feature inheritance has established the relation between two lexical items.

(19) {John, {<√hit, v*>, {$_{\delta}$ Mary, {$_{\gamma}$ $t_{\sqrt{\text{hit}}}$, t_{Mary}}}}}

i. Merge externally forms {$_{\gamma}$ √hit, Mary}.

ii. Merge internally forms {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}}.

iii. Merge externally forms {John, {v*, {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}}}}.

iv. **The uPhi-set on v* minimally searches for *Mary* and √hit: √hit inherits uPhi-set from v*** (feature inheritance).

v. Labeling and feature valuation take place by minimal search: δ and γ are labeled as <Phi, Phi> and √P respectively.

vi. Pair-Merge internally forms <√hit, v*> **thanks to the relation established via minimal search in (19iv)**: such a v* becomes invisible, and the phase-hood is activated on $t_{\sqrt{\text{hit}}}$.

vii. The complement of $t_{\sqrt{\text{hit}}}$ gets transferred.

The generalization in (18) is quite similar to the condition on internal PM by SM discussed in chapter 2, which is repeated here as (20):

(20) Merge only sees relevant heads iff feature inheritance has established the relation between two lexical items.

The generalization in (18) is, therefore, finally completed by (20) and PM by SM as shown below:

(21) Minimal search for Merge does not apply iff minimal search for feature inheritance has established the relation between two lexical items.

Rigidly assuming that feature inheritance is minimal search makes the condition on internal PM by SM much more advanced. Although we do not know why the original statement in (20) can hold, we can deduce the final version of the condition in (21) as discussed above. The deduction is summarized as follows:

(22) Deductive Steps to the Condition on Internal PM by SM of Heads:

i. The Binary Requirement of Minimal Search for Merge: n = 2/Merge $(X_1, \ldots, X_n) = \{X_1, \ldots, X_n\}$.

ii. Feature Inheritance = Minimal Search, which establishes the relation between two lexical items.

iii. The relation established by Feature Inheritance = Minimal Search ensures binarity of minimal search for Merge.

iv. Internal Merge of two heads applies successfully because the binarity has been ensured by minimal search for feature inheritance.

v. The condition on internal PM by SM is deduced.

The typical structure and derivation for the v*P phase in (17) is refined, eliminating a primitive operation, pair-Merge, as argued elsewhere:

(23) {John, {{√hit, {√hit, v*}}, $\{_\delta$ Mary, $\{_\gamma$ $t_{\surd hit}$, t_{Mary}}}}}

i. Merge externally forms $\{_\gamma$ √hit, Mary}.

ii. Merge internally forms $\{_\delta$ Mary, $\{_\gamma$ √hit, t_{Mary}}}.

iii. Merge externally forms {John, {v*, $\{_\delta$ Mary, $\{_\gamma$ √hit, t_{Mary}}}}}.

iv. √hit inherits features from v*.

v. PM by SM internally forms {√hit, {√hit, v*}}: such a v* becomes invisible, and the phase-hood is activated on $t_{\surd hit}$.

vi. Labeling and Agree take place by minimal search: δ and γ are labeled as <Phi, Phi> and √P respectively.

vii. The complement of $t_{\sqrt{\text{hit}}}$ gets transferred.

(Cf. (9))

The application of internal Merge of heads in (23v) is possible thanks to the condition in (21).

4.4 Resolving an XP-YP Problem

4.4.1 An XP-YP Problem

The proposals in (13) and (14) provide the principled explanation for a symmetric XP-YP problem in *there* constructions. In the copular structure {*be* {$_{\delta}$ XP, YP}}, XP or YP internally merges to SPEC-T (i.e. SPEC-*be*) for labeling δ (see Chomsky (2013: 43–44) and Moro (1997)) unless X agrees with Y:

(24) XP copula {$_{\delta}$ t_{XP}, YP}

As a result, δ is labeled as YP because a lower copy left by set-Merge (but not pair-Merge) is, by definition, invisible. For instance, the underlying structure of *the cause of the fire is lightning* or *lightning is the cause of the fire* is as follows:

(25) [be [lightning, the cause of the fire]]

The resulting sentences depend on which SO internal Merge takes.

This readily predicts that internal Merge of XP or YP does not take place in *there* constructions because *there* has externally merged to SPEC-T (i.e. *there* insertion). The prediction leads the derivation to crash at the interface(s) since δ is not labeled (the XP-YP problem).

(26) An Empirical Problem (The XP-YP Problem):

In a *there* construction, such as *there are* [$_{\delta}$ [$_{nP}$ *three men*] [$_{PP}$ *in the room*]], label δ is not determined.

In the next section, I show that the proposals in (11) and (14) resolve the empirical XP-YP problem.

4.4.2 *There Be* Constructions

The proposals in (13) and (14) resolve the XP-YP problem above, assuming that the copula is virtually the derivational ordered pair {√be, {√be, T}} or the representational ordered pair <√be, T>. The structure and derivation of *there are three men in the room* under the current minimalist framework is as follows (bold parts below are new analyses entailed by the proposals):[1]

(27) {C, {$_{\gamma}$ there, {{√be, {√be, T}}, {$_{\delta}$ {three, men}, {$_{\varepsilon}$ $t_{\surd be}$, {$t_{\text{three men}}$, {in, {the, room}}}}}}}}

i. Merge externally forms {√be, {{three, men}, {in, {the, room}}}}.

ii. Merge internally forms {{three, men}, {√be, {$t_{\text{three men}}$, {in, {the, room}}}}}.

iii. Merge externally forms {C, {there, {T, {{three, men}, {$_{\varepsilon}$ √be, {$t_{\text{three men}}$, {in, {the, room}}}}}}}}

iv. **The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C** (partial feature inheritance).

v. **The uPhi-set continues to search for *three men* and √be: √be inherits the full uPhi-set from C** (full feature inheritance).

vi. PM by SM internally forms {√be, {√be, T}}: such a T becomes invisible.

---

[1] Richards and Biberauer (2005) and Richards (2006) argue that the expletive *there* generally occurs at SPEC-v (see also Alexiadou and Schaefer (2011)). However, I adopt the standard assumption that the expletive *there* externally merges to SPEC-T. For another possibility, see Nakajima (1997).

vii. **Labeling and feature valuation take place by minimal search: γ, δ and ε are labeled <Phi, Phi>, <Phi, Phi> and √P respectively**.

viii. The complement of C gets transferred.

In (27vii), the δ is labeled as <Phi, Phi> due to the feature inheritance implemented by minimal search in (27v). Accordingly, the XP-YP problem dissolves. I assume that *there* is not an X^0-level object but rather an XP-level object which is perhaps {n, pro} or {d, pro} (see Uriagereka (1988); Epstein, Kitahara, Seely (2014a)). Additionally, I assume that n or d in *there* bears the defective phi-set which contains [3rd-person] only (see Richards and Biberauer (2005) and Richards (2007)).[2] Therefore, the visible √be in {√be, {√be, T}} bears the full uPhi-set but defectively agrees with *there* under the proposal in (14), which is Labeling = Feature Valuation = Minimal Search. One might wonder why the defective agreement does not cause the derivation to crash at the CI interface because the defective agreement leaves the other uPhi-features on the √be (i.e. [u-number] and [u-gender]) unvalued.[3] All of the remaining uPhi-features, however, are valued on the lower copy of √be. I assume that the derivation does not crash if the valuation of one copy is completed regardless of whether it is the lower copy or the higher copy (cf. Epstein, Kitahara and Seely (2017) and Kitahara (2017)). This is quite plausible because copies are parts of a *discontinuous object* or a *chain* (i.e. sequence) of occurrences. That is, the higher copy of √be is the same as the lower copy of √be. This essential identity is simply ensured by internal Merge in the current minimalist framework. As long as {√be, {√be, T}} is formed by internal Merge, the complete-phi-valuation of √be is sufficient for one copy. For more on this topic, see Epstein, Kitahara and Seely (2017) and Kitahara (2017).

It is logically possible that PM by SM forms {T, {T, √be}}, the reverse derivational ordered pair, to the extent that we assume freely applying Merge. In this case, we do

[2] Chomsky (2000, 2001, 2004) argues that the *there* expletive is a simple head (i.e. X^0), which bears the uninterpretable-person feature only.

[3] The defective agreement might leave all three Phi-features (i.e. [u-person], [u-number] and [u-gender]) unvalued as discussed in Chomsky (2001), although I do not select this option.

not need to wonder if the phi-valuation of √be is sufficient for one copy, as it is invisible. Instead, 3rd-person on *there* and u-person on T, under the proposal in (13), share the person feature. Hence, the label of γ is <Phi, Phi> or <Person, Person>. This pattern predicts that the default agreement is the instantiation of {T, {T, √be}} (or <T, √be >). The prediction seems right because we observe the default agreement effects in the list *there* construction as illustrated below:[4]

(28) Q: Who is still here (to do the work)?

a. A: There {is/*am} only me.

b. A: There {remains/*remain} only me.

c. A: There {is/are} only {us/John and Bill}.

(adapted from Chomsky (2000: 149, fn. 90))

(29) And there's two components in [Division H], which is the operations division: the people that do the flight activity planning procedures work, provide for the crew activity planning and the time line support and integrated procedures development and overall flight data file management; <u>*and then there is the payload support folks*</u>, who provide for customer operations integration and support of their onboard interfaces.

(Ward and Birner (1995: 734))

These facts support the partial feature inheritance illustrated in (27iv) and the free creation of reverse order {T, {T, √be}} by PM by SM.

---

[4] In list reading, *be* verbs are often contracted as shown below:

(i) There's only me.

(ii) There's the payload support folks.

I do not enter into the discussion of the contraction, but the reverse order of the derivational ordered pair might involve this phenomenon. Also, note that the definiteness restriction disappears in list reading. See section 4.6 for this restriction.

4.4.3 Unaccusative *There* Constructions

Even if unaccusative verbs such as *appear* occur in *there* constructions, the XP-YP problem dissolves under the proposals. According to Epstein, Kitahara and Seely (2016), external pair-Merge of heads freely applies in any order. Recall that they argue that this type of application of pair-merge occurs in passive/unaccusative structures (see chapter 2). Adapting their analysis, I argued in chapter 2 that PM by SM externally forms {√root, {√root, v*}} in passive/unaccusative constructions. Following the analysis of the external application of PM by SM and assuming abstract √do in the lexicon, the structure and derivation of *there appeared a ship on the horizon* is as follows:

(30) {C, {$_{\gamma}$ there, {{√do, {√do, T}}, {{√appear, {√appear, v*}}, {$_{\delta}$ {a, ship}, {$_{\varepsilon}$ t$_{\surd do}$, {t$_{a\ ship}$, {on, {the, horizon}}}}}}}}}

i. PM by SM externally forms {√appear, {√appear, v*}}.

ii. Merge externally/internally forms {C, {$_{\gamma}$ there, {T, {{√appear, {√appear, v*}}, {$_{\delta}$ {a, ship}, {$_{\varepsilon}$ √do, {t$_{a\ ship}$, {on, {the, horizon}}}}}}}}}.

iii. **The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C** (partial feature inheritance).

iv. **The uPhi-set continues to search for *a ship* and √do: √do inherits the full uPhi-set from C** (full feature inheritance).

v. PM by SM internally forms {√do, {√do, T}}: such a T becomes invisible to minimal search.

vi. **Labeling and feature valuation take place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and √P, respectively**.

vii. The complement of C gets transferred.

Again, the XP-YP problem dissolves thanks to (30iii, iv). Note that how √do achieves a phonological realization is a matter of externalization. I do not discuss it in detail because this is beyond the scope of this paper. However, it is plausible to assume that

√do merges with T because √do overtly appears at T when the sentence is emphasized or negativized.

(31) a. We DO/DID want to help you.
b. We do/did not want to help you.

What meaning √do has is a problem if we assume √do in the lexicon, as heads (e.g. Agr) cannot exit only for the purely syntactic operations, following the SMT. Notice that √do is unambiguously necessary for the SM system, but its existence is not essential, considering the primacy of the CI system (i.e. Language of Thought). Once we, however, assume √do can be not only the root of auxiliary *do* but also that of main verb *do*, the problem dissolves. The root of main verb *do* is undoubtedly √do, and the verb *do* probably is comprised of {√do, v*} or <√do, v*>. Undoubtedly, everyone agrees that the verb *do* has some meaning. Thus, the true problem is what meaning auxiliary *do* has, and this is clearly beyond the scope of this paper.

In unaccusative *there* constructions, it is also logically possible that PM by SM forms the reverse order of {√do, {√do, T}}, which is {T, {T, √do}}. Again, this possibility is supported by the list reading:

(32) Q: Who is still here (to do the work)?
A: There remains/*remain only me.
(=(28b))

Section 4.4, under the proposals in section 4.3, has provided the new analyses of two types of *there* constructions: the *there be* construction and the unaccusative *there* construction. In the next section, I will show that the proposals and the analyses further explain peripheral *there* constructions.

4.5. "Unaccusativized" *There* Constructions

4.5.1 Unaccusativized Unergatives Verbs

Given the unaccusativization of unergatives and the basic analysis in unaccusative *there* constructions in section 4.4.3, I propose the following structure and derivation of an unergative *there* construction, *there walked into the room three men*:[5]

(33) {C, {$_{\gamma}$ there, {<√do, T>, < {<<√walk, into the room>, v*>, {$_{\delta}$ ~~{three, men}~~, {$_{\varepsilon}$ $t_{\surd\text{do}}$, $t_{\text{three men}}$}}}, {three, men} >}}}

i. PM by SM externally forms <into, <room, the>>.

ii. PM by SM externally forms <<√walk, into the room>, v*> (i.e. {{√walk, {√walk, into the room}}, {{√walk, {√walk, into the room}}, v*}}).

iii. Merge forms {<<√walk, into the room>, v*>, {$_{\delta}$ {three, men}, {$_{\varepsilon}$ √do, $t_{\text{three men}}$}}}.

iv. PM by SM internally forms < {<<√walk, into the room>, v*>, {$_{\delta}$ ~~{three, men}~~, {$_{\varepsilon}$ √do, $t_{\text{three men}}$}}}, {three, men} >.

v. Merge forms {C, {$_{\gamma}$ there, {T, < {<<√walk, into the room>, v*>, {$_{\delta}$ ~~{three, men}~~, {$_{\varepsilon}$ √do, $t_{\text{three men}}$}}}, {three, men} >}}}.

vi. **The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C** (partial feature inheritance).

vii. **The uPhi-set continues to search for *three men* and √do: √do inherits the full uPhi-set from C** (full feature fnheritance).

viii. PM by SM internally forms <√do, T> (i.e. {√do, {√do, T}}): such a T becomes invisible.

ix. **Labeling and feature valuation take place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and √P, respectively**.

x. The complement of C gets transferred.

---

[5] The angle brackets for *into the room* are omitted for expository purposes; *into the rooms* is assumed to be a complex head formed by PM by SM. Angle brackets for the representational ordered pair are adopted for making the notation simpler. In addition, the strike-through elements indicate that they are lower copies formed by internal PM by SM of phrases.

First, the unergative verb *walk* (*into the room*) is unaccusativized by external PM by SM of √walk and v*.[6] The unaccusativization in *there* constructions is supported by the following empirical facts (cf. Omune (2016)).

(34) a. There walked into the room a fierce-looking tomcat.
b. A fierce-looking tomcat came into the room by walking.
c. There ambled into the room a frog.
d. A frog came into the room by ambling.
e. Suddenly, there ran out of the bushes a grizzly bear.
f. A grizzly bear came out of the bushes by running.

(a, c: Milsark (1974: 155, 246))
(e: Lumsden (1988: 38))

---

[6] I put aside the problem of whether Merge, PM by SM or the combination of both forms *into the room*. Each option is logically possible under the conception of freely applying Merge. In addition, note that a locative expression such as *into the room* is often part of the unaccusativized verb in the unaccusativized unergative *there* construction, but it is not necessary.

(i) * There walked a well-known actor through passport control.
(ii) Into the courtroom there walked two people I had thought were dead.

(i: Lumsden (1988: 38))
(ii: Kuno and Tkami (2004: 45))

In the latter case, external or internal PM by SM seems to apply to the phrase *into the room*. Then, the phrase becomes part of the derivational ordered pair of the matrix clause. Ultimately, it largely depends on the CI interpretation whether linguistic expressions are interpretable. Thus, unaccusataivized interpretation is allowed in the *there* construction as long as the resulting expressions satisfy the suitable condition. I do not enter into what condition it is because it is a matter of the CI semantics. Whatever this condition is, the examples above, to some extent, support freely applying Merge. Essentially, any expression can be generated via Merge, regardless of whether it is deviant or not. Merge can therefore yield the expression of (i), but the CI system regards it as deviant. In contrast, the CI system regards the expression of (i) formed through Merge as natural under the relevant interpretive condition.

Sentences in (34b, d, f) are roughly paraphrased versions of those in (34a, c, e) respectively. The theta-role of all subjects in the paraphrased sentences are the theme assigned by the main verb *come*. The semantic flavor, such as the agent, is assigned by the verbs in the *by*-phrases (i.e. *by walking*, *by ambling* and *by running*). The unergative verbs thus simply "modify" the events in unergative *there* constructions. Crucially, the leading event is expressed via the unaccusative structure. These facts therefore support the analysis of the unaccusativization of unergative verbs in (33).[7] The √walk just modifies the event of v* in <<√walk, into the room>, v*>, and the entire structural configuration, including <<√walk, into the room>, v*>, denotes the main unaccusative event (for more on this topic, see Acedo-Matellán (2010) and Acedo-Matellán and Mateu (2014), among others).

Second, assuming that the abstract √do occurs in a relevant structure, labeling δ takes place successfully. In the proposed unaccusativized structure, √do can be, by hypothesis, introduced in the relevant course of the derivation. Minimal search for labeling therefore labels δ as <Phi, Phi> through the relation Agree (i.e. feature sharing) for the same reason as the typical unaccusative case. Eventually, labeling takes place

---

[7] Unergative verbs such as *frown*, *smile*, *breathe* do not occur in *there* constructions because they are not verbs of existence/appearance (cf. Levin (1993)):

(i) * There frowned/smiled/breathed into the room a stranger I met yesterday.
(ii) * Into the room there frowned/smiled/breathed a stranger I met yesterday.

Interestingly, these verbs are unacceptable in the *by*-phrase sentences, but acceptable without *by*:

(i) * A stranger came into the room by frowning/smiling/breathing.
(ii) A stranger came into the room, frowning/smiling/breathing.

As an informant points out, the former sentence is unacceptable because the *by*-phrase modifies the way a stranger comes into the room, but the meaning of *frown*/*smile*/*breathe* is not directly related to the verbal event of *come*. In contrast, the latter sentence is acceptable because *frowning*/*smiling*/*breathing* just describes the behavior of a stranger.

successfully in the unergative *there* construction under the proposed reformulation of feature inheritance and feature valuation.

Third, PM by SM of *phrases* applies at the end of the cycle in the cancelled-v*P phase, and the associate adjoins to δ, which is informally called the right edge of vP. I adopt this landing site, which is proposed in Chomsky's (2001) Th/Ex (Thematizaion/Extraction), but do not adopt Th/Ex itself because it is not only conceptually dubious but also empirically problematic. In the former, as Chomsky (2016a) recently mentions, Spell-Out does not exist in the current minimalist framework (see also Obata (2010)). Th/Ex is therefore untenable simply because it is an operation of the phonological component and applies at Spell-Out within the weak phase. In the latter, because it is an operation of the phonological component, Th/Ex (i.e. rightward movement) does not change meanings.[8] However, it apparently causes surface semantic effects.

(35) *be* verbs
 a. *There is the man in the room.
 unaccusatives
 b. *There appeared all ships on the horizon.
 c. *There arrived the man in the room.

