A Theory of Labeling in the Minimalist Program:
Valuation in Merge and Its Application

A Dissertation

Submitted to the University of Tsukuba

In Partial Fulfillment of the Requirements for

the Degree of

Doctor of Philosophy in Linguistics

Akihiko　SAKAMOTO

2013

A Theory of Labeling in the Minimalist Program:
Valuation in Merge and Its Application

A Dissertation

Submitted to the University of Tsukuba

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Linguistics

Akihiko   SAKAMOTO

2013

Acknowledgments

I have spent my (under)graduate years studying linguistics. There have always been many teachers inspiring my intellectual curiosity and colleagues with whom I have competed. I have learned a great deal from these valuable people.

I am deeply indebted to the members of my thesis committee: Nobuhiro Kaga, Yukio Hirose, Koichi Takezawa, Masaharu Shimada, and Kunio Nishiyama. Nobuhiro Kaga is the chair, from whom I learned stubbornness in a good sense. He has a strong belief in research and sticks to his position no matter what. His stubbornness guided my investigation in the right direction many times.

Yukio Hirose is the holder of a penetrating insight into linguistic phenomena, which always surprised me. Although his study field and mine are worlds apart, his comments go straight to the point. Such an insight brought about unexpected findings in many cases.

Koichi Takezawa is a highly-knowledgeable person who can easily explain any abstract concept in clearly understandable terms. I acquired a fundamental way of thinking theoretically in his class.

Masaharu Shimada has quite an ability to detect the true nature of what students, including me, are trying to say in articles and presentations. No matter how busy he was, he never neglected the responsibility of instructing us, regardless of becoming underslept. He helped to improve my study in an empathetic manner.

My academic life has its roots at Ibaraki University, where Kunio Nishiyama patiently trained an ignorant student like me from the beginning. After proceeding to the University of Tsukuba, he gave me regular instruction for my study. His contributions have been absolutely indispensable in forming who I am today.

I am also grateful to Norio Yamada, Masao Okazaki, Naoaki Wada, Akiko Nagano, Ta-

Table of Contents

v

List of Abbreviations

ACD	Antecedent-contained deletion	IA	Internal argument
AF	Agree feature	IM	Internal merge
ATB	Across-the-board	LA	Labeling algorithm
CED	Condition on extraction domain	LI	Lexical item
CI	Conceptual-intentional system	MP	Minimalist program
DM	Distributed morphology	PIC	Phase impenetrability condition
EA	External argument	PISH	Predicate-internal subject hypothesis
EF	Edge feature	QR	Quantifier raising
ECM	Exceptional case-marking	SM	Sensorimotor system
ECP	Empty category principle	SMT	Strong minimalist thesis
EM	External merge	SO	Syntactic object
EPP	Extended projection principle	UG	Universal grammar
FI	Full interpretation	VMH	Vacuous movement hypothesis
FL	Faculty of language		

Chapter 1

Preface

1.1.   Overview

1.1.1.   Objects, Methods, and Goals in Linguistic Research

Modern (natural) science has often made striking progress in the form