In *there* constructions, a *strong* determiner (including a null determiner) cannot occur as an associate's partner. Thus, as the *strong* reading, we cannot use definites (e.g. *the*, demonstratives, personal pronouns and possessives) and universals (e.g. *all* and *every*) as a determiner of an associate nP.[9] We call this phenomenon the definiteness restriction (regardless of the use of actual definites (cf. Milsark (1974, 1977)). The definiteness

8 Th/Ex is not restricted to the *rightward* movement. It also moves the associate *leftward*, for example, in passives (see Chomsky (2001)). However, I focus on the rightward case for expository purposes.

9 Recall that the definiteness restriction has nothing to do with the list *there* construction (see (32)).

restriction is, however, relaxed when the associates are extraposed rightward, regardless of types of verbs.

(36) *unaccusatives*

a. There came to his mind her beautiful and intelligent face.

unergatives

b. There {ambled/hopped} into the room my neighbor's frog.

transitives

c. There entered the room the new professor from Sweden.

(a: Quirk et al. (1985: 1409))

(b: adapted from Milsark (1974: 246))

(c: Julien (2002:13))

We cannot regard Th/Ex as an operation of the phonological component. The rightward movement of Th/Ex clearly causes the surface semantic effect.[10] Notice that Chomsky (2001) himself argues that Th/Ex applies when an argument structure includes v, a light verb marking unaccusativity or passive voice. If so, we can assume that something like Th/Ex may apply in unergative *there* constructions because unergatives are unaccusativized in *there* constructions (see (34) and the relevant discussions above). In the structure in (33), internal PM by SM of phrases replaces Th/Ex. The replacement by PM by SM is reasonable because PM by SM is an operation in the narrow syntax, and therefore it can cause surface semantic effects. Accordingly, something like focus interpretations in (36) is caused by the structural configuration yielded by internal PM by SM. Furthermore, internal PM by SM obtains the surface word order such as in (34a, c,

[10] Chomsky, Gallogo and Ott (to appear) state that the notion of *meaning in the broad sense* should be captured by the holistic interpretation of <SEM, PHON> pairs. That is, SEM provides *meaning in the narrow sense*. If we follow their assumption, discourse-related interpretations such as Focus and Topic may be relegated to PHON. As a result, Th/Ex can change meanings in the broad sense. I do not pursue this possibility in this paper, but I believe that it is worth exploring.

e) and (36b, c). The analysis based on PM by SM is thus better than analyses based on Th/Ex or phonological movement.

For the case of the transitive *there* construction as in (36), we will see relevant analyses in the next section.

4.5.2 Unaccusativized Transitive Verbs

One of the marginal constructions in English is the TEC (transitive expletive construction). I will show in this section that the TEC is also explained by the similar analysis of the structure and derivation of unergative *there* constructions.

Given the basic analysis of unaccusative/unergative *there* constructions, I propose the following structure and derivation of a transitive *there* construction, *there entered the room three men from Sweden*:

(37) {C, {$_{\gamma}$ there, {<√do, T>, <{<<√enter, the room>, v*>, {$_{\delta}$ ~~three men from Sweden~~, {$_{\varepsilon}$ $t_{\sqrt{do}}$, $t_{\text{three men from Sweden}}$}}}, three men from Sweden>}}}

i. PM by SM externally forms <room, the>.

ii. PM by SM externally forms <<√enter, the room>, v*> (i.e. {{√enter, {√enter, the room}}, {{√enter, {√enter, the room}}, v*}}).

iii. Merge forms {<<√enter, the room>, v*>, {$_{\delta}$ three men from Sweden, {$_{\varepsilon}$ √do, $t_{\text{three men from Sweden}}$}}}.

iv. PM by SM internally forms <{<<√enter, the room>, v*>, {$_{\delta}$ ~~three men from Sweden~~, {$_{\varepsilon}$√do, $t_{\text{three men from Sweden}}$}}}, three men from Sweden>.

v. Merge forms {C, {$_{\gamma}$ there, {T, <{<<√enter, the room>, v*>, {$_{\delta}$ ~~three men from Sweden~~, {$_{\varepsilon}$√do, $t_{\text{three men from Sweden}}$}}}, three men from Sweden>}}}.

vi. **The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C** (partial feature inheritance).

vii. **The uPhi-set continues to search for *three men* and √do: √do inherits the full uPhi-set from C** (full feature inheritance).

viii. PM by SM internally forms <√do, T> (i.e. {√do, {√do, T}}): such a T becomes invisible.

ix. **Labeling and feature valuation take place by minimal search: γ, δ and ε are labeled <Phi, Phi>, <Phi, Phi> and √P respectively**.

x. The complement of C gets transferred.

As already shown in the cases of unaccusativized unergative *there* constructions, I assume that the verbal root in the TEC gets unaccusativized, too. Because roots are generally assumed to be universally underspecified as to categories, there is virtually no substantial element like transitive, unergative, and unaccusative roots (cf. Marantz (1997); Halle and Marantz (1993)). That is, there is no reason to preclude external PM by SM of v* and a root unless stipulated. The unaccusativization of a transitive verb is, therefore, logically available in the TEC. As seen above, unaccusativization of unergative verbs was in effect supported by the empirical facts in section 4.5.1. Unaccusativization of transitive verbs is also empirically supported by facts of Japanese.

In a Hiroshima dialect of Japanese, we can observe the phenomena of unaccusativization of transitive verbs. In Standard Japanese and many other Japanese dialects, the verb *noku* can be both intransitive (*noku*) and transitive (*nokeru*) as shown below.

(38) Many Dialects in Japanese:

a. Taro, soko noite/*nokete.
Taro, there √move-I-TE/√move-E-TE
'Taro, get out of the way.'

b. Taro, sore *noite/nokete.
Taro, it √move-I-TE/√move-E-TE
'Taro, get it out of the way'

The examples show that *nokete* cannot be interpreted as the intransitive meaning. On the contrary, *nokete* can be interpreted as the intransitive in a Hiroshima dialect.[11]

(39) A Hiroshima Dialect:

a. Taro, soko noite.
   Taro, there √move-I-TE
   'Taro, get out of the way.'

b. Taro, sore *noite
   Taro, it √move-I-TE
   'Taro, get it out of the way'

c. Taro, sore nokete.
   Taro, it √move-E-TE
   'Taro, get it out of the way'

d. Taro, soko nokete.
   Taro, there √move-E-TE
   'Taro, you disappear from here.'

The *nokete* in (39d) could be the instantiation of unaccusativization of the transitive verb *nokeru*. This analysis is further support by the following facts:

(40) Many Dialects in Japanese:

a. nande wazato noitan (*intransitive*)
   why intentionally √move-I-PAST-Q
   'Why did you intentionally get out of the way?'

c. nande wazato noketan (*transitive*)
   why intentionally √move-E-PAST-Q
   'Why did you intentionally get it out of the way?'

---

[11] In Japanese (including the Hiroshima dialect), there are verbs *doku* and *dokeru* which have the same meaning as *doku* and *dokeru*. Interestingly, *doku* and *dokeru* also show the same distributions, which are shown in (38)–(41), as *noku* and *nokeru*.

(41) A Hiroshima Dialect:

a. nande wazato noitan (*intransitive*)
why intentionally √move-I-PAST-Q
'Why did you intentionally get out of the way?'

b. *nande wazato noketan (*intransitive*)
why intentionally √move-E-PAST-Q
'Why did you intentionally disappear from here?'

c. nande wazato noketan (*transitive*)
why intentionally √move-E-PAST-Q
'Why did you intentionally get it out of the way?'

Generally, unergative verbs are volitional, but accusative verbs are not. Thus, *wazato* (intentionally) can modify the unergative verb *noita* (got out of the way) as shown in (40)–(41). Because transitive verbs can be volitional, too, *wazato* (intentionally) can modify *noketa* (got it out of the way) as in (40)–(41). However, in (41b), *wazato* (intentionally) cannot modify the intransitive verb *noketa* (got out of the way) in the Hiroshima dialect. This fact suggests that the intransitive verb *nokeru* in the Hiroshima dialect is not unergative but unaccusative. Therefore, these Japanese facts imply that unaccusativization of transitive verbs are empirically possible. If this analysis is on the right track, it is not dubious that PM by SM externally forms {√nok, {√nok, v*}}/<√nok, v*>.

In (37), it is logically plausible that external PM by SM of √enter and *the room* is possible as long as *the room* is an X^0-level object. That is, PM by SM yields <√enter, the room>. We will also discuss later the empirical advantage of the creation of an ordered pair <<√root, n>, p> or <√root, n>.[12] In addition, it is assumed that the abstract null √do occurs in the structure in (37). If it is on track, then labeling δ successfully takes place in transitive *there* constructions too. Internal PM by SM of phrases applies

[12] In this notation of representational ordered pairs, the head d is omitted.

at the end of the cycle of the unaccusativized v*P phase in the proposed structure and derivation.

4.6. Explaining Further Empirical Facts

4.6.1 Non-(sub)extractability of Associates

The following data show non-(sub)extractability of associates in unaccusativized *there* constructions (i.e. unergative/transitive *there* constructions).

(42) *be verbs*

a. Who was there t_{wh} in the kitchen?

b. You remember a guy who there was t_{wh} for you.

(43) *unaccusatives (IVES)*

a. ? I saw a ship which there appeared t_{wh} on the horizon.

b. ? I remember several new facts which there emerged t_{wh} at the meeting.

(44) *unergatives/transitives (OVES)*

a. * I saw a fierce-looking tomcat which there walked into the room t_{wh}.

b. * I saw a frog which there ambled into the room t_{wh}.

c. * I saw a strange man which there entered the room t_{wh}.

The data are examples of *wh*-extraction in the *there be* construction, the IVES (Inside Verbal Existential Sentence) and the OVES (Outside Verbal Existential Sentence).[13] In the first example, *wh*-associates are extractable. In the second example, though the resultant sentences are not completely acceptable, the *wh*-associates are extractable. However, *wh*-associates are non-extractable in the last example because internal PM by SM of phrases applies to associates in the OVES. After this application of PM by SM, associates become invisible, being part of the derivational and representational ordered pairs. As a result, the associates cannot further internally merge.

---

[13] I follow the classification of *there* constructions in Milsark (1974). The words *Inside Verbal* and *Outside Verbal* indicate that the final landing site of associates is v*P internal or external, respectively.

We have one theory-internal concern about the timing and the domain of the application of labeling and valuation by minimal search. Considering the facts in (42)–(44), minimal search should apply first *in the cancelled v*P phase* after the next CP phase has been formed. The steps of the derivation for the IVES are shown below (the IVES examples below are from Levin (1993: 89)):

(45) There appeared a ship on the horizon.

i. Merge forms {$_{\iota}$ C, {$_{\gamma}$ there, {T, {$_{\kappa}$ {√appear, {√appear, v*}}, {$_{\delta}$ {a, ship}, {$_{\varepsilon}$ √do, {$t_{\text{a ship}}$, {on, {the, horizon}}}}}}}}}.

ii. Feature inheritance by minimal search takes place.

iii. PM by SM internally forms {√do, {√do, T}}.

iv. Labeling and feature valuation take place **in κ**: κ, δ and ε are labeled as v*P, <Phi, Phi> and √P, respectively.

v. Labeling and feature valuation take place **in ι**: ι and γ are labeled as CP and <Phi, Phi>, respectively.

vi. The complement of C gets transferred.

The *first* application of labeling and valuation begins within the smallest proposition κ (see (45iv)), regardless of the cancellation of the v*P phase. The *second* application of labeling and valuation finally covers the remaining area including the CP phase (see (45v)). This assumption is crucial because we cannot explain the facts in (43) without it. Suppose labeling and valuation start from the topmost area, ι in (45). Then, we cannot extract the associate because it is transferred immediately after the relevant operations have applied.

The analysis further suggests that structures of *there be* constructions also contain v* whose phase-hood has been cancelled by external PM by SM. Accordingly, the structure and derivation for the *there be* construction is revised as follows (cf. (27)):

(46) There are three men in the room.

{$_\iota$ C, {$_\gamma$ there, {{{√be, {√be, v*}}, {{√be, {√be, v*}}, T}}, {$_\delta$ {three, men}, {$_\varepsilon$ $t_{\{\surd\text{be}, \{\surd\text{be}, \text{v*}\}\}}$}, {$t_{\text{three men}}$, {in, {the, room}}}}}}}}

i. **PM by SM externally forms {√be, {√be, v*}}, which is representationally <√be, v*>**.

ii. Merge externally forms {{√be, {√be, v*}}, {{three, men}, {in, {the, room}}}}.

iii. Merge internally forms {{three, men}, {{√be, {√be, v*}}, {$t_{\text{three men}}$, {in, {the, room}}}}}.

iv. Merge externally forms {C, {there, {T, {{three, men}, {$_\varepsilon$ {√be, {√be, v*}}, {$t_{\text{three men}}$, {in, {the, room}}}}}}}}.

v. The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C (partial feature inheritance).

vi. The uPhi-set continues to search for *three men* and {√be, {√be, v*}: {√be, {√be, v*} inherits the full uPhi-set from C (full feature inheritance).

vii. **PM by SM internally forms {{√be, {√be, v*}}, {{√be, {√be, v*}}, T}}, which is representationally <<√be, v*>, T>**.

viii. Labeling and feature valuation take place in δ: δ and ε are labeled as <Phi, Phi> and √P respectively.

ix. Labeling and feature valuation take place in ι: ι and γ are labeled CP and <Phi, Phi>, respectively.

x. The complement of C gets transferred.

Given the revised version, the extraction in (42) is explained. In (46), labeling and feature valuation start to apply in ε, the cancelled v*P phase. The associates are extractable because of this analysis. In other words, they were not extractable if the relevant operations started to apply to ι and ε all at once. Recall that the complement of a phase head is transferred immediately after the application of labeling and valuation, which take place at the timing of the Transfer operation.

It is obscure why the relevant operations apply first in the cancelled v*P phase since the phase-hood has been cancelled. Crucially, the substantial status of v* is still visible though it is invisible with respect to uPhi. As argued elsewhere, it is uPhi that triggers Transfer. Therefore, it is plausible that the invisible uPhi cancels Transfer. In contrast, it seems that minimal search can see the other substantial features on v* (e.g. the categorical feature) since <√root, v*> is able to be labeled. When minimal search applies, its staring point is determined by the substantial features that mark the verbal event yielding the smallest proposition, the argument structure. Thus, v* is essential even for the *there* be construction. If this reasoning is on track, the timing and the domain of the application of labeling and valuation by minimal search is deduced. Furthermore, they are induced by the data concerning extraction of associates in (42)–(44). The deduction and the induction lead us to the following conclusion: Minimal search for labeling and valuation applies *phase by phase* even if phase cancellation has made the uPhi invisible. Namely, phase cancellation, in effect, does not cancel phases but cancels Transfer.

As shown below, subextraction from associates shows a pattern similar to the cases of extraction in (42)–(44).

(47) *be verbs*

a. Of which artist was there a portrait t_{wh} on the wall?

b. How many ancestors were there a portrait/portraits of t_{wh} on the wall?

(48) *unaccusatives (IVES)*

a. Of which artist did there hang a portrait t_{wh} on the wall?

b. How many ancestors did there hang a portrait/portraits of t_{wh} on the wall?

(a, b: adapted from Nishihara (1999: 392, fn. 10))

(49) *unergatives/transitives (OVES)*

a. *Which community did there walk into the room a member of t_{wh}?

b. *Of which community did there walk into the room a member t_{wh}?

c. *Which community did there enter the room a member of t_{wh}?

d. *Of which community did there enter the room a member t_{wh}?

(a, b: adapted from Nishihara (1999: 394))

These facts are examples of *wh*-subextraction in the *there be* construction, the IVES and the OVES. The facts indirectly support the analysis of extraction of associates and the timing and the domain of minimal search for labeling and valuation shown in (46) and the relevant discussions.

However, the facts of subextraction in (47)–(49) do not necessarily induce the timing and the domain of minimal search for labeling and valuation because an n head in *a portrait of which*/*how many ancestors* is what the minimal search needs. In other words, the wh-phrases do not take part in labeling and valuation with uPhi inherited from C. Unless agreeing with uF, objects can escape from Transfer by recurring internal Merge. First, *which*/*how many ancestors*, for example, internally merges before the application of labeling and valuation. Second, it merges with {C_Q, {Expl, {T, …}}}. Finally, labeling and valuation take place successfully even if the search starts from the topmost-CP-phase area since there remains an n head in *a portrait of* for agreeing with uPhi inherited from C. Therefore, the subextraction phenomena do not induce the timing and domain of labeling and valuation. If we assume the almost same derivation as (45) and (46), the facts in (47) and (48), nevertheless, support the analysis of subextraction of associates.

More crucially, it is also supported that internal PM by SM of phrases does not have to apply to associates of *there be* constructions and the IVES, but it must apply to those of the OVES. After the application of internal PM by SM of phrases, parts of associates become non-extractable because the associates have been invisible; hence (49).

In sum, associates must stay at SPEC-δ for labeling/valuation. In the *there be* construction and the IVES, associates can further internally merge because Transfer does not apply in the cancelled v*P phase when minimal search for labeling and valuation takes place in that cancelled phase. In the OVES, associates, in contrast, must experience internal PM by SM of phrases for correct interpretation. After the application of the PM by SM, associates become invisible to (sub)extraction in the narrow syntax; hence the structure and derivation of (30), (33), (37), (45), (46).

4.6.2 Scopal Phenomena and Binding Effects

Theoretically speaking, there is no LF covert movement on purely conceptual grounds in the minimalist framework because covert movement is counter-cyclic and too complex. Chomsky (1995b: 254) stated:

> "The computational system C_{HL} is based on two operations, Merge and Move. We have assumed further that Merge always applies in the simplest possible form: at the root. What about Move? The simplest case again is application at the root: if the derivation has reached the stage Σ, then Move selects α and targets Σ, forming {γ, {α, Σ}}. But covert movement typically embeds α, and therefore takes a more complex form: given Σ, select K within Σ and raise α to target K, forming {γ, {α, K}}, which substitutes for K in Σ."

As indicated by Epstein, Kitahara and Seely (2013), such substitution is counter-cyclic. Hence, the violation of the binary requirement of minimal search for Merge, which is one of third-factor principles, occurs (see (15)). Formally, Epstein, Kitahara and Seely (2013: 80) represented this counter-cyclic operation as follows:

(50) Applied to α and β within Σ, where α is a term of β, Merge
  i. takes α and β, forming {α, β}, and
  ii. takes H(β) and {α, β}, forming {H(β), {α, β}}, and
  iii. replaces β in Σ by {H(β), {α, β}}.

The third step is the replacement by counter-cyclic internal Merge which should be eliminated for conceptual reasons. Additionally, Chomsky (2007: 16) mentioned:

> “Merge yields compositional/cyclic properties of the kind that have repeatedly been found. Optimally, there should be only a single cycle of operations. EST postulated five separate cycles: X-bar theory projecting D-structure, overt operations yielding S-structure, covert operations yielding LF, and compositional mappings to the SM and CI interfaces. With the elimination of D- and S-structure, what remains are three cycles: the narrow-syntactic operation Merge (now with overt and covert operations intermingled), and the mappings to the interfaces. As noted earlier, optimal computation requires some version of strict cyclicity. That will follow if at certain stages of generation by repeated Merge, the syntactic object constructed is sent to the two interfaces by an operation Transfer, and what has been transferred is no longer accessible to later mappings to the interfaces (the phase-impenetrability condition PIC).”

Therefore, there is no LF covert movement in the technical sense. Furthermore, empirical facts also support this view in the labeling theory. Epstein, Kitahara and Seely (2015: 236, fn. 6) point out that covert movement could overgenerate the following sentence (see also Chomsky (2013, 2015a) for the relevant data and analysis):

(51) *They thought [in which Texas city [C [TP the man was assassinated t_{wh}]]]?
(Epstein, Kitahara and Seely (2015: 225))

If there were LF covert movement, *in which Texas city* could raise to SPEC-C in LF. Under labeling by minimal search (Chomsky (2013, 2015a)), this sentence is correctly excluded by labeling failure. The uQ on *in which Texas city* (see Epstein, Kitahara and Seely (2017), Kitahara (2017)) does not agree with C because C does not bear Q.

Further empirical facts concerning the expletive (Expl) also support no covert movement. Expl has been assumed to be a meaningless element at LF or the CI interface. On one hand, in Chomsky (1986: 179), Expl is deleted and is substituted for an associate at LF. On the other hand, an associate adjoins to Expl at LF in Chomsky (1995b: 155). Hartmann (2008) argues against these analyses due to the following counter-examples:

(52) a. Someone must be in his house.

i. someone > must   ii. must > someone

b. There must be someone in his house.

i. *someone > must   ii. must > someone

(a, b: adapted from Williams (1984: 152))

(53) a. Many women aren't sick.

i. many > not   ii. not > many

b. There aren't many women sick.

i. *many > not   ii. not > many

(54) a. Many ships seem to be in the harbor.

i. many > seem   ii. seem > many

b. There seem to be many ships in the harbor.

i. *many > seem   ii. seem > many

(a, b: adapted from Hartmann (2008: 19))

As shown above, the associates in the *there* construction must have narrow scopes. In other words, the associates are expected to have both scope options (i, ii) available if the domain of scope is determined by LF covert movement. However, only one option (i.e. (ii)) is available in all the three cases in (52b), (53b) and (54b). Thus, these facts show that associates do not substitute or adjoin to Expl in the *there* construction.

The data about negative polarity items and anaphors on *there* constructions also lead us to the same conclusion:

(55) a. *There seems to any European team to be no NBA team beatable.

b. No NBA team seems to any European team to be beatable.

(a, b: Bošković (1997: 98))

(56) a. *There seems to himself to be someone in the garden.

b. Someone seems to himself to be in the garden.

(a, b: Bošković (1997: 98))

The sentence in (55a) is unacceptable; this leads us to predict that the negative polarity item *any* is not asymmetrically c-commanded by *no* (I use the technical term *c-command* only for expository purposes). This fact is not explained by the associate-substituting analysis or associate-adjoining analysis because the representation of (55a) at LF (or the CI) should be almost identical to that of (55b). Similarly, (56a) should not cause any problem if associates were substituted or adjoined to Expl, since each analysis makes the representation of (56a) at LF virtually identical to that of (56b) at LF (or the CI). Eventually, the facts about negative polarity items and anaphors above all indicate that LF-substitution and LF-adjunction of associates are not tenable. Therefore, associates are interpreted (at least) at the originating position, feature-sharing position (i.e. SPEC-δ in (30), (33), (37) and (46)) and/or extraposed positon (i.e. the so called right periphery of v*P) in the *there* constructions.

It is rational to also expect associates in OVES to be interpreted at the originating position, the feature-sharing position and/or at the extraposed positon. In fact, this prediction is supported by the following facts:

(57) a. There walked into the classroom three students from my department.
b. There walked into the classroom no one from my department.
c. *There walked into any classroom no one from my department.
(c: Deal (2009: 27, fn. 48))

(58) a. There entered the classroom three students from my department.
b. There entered the classroom no one from my department.
c. *There entered any classroom no one from my department.

The sentences above are examples of the negative-polarity items in the OVES. These facts also show that associates in OVES do not asymmetrically c-command a locative expression, which is assumed to be a complex p head or n head. This supports the proposed structures and derivations in (33) and (37), which are repeated as (59)–(60), because p and n are invisible by being part of <<√root, n>, p> or <√root, n> formed by external PM by SM.

(59) {C, {$_{\gamma}$ there, {<√do, T>, < {<<√walk, into the room>, v*>, {$_{\delta}$ ~~{three, men}~~, {$_{\varepsilon}$ $t_{\sqrt{}\text{do}}$, $t_{\text{three men}}$}}}, {three, men} >}}}

i. PM by SM externally forms <into, <room, the>>.

ii. PM by SM externally forms <<√walk, into the room>, v*> (i.e. {{√walk, {√walk, into the room}}, {{√walk, {√walk, into the room}}, v*}}).

iii. Merge forms {<<√walk, into the room>, v*>, {$_{\delta}$ {three, men}, {$_{\varepsilon}$ √do, $t_{\text{three men}}$}}}.

iv. PM by SM internally forms < {<<√walk, into the room>, v*>, {$_{\delta}$ ~~{three, men}~~, {$_{\varepsilon}$ √do, $t_{\text{three men}}$}}}, {three, men} >.

v. Merge forms {C, {$_{\gamma}$ there, {T, < {<<√walk, into the room>, v*>, {$_{\delta}$ ~~{three, men}~~, {$_{\varepsilon}$ √do, $t_{\text{three men}}$}}}, {three, men} >}}}.

vi. **The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C** (partial feature inheritance).

vii. **The uPhi-set continues to search for *three men* and √do: √do inherits the full uPhi-set from C** (full feature inheritance).

viii. PM by SM internally forms <√do, T> (i.e. {√do, {√do, T}}): such a T becomes invisible.

ix. **Labeling and feature valuation take place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and √P, respectively**.

x. The complement of C gets transferred.

(60) {C, {$_{\gamma}$ there, {<√do, T>, <{<<√enter, the room>, v*>, {$_{\delta}$ ~~three men from Sweden~~, {$_{\varepsilon}$ $t_{\sqrt{}\text{do}}$, $t_{\text{three men from Sweden}}$}}}, three men from Sweden>}}}

i. PM by SM externally forms <room, the>.

ii. PM by SM externally forms <<√enter, the room>, v*> (i.e. {{√enter, {√enter, the room}}, {{√enter, {√enter, the room}}, v*}}).

iii. Merge forms {<<√enter, the room>, v*>, {$_\delta$ three men from Sweden, {$_\varepsilon$ √do, $t_{\text{three men from Sweden}}$}}}.

iv. PM by SM internally forms <{<<√enter, the room>, v*>, {$_\delta$ ~~three men from Sweden~~, {$_\varepsilon$√do, $t_{\text{three men from Sweden}}$}}}, three men from Sweden>.

v. Merge forms {C, {$_\gamma$ there, {T, <{<<√enter, the room>, v*>, {$_\delta$ ~~three men from Sweden~~, {$_\varepsilon$√do, $t_{\text{three men from Sweden}}$}}}, three men from Sweden>}}}.

vi. **The uPhi-set on C minimally searches for *there* and T: T inherits u-person from C** (partial feature inheritance).

vii. **The uPhi-set continues to search for *three men* and √do: √do inherits the full uPhi-set from C** (full feature inheritance).

viii. PM by SM internally forms <√do, T> (i.e. {√do, {√do, T}}): such a T becomes invisible.

ix. **Labeling and feature valuation take place by minimal search: γ, δ and ε are labeled <Phi, Phi>, <Phi, Phi> and √P, respectively**.

x. The complement of C gets transferred.

The associates are interpreted at the originating position, SPEC-δ and/or the extraposed positon. The extraposed positon is for the discourse-related interpretation, such as the relaxation of the definiteness restriction. Crucially, the locative expressions above do not c-command the associates at any position because they are part of complex heads.

4.7. Summary

In this chapter, I have reformulated feature inheritance as a minimal search by uPhi. In addition, feature valuation has been reformulated to be the same minimal search as labeling. Minimal search itself is part of Minimal Computation, a third factor principle, which belongs to the law of nature and is not specific to language. The reformulations are, therefore, conceptually ideal in order for eliminating and reducing language-specific operations from the core system of language, following the SMT. Furthermore, it has been shown that the reformulations explain derivations of various *there* constructions. Specifically, the reformulations explain not only familiar constructions—the *there be* construction and the unaccusative *there* construction—but also somewhat peripheral constructions, including the unergative *there* construction and the TEC. Assuming that the peripheral ones are unaccusativzed, I have shown that the analyses further explain the facts about (sub)extraction of the associate and scopal and binding phenomena.

Chapter 5

Alternatives

5.1 Introduction

In this chapter, I will explore alternative approaches to solving an XP-YP problem in *there* constructions and to explaining the (non-)extractability of the indirect object in double object constructions, based on the current framework of the Minimalist Program (Chomsky (2013, 2015a)).

T. Daniel Seely (p.c.), Epstein, Kitahara and Seely (2014b) and Kitahara (2017) argue that feature inheritance is an unnecessary operation that therefore should be eliminated. To eliminate feature inheritance, they capitalized on freely applying Merge. That is, they present a new type of application of external Merge of heads. In their approach, Merge (C, T_{uPhi}) and Merge (v*, $\sqrt{}root_{uPhi}$) apply externally. Then, this application of Merge establishes C-T and v*-√root relations. Importantly, the phase heads do not bear uPhi, whereas T and √root inherently bear uPhi in their approach (see Kitahara (2017)). Consequently, feature inheritance does not take place because the external application of Merge of heads ensures the C-T relation and the v*-√root relation, and both T and √root, which inherently bear uPhi, do not need to inherit uPhi from the phase heads. I will show that this approach solves the XP-YP problem in *there* constructions.

In addition, I will show that labeling (Chomsky (2013, 2015a)) and maximality (Rizzi (2015a, b, 2016)) explain the (non-)extractability of the indirect object in double object constructions. Rizzi (2015a, b, 2016) proposed the maximality principle, which makes the maximal objects with a given label unmovable or frozen in place. I propose that P_{HAVE}, which was proposed and supported by Harley (1995, 2002) and Harley and Jung (2015), can be a criterial head that yields a possible halting site; P_{HAVE} makes the indirect object unmovable. Namely, the indirect object and the head P_{HAVE} share the same criterial feature P_{HAVE} under labeling (see Chomsky (2013, 2015a)). Then, maximality (see Rizzi (2015a, b, 2016)) freezes the indirect object in place. It will also be shown that this analysis explains the non-sub-extractability of the indirect object.

The first alternative approach may be superior to the analyses in chapter 4 in that it eliminates feature inheritance and makes computational system much simpler. However, this alternative contains technical problems. I will point out this problem later, and argue that this alternative should be discarded. The second alternative covers various empirical phenomena as well as the proposed analyses in the earlier chapters. The alternative can explain even new facts that the earlier analyses cannot handle. For instance, Rizzi's (2015a, b, 2016) analysis based on the cartographic approach can explain non-sub-extractability of the indirect object in the double object construction. However, the main point of this chapter is to show the conceptual advantages of the analyses proposed in the earlier chapters. The alternative approach may explain more empirical facts, but it just points out the problems in the earlier analyses or in the current model. As Chomsky, Gallego and Ott (to appear) point out, the cartographic approach violates the Inclusiveness Condition (cf. note 5). Recall that the ultimate goal of the Minimalist Program is to understand C_{HL}, the faculty of language or the language capacity. Explaining empirical facts is important but not the primary concern.

5.2 Alternative 1 (for the XP-XP Problem)

5.2.1 Analytic Assumptions

In earlier versions of the minimalist framework such as Chomsky (2008), the structure and derivation of *John hit Mary* is as follows:

(1) [$_{CP}$ C [$_{TP}$ John T [$_{v*P}$ t_{John} hit-v* [$_{VP}$ Mary t_{hit} t_{Mary}]]]]

i. External Merge forms [$_{v*P}$ John v* [$_{VP}$ hit Mary]].

ii. The following computations of v*P phase simultaneously take place: feature inheritance, Probe-Goal Agree, Internal Merge/Move (triggered by EPP), V raising and Transfer. VP is transferred.

iii. [$_{v*P}$ John hit-v* [$_{VP}$ Mary t_{hit} t_{Mary}]] is left in the workspace.

iv. External Merge forms [$_{CP}$ C [$_{TP}$ T [$_{v*P}$ John hit-v* [$_{VP}$ Mary t_{hit} t_{Mary}]]]].

v. The following computations of CP phase simultaneously take place: feature inheritance, Probe-Goal Agree, internal Merge/Move (triggered by EPP) and Transfer. TP is transferred.

vi. [$_{CP}$ C [$_{TP}$ John T [$_{v*P}$ t_{John} hit-v* [$_{VP}$ Mary t_{hit} t_{Mary}]]]]

As argued in chapter 4, feature inheritance takes place unproblematically in this derivation. (i) External Merge forms the set relation [$_{CP}$ C [$_{TP}$ T [$_{v*P}$ EA v*…]]] (EA: external argument). (ii) T inherits features from C (feature inheritance) because external Merge has established the direct relation of C-T. In other words, there is no intervening object between C and T. (iii) Internal Merge forms [$_{CP}$ C [$_{TP}$ EA T [$_{v*P}$ t_{EA} v*…]]] (*t*: the lower copy). However, this third step is counter-cyclic, since internal Merge tucks EA into C-TP. This is clearly problematic in the strictly cyclic application of simplest Merge (Chomsky (2013, 2015a)). Simplest Merge just takes two elements, α, β, and forms the simplest unordered set {α, β}, regardless of whether the application of simplest Merge is external or internal. The strictly cyclic application of Merge is possible because labeling by minimal search and the weakness of heads replace EPP (see Chomsky (2013, 2015a)). Therefore, Merge is a strictly cyclic operation in the present minimalist framework. For example, the structure and derivation of *John hit Mary* is as follows:

(2) $\{\text{C}, \{\text{John}, \{\text{T}, \{t_{\text{John}}, \{\langle\sqrt{\text{hit}}, \text{v}^*\rangle, \{_{\delta}\ \text{Mary}, \{_{\gamma}\ t_{\sqrt{\text{hit}}}, t_{\text{Mary}}\}\}\}\}\}\}\}$

i. Merge externally forms $\{_{\gamma}\ \sqrt{\text{hit}}, \text{Mary}\}$.

ii. Merge internally forms $\{_{\delta}\ \text{Mary}, \{_{\gamma}\ \sqrt{\text{hit}}, t_{\text{Mary}}\}\}$.

iii. Merge externally forms $\{\text{John}, \{\text{v}^*, \{_{\delta}\ \text{Mary}, \{_{\gamma}\ \sqrt{\text{hit}}, t_{\text{Mary}}\}\}\}\}$.

iv. $\sqrt{\text{hit}}$ inherits features from v*.

v. Labeling and Agree take place by minimal search

vi. Pair-Merge internally forms $\langle\sqrt{\text{hit}}, \text{v}^*\rangle$ with v* affixed: such a v* becomes invisible, and the phase-hood is activated on $t_{\sqrt{\text{hit}}}$.

vii. The complement of $t_{\sqrt{\text{hit}}}$ gets transferred.

viii. Merge externally forms $\{\text{T}, \{\text{John}, \{\langle\sqrt{\text{hit}}, \text{v}^*\rangle, \{_{\delta}\ \text{Mary}, \{_{\gamma}\ t_{\sqrt{\text{hit}}}, t_{\text{Mary}}\}\}\}\}\}$.

ix. Merge internally forms $\{\text{John}, \{\text{T}, \{t_{\text{John}}, \{\langle\sqrt{\text{hit}}, \text{v}^*\rangle, \{_{\delta}\ \text{Mary}, \{_{\gamma}\ t_{\sqrt{\text{hit}}}, t_{\text{Mary}}\}\}\}\}\}\}$.

x. Merge externally forms $\{\text{C}, \{\text{John}, \{\text{T}, \{t_{\text{John}}, \{\langle\sqrt{\text{hit}}, \text{v}^*\rangle, \{_{\delta}\ \text{Mary}, \{_{\gamma}\ t_{\sqrt{\text{hit}}}, t_{\text{Mary}}\}\}\}\}\}\}\}$.

xi. T inherits features from C.

xii. Labeling and Agree take place by minimal search.

xiii. The complement of C gets transferred.

As we have also seen also in the previous chapter, the C-T relation has never been established, since *John* intervenes with the C-T relation. Epstein, Kitahara and Seely (2014b), Kitahara (2017) and T. Daniel Seely (p.c.) resolve this problem by proposing a new analysis. In their assumption, Merge forms {EA, {C, T}, …} or {IA, {v*, √root}, …}, and then internally forms $\{\text{C}, \{\text{EA}, \{t_{\text{C}}, \text{T}\}\ldots\}\}$ or $\{\text{v}^*, \{\text{IA}, \{t_{\text{v}^*}, \sqrt{\text{root}}\}\ldots\}\}$ in the normal course of derivation (EA: external argument, IA: internal argument). If this is tenable, the C-T/v*-√root relation is established by the first external Merge of C-T or v*-√root. The problem of no C-T/v*-√root relation is thus solved. Furthermore, Kitahara (2017) assumes that T and √root inherently bear uPhi, so their approach eliminates feature inheritance. Following their analysis, the derivation of *John hit Mary* proceeds as follows:

(3) {C, {$_{\delta}$ John, {$_{\gamma}$ {t_{C}, T}, {t_{John}, {<√hit, v*>, {$_{\kappa}$ Mary, {$_{\iota}$ {$_{\varepsilon}$ t_{v*}, $t_{\surd hit}$}, t_{Mary}}}}}}}}

i. Merge externally forms {$_{\iota}$ {$_{\varepsilon}$ v*, √hit}, Mary}.

ii. Merge internally forms {v*, {$_{\kappa}$ Mary, {$_{\iota}$ {$_{\varepsilon}$ t_{v*}, √hit}, t_{Mary}}}}.

iii. Merge externally forms {John, {v*, {$_{\kappa}$ Mary, {$_{\iota}$ {$_{\varepsilon}$ t_{v*}, √hit}, t_{Mary}}}}}.

iv. Labeling takes place by minimal search: κ, ι and ε are labeled as <Phi, Phi>, √P and √hit, respectively.

v. Pair-Merge internally forms <√hit, v*> with v* affixed: such a v* becomes invisible, and the phase-hood is activated on {$_{\surd hit}$ t_{v*}, $t_{\surd hit}$}.

vi. The complement of {$_{\surd hit}$ t_{v*}, $t_{\surd hit}$} gets transferred.

vii. Merge externally forms {{C, T}, {John, {<√hit, v*>, {$_{\delta}$ Mary, {$_{\gamma}$ {$_{\varepsilon}$ t_{v*}, $t_{\surd hit}$}, t_{Mary}}}}}}.

viii. Merge internally forms {C, {$_{\delta}$ John, {$_{\gamma}$ {t_{C}, T}, {t_{John}, {<√hit, v*>, {$_{\kappa}$ Mary, {$_{\iota}$ {$_{\varepsilon}$ t_{v*}, $t_{\surd hit}$}, t_{Mary}}}}}}}}.

ix. Labeling takes place by minimal search: δ and γ are labeled as <Phi, Phi> and TP, respectively.

x. The complement of the higher C gets transferred.

It follows that Probe-Goal Agree as proposed by Chomsky (2000, 2001), does not exist, but agreement or feature valuation takes place as an instantiation of feature sharing by minimal search (see Chomsky (2013, 2015a), Epstein, Kitahara and Seely (2014b), Kitahara (2017)). The approach in (3), therefore, does not postulate both (i) feature inheritance discussed by Chomsky (2013, 2015a) and (ii) Probe-Goal Agree proposed by Chomsky (2000, 2001). Consequently, this approach maximizes syntactic minimization; it maximally takes advantage of the effects of Merge (see Epstein, Kitahara and Seely (2012, 2013, 2014a, b, 2015)).

Crucially, their approach seems to be conceptually more ideal than the proposed analyses in chapter 4 because it eliminates feature inheritance and makes computation simpler, in the spirit of a Merge-only system. Thus, it is valuable to show that their approach can also explain the empirical facts, which the previous chapter accounted for.

If the alternative analysis based on their approach can cover the same amount of empirical facts, we should replace the analysis in chapter 4 with the alternative. However, I would like to argue, in this chapter, that the alternative can cover the same facts, but has at least two technical problems. The first problem is that the alternative discards the analysis of external pair-Merge of heads by Epstein, Kitahara and Seely (2016). The second problem is that it is obscure how we assume that {H, H} is a single head. Therefore, the alternative is not a replacement (see section 5.2.2.4 for the detail of two problems and this conclusion). In what follows, I introduce additional assumptions essential for alternative 1.

Importantly, following Chomsky (2015a), Epstein, Kitahra and Seely (2014b) and Kitahara (2017) assume that both $\{H_1, t_{H2}\}$ and {H, √root} (where H is a head) are not sets but heads.

(4) $\{H_1, H_2\}$ is a head if minimal search can identify its label.

We, for example, cannot identify the label of ι in (3) without this assumption. Yet, this assumption is incompatible with the proposed analyses in the earlier chapters. I will discuss this matter later.

As discussed in the previous chapters, particularly in chapter 4, minimal search requires *binary* pair-Merge.

(5) The Binary Requirement of Minimal Search for pair-Merge:
$n = 2$/pair-Merge $(X_1, \ldots, X_n) = \langle X_1, \ldots, X_n \rangle$

In chapters 2 and 4, I argued that the traditional head-raising violated this binary requirement as far as internal pair-Merge was a syntactic operation (contrary to Chomsky, as reported in Nomura (2017), and Chomsky, Gallego and Ott (to appear)). Namely, (3v) violates the binary requirement. To make the head-raising apply, we need something like the following condition:

(6) The Condition on Internal Pair-Merge of Heads:

Pair-Merge only sees syntactic sets between two heads iff the relation between two lexical items has been established.

The external application of Merge (C, T) and Merge (v*, √root) can deduce the condition, although Epstein, Kitahara and Seely (2014b) and Kitahara (2017) did not mention this.

(7) Deductive Steps to the Condition on Internal Pair-Merge of Heads:

i. The Binary Requirement of Minimal Search for Merge: n = 2/pair-Merge (X_1, …, X_n) = <X_1, …, X_n>.

ii. The external application of Merge (C, T) and Merge (v*, √root) establishes the C-T/v*-√root relation.

iii. The relation established by the external Merge of heads ensures the binarity of minimal search for Merge.

iv. Internal pair-Merge of two heads applies successfully because the binarity has been ensured by minimal search for feature inheritance.

v. The condition on internal pair-Merge of heads is deduced.

Despite the theoretical advantages, the derivational system in (3) leaves a question about simple empirical data of *there* constructions that cause the XP-YP problem. In a copular structure {*be* {$_δ$ XP, YP}}, XP or YP internally merges to SPEC-T (i.e. SPEC-*be*) for labeling δ (see Chomsky (2013: 43–44)) unless X agrees with Y (see chapter 4). This readily predicts that the internal merger of XP or YP to SPEC-T does not take place in *there* constructions because *there* merges to SPEC-T. The prediction leads the derivation to crash at the interface(s), since δ is not labeled (the XP-YP problem).

(8) The XP-YP Problem:

In a *there* construction such as *there are* [$_δ$ [$_{NP}$ *three men*] [$_{PP}$ *in the room*]], the label δ is not determined.

In this section, I resolve this problem by adopting the derivational system in (3), which eliminates feature inheritance and maximizes the effects of freely applying simplest Merge.

5.2.2 Structures and Derivations of *There* Constructions

5.2.2.1 *There Be*/Unaccusative *There* Constructions

Given the basic system as in (3), I propose the following structure and derivation of *there are three men in the room*:

(9) $\{\mathrm{C}, \{_{\gamma}\ \text{there}, \{<<\sqrt{}\text{be}, \mathrm{T}>, \{_{\delta}\ \{t_{\text{there}}, \{\text{three}, \text{men}\}\}, \{_{\varepsilon}\ \{t_{\mathrm{C}}, \{t_{\mathrm{T}}, t_{<\sqrt{}\text{be}, \text{v}^*>}\}\}, \{t_{\{\text{there}, \{\text{three}, \text{men}\}\}}, \{\text{in}, \{\text{the}, \text{room}\}\}\}\}\}\}\}\}$

i. Pair-Merge externally forms <√be, v*>: such a v* becomes invisible.

ii. Merge externally forms {{C, {T, <√be, v*>}}, {{there, {three, men}}, {in, {the, room}}}. Note that **both T and √be inherently bear uPhi**.

iii. Merge internally forms $\{\mathrm{C}, \{\text{there}, \{\mathrm{T}, \{\{t_{\text{there}}, \{\text{three}, \text{men}\}\}, \{\{t_{\mathrm{C}}, \{t_{\mathrm{T}}, <\sqrt{}\text{be}, \text{v}^*>\}\}, \{t_{\{\text{there}, \{\text{three}, \text{men}\}}, \{\text{in}, \{\text{the}, \text{room}\}\}\}\}\}\}\}\}\}$.

iv. Pair-Merge internally forms <<√be, v*>, T> with T affixed: such a T becomes invisible.

v. Labeling takes place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and √P, respectively.

vi. The complement of (the highest) C gets transferred.

Notice that the operation Merge including pair-Merge, is able to apply freely in any order under the conception of freely applying Merge (see Epstein, Kitahara and Seely (2016), Nomura (2017), Epstein, Obata and Seely (to appear), among others). It is thus logically possible for pair-Merge to apply just before minimal search for labeling. In addition, because pair-Merge was reformulated in chapters 2 and 3, the derivation above should be revised as follows:

(10) {C, {$_{\gamma}$ there, {$_{\kappa}$ {{√be, {√be, v*}}, {{√be, {√be, v*}}, T}}, {$_{\delta}$ {t_{there}, {three, men}}, {$_{\varepsilon}$ {t_{C}, {t_{T}, $t_{\{\surd\text{be}, \{\surd\text{be}, \text{v*}\}\}}$}}, {$t_{\{\text{there}, \{\text{three, men}\}\}}$, {in, {the, room}}}}}}}

i. PM by SM externally forms {√be, {√be, v*}}: such a v* becomes invisible.

ii. Merge externally forms {{C, {T, {√be, {√be, v*}}}}, {{there, {three, men}}, {in, {the, room}}}}}}. Note that both T and √be inherently bear uPhi.

iii. Merge internally forms {C, {there, {T, {{t_{there}, {three, men}}, {{t_{C}, {t_{T}, {√be, {√be, v*}}}}, {$t_{\{\text{there}, \{\text{three, men}\}\}}$, {in, {the, room}}}}}}}}.

iv. PM by SM internally forms {{√be, {√be, v*}}, {{√be, {√be, v*}}, T}}: such a T becomes invisible.

v. Labeling takes place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and √P, respectively.

vi. The complement of (the highest) C gets transferred.

There are four notable points in this derivation. First, recall that PM by SM is just Merge. PM by SM should apply before labeling since labeling applies when Transfer applies. After the application of Transfer is completed, the complement of the phase head is impossible to modify in terms of the PIC. The first application of PM by SM forms {√be, {√be, v*}}, which cancels the phasehood of v* (see Epstein, Kitahara and Seely (2016)). As already shown in chapter 4, √be is required to be part of the derivational ordered pair for empirical reasons which I will discuss again later in this chapter.

Second, following Abe (2016) and Goto (2017), *there* bearing uPhi and the associate nP *three men* externally merge and establish the agreement relation (although its valuation is optional at this point (see Abe (2016) and Goto (2017))). It follows that this relation between *there* and *three men* ensures the expletive-associate relation (see Chomsky (1986: 179, 1995b: 155), among others, for the expletive-associate relation). Assuming that *there* expletive bears (defective) uPhi (see Chomsky (2001)), it is reasonable that *there* receives the identical value from its associate thanks to the expletive-associate relation formed by external Merge of the heads. Therefore, after *there* internally merges to κ, γ is labeled as <Phi, Phi> under feature sharing between the valued *there* and <<√be, v*>, T>/{{√be, {√be, v*}}, {{√be, {√be, v*}}, T}} bearing uPhi. Note that Abe (2016), Chomsky (2000, 2004) and Goto (2017) discuss the possibility that *there* expletive is a head D, but I adopt the traditional view, under which the expletive is an XP object (perhaps DP) following Chomsky (1981); Epstein, Kitahara and Seely (2014a); and Uriagereka (1988), among others.

Third, the new analysis is that, following Epstein, Kitahara and Seely (2014b), Kitahara (2017) and T. Daniel Seely (p.c.), C and T need to externally merge in the first place, eliminating the problematic application of feature inheritance in Chomsky (2013, 2015a). Adapting their proposal, I assume that T merges with √be before merging with C. C and {T, √be} then merge and directly establish the substantial relation. Crucially, T and √be inherently bear uPhi. After the T internally merges to δ, *there* internally merges to SPEC-T, and the C internally merges to γ, the √be raises (i.e. internally pair-merges) to T (the directionality or asymmetricity of "merge X *to* Y" is only for expository purposes). Consequently, the symmetric XP-YP problem dissolves since minimal search can identify δ as <Phi, Phi>. Namely, nP *three men* and √be bearing uPhi share the Phi-set [third person, plural, masculine], and it becomes the label for δ under labeling by minimal search. Note that the lower copy √be is, by definition, *visible* when it has been left by internal PM by SM. Therefore, in {t_{C}, {t_{T}, $t_{\sqrt{be}}$}}, the visible lower copy is only $t_{\sqrt{be}}$. The visible lower copy √be agrees with the associate. (See Abe (2016) and Goto (2017) for the agreement relation established by external Merge.)

Fourth, it is important to note that the condition of internal PM by SM holds in the proposed analysis (see chapter 4):

(11) The Condition on Internal PM by SM:
Minimal search for Merge does not apply iff minimal search for feature inheritance has established the relation between two lexical items.

Its deduction is similar to (7).

(12) Deductive Steps to the Condition on Internal PM by SM:

i. The Binary Requirement of Minimal Search for Merge: n = 2/pair-Merge (X_1, …, X_n) = <X_1, …, X_n>.

ii. The external application of Merge (C, T) and Merge (v*, √root) establishes the C-T/v*-√root relation.

iii. The relation established by the external Merge of heads ensures the binarity of minimal search for Merge.

iv. Internal PM by SM of heads successfully applies because the binarity has been ensured by minimal search for feature inheritance.

v. The condition on internal PM by SM of heads is deduced.

Therefore, the proposed analysis not only solves the XP-YP problem but also deduces the condition on internal PM by SM.

The proposal further suggests that the structure and derivation of *there* constructions with unaccusative verbs is almost identical to that of *there be* constructions. I propose the following structure and derivation of an unaccusative *there* construction, *there appeared a ship on the horizon* (hereafter, I use the notation of the representational ordered pair, instead of that of derivational ordered pairs, in this section):

(13) {C, {$_{\gamma}$ there, {<√do, T>, {<√appear, v*>, {$_{\delta}$ {t_{there}, {a, ship}}, {$_{\varepsilon}$ {t_{C}, {t_{T}, $t_{\sqrt{\text{do}}}$}}, {{t_{there}, $t_{\text{a ship}}$}, {on, {the, horizon}}}}}}}}}

i. PM by SM externally forms <√appear, v*>.

ii. Merge externally forms {$_{\varepsilon}$ {C, {T, √do}}, {{there, {a, ship}}, {on, {the, horizon}}}}.

iii. Merge internally forms {$_{\delta}$ {there, {a, ship}}, {$_{\varepsilon}$ {C, {T, √do}}, {{t_{there}, $\text{t}_{\text{a ship}}$}, {on, {the, horizon}}}}}.

iv. Merge externally forms {<√appear, v*>, {$_{\delta}$ {there, {a, ship}}, {$_{\varepsilon}$ {C, {T, √do}}, {{t_{there}, $\text{t}_{\text{a ship}}$}, {on, {the, horizon}}}}}}.

v. Merge internally forms {C, {there, {T, {<√appear, v*>, {$_{\delta}$ {t_{there}, {a, ship}}, {$_{\varepsilon}$ {t_{C}, {t_{T}, √do}}, {{t_{there}, $t_{\text{a ship}}$}, {on, {the, horizon}}}}}}.

vi. PM by SM internally forms <√do, T> with T affixed: such a T becomes invisible.

vii. Labeling takes place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and √P, respectively.

viii. The complement of (the highest) C gets transferred.

First, following Epstein, Kitahara and Seely (2016), I assume that unaccusative verbs are the instantiation of an ordered pair <√root, v*> that is formed by external pair-Merge of heads. Namely, external PM by SM takes √appear and v* and forms the ordered pair <√appear, v*> in the structure above. The representational ordered pair <√appear, v*> then externally merges to δ at the relevant point of the derivation.

Second, I assume that √do occurs in unaccusative constructions without the assignment of phonological features to it. Following Distributed Morphology (Halle and Marantz (1993), Marantz (1997), Embick and Marantz (2008), among others; see also Kitahara (2017)), I assume that phonological features are assigned at the externalization process. In (9), {√be, {√be, T}} is externalized as *are* because it is the only possible element for "a verb." In (13), <√do, T> is, in contrast, externalized as a null element (except for situations like emphasis, negatives and interrogatives), since <√appear, v*> is regarded as the possible verb. That is, the externalization process does not assign any phonological feature to √do in (13).

5.2.2.2 *Unaccusativized There* Constructions

Among *there* constructions in English, *there be* NP and *there* $\mathrm{V}_{\text{unaccusative}}$ NP are often discussed in the literature. The analyses based on *freely applying Merge*, however, also explain the peripheral "unergative" *there* constructions without postulating any redundant stipulation, thus strictly conforming to the SMT.

Given the basic analysis in unaccusative *there* constructions, I propose the following structure and derivation of an unergative *there* construction, *there walked into the room three men* (the brackets for *into the room* are omitted for expository purpose):

(14) {C, {$_{\gamma}$ there, {<√do, T>, <{<<√walk, into the room>, v*>, {$_{\delta}$ {t_{there}, $t_{\text{three men}}$}, {$_{\varepsilon}$ {t_{C}, {t_{T}, $t_{\sqrt{}\text{do}}$}}, $t_{\{\text{there, \{three men\}}\}}$}}}, {three, men}>}}}

i. Merge externally forms *into the room*.

ii. PM by SM externally forms <<√walk, into the room>, v*>; *walk* becomes "unaccusativized."

iii. Merge externally forms {$_{\varepsilon}$ {C, {T, √do}}, {there, {three, men}}}.

iv. Merge internally forms {$_{\delta}$ {there, {three, men}}, {$_{\varepsilon}$ {C, {T, √do}}, $t_{\{\text{there, \{three men\}}\}}$}}.

v. Merge externally forms {<<√walk, into the room>, v*>, {$_{\delta}$ {there, {three, men}}, {$_{\varepsilon}$ {C, {T, √do}}, $t_{\{\text{there, \{three men\}}\}}$}}}.

vi. Merge internally forms {C, {there, {T, {<<√walk, into the room>, v*>, {$_{\delta}$ {t$_{\text{there}}$, {three, men}}, {$_{\varepsilon}$ {t_{C}, {t_{T}, √do}}, $t_{\{\text{there, \{three men\}}\}}$}}}}}}.

vii. PM by SM internally forms <√do, T> with T affixed: such a T becomes invisible to minimal search.

viii. Labeling takes place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and vP (which is the label of √walk-*into the room*-v*), respectively.

ix. PM by SM internally forms {C, {there, {<√do, T>, <{<<√walk, into the room>, v*>, {$_{\delta}$ {t_{there}, $t_{\text{three men}}$}, {$_{\varepsilon}$ {t_{C}, {t_{T}, $t_{\sqrt{}\text{do}}$}}, $t_{\{\text{there, \{three men\}}\}}$}}}, {three, men}>}}}.

x. The complement of (the highest) C gets transferred.

First, the unergative verb *walk* (*into the room*) is unaccusativized by external PM by SM of √walk and *into the room* and <√walk, into the room> and v*.[1] As argued in chapter 4, this unaccusativization is supported by the following empirical facts (see Omune (2016)):

[1] I put aside the problem of whether Merge, PM by SM or the combination of both forms *into the room*. Each option is logically possible under the conception of freely applying Merge (see also fn. 5 in chapter 4).

(15) a. There walked into the room a fierce-looking tomcat.
b. A fierce-looking tomcat came into the room by walking.
c. There ambled into the room a frog.
d. A frog came into the room by ambling.
e. Suddenly, there ran out of the bushes a grizzly bear.
f. A grizzly bear came out of the bushes by running.

(a, c: Milsark (1974: 155, 246)))
(e: Lumsden (1988: 38))

(15b, d and f) are the (roughly) paraphrased versions of (15a, c and e), respectively. The theta-role of all of the subjects in the paraphrased sentences are theme assigned by the main verb *come*. The semantic flavor, such as the agent, is assigned by the *by*-phrases (i.e. *by walking*, *by ambling* and *by running*). The unergative verbs, thus, just "modify" the events in unergative *there* constructions. The leading event is expressed by the unaccusative structure. These facts, therefore, support the analysis of unaccusativization of an unergative verb in (9). The √walk just modifies the event of v* in <<√walk, into the room>, v*>, and the whole structural configuration, including <<√walk, into the room>, v*>, denotes the main "unaccusative" event.

Second, assuming that the abstract √do occurs in an unaccusative structure, δ is successfully labeled. In the proposed unaccusativized structure, √do can be, by hypothesis, introduced in the relevant course of derivation. Minimal search for labeling, therefore, labels δ as <Phi, Phi> through the relation Agree (i.e. feature sharing) for the same reason as the typical unaccusative case. Eventually, the symmetric XP-YP problem in (8) is also solved in this way in an unergative *there* construction.

Third, PM by SM of phrases applies at the end of the cycle, and the associate adjoins to δ, which is informally the right edge of vP. I adopt this landing site based on Chomsky's (2001) Th/Ex (Thematizaion/Extraction) but do not adopt Th/Ex itself because it is not only conceptually dubious but also empirically problematic. For conceptual reasons, as Chomsky (2016a) has recently mentioned, Spell-Out does not exist in the current minimalist framework. Th/Ex is therefore untenable, simply because it is an operation of the phonological component and applies at Spell-Out for the weak phase. For empirical reasons, because it is an operation of the phonological component, Th/Ex does not change meaning. However, it apparently causes the surface semantic effect as shown below:

(16) <u>*be* verbs</u>
 a. *There is the man in the room.
 <u>*unaccusatives*</u>
 b. *There appeared all ships on the horizon.
 c. *There arrived the man in the room.

In *there* constructions, a "strong" determiner (including a null determiner) cannot occur as an associate's partner. Thus, as the "strong" reading, we cannot use definites (e.g. *the*, demonstratives, personal pronouns and possessives) or universals (e.g. *all* and *every*) as determiners of an associate nP. We call this phenomenon as the definiteness restriction (regardless of the use of actual definites). However, the definiteness restriction is relaxed when the associates are extraposed rightward, regardless of the type of verbs:

(17) <u>*unaccusatives*</u>
 a. There came to his mind her beautiful and intelligent face.
 <u>*unergatives*</u>
 b. There ambled into the room my neighbor's frog.
 <u>*transitives*</u>
 c. There entered the room the new professor from Sweden.

(a: Quirk et al. (1985: 1409))
(b: Milsark (1974: 246))
(c: Julien (2002: 13))

Therefore, Th/Ex causes the surface semantic effect. Notice that Chomsky (2001) himself argues that Th/Ex applies when an argument structure includes v, a light verb marking unaccusative/passive. If so, we can assume that Th/Ex also applies in unergative *there* constructions because unergatives are "unaccusativized" in *there* constructions (see (15) and the relevant discussions above). We will see the relevant analysis below for the case of the transitive *there* construction or the TEC, as in (17).

Given the basic analysis of unaccusative/unergative *there* constructions, I propose the following structure and derivation of a transitive *there* construction, *there entered the room three men from Sweden*:

(18) {C, {$_{\gamma}$ there, {<√do, T>, <{<<√enter, the room>, v*>, {$_{\delta}$ {t_{there}, $t_{\text{three men from Sweden}}$}, {$_{\varepsilon}$ {t_{C}, {t_{T}, $t_{\surd\text{do}}$}}, $t_{\{\text{there, }\{\{\text{three men}\}, \{\text{from Sweden}\}\}\}}$}}}, {three men from Sweden}>}}}

i. Merge externally forms *the room*.

ii. PM by SM externally forms <<√enter, the room>, v*>; *enter* becomes "unaccusativized."

iii. Merge externally forms {$_{\varepsilon}$ {C, {T, √do}}, {there, {three men from Sweden}}}.

iv. Merge internally forms {$_{\delta}$ {there, {three men from Sweden}}, {$_{\varepsilon}$ {C, {T, √do}}, $t_{\{\text{there, }\{\text{three men from Sweden}\}\}}$}}.

v. Merge externally forms {<<√enter, the room>, v*>, {$_{\delta}$ {there, {three men from Sweden}}, {$_{\varepsilon}$ {C, {T, √do}}, $t_{\{\text{there, }\{\text{three men from Sweden}\}\}}$}}}.

vi. Merge internally forms {C, {$_{\gamma}$ there, {T, {<<√enter, the room>, v*>, {$_{\delta}$ {t_{there}, {three men from Sweden}}, {$_{\varepsilon}$ {t_{C}, {t_{T}, $t_{\surd\text{do}}$}}, $t_{\{\text{there, }\{\text{three men from Sweden}\}\}}$}}}}}}.

vii. PM by SM internally forms <√do, T>: such a T becomes invisible to minimal search.

viii. Labeling takes place by minimal search: γ, δ and ε are labeled as <Phi, Phi>, <Phi, Phi> and vP (which is the label of √enter-*the room*-v*), respectively.

ix. PM by SM internally forms {C, {there, {<√do, T>, <{<<√enter, the room>, v*>, $\{\{t_{\text{there}}, t_{\text{three men from Sweden}}\}, \{\{t_{\text{C}}, \{t_{\text{T}}, t_{\surd\text{do}}\}\}, t_{\{\text{there}, \{\{\text{three men}\}, \{\text{from, Sweden}\}\}\}}\}\}$, {three men from Sweden}>}}}.

x. The complement of (the highest) C gets transferred.

First, as already shown in the case of unaccusativized unergative *there* constructions, the verbal root √enter also becomes unaccusativized. Because roots are universally underspecified as to categories, there is virtually no substantial element like a "transitive, unergative, or unaccusative" root. The unaccusativization of a transitive verb is therefore logically available unless stipulated. In addition, external PM by SM of a v* and an unergative root is in effect supported by the empirical facts shown in chapter 4. External PM by SM of √enter and *the room* is therefore possible. There is no reason to preclude external PM by SM of a v* and a root used for transitive verbs.

Second, the abstract null √do also occurs in this structure. As discussed earlier, the existence of the element is conceptually unproblematic in (13). If it is on track, δ is successfully labeled. The XP-YP problem thus also dissolves in transitive *there* constructions.

5.2.2.3 Accounting for Empirical Facts

The proposed analyses can also account for the same facts shown in chapter 4, which are repeated below.

(19) Extraction (*be verbs*):

a. Who was there t_{wh} in the kitchen?

b. You remember a guy who there was t_{wh} for you.

(20) Extraction (*unaccusatives* (*IVES*)):

a. ? I saw a ship which there appeared t_{wh} on the horizon.

b. ? I remember several new facts which there emerged t_{wh} at the meeting.

(21) Extraction (*unergatives/transitives* (*OVES*)):

a. * I saw a fierce-looking tomcat which there walked into the room t_{wh}.

b. *I saw a frog which there ambled into the room t_{wh}.

c. *I saw a strange man [which] there entered the room t_{wh}.

(22) Sub-extraction (*be verbs*)

a. Of which artist was there a portrait t_{wh} on the wall?

b. How many ancestors was there a portrait of t_{wh} on the wall?

(23) Sub-extraction (*unaccusatives* (*IVES*))

a. Of which artist did there hang a portrait t_{wh} on the wall?

b. How many ancestors did there hang a portrait of t_{wh} on the wall?

(a, b: adapted from Nishihara (1999: 392, fn. 10))

(24) Sub-extraction (*unergatives/transitives* (*OVES*))

a. *Which community did there walk into the room a member of t_{wh}?

b. *Of which community did there walk into the room a member t_{wh}?

c. *Which community did there enter the room a member of t_{wh}?

d. *Of which community did there enter the room a member t_{wh}?

(a, b: adapted from Nishihara (1999: 394))

(25) Scope (*quantifier-auxiliary*)

a. Someone must be in his house.

i. someone > must   ii. must > someone

b. There must be someone in his house.

i. *someone > must   ii. must > someone

(a, b: adapted from Williams (1984: 152))

(26) Scope (*quantifier-negative*)

a. Many women aren't sick.

i. many > not   ii. not > many

b. There aren't many women sick.

i. *many > not   ii. not > many

(27) Scope (*quantifier-verb*)

a. Many ships seem to be in the harbor.

i. many > seem   ii. seem > many

b. There seem to be many ships in the harbor.

i. *many > seem ii. seem > many

(a, b: adapted from Hartmann (2008: 19))

(28) NPI

a. *There seems to any European team to be no NBA team beatable.

b. No NBA team seems to any European team to be beatable.

(a, b: Bošković (1997: 98))

(29) Anaphor

a. *There seems to himself to be someone in the garden.

b. Someone seems to himself to be in the garden.

(a, b: Bošković (1997: 98))

(30) NPI

a. There walked into the classroom three students from my department.

b. There walked into the classroom no one from my department.

c. *There walked into any classroom no one from my department.

(c: Deal (2009: 27, fn. 48))

(31) NPI

a. There entered the classroom three students from my department.

b. There entered the classroom no one from my department.

c. *There entered any classroom no one from my department.

In the *there be* construction and the IVES, the associates are easy to (sub-)extract because the phasehood of v* has been canceled. Recall that labeling takes place in the canceled v*P phase first and the CP phase second (see the relevant discussion in chapter 4). Thus, v* is essential even in the *there* be construction. In the OVES, the associates are impossible to (sub-)extract, since they are adjoined to the right peripheral of v*P by internal PM by SM. That is, the associates are invisible, being part of the derivational ordered pair or the representational ordered pair.

The facts in (25)–(31) suggest that LF covert movement empirically does not exist. Furthermore, the covert movement conceptually does not exist, as argued in chapter 4. The proposed analyses are compatible with these facts. In particular, the locative expressions in the OVES are part of unaccusativized verbs, which makes the locative expression invisible. Associates therefore cannot c-command the expressions in (30).

5.2.2.4 Two Reasons for Discarding Alternative 1

As we have seen in this chapter, the alternative analysis, which is based on the elimination of feature inheritance by external Merge of heads (Epstein, Kitahara and Seely (2014b) and Kitahara (2007)), resolves the XP-YP problem in *there* constructions. The alternative, moreover, covers the same amount of empirical facts, which the analyses in chapter 4 explained. However, the proposed alternative has at least two conceptual problems. First, T's or √root's inherent bearing of uPhi makes the phase cancellation of external pair-Merge (or PM by SM) of heads meaningless. Second, it is controversial to assume that {H_1, H_2} can be regarded as a head. Recall that the phase cancellation process in the bridge verb construction is as follows:

(32) {John, {<√think, v*>, {that, …}}}
  i. Pair-Merge of heads externally forms <√think, v*>: v* becomes invisible with respect to both its uPhi and its phase-hood.
  ii. Merge externally forms {John, {<√think, v*>, {that, …}}}.

The phase cancellation can proceed under the assumption of v* inherently bearing uPhi. This uPhi becomes invisible when pair-Merge externally forms the ordered pair <√think, v*>. If √think inherently bears uPhi, the derivation crashes because the uPhi never becomes invisible and cannot be valued.

Solving the problem, we can assume uPhi is inherently on phase heads, and uPhi transmits into T/√root through the C-T/v*-√root relation established by external Merge of heads. In fact, Kitahara (2017: 245) considers this option:

(33) a. C/v* bears uPhi inherently, and T/R inherits uPhi from C/v* derivationally.
  b. T/√root bears uPhi inherently.

(adapted from Kitahara (2017: 245))

If we select the first option, however, we cannot eliminate feature inheritance. Thus, the second option is conceptually better for maximizing syntactic minimization in the spirit of a Merge-only system. The first option, therefore, weakens the conceptual advantages. As long as the first option is selected, the alternative is conceptually not better than the proposed analyses in chapter 4.

As for the second problem, how do we regard $\{H_1, H_2\}$ as a head? Chomsky (2015a) and Chomsky, Gallego and Ott (to appear) argue that the label of $\{H_1, H_2\}$ becomes H_1 if H_2 is weak. Suppose that there is a head-head symmetric structure {n, √book} for the nominal expression *book*. When minimal search labels the structure, its label becomes n, which is equivalent to the head n under the Bare Phrase Structure. I agree with this type of labeling. However, is the labeled *structure/set* equivalent to the head? As long as labels are irrelevant to syntactic computation in the spirit of simplest Merge, the structure {n, √book} itself should not be the head but the set containing two heads. If we want to make the head from n and √book, we have PM by SM, the reformulated version of pair-Merge. The same scenario holds when one member in $\{H_1, H_2\}$ is the lower copy. That is, $\{t_C, T\}/\{t_{v*}, \sqrt{root}\}$ should be regarded as the set. Even if the lower copy is invisible, it exists in syntax and is interpreted by the CI system. If it is tenable, the entire analysis in this section, alternative 1, should be discarded.

5.3 Alternative 2 (for the Double Object Construction)

5.3.1 Introduction

In this section, I argue that Chomsky's (2013, 2015a) labeling theory and Rizzi's (2015a, b, 2016) maximality principle explain the non-extractability of the indirect object in the double object construction. The non-extractability has been analyzed within the earlier framework of syntactic theory (e.g. Oba (2005, 2016) and Hallman (2015)). However, I have not adopted the earlier framework because the recent framework is conceptually better in terms of the SMT, evolvability and computational efficiency. The recent framework has even explained various linguistic phenomena empirically (see

Chomsky (2013, 2015a), Rizzi (2015a, b, 2016), Epstein, Kitahara and Seely (2016), etc.). However, the non-extractability of the indirect object has not truly been explained using the current minimalist framework, and it should, therefore, be reanalyzed using this current framework.[2]

Chomsky's (2013, 2015a) labeling algorithm accounts for the obligatory exit of a syntactic object such as a vP-internal subject (i.e. an external argument (EA) at SPEC-vP). In contrast, Rizzi's (2015a, b, 2016) maximality principle or maximality provides a justification for permissible "halting sites" for movement. Let us consider the following example:

(34) *[$_{Q}$ Which book] does John wonder [$_{Q}$ t was published last year]?

The wh-phrase is not extractable under labeling and maximality. Namely, *which book*, the non-maximal syntactic object with the Q label, is frozen in place following maximality under which only maximal objects with a given label can be moved. Assuming this analysis is tenable, I will propose that it extend to the case of the double object construction.

This section is structured as follows. Section 5.3.2 briefly reviews Chomsky's (2013, 2015a) labeling algorithm and Rizzi's (2015a, b, 2016) maximality. In Section 5.3.3, I will show the semantic and syntactic properties of the double object construction. Section 5.3.4 proposes that P$_{\text{HAVE}}$ be a criterial feature. In Section 5.3.5, the proposal under labeling and maximality accounts for the non-extractability of the indirect object. In addition, the analysis explains the extractable case of the indirect object. Section 5.3.6 shows that the proposed analysis also explains a phenomenon of sub-extraction with respect to the double object construction. Section 5.3.7 concludes Section 5.3.

5.3.2 Labeling, Maximality and the Halting Problem

[2] The analysis in chapter 2 explains non-extractability of the indirect object but not its non-*sub*-extractability of that.

Chomsky (2013, 2015a) argues that labels on syntactic structures are essential for interpreting the syntactic objects at the process of externalization and at the conceptual-intentional (CI) interface. Informally, externalization is phonological realization, and the CI system is for semantics. Accordingly, Chomsky proposes the labeling algorithm as an essential computation in syntax. Labeling occurs as a result of minimal search conforming to a third-factor principle, Minimal Computation which is not unique to language (for more on the third factor, see Chomsky (2005, 2013) among others). The labeling algorithm is, therefore, not an arbitrary stipulation but a naturally deduced algorithm that is appropriate for normal science.

The labeling algorithm explains the obligatory exit of the EA from vP. Thus, we no longer assume EPP (the Extended Projection Principle) or its descendant, the EPP feature. Informally, EPP states that SPEC-T or the subject position should be overtly occupied by an element.

(35) [C [$_{\text{F}}$ John$_{\text{F}}$ T$_{\text{F}}$ [t_{JohnF} like Mary]]] (John likes Mary.)

In the derivation of (35), the EA *John* externally merges with (i.e. is base-generated in) vP. Then, *John* internally merges with (i.e. moves to) T because *John* and *like* do not share any features, but *John* shares Phi-features with T (i.e. agreement). EPP, therefore, reduces to labeling which itself is just minimal search.[3] According to Rizzi (2015a, b, 2016), a shared F can be "a criterial feature" such as Q(uestion), Foc(us), Top(ic) and Subj(ect).

Adopting the labeling algorithm, Rizzi (2015a, b, 2016) postulates the maximality principle:

[3] To completely eliminate EPP, it is essential to assume the weakness of T in English (and that of the verbal roots in any language). Chomsky (2015a) assumes that weak heads cannot serve as labels and therefore must agree to be strong and serve as labels. Accordingly, expletives (e.g. *it* and *there*) must merge and agree with T in expletive constructions. What features expletive *there* bears are controversial, but it arguably has at least one Phi-feature: [person] (see Chomsky (2000, 2001, 2004), Richards and Biberauer (2005), among others).

(36) Maximality: only maximal objects with a given label can be moved.

(Rizzi (2015b: 327))

The maximality principle readily deduces a stipulation of the traditional X-bar theory (i.e., XP objects can be moved, but X-bar objects cannot) under labeling that conforms to a third-factor principle not specific to human language. Adhering to the SMT, the theory of labeling and maximality is thus superior to the X-bar theory or the other theories largely specific to human language.

Rizzi (2010, 2015a, b, 2016) and others show "the halting problem" for the wh-movement. That is, a wh-phrase at SPEC-Q, in which Q is selected by a verb like *wonder*, is impossible to extract, hence, "criterial freezing" occurs.

(37) The Halting Problem for Wh-Movement:

a. Bill wonders [$_{Q}$ [$_{Q}$ which candidate] Q you voted for t_{wh}]

b. *[$_{Q}$ Which candidate] does Bill wonder [$_{Q}$ t_{wh} Q you voted for t_{wh}]

(adapted from Rizzi (2010: 20))

This halting problem has been explained in many ways, and it dates back to Lasnik and Saito (1992). In the current minimalist theory, the problem is recaptured under labeling and maximality. As Rizzi (2016: 116) states, "[i]n terms of BPS, a maximal projection must be understood dynamically, as the maximal node with a given label." If [$_{Q}$ which candidate] is moved from SPEC-Q as in (37b), it violates maximality since the phrase is the non-maximal object with the given label Q. The maximal object with the label Q is [$_{Q}$ which candidate Q [you voted for t_{wh}]]. Thus, [$_{Q}$ which candidate] (i.e. X-bar object) is unmovable, but [$_{Q}$ which candidate Q [you voted for t_{wh}]] (i.e. XP object) is movable.

An anonymous reviewer mentioned that *which candidate* is not the X-bar object but the maximal object DP bearing [+Q]. According to Rizzi (2015a, b), criterial features become labels when they are shared. Following this assumption, the label of *which candidate* is not DP but just Q. Therefore, the maximal object bearing the Q label is [$_{Q}$

[$_{Q}$ which candidate] Q you voted for t_{wh}] in (37). In other words, [$_{Q}$ which candidate] is *informally* the X-bar object bearing the Q label. Furthermore, Rizzi (2015a, b, 2016) does not postulate C_Q (i.e. C bearing [+Q]), contrary to Chomsky (2013, 2015a). Instead, he assumes that the dedicated Q head is present. Note that the word *X-bar object* is used as an informal term to denote an unmovable syntactic object. There is no X-bar object, which was defined in the X-bar theory, in the minimalist framework since it violates the Inclusiveness Condition (see Chomsky (1995a, b)).

5.3.3 The Subj Head

While the previous section explained the freezing effect in an A-bar position (i.e. SPEC-Q), this section gives a general review of the freezing effect in an A position. Rizzi (2015b) argues that Subj appears in the TP space under his cartographic approach dating back to Rizzi (1997) (the head Phi and the head T may be a single head (see Rizzi 2015b)).

(38) … Fin … Subj … Phi … T …

(Rizzi (2015a: 26))

The Subj head above is, in his terms, a criterial feature/head because "the subject is the argument that is taken as 'being about' that argument" (Rizzi (2015a: 24)). One might say that the Subj head yielding this aboutness property does not exist since subjects can be regarded as topic-like elements in some sense. In other words, Topic is enough to explain the aboutness property. Rizzi (2006), however, argues against this view. Even in a language like Italian which permits null subjects, a subject is feasible as an answer to the question *what happened*.

(39) The Aboutness Property of Subjects:
Che cosa è successo?
'What happened?'
a. Un camion ha tamponato l'autobus per Roma

a truck has bumped-into the bus for Rome
'A truck bumped into the bus for Rome.'

b. L'autobus per Roma è stato tamponato da un camion
the bus for Rome is been bumped-into by a truck
'The bus for Rome was bumped into by a truck.'

c. #L'autobus per Roma, un camion lo ha tamponato
the bus for Rome a truck it has bumped-into
'The bus for Rome, a truck bumped into it.'

(a–c: Rizzi (2006:122))

In the first two answers above, the subjects can be answers to the question *Che cosa è successo?* ('what happened?') but in the last one, the topic *L'autobus* cannot be a possible answer.[4] The subject criterion or Aboutness needs to be distinct from the Topic.[5]

Furthermore, Rizzi (2015a: 26) states that the functional head, Subj, structurally defines the subject-predicate articulation. It follows that SPEC-Subj, Subj and COMPL-Subj cause the subject-predicate interpretation as far as the structural configuration contributes to the semantic interpretation of the CI system. The canonical position for subjects is, therefore, not SPEC-T but SPEC-Subj under the system in Rizzi (2015a, b, 2016).

(40) a. Subj attracts a nominal element to its SPEC.

---

[4] An anonymous reviewer stated that (39c) is unacceptable because *L'autobus* is focused. If it were focused, however, the comma after *L'autobus per Roma* should be removed from the sentence.

[5] An anonymous reviewer wondered whether postulating Aboutness is compatible with the minimalist theory. This is an important question because the postulation of discourse related features (e.g. Foc, Top and Subj) in lexicon violates the Inclusiveness Condition (see Chomsky, Gallego and Ott (to appear)). Conceptually, the minimalist theory should not assume those features. However, without postulating those features, the minimalist theory could lose its empirical coverage of explaining various phenomena such as topicalization, focalization and Aboutness. To reduce the tension between conceptual desiderata and empirical coverage, it is necessary to await the further refinement of the theory itself.

b. Subj triggers the aboutness interpretation and the subject-predicate interpretation at the interface.

(a, b: adapted from Rizzi (2015a: 26))

The syntactic property (40a) reminds us of EPP. However, there is no EPP in the spirit of the labeling theory (see note 3). It is necessary to assume a [+Subj] feature on the nominal so that Subj attracts a nominal element.

Assuming this subject criterion, the syntactic object at SPEC-Subj is frozen. Rizzi (2015a, b) shows that the *that*-trace effect can be explained under labeling and maximality.

(41) *That*-Trace Effect:

a. *Who$_{[+Subj]}$ do you think [that [$_{Subj}$ $t_{wh[+Subj]}$ Subj will come]]?

b. Who$_{[+Q]}$ do you think [that [$_{subj}$ Mary$_{[+Subj]}$ Subj will meet $t_{wh[+Q]}$]]?

(a, b: adapted from Rizzi (2015a: 27))

In (41a), *who* is impossible to extract since [$_{Subj}$ who Subj will come] is the maximal object with the given label Subj in the embedded clause, which violates the maximality principle. One might say that the label/projection of *who* is not Subj but Q. Recall that the label/projection of the syntactic object becomes the shared criterial feature when feature sharing occurs. The label of *who* becomes Subj because *who* shares the feature with the Subj head. See the relevant discussions in section 5.2.2. In (41b), *who* can be extracted because it is the maximal object with the given label Q or D.[6] Thus, only maximal objects with a given label can be moved, regardless of a Q label.

[6] Theoretically speaking, [$_{Subj}$ who Subj will come] can be extracted from the embedded CP in (41a) as long as we follow the maximality principle because the syntactic object is maximal with respect to the label Subj. Nevertheless, the extraction seems to be empirically impossible. I tentatively assume that the phase impenetrability condition (PIC) (see Chomsky (2000, 2001)) bars this extraction. The PIC states that the complement of a phase head (i.e. C and v) is not accessible. Therefore, the complement of *that* is non-extractable. The compatibility between PIC and the maximality principle remains to be investigated.

If the complementizer *that* is omitted, then the *that*-trace effect disappears. As Rizzi (2015b: 335, fn. 16) notes, when C is omitted, Subj is also omitted.[7] Therefore, freezing is not triggered.

(42) Who do you think ~~[C~~ [t_{wh} ~~Subj~~ will come]~~]~~?

In this case, the wh-phrase *who* is maximal with respect to its Q label under labeling and maximality. Notice that the label of the entire embedded clause becomes the label of *will*, which is T (or another head for auxiliaries), because *who* has moved from SPEC-T. Also, even if C is not omitted, the label of the whole embedded clause should not be Q for two reasons: C and Q are different heads in Rizzi's (2015a, b, 2016) system, and *think* cannot select Q.

5.3.4 Properties of the Double Object Construction

Since Larson (1988), it has been widely accepted that an asymmetric relation exists between the indirect object and the direct object. That is, the indirect object should asymmetrically c-command the direct object in the double object construction.

(43) a. I showed Mary herself.
b. * I showed herself Mary.
c. I showed no one anything.
d. * I showed anyone nothing.

(a–d: Larson (1988: 336–337))

The first and the third examples are unproblematic since *Mary* and *no one* asymmetrically c-command *herself* and *anything* respectively. In contrast, the second and the fourth examples are problematic since *herself* and *anyone* asymmetrically c-command *Mary* and

7 Rizzi (2015b) does not explain in detail why the C deletion triggers the Subj deletion. However, he suggests that the C and Subj are in a special relation that is motivated by feature inheritance. For more on feature inheritance, see Richards (2007) and Chomsky (2007).

nothing respectively. As long as the phenomena concerning anaphors and the negative polarity items are explained by the hierarchical relations, the indirect object occurs hierarchically higher than the direct object.

In the double object construction, it is widely assumed that there is prospective possession between the indirect object and the direct object (see Green (1974), Pinker (2013), Harley (1995, 2002), Harley and Jung (2015), and works cited therein). *John gives Mary the book* is roughly paraphrased as *John* "*causes*" *Mary to* "*have*" *the book*. The verb *give* can be decomposed into at least two elements, CAUSE and HAVE.

(44) a. John sent a package to the {border/boarder}.
b. John sent the {boarder/*border} a package.

(a, b: adapted from Pinker (2013: 56))

In the dative construction, there is no such entailment as the one shown in (44a). The indirect object can, therefore, be an inanimate nominal such as *border*. In contrast, because prospective possession is denoted in the double object construction, the inanimate nominal *border* cannot occur as an indirect object (see (44b)). Note that possession does not have to be literal, but the referent of the indirect object can metaphorically possess the referent of the direct object.

(45) a. John told Mary the story.
b. John asked Mary a question.
c. John showed Mary the answer.

The examples above all show the recipient's metaphorical possession. The properties reviewed in this section depend on semantics. Harley (1995, 2002) and Harley and Jung (2015) capture the semantic properties through the dedicated feature/head P_{HAVE}. Specifically, see Harley and Jung (2015) for a detailed discussion on how the syntactic head P_{HAVE} is amenable to the semantic properties.

5.3.5 Structures of the Double Object Construction

Under a version of Distributed Morphology, Harley and Jung (2015) capture this semantic peculiarity by using the dedicated head P_{HAVE} which is one of the prepositional heads proposed by Harley (1995, 2002) (for more on Distributed Morphology, see Halle and Marantz (1997), Marantz (1997, 2013) and Embick and Marantz (2008)). According to Harley and Jung (2015), the structure of the double object construction is as follows (the irrelevant parts of the structure are omitted):

(46) John sent [the boarder P_{HAVE} a package]

The prospective possession shown in (44)–(45) is explained by this structure; P_{HAVE} denotes prospective possession in (46).[8] Furthermore, this structure also captures the asymmetric relation shown in (43).

The structure proposed by Harley and Jung (2015) has a significant advantage concerning the empirical facts on the double object construction. However, they do not mention the non-extractability of the indirect object. As discussed in chapter 2, the indirect object is impossible to extract.

(47) Wh-Movement:
  a. What did Mary give/send Mary *t*?
  b. *Who did Mary give/send *t* the computer?

(48) Relativization:
  a. This is the computer which Mary gave/sent the friend *t*.
  b. *This is the friend who Mary gave/sent *t* the computer.

(49) Clefting:
  a. It is the computer that Mary gave/sent John *t*.
  b. *It is John that Mary gave/sent *t* that computer.

[8] The detailed argument structure of (46) is [EA v_{cause}-R_{send} [IA_1 P_{HAVE} IA_2]]. Note that v_{cause} is for a transitive/causative event, R_{send} denotes the verbal root of *send*, and IA is an internal argument. In addition, R_{send} directly adjoins to v_{cause} in this structure.

(50) *Tough* Movement:

a. That computer is easy to give/send John *t*.

b. *John is easy to give/send *t* that computer.

(51) Topicalization:

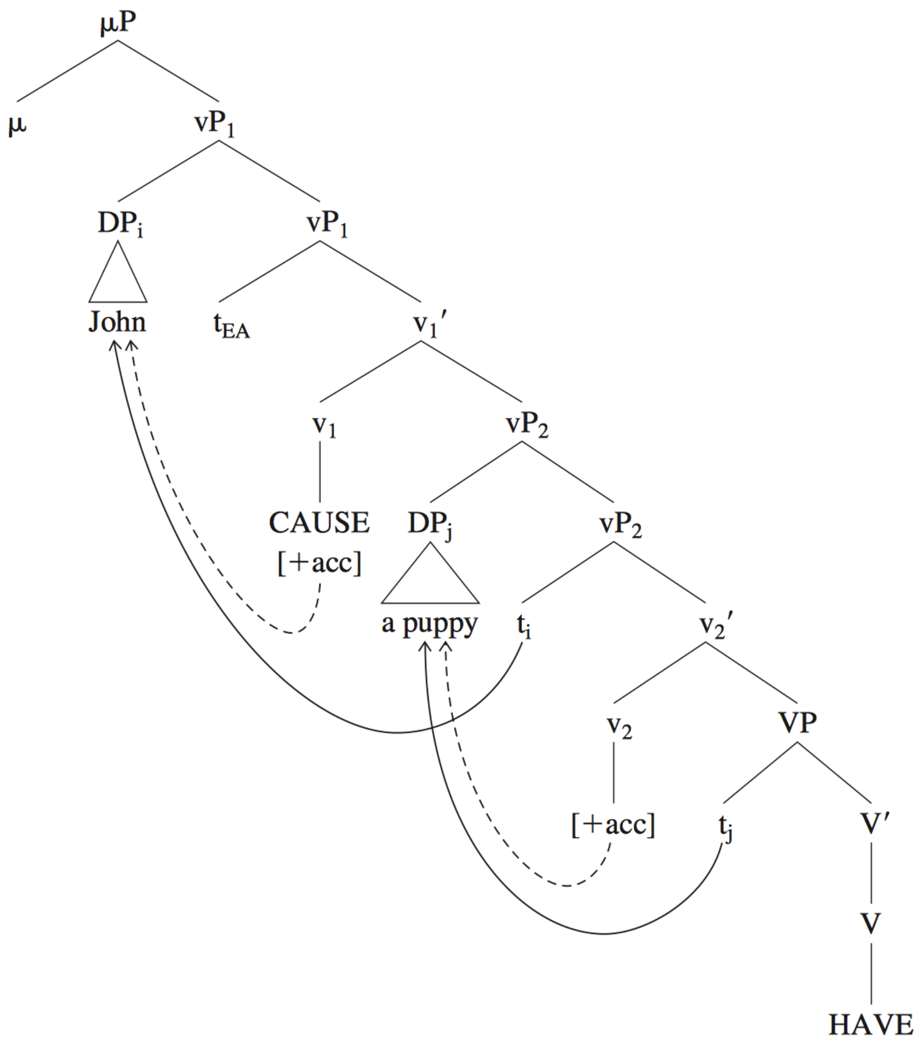
a. That computer, Mary gave/send John *t*.

b. *John, Mary gave/send *t* that computer.

((47)–(51): cf. Oba (2005: 61))

Oba (2005, 2016) and Hallman (2015) explain the non-extractability by assuming the elaborated syntactic structures. However, their structures are proposed under the earlier framework. The non-extractability should, therefore, be reanalyzed under the current framework. For instance, Oba (2005) argues that Thematization/Extraction (Th/Ex), which was originally proposed by Chomsky (2001), applies to the double object structure. When Th/Ex applies to the indirect object, it becomes unmovable because its phonological features have been spelled out to the phonological component. However, the postulation of Th/Ex is dubious for empirical reasons. See Julien (2002), Omune (2016) and Sobin (2014) for relevant discussions. Oba (2016) adopts Rizzi's (2010) approach of criterial freezing and proposes a new structure of the double object construction. His analysis successfully explains the non-extractability of the indirect object but is still based on earlier frameworks such as those of Chomsky (2000, 2001) and Rizzi (2010). Hallman (2015) adapts the traditional Larsonian structure and proposes the following structure for the double object construction:

(52) Hallman's (2015) Structure of the Double Object Construction:



(Hallman (2015: 396))

In this structure, DP_i *John* is the indirect object, and DP_j *a puppy* is the direct object. Hallman (2015: 417) argues that "[A-bar] movement may proceed from vP_2 but not vP_1." The indirect object is, therefore, non-extractable. However, this is not the explanation but the mere generalization from the familiar phenomena (e.g. Chomsky (1981), Lasnik and Saito (1984, 1992) and Diesing (1992)), including (47)–(51). What we want is not the generalization but the explanation. *Why can't A-bar movement proceed from vP_1?* Hallman's (2015) analysis does not answer the question. Furthermore, as mentioned, the structure is proposed in the earlier framework. In the following section, I propose that structure (46) accounts for both the non-extractability of the indirect object and the extractability of the direct object under labeling and maximality. Additionally, the indirect object becomes extractable in passives.

(53) John was given the book.

I will also show that this extractability is explained by structure (46).

5.3.6 [+P_{HAVE}] = a criterial feature

Adapting Rizzi's (2006, 2015a, b, 2016) subject criterion (and maximality) discussed in section 5.3.3, I propose that P_{HAVE} is a criterial feature like Subj.

(54) A Proposal:
[+P_{HAVE}] is a criterial feature.

The proposal is straightforward when we consider prospective possession which is the core property of P_{HAVE}. Since [X [P_{HAVE} Y]] is roughly interpreted as X HAVE Y, the subject-predicate interpretation holds. Namely, X is interpreted as something like the subject, which has the aboutness property. This view is supported by the following facts:

(55) a. John sent the girl a letter.
b. #John sent a girl the letter.

(Oba (2016: 207))

When we compare the two examples above, the former is more natural than the latter since the indirect object is generally interpreted as old information in the double object construction. This is explained by the aboutness property. The indirect object is taken as "being about" that argument. Because criterial features are related to discourse properties, it is reasonable to assume that (54) is true.

Additionally, as reviewed in section 5.3.3, Subj occurs lower than C, although CP's left periphery is typically rich in the standard cartographic approach (see Rizzi (1997), among others). Given the parallelism between the CP and vP phases, it is plausible that the criterial head P_{HAVE} occurs lower than v.

(56) CP phase:
… C … Subj …

(57) vP phase:
… v … P_{HAVE} …

In sum, because P_{HAVE} denotes the subject-predicate interpretation bearing prospective possession, it qualifies as a criterial feature like [+Subj]. In addition, the parallelism between vP and CP holds by assuming P_{HAVE} in vP.

5.3.7 A New Analysis

As shown in section 5.3.5, it is impossible to extract the indirect object in the double object construction. This non-extractability is readily explained by proposal (54). Adopting Harley and Jung's (2015) analysis, the structure of the vP phase for the double object construction is as follows (IO = indirect object, DO = direct object, and R is a verbal root whose categories are underspecified):

(58) [EA [$_{v}$ v-R [$_{\alpha}$ IO [$_{P_{HAVE}}$ P_{HAVE} DO]]]]

I assume that IO bears [$+P_{HAVE}$] in terms of labeling. Without assuming this, α cannot be labeled because both IO and [$_{P_{HAVE}}$ P_{HAVE} DO] are maximal objects. Recall that labeling [XP YP] fails unless either XP or YP moves, or unless X and Y share the same feature. The assumption is therefore tenable.

(59) [EA [$_{v}$ v-R [$_{P_{HAVE}}$ IO$_{[+P_{HAVE}]}$ [$_{P_{HAVE}}$ P_{HAVE} DO]]]]

Crucially, proposal (54) indicates that [+$\mathrm{P}_{\mathrm{HAVE}}$] on IO is a criterial feature. Given (59), the structure for *John sent the boarder a package* is as follows (*the boarder*$_{[+\mathrm{P_{HAVE}}]}$ means that the nominal bears the feature [+$\mathrm{P}_{\mathrm{HAVE}}$]):

(60) [C [John … T … [send [$_{\mathrm{P_{HAVE}}}$ [$_{\mathrm{P_{HAVE}}}$ the boarder$_{[+\mathrm{P_{HAVE}}]}$] [$_{\mathrm{P_{HAVE}}}$ $\mathrm{P}_{\mathrm{HAVE}}$ [$_{\mathrm{N}}$ a package]]]]]]

In this structure, a criterial feature [+$\mathrm{P}_{\mathrm{HAVE}}$] is assigned to the indirect object *the boarder*. When *the boarder*$_{[+\mathrm{P_{HAVE}}]}$ merges with [$_{\mathrm{P_{HAVE}}}$ $\mathrm{P}_{\mathrm{HAVE}}$ [$_{\mathrm{N}}$ a package]], they share the criterial feature [+$\mathrm{P}_{\mathrm{HAVE}}$]. After a criterial feature is shared, the highest object with the label of the shared criterial feature becomes the maximal object. In contrast, feature sharing degrades each object, which was maximal, with the label of the criterial feature to the X-bar objects. Consequently, [$_{\mathrm{P_{HAVE}}}$ [$_{\mathrm{P_{HAVE}}}$ the boarder$_{[+\mathrm{P_{HAVE}}]}$] [$_{\mathrm{P_{HAVE}}}$ $\mathrm{P}_{\mathrm{HAVE}}$ [$_{\mathrm{N}}$ a package]]]] becomes the maximal object with the given label $\mathrm{P}_{\mathrm{HAVE}}$ under the labeling algorithm and maximality. In other words, both *the boarder*$_{[+\mathrm{P_{HAVE}}]}$ and [$_{\mathrm{P_{HAVE}}}$ $\mathrm{P}_{\mathrm{HAVE}}$ [$_{\mathrm{N}}$ a package]] become X-bar objects with the given label $\mathrm{P}_{\mathrm{HAVE}}$. This means that *the boarder* (and [$_{\mathrm{P_{HAVE}}}$ $\mathrm{P}_{\mathrm{HAVE}}$ [$_{\mathrm{N}}$ a package]]) cannot move further under the maximality principle. Namely, *the boarder* is frozen because it is not the maximal object with the given label $\mathrm{P}_{\mathrm{HAVE}}$. All of the bad cases in (47)–(51) are therefore explained under labeling and maximality without postulating a new mechanism. The indirect object is non-extractable from the structure (59).

Contrary to the case of the indirect object, the direct object is extractable as in (47)–(51). This extractability is also explained by structures (59)–(60) under labeling and maximality. The direct object *a package* in (60) is the maximal object with the given label N.[9] Given the maximality principle, such objects are extractable. All of the good cases in (59)–(60) are therefore explained.

---

[9] This label may not be N but may be other features such as D, Subj and Top, depending on which features the direct object bears. As for nominal structures under the labeling theory, see Oishi (2015).

As mentioned, it has been argued in the literature that the indirect object becomes extractable in passive voice.[10]

(61) John was given the book.

This case could be a counter-example to the analysis, but it is not necessarily so. If we assume that P_{HAVE} raises (or internally pair-merges) to v-R, this case becomes unproblematic.

(62) $[_{v}\ P_{HAVE}\text{-v-R}\ [IO_{[+P_{HAVE}]}\ t_{P_{HAVE}}\ DO]]$

There is no feature sharing between $IO_{[+P_{HAVE}]}$ and $t_{P_{HAVE}}$ since the lower copy is, by definition, invisible.[11] Hence, the IO can be moved further because it is the maximal object with the given label P_{HAVE}.

---

[10] In passive voice, the IO should move from the predicate internal position in terms of labeling. If the IO remains in situ, labeling cannot correctly take place in the TP space. That is, the IO should move to SPEC-T because T alone is too weak to serve as a label. Given the weakness of T, Subj could be weak since SPEC-Subj is the canonical position for subjects under Rizzi's (2015a, b, 2016) cartographic approach. See also note 3. In addition to passives, we can extract the IO but not the DO in the multiple wh-question (cf. Larson (1988: 336–337) and Hornstein (1995: 127)).

(i) Who did you give what/which check?
(ii) * What/which check did you give who?

(i, ii: Hornstein (1995: 127))

It is clear that the multiple wh-phrases affect the extractability. Accordingly, it seems that the extractability, in this case, is not explained by the structures of the double object construction and the dative construction alone. Rather, it should be explained by the structures and the mechanism of the multiple wh-question. Additional research is required on this topic.

[11] According to Epstein, Kitahara and Seely (2016), the lower copy left by internal pair-Merge of heads (i.e. head-raising) is visible. However, following Nomura (2017), I simply assume that all lower copies are invisible under the definition: "α to be 'in the domain D' if and only if every occurrence of α is a term of D" (Chomsky (2013: 44)). See Chomsky (2015a) for more on this matter.

As is well known, the IO and the DO are both extractable from the dative construction.

(63) $[_{v}$ P_{HAVE}-v-R $[IO_{[+P_{HAVE}]}$ $t_{P_{HAVE}}$ DO]]

It is straightforward because P_{HAVE} does not occur in the construction, but the typical or standard P (i.e. *to* in (63)) does. Neither SPEC-P nor COMPL-P are criterial positions.

5.3.8 Featural Relativized Minimality and Sub-Extraction

Runner (2001) observes that the indirect object cannot be sub-extracted.

(64) * Who did you say Cindy sent a friend of t_{wh} a picture?

(Runner (2001: 40))

This empirical fact leads to an interesting consequence about the formalization of $[+P_{HAVE}]$ based on fRM (featural Relativized Minimality). That is, $[+P_{HAVE}]$ belongs to the Operator class.

As Rizzi (2015a: 28) notes, fRM "assumes that relevant morphosyntactic features triggering movement are organized into feature classes along the following lines:"

(65) Argumental: **Subj**, person, number, gender, case, …
Operator: **Q**, Foc, Neg, Quantificational adverbials, …
Modifier: …
Topic: …

(adapted from Rizzi (2015a: 28))

According to Rizzi (2015a), fRM explains the locality problem: the closest nominal would be attracted. As Rizzi (2015a: 27) notes, "[o]ne salient property of chains terminating in Spec Subj is that they are strictly local, with the closest nominal element systematically attracted to Spec Subj." Let us consider the following structures with respect to Foc and Subj:

(66) a. Foc … [DP v [V $DP_{[+Foc]}$]]

b. Subj … [DP v [V $DP_{[+Subj]}$]]

(Rizzi (2015a: 28))

In (66a), Foc can attract $DP_{[+Foc]}$ since there is no intervening element between the heads. On the other hand, Subj cannot attract $DP_{[+Subj]}$ because DP at SPEC-v is a possible intervener that bears "Argumental" features (e.g. person, number, gender). The locality of subjects is thus explained by fRM.

As shown in the previous section, [+Subj] belongs to the Argumental class in (65). However, what class does [+P_{HAVE}] belong to? It seems straightforward that it belongs to the same class as [+Subj] because its semantic property is similar to that of [+Subj]. This feature classification, however, leads us to an unwelcome result concerning the case of sub-extraction. Let us consider the structure of (64) under the proposed analysis:

(67) *[$_{Q}$ $Who_{[+Q]}$ [Q you say Cindy sent [$_{P_{HAVE}}$ [$_{P_{HAVE}}$ a friend of $t_{wh[+Q]}$] [P_{HAVE} a picture]]]]

(*Who did you say Cindy sent a friend of t_{wh} a picture?)

Here, *who* must be sub-extractable if the feature classification is tenable, but the fact is opposite; it is no longer tenable. Nevertheless, the structure of (67) still seems correct. The problem of sub-extraction can be solved by assuming that the feature class of [+P_{HAVE}] is not Argumental but Operator. That is, I assume [+P_{HAVE}] and Q belong to the same feature class.

(68) Argumental: **Subj**, person, …

Operator: **P_{HAVE}**, **Q**, Foc, …

The revised classification correctly explains the derivation of (67). Q fails when it tries to attract *who*$_{[+Q]}$. The label P_{HAVE} on *a friend of who*$_{[+Q]}$ blocks Q's attraction because P_{HAVE} and Q both belong to the same "Operator" class.

The semantic nature of Operator [+P_{HAVE}] is obscure, but it is possible to interpret the argument structure of the double object construction as follows (R_{send} is the verbal root of *send*):

(69) [John [v_{cause}-R_{send} [$_{P_{HAVE}}$ [$_{P_{HAVE}}$ Mary$_{[+P_{HAVE}]}$] [$_{P_{HAVE}}$ P_{HAVE} [the book]]]]]

↓

for PROSPECTIVE POSSESSOR *x*, *x* Mary, John CAUSE *x* to HAVE the book by sending.

I will leave further refinement of the semantic nature of Operator [+P_{HAVE}] to future research.

In short, the following generalization holds: P_{HAVE} is semantically similar to Subj, but its feature class is the same as Q (and other features like Foc, Neg). Although how this feature classification affects the semantic nature is vague, the problem not only applies to P_{HAVE} but also to other heads such as Foc, Neg and Quantificational adverbials. We need further studies to unveil the detailed properties of P_{HAVE} and the feature classes.

5.3.9 Problems of Alternative 2

Section 5.3 briefly reviewed the recent theoretical framework of syntax and discussed its consequences, particularly focusing on the double object construction in English. The main proposal in this paper is that [+P_{HAVE}] is a criterial feature. Given the proposal and structure by Harley and Jung (2015), labeling and maximality yield a new analysis of the non-extractability of IO. In addition, it has been suggested that P_{HAVE} is a criterial head belonging to the feature class, Operator. This suggestion eventually accounts for the non-sub-extractability of IO.

Despite these empirical advantages, alternative 2 in this section seriously violates the Inclusiveness Condition as mentioned in note 5. Again, criterial features such as Subj and P_{HAVE}, which are discourse-related, should not exist in the lexicon because discourse related properties seem to be obtained through structural configurations (see Chomsky, Gallego and Ott (to appear)). For the true explanation, alternative 2 should therefore be discarded. It is very challenging to explain the non-sub-extractability of indirect object without analogous analyses, but I think optimistically that additional research will solve this problem.

Chapter 6

Conclusion

This paper has explored the possibility of eliminating pair-Merge, feature inheritance and feature valuation as primitive or elementary operations in the quest for a minimalist explanation of language. Such operations must be removed for two reasons: simplicity for the methodological minimalism and evolvability for the SMT. The former is just a normal scientific procedure like Ockham's razor. The latter is more remarkable in the Minimalist Program for linguistic theory. Operations in core syntax are specific to human language as long as they are primitive. Many primitive operations raise a problem of evolvability. Namely, many primitive operations are too complex to explain the sudden emergence of (the basic properties of) language. To hold a simple picture, all we can do is postulate a sole and simple operation, Merge. Once we postulate Merge, it can be used to explain evolvability. Moreover, Merge ensures the basic properties of discrete infinity and displacement. The number of operations conforming to the first-factor genetic endowment should therefore be as few as possible if we assume that language is designed to be *perfect*.

Pair-Merge of heads, which was called head-adjunction in the earlier framework, needs to be reformulated because it is a primitive operation. In chapter 2, I argued that it was just Merge of heads. To reformulate it, I capitalized on a definition of ordered pairs in set theory. Building this definition into definitions of pair-Merge and simplest Merge, I proposed the following reformulation of pair-Merge:

(1) Pair-Merge of heads formulated by Simplest Merge (PM by SM):

a. $\text{Pair-Merge}(\alpha, \beta) = \text{Merge}(\alpha, (\alpha, \beta)) = \text{Merge}(\alpha, \{\alpha, \beta\}) = \{\alpha, \{\alpha, \beta\}\} = <\alpha, \beta>$

b. $\text{Pair-Merge}(\alpha, \beta) = \text{Merge}(\beta, (\alpha, \beta)) = \text{Merge}(\beta, \{\alpha, \beta\}) = \{\beta, \{\alpha, \beta\}\} = <\beta, \alpha>$

This reformulation, *PM by SM*, eliminates the primitive operation pair-Merge because PM by SM is essentially the double application of Merge. The problem of evaluability caused by the primitive operation, pair-Merge, is thus solved by this proposed reformulation. PM by SM, however, raises a new problem: there is no reason why PM by SM would be restricted to apply to heads or computational atoms. It cannot apply to other elements for empirical reasons. Such stipulations should be removed, and so they were removed in chapter 3.

Conceptually, PM by SM should apply to any element because it is, by definition, equal to simplest Merge. As mentioned above, an empirical reason bars this possibility. If we assume that PM by SM could apply to anything, it can would apply to an X^0-level head and an XP-level phrase as shown below:

(2) {John, {{√hit, {√hit, v*}}, {$_{\delta}$ Mary, {$_{\gamma}$ $t_{\sqrt{}\text{hit}}$, t_{Mary}}}}}

i. Merge externally forms {$_{\gamma}$ √hit, Mary}.

ii. Merge internally forms {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}}. (I.e., {$_{\delta}$ Mary, {$_{\gamma}$ √hit, t_{Mary}}} formed by PM by SM is representationally interpreted as the ordered pair <Mary, √hit>.)

iii. The derivation continues, but it crashes because of the derivational ordered pair $\{_\delta$ Mary, $\{_\gamma \sqrt{}$hit, $t_{Mary}\}\}$.

The first and second steps could be equal to PM by SM, since they can be assumed to be the double application of simplest Merge. However, it is trivial that $\{_\delta$ Mary, $\{_\gamma \sqrt{}$hit, $t_{Mary}\}\}$ is not an adjunct structure. For this reason, it seems that we should give up on applying PM by SM to phrases. Nevertheless, when we consider the *levelness* of members, pair-Merge of phrases can be reformulated under proposal (1). Trivially, the following generalization holds in set theory:

(3) A and α are *different-level* members where A is a set, but α is not.

Capitalizing on this generalization, PM by SM can apply to both two heads and to two phrases but not to a head and a phrase. Pair-Merge is, therefore, finally reformulated as follows:

(4) PM by SM (the final version):

<u>*head-head adjunction*</u>:

a. Pair-Merge (X^0, Y^0) = Merge $(X^0, (X^0, Y^0))$ = Merge $(X^0, \{X^0, Y^0\})$ = $\{X^0, \{X^0, Y^0\}\}$ = $\langle X^0, Y^0 \rangle$

b. Pair-Merge (Y^0, X^0) = Merge $(Y^0, (Y^0, X^0))$ = Merge $(Y^0, \{Y^0, X^0\})$ = $\{Y^0, \{Y^0, X^0\}\}$ = $\langle Y^0, X^0 \rangle$

<u>*phrase-phrase adjunction*</u>:

c. Pair-Merge (XP, YP) = Merge (XP, (XP, YP)) = Merge (XP, {XP, YP}) = {XP, {XP, YP}} = <XP, YP>

d. Pair-Merge (YP, XP) = Merge (YP, (YP, XP)) = Merge (YP, {YP, XP}) = {YP, {YP, XP}} = <YP, XP>

This version of PM by SM correctly explains not only the simple derivation in (2) but also a variety of empirical facts. PM by SM for head-head adjunction, for example, explains the (non)extractability of grammatical objects in double object constructions, cognate object constructions, small clause constructions and ECM constructions in chapter 2. PM by SM for phrase-phrase adjunction, for example, accounts for the (anti)-adjunction condition effects, the Condition C effects and the Specificity Effects in chapter 3.

Other operations remain to be reformulated: feature inheritance and feature valuation. In the original definition, feature inheritance and feature valuation are primitive operations. In the linguistic literature, e.g. Nomura (2017), Epstein, Kitahara and Seely (2014a) and Kitahara (2017), these operations are assumed not to be primitive but to be reducible to third-factor principles or Merge. Even if they are not primitive, feature inheritance cannot correctly occur under the current minimalist model adopting strictly cyclic Merge. That is feature inheritance is technically problematic since Merge does not establish C-T and v*-√root relations in the standard course of derivations (see (2)). In chapter 4, I reduced both operations to minimal search conforming to the third-factor Minimal Computation, and solved a technical problem. In this reformulation, I argued that the equation Feature Inheritance = Matching = Minimal Search holds because of the following reason:

(5) Feature Inheritance = Matching = Minimal Search:
Feature inheritance is minimal search by uPhi-set, which identifies the most appropriate goal; hence matching.

This proposal solves the technical problem mentioned above since minimal search for feature inheritance can establish the appropriate relations. Furthermore, I assume that feature valuation is equal to labeling by minimal search:

(6) Labeling = Feature Valuation = Minimal Search:
Labeling is minimal search, which identifies unlabeled syntactic objects and unvalued features; hence valuation.

This type of assumption is essential for eliminating feature valuation as a primitive operation. Therefore, both feature inheritance and feature valuation are minimal search. Chapter 4 showed that PM by SM and these reformulations can explain various facts involving *there* constructions such as long-standing agreements, (non)-extractability of associates and scopal and binding phenomena.

These reformulations of pair-Merge, feature inheritance and feature valuation are not only conceptually significant in terms of the SMT but also empirically important for its empirical coverage. Some of the empirical facts, however, can be explained by other possible alternatives. In chapter 5, I also showed that two alternatives can explain some or more facts explained by the reformulations. However, the main argument of this chapter was to show that the analyses, based on PM by SM and the other reformulations, were superior to these alternatives by refuting both alternatives (though they are, to some extent, appealing).

We undoubtedly need additional research to clarify the true nature of language, , but things like the reformulations proposed in this paper are essential in terms of simplicity, evolvability and the SMT. For example, the newest talk given by Chomsky (2017b) reformulated *Merge*. Further developments of minimalist theory are clearly required to approach the ultimate goal of the Minimalist Program.

References

Abe, Jun (2016) "How to Probe Expletives," *Studia Linguistica*. <doi: 10.1111/stul.12057>

Acedo-Matellán, Víctor (2010) *Argument Structure and the Syntax-Morphology Interface: A Case Study in Latin and Other Languages*, Doctoral dissertation, Universitat de Barcelona.

Acedo-Matellán, Víctor and Jaume Mateu (2014) "From Syntax to Roots: A Syntactic Approach to Root Interpretation," *The Syntax of Roots and the Roots of Syntax*, ed. by Artemis Alexiadou, Hagit Borer and Florian Schäfer, 14–32, Oxford University Press, Oxford.

Alexiadou, Artemis and Florian Schaefer (2011) "An Unaccusativity Diagnostic at the Syntax-Semantics Interface: *There*-Insertion, Indefinites and Restitutive *again*," *Proceedings of Sinn and Bedeutung* 15, 101–115.

Bagaria, Joan (2014) "Basic Set Theory," *Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta, The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, Stanford. Retrieved October 5, 2016, from <http://plato.stanford.edu/entries/set-theory/basic-set-theory.html>

Borer, Hagit (2003) "Exo-Skeletal vs. Endo-Skeletal Explanations: Syntactic Projections and the Lexicon," *The Nature of Explanation in Linguistic Theory*, ed. by John Moore and Maria Polinsky, 31–67, CSLI Publications, Stanford, California.

Borgonovo, Claudia and Ad Neeleman (2000) "Transparent Adjuncts," *Canadian Journal of Linguistics* 45, 199–224.

Bošković, Željko (1997) *The Syntax of Nonfinite Complementation: An Economy Approach*, MIT Press, Cambridge, MA.

Bošković, Željko (2016) "On the Timing of Labeling: Deducing Comp-Trace Effects, the Subject Condition, the Adjunct Condition, and Tucking in from Labeling," *The Linguistic Review* 33, 17–66.

Bowers, John (1987) "Extended X-Bar Theory, the ECP and the Left Branch Condition," *WCCFL* 6, 47–62.

Carstens, Vicki, Norbert Hornstein and T. Daniel Seely (2016) "Head-Head Relations in Problems of Projection," *The Linguistic Review* 33, 67–86.

Cecchetto, Carlo and Caterina Donati (2015) *(Re)labeling*, MIT Press, Cambridge, MA.

Chomsky, Noam (1973) "Conditions on Transformations," *A Festschrift for Morris Halle*, ed. by Stephen R. Anderson and Paul Kiparsky, 232–286, Holt, Rinehart and Winston, New York.

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

Chomsky, Noam (1986) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.

Chomsky, Noam (1995a) "Bare Phrase Structure," *Government and Binding Theory and the Minimalist Program*, ed. by Gert Webelhuth, 385–439, Blackwell, Oxford.

Chomsky, Noam (1995b) *The Minimalist Program*, MIT Press, Cambridge, MA.

Chomsky, Noam (2000) "Minimalist Inquiries: The Framework," *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels and Juan Uriagereka, 89–155, MIT Press, Cambridge, MA.

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

Chomsky, Noam (2004) "Beyond Explanatory Adequacy," *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*, ed. by Adriana Belletti, 104–131, Oxford University Press, Oxford.

Chomsky, Noam (2005) "Three Factors in Language Design," *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam (2007) "Approaching UG from Below," *Interface + Recursion = Language?: Chomsky's Minimalism and the View from Syntax-Semantics*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29, Mouton de Gruyter, Berlin.

Chomsky, Noam (2008) "On Phases," *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, MIT Press, Cambridge, MA.

Chomsky, Noam (2010) "Some Simple Evo Devo Theses: How True Might They Be for Language?," *The Evolution of Human Language: Biolinguistic Perspectives*, ed. by Richard K. Larson, Viviane Déprez and Hiroko Yamakido, 45–62, Cambridge University Press, New York.

Chomsky, Noam (2012) *The Science of Language: Interviews with James McGilvray*, Cambridge University Press, New York.

Chomsky, Noam (2013) "Problems of Projection," *Lingua* 130, 33–49.

Chomsky, Noam (2014) "Minimal Recursion: Exploring the Prospects," *Recursion: Complexity in Cognition*, ed. by Tom Roeper and Margaret Speas, 1–15, Springer, New York.

Chomsky, Noam (2015a) "Problems of Projection: Extensions," *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann and Simona Matteini, 3–16, John Benjamins, Amsterdam.

Chomsky, Noam (2015b) "A Discussion with Naoki Fukui and Mihoko Zushi," *Sophia Linguistica* 64, 69–97, Sophia University.

Chomsky, Noam (2016a) "The Language Capacity: Architecture and Evolution," *Psychonomic Bulletin & Review*. <doi:10.3758/s13423-016-1078-6> [Chomsky, Noam (2017) "The Language Capacity: Architecture and Evolution," *Psychonomic Bulletin & Review* 24, 200–203.]

Chomsky, Noam (2016b) "Puzzles about Phases," ms., MIT.

Chomsky, Noam (2016c) "A Dialogue with Noam Chomsky," A talk given at Residència d' Investigadors, Centrede Lingüística Teòrica de la Universitat Autònoma de Barcelona.
< https://www.youtube.com/watch?v=ZQa2X88PtYY >

Chomsky, Noam (2017a) "Language Architecture and Its Import for Evolution," *Neuroscience and Biobehavioral Reviews*. <http://dx.doi.org/10.1016/j.neubiorev.2017.01.053>

Chomsky, Noam (2017b) "Chomsky's READING Lecture," a talk given at the University of Reading.
<https://www.facebook.com/search/top/?init=quick&q=University%20of%20Reading%20chomsky&tas=0.21866556606255472>

Chomsky, Noam, Ángel J. Gallego and Dennis Ott (to appear) "Generative Syntax: Questions, Crossroads, and Challenges," Special Issue of *Catalan Journal of Linguistics*. Retrieved June 30, 2017, from <https://ling.auf.net/lingbuzz/003507>

Citko, Barbara (2005) "On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge," *Linguistic Inquiry* 36, 475–496.

Collins, Chris (2002) "Eliminating Labels," *Derivation and Explanation in the Minimalist Program*, ed. by Samuel D. Epstein and T. Daniel Seely, 42–64, Blackwell, Oxford.

Collins, Chris (2014) "Merge (X, Y) = {X, Y}," ms., New York University. Retrieved July 6, 2016, from <http://ling.auf.net/lingbuzz/002186>

Deal, Amy Rose (2009) "The Origin and Content of Expletives: Evidence from 'Selection,'" *Syntax* 12, 285–323.

den Dikken, Marcel (2012) "The Phase Impenetrability Condition and Successive Cyclicity: A Reconsideration," ms., City University of New York.

Diesing, Molly (1992) *Indefinites*, MIT Press, Cambridge, MA.

Embick, David and Alec Marantz (2008) "Architecture and Blocking," *Linguistic Inquiry* 39, 1–53.

Embick, David and Rolf Noyer (2007) "Distributed Morphology and the Syntax-Morphology Interface," *The Oxford Handbook of Linguistic Interfaces*, ed. by Gillian Ramchand and Charles Reiss, 289–324, Oxford University Press, Oxford.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2012) "Structure Building That Can't Be," *Ways of Structure Building*, ed. by Myriam Uribe-Etxebarria and and Vidal Valmala, 253–270, Oxford University Press, Oxford. [Reprinted in *Explorations in Maximizing Syntactic Minimization*, ed. by Samuel D. Epstein, Hisatsugu Kitahara and T. Daniel Seely, 2015, 155–174, Routledge, New York.]

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2013) "Simplest Merge Generates Set Intersection: Implications for Complementizer-Trace Explanation," *The Proceedings of GLOW in Asia XII*, 77–92. [Reprinted in *Explorations in Maximizing Syntactic Minimization*, ed. by Samuel D. Epstein, Hisatsugu Kitahara and T. Daniel Seely, 2015, 175–194, Routledge, New York.]

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2014a) "Labeling by Minimal Search: Implications for Successive-Cyclic A-Movement and the Conception of the Postulate 'Phase,'" *Linguistic Inquiry* 45, 463–481. [Reprinted in *Explorations in Maximizing Syntactic Minimization*, ed. by Samuel D. Epstein, Hisatsugu Kitahara and T. Daniel Seely, 2015, 201–221, Routledge, New York.]

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2014b) "Agreement Is the Shared Prominent Feature Option of POP's Labeling Algorithm," ms., University of Michigan, Keio University and Eastern Michigan University.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2015) "What Do We Wonder Is Not Syntactic?," *Explorations in Maximizing Syntactic Minimization*, ed. by Samuel D. Epstein, Hisatsugu Kitahara and T. Daniel Seely, 222–239, Routledge, New York.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2016) "Phase-Cancellation by Pair-Merge of Heads," *The Linguistic Review* 33, 87–102.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2017) "Successive Cyclic Wh-Movement without Successive Cyclic Crashing," paper presented at ELSJ 10th International Spring Forum.

Epstein, Samuel D., Miki Obata and T. Daniel Seely (to appear) "Is Linguistic Variation Entirely Linguistic?," *Linguistic Analysis*.

Fiengo, Robert and James Higginbotham (1981) "Opacity in NP," *Linguistic Analysis* 7, 395–421.

Frampton, John and Sam Gutmann (2002) "Crash-Proof Syntax," *Derivation and Explanation in the Minimalist Program*, ed. by Samuel D. Epstein and T. Daniel Seely, 90–105, Blackwell, Oxford.

Gallego, Ángel J. (2009) "On Freezing Effects," *Iberia* 1, 33–51.

Goto, Nobu (2016) "A Generalization of Non-Phase Heads and a New Rationale for Feature Inheritance: In Pursuit of the Universality of Language," paper presented at the workshop "Exploring the Notion of Parameters in the Minimalist Program" in the 153th conference of the English Linguistic Society of Japan.

Goto, Nobu (2017) "How to Label *There*-Constructions," *Studies in English Literature: Regional Branches Combined Issue* 9, 33–43. <http://doi.org/10.20759/elsjregional.9.0_33>

Green, Georgia M. (1974) *Semantics and Syntactic Regularity*, Indiana University Press, Indiana.

Hale, Ken and Samuel J. Keyser (1993) "On Argument Structure and the Lexical Expression of Syntactic Relations," *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 53–109, MIT Press, Cambridge, MA.

Hale, Ken and Samuel J. Keyser (2002) *Prolegomenon to a Theory of Argument Structure*, MIT Press, Cambridge, MA.

Halle, Morris and Alec Marantz (1993) "Distributed Morphology and the Pieces of Inflection," *The View from Building 20*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176, MIT Press, Cambridge, MA.

Hallman, Peter (2015) "Syntactic Neutralization in Double Object Constructions," *Linguistic Inquiry* 46, 389–424.

Harley, Heidi (1995) *Subjects, Events and Licensing*, Doctoral dissertation, MIT.

Harley, Heidi (2002) "Possession and the Double Object Construction," *Linguistic Variation Yearbook* 2, 29–68.

Harley, Heidi (2011) "A Minimalist Approach to Argument Structure," *The Oxford Handbook of Linguistic Minimalism*, ed. by Cedric Boeckx, 426–447, Oxford University Press, Oxford.

Harley, Heidi and Hyun Kyoung Jung (2015) "In Support of the P_{HAVE} Analysis of the Double Object Construction," *Linguistic Inquiry* 46, 703–730.

Hartmann, Jutta (2008) *Expletives in Existentials: English there and German da*, Doctoral dissertation, Universiteit van Tilburg.

Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch (2002) "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?," *Science* 298, 1569–1579. [Reprinted in *The Evolution of Human Language: Biolinguistic Perspectives*, ed. by Richard K. Larson, Viviane Déprez and Hiroko Yamakido, 2010, 14–42, Cambridge University Press, New York.]

Hauser, Marc D., Charles Yang, Robert C. Berwick, Ian Tattersall, Michael J. Ryan, Jeffrey Watumull, Noam Chomsky and Richard C. Lewontin (2014) "The Mystery of Language Evolution," *Frontiers in Psychology* 5, 1–12.

Horita, Yuko (1996) "English Cognate Object Constructions and Their Transitivity," *English Linguistics* 13, 221–247.

Hornstein, Norbert (1995) *Logical Form: From GB to Minimalism*, Blackwell, Oxford.

Hornstein, Norbert (2009) *A Theory of Syntax: Minimal Operations and Universal Grammar*, Cambridge University Press, New York.

Huang, C.-T. James (1982) *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT.

Jackendoff, R. (1977) *X Syntax: A Study of Phrase Structure*, MIT Press, Cambridge, MA.

Johnson, Kyle (2003) "Towards an Etiology of Adjunct Islands," *Nordlyd* 31: 187–215.

Julien, Martin (2002) "On the Syntax of 'Th/Ex'," paper presented at the conference "The Displacement Property of Human Language" in University Tromsø.

Karimi, Simin (1999) "Specificity Effect: Evidence from Persian," *The Linguistic Review* 16, 125–141.

Kasai, Hironobu, Takaomi Kato, Hiroki Narita, Naoki Fukui and Mihoko Zushi (2016) "Remarks on the Primitive Operations and Interface Conditions of Syntax," paper presented at the symposium "Tojihou no Genshienzan to Interface (The Primitive Operations of Syntax and Interfaces)" in the 88th Conference of the English Literary Society of Japan.

Kato, Takaomi, Hiroki Narita, Hironobu Kasai, Mihoko Zushi and Naoki Fukui (2016) "On the Primitive Operations of Syntax," *Advances in Biolinguistics*, ed. by Koji Fujita and Cedric Boeckx, 29–45, Routledge, New York.

Kawashima, Ruriko and Hisatsugu Kitahara (2015) "On the Core Ideas of Generative Grammar: Then and Now," *In Untiring Pursuit of Better Alternatives*, ed. by Hiroki Egashira, Hisatsugu Kitahara, Kazuo Nakazawa, Tadao Nomura, Masayuki Oishi, Akira Saizen and Motoko Suzuki, 198–208, Kaitakusha, Tokyo.

Kayne, Richard (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, MA

Kayne, Richard (2004) "Prepositions as Probes," *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*, ed. by Adriana Belletti, 192–212, Oxford University Press, Oxford.

Kayne, Richard (2011) "Antisymmetry and the Lexicon," *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, ed. by Anna Maria Di Sciullo and Cedric Boeckx, 329–353, Oxford University Press, Oxford.

Kitahara, Hisatsugu (2017) "Maximizing Syntactic Minimization: Five Preliminary Studies," *Reports of the Keio Institute of Cultural and Linguistic Studies* 48, 239–265.

Kratzer, Angelica (1996) "Severing the External Argument from Its Verb," *Phrase Structure and the Lexicon*, ed. by Johan Rooryck and Laurie Zaring, 109–137, Kluwer Academic Publishers, Dordrecht.

Kuno, Susumu and Ken-ichi Takami (2004) *Functional Constraints in Grammar: On the Unergative-Unaccusative Distinction*, John Benjamin, Amsterdam.

Kuratowski, Casimir (1921) "Sur la Notion de l'Ordre dans la Théorie des Ensembles," *Fundamenta Mathematicae* 2, 161–171.

Larson, Richard (1988) "On the Double Object Construction," *Linguistic Inquiry* 19, 335–391.

Lasnik, Howard, and Mamoru Saito (1984) "On the Nature of Proper Government," *Linguistic Inquiry* 15, 235–289.

Lasnik, Howard and Mamoru Saito (1992) *Move α: Conditions on Its Application and Output*, MIT Press, Cambridge, MA.

Lebeaux, David (2000) *Language Acquisition and the Form of the Grammar*, John Benjamins, Amsterdam.

Levin, Beth (1993) *English Verb Classes and Alternations: A Preliminary Investigation*, the University of Chicago Press, Chicago.

Lumsden, Michael (1988) *Existential Sentences*, Croom Helm, London.

Macfarland, Talke (1995) *Cognate Objects and the Argument/Adjunct Distinction in English*, Doctoral dissertation, Northwestern University.

Marantz, Alec (1981) *On the Nature of Grammatical Relations*, Doctoral dissertation, MIT.

Marantz, Alec (1997) "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon," *University of Pennsylvania Working Papers in Linguistics* 4, 201–225.

Marantz, Alec (2005) "Objects out of the Lexicon!: Argument-Structure in the Syntax," ms., MIT.

Marantz, Alec (2013) "Verbal Argument Structure: Events and Participants," *Lingua* 130, 152–168.

Mateu, Jaume (1997) "Universals of Semantic Construal for Lexical Syntactic Relations," GGT-94 Research Report, Universitat Autònoma de Barcelona. Retrieved April 14, 2015, from <http://filcat.uab.cat/clt/publicacions/reports/pdf/GGT-99-4.pdf>

Mateu, Jaume (2002) *Argument Structure: Relational Construal at the Syntax-Semantics Interface*, Doctoral dissertation, Universitat Autònoma de Barcelona.

Mateu, Jaume (2010) "On the L-Syntax of Mannar and Causation," GGT-09-09 Research Report, Universitat Autònoma de Barcelona. Retrieved March 14, 2015, from <http://filcat.uab.cat/clt/publicacions/reports/pdf/GGT-09-09.pdf>

Mateu, Jaume (2012) "Conflation and Incorporation Processes in Resultative Constructions," GGT-11-03 Research Report, Universitat Autònoma de Barcelona. Retrieved March 14, 2015, from <http://filcat.uab.es/clt/publicacions/reports/pdf/GGT-11-03.pdf>

Mateu, Jaume (2014) "Argument Structure," *The Routledge Handbook of Syntax*, ed. by Andrew Carnie, Dan Siddiqi and Yosuke Sato, 24–41, Routledge, New York.

May, Robert (1985) *Logical Form: Its Structure and Derivation*, MIT Press, Cambridge, MA.

Milsark, Gary (1974) *Existential Sentences in English,* Doctoral dissertation, MIT.

Milsark, Gary (1977) "Toward an Explanation of Certain Peculiarities of the Existential Construction in English," *Linguistic Analysis* 3, 1–29.

Miyamoto, Yoichi (2012) "On Transparent Adjuncts in Japanese," *Ways of Structure Building*, ed. by Myriam Uribe-Etxebarria and Vidal Valmala, 330–365, Oxford University Press, Oxford.

Mizuguchi, Manabu (2016) "De-activation of *Φ* through Pair-Merge," *English Linguistics* 33, 108–118.

Moro, Andrea (1997) "Dynamic Antisymmetry: Movement as a Symmetry-Breaking Phenomenon," *Studia Linguistica* 51, 50–76.

Munataka, Takashi (2017) "Why Is Labeling Important?," ms., Yokohama National University.

Nakajima, Heizo (1997) "A Generative View of the Cognitive Analysis of Raising," *Studies in English Linguistics*, ed. by Masato Ukaji, Toshio Nakao, Masaru Kajita, and Syuji Chiba, 474–491, Taishukan, Tokyo.

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

Nishihara, Toshiaki (1999) "On Locative Inversion and *There*-Construction," *English Linguistics* 16, 381–404.

Nomura, Masashi (2003) "Expletive Move!," *Proceedings of the 32nd Western Conference on Linguistics*, 207–220.

Nomura, Masashi (2017) “Pair-Merge and Feature-Valuation via Minimal Search: Evidence from Icelandic,” *WCCFL* 34, 395–403.

Nunes, Jairo (1995) *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*, Doctoral dissertation, University of Maryland.

Nunes, Jairo and Juan Uriagereka (2000) “Cyclicity and Extraction Domains,” *Syntax* 3, 20–43.

Oba, Yukio (1999) “Nominal Expressions as a Phase,” *Currents in Linguistic Research: A Festschrift for Professor Kazuyuki Yamamoto on the Occasion of His Retirement from Yamaguchi University*, ed. by Toshiaki Inada, Kozo lwabe, Yukio Oba, Masanori Suiko, Masashi Takemoto and Hdeo Nishimura, 21–36, Kaitakusha, Tokyo.

Oba, Yukio (2003) “Tokuteiseikouka to Phase fukashinjouken (The Specificity Effect and the Phase Impenetrability Condition),” *Ichikawashou 36nen no Kiseki*, ed. by Yuzaburo Murata, 202–210, Kaitakusha, Tokyo.

Oba, Yukio (2005) “The Double Object Construction and Thematization/Extraction,” *English Linguistics* 22, 56–81.

Oba, Yukio (2013) “Eigo no Douzoku Mokutekigo Koubun no Imitougoteki na Tokucho ni tsuite (On Syntactic and Semantic Properties of the Cognate Object Construction in English),” *Gengogaku karano Tenbou 2013*, ed. by Fukuoka Linguistic Circle, 66–80, Kyushu University Press, Fukuoka.

Oba, Yukio (2016) “Nijuu Mokutekigo Koubun no Kansetsu Mokutekigo no Tokusei to Sono Kouzou ni tsuite (The Structure of the Double Object Construction and Properties of the Indirect Object),” *Gengogaku no Ima wo Shiru 26 Kou* (*26 Essays in Current Linguistic Research*), ed. by Akira Kikuchi, Takamichi Aki, Toru Suzuki, Naoto Tomizawa, Tatsuya Yamagishi and Shin-ichi Kitada, 199–209, Kenkyusha, Tokyo.

Obata, Miki (2010) *Root, Successive-Cyclic and Feature-Splitting Internal Merge: Implications for Feature-Inheritance and Transfer*, Doctoral dissertation, University of Michigan.

Oishi, Masayuki (2015) “The Hunt for a Label,” *In Untiring Pursuit of Better Alternatives*, ed. by Hiroki Egashira, Hisatsugu Kitahara, Kazuo Nakazawa, Tadao Nomura, Masayuki Oishi, Akira Saizen and Motoko Suzuki, 322–334, Kaitakusha, Tokyo.

Omune, Jun (2016) “On Syntactic Structures of Inside and Outside Verbal Existential Sentences,” *JELS* 33, 114–120.

Omune, Jun (2017) “Reformulating Feature Inheritance,” paper presented at the 2nd regular meeting of Fukuoka Linguistic Circle.

Omune, Jun (to appear-a) “Reformulating Pair-Merge of Heads,” *English Linguistics* 34.

Omune, Jun (to appear-b) “Some Consequences of Labeling and Maximality,” *Kinki University (Kindai University), Center for Liberal Arts and Foreign Language Education Journal, Foreign Language Edition* 8.

Oseki, Yohei (2015) “Eliminating Pair-Merge,” *WCCFL* 32, 303–312.

Ott, Dennis (2011) *Local Instability*, Doctoral dissertation, Harvard University.

Pinker, Steven (2013) *Learnability and Cognition: The Acquisition of Argument Structure, New Edition*, MIT Press, Cambridge, MA.

Postal, Paul (1974) *On Raising: An Inquiry into One Rule of English Grammar and Its Theoretical Implications*, MIT Press, Cambridge, MA.

Postal, Paul (1993) “Some Defective Paradigms,” *Linguistics Inquiry* 24, 347–364.

Quirk, Randolph et al. (1985) *A Comprehensive Grammar of the English Language*, Longman, London.

Richards, Marc (2006) “Object Shift, Phases, and Transitive Expletive Constructions in Germanic,” *Linguistic Variation Yearbook* 6, 139–159.

Richards, Marc (2007) “On Feature Inheritance: An Argument from the Phase Impenetrability Condition,” *Linguistic Inquiry* 38, 563–572.

Richards, Marc (2009) “Internal Pair-Merge: The Missing Mode of Movement,” *Catalan Journal of Linguistics* 8, 55–73.

Richards, Marc (2012) “On Feature Inheritance, Defective Phases, and the Movement-Morphology Connection,” *Phases: Developing the Framework*, ed. by Angel J. Gallego, 195–232, Mouton de Gruyter, Berlin.

Richards, Marc and Theresa Biberauer (2005) “Explaining Expl,” *The Function of Function Words and Functional Categories*, ed. by Marcel den Dikken and Christina M. Tortora, 115–153, John Benjamins, Amsterdam.

Rizzi, Luigi (1997) “The Fine Structure of the Left Periphery,” *Elements of Grammar: Handbook in Generative Syntax*, ed. by Liliane Haegeman, 281–337, Kluwer, Dordrecht.

Rizzi, Luigi (2006) “On the Form of Chains: Criterial Positions and ECP Effects,” *Wh-Movement: Moving On*, ed. by Lisa Cheng and Norbert Carver, 97–133, MIT Press, Cambridge, MA.

Rizzi, Luigi (2010) “On Some Properties of Criterial Freezing,” *The Complementizer Phase: Subjects and Operators*, ed. by E. Phoevos Panagiotidis, 51–75, Oxford University Press, Oxford.

Rizzi, Luigi (2015a) “Notes on Labeling and Subject Positions,” *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann and Simona Matteini, 17–46, John Benjamins, Amsterdam.

Rizzi, Luigi (2015b) “Cartography, Criteria, and Labeling,” *Beyond Functional Sequence: The Cartography of Syntax, Volume 10*, ed. by Ur Shlonsky, 314–338, Oxford University Press, New York.

Rizzi, Luigi (2016) “Labeling, Maximality and the Head-Phrase Distinction,” *The Linguistic Review* 33, 103–127.

Rodman, Robert (1977) “Concerning the NP Constraint,” *Linguistic Inquiry* 8, 181–184.

Runner, Jeffrey T. (2001) “The Double Object Construction at the Interfaces,” *University of Rochester Working Papers in Language Sciences* 2, 23–51.

Saito, Mamoru (2016) “(A) Case for Labeling: Labeling in Languages without ϕ-Feature Agreement,” *The Linguistic Review* 33, 129–175.

Seely, T. Daniel (2006) “Merge, Derivational C-Command, and Subcategorization in a Label-Free Syntax,” *Minimalist Essays*, ed. by Cedric Boeckx, 182–217, John Benjamins, Amsterdam.

Sobin, Nicholas (2014) “Th/Ex, Agreement, and Case in Expletive Sentences,” *Syntax* 17.

Speas, Margaret (1991) “Generalized Transformations and the D-Structure Position of Adjuncts,” *Syntax and Semantics, Volume 25, Perspectives on Phrase Structure: heads and licensing*, ed. by Susan Rothstein, 241–257, Academic Press.

Sugimoto, Yushi (2016) “De-phased by External Pair-Merge of C to T and Its Consequences,” *Proceedings of Conference of the Student Organization of Linguistics in Europe XXIV*, 237–253.

Tourlakis, George (2003) *Lectures in Logic and Set Theory: Volume 2, Set Theory*, Cambridge University Press, New York.

Truswell, Robert (2007) “Extraction from Adjuncts and the Structure of Events,” *Lingua* 117, 1355–1377.

Ura, Hiroyuki (1993) “On Feature-Checking for *Wh*-Traces,” *MIT Working Papers in Linguistics* 18*: Papers on Case and Agreement I*, 243–280.

Uriagereka, Juan (1988) *On Government*, Doctoral dissertation, University of Connecticut.

Ward, Gregory and Betty Birner (1995) “Definiteness and the English Existential,” *Language* 71, 722–742.

Williams, Edwin (1984) “*There*-Insertion,” *Linguistic Inquiry* 15, 131–154.

Wood and Marantz (2015) “The Interpretation of External Arguments,” ms., NYU. Retrieved March 18, 2015, from <http://psych.nyu.edu/morphlab/publications/Wood%20&%20Marantz%202015.pdf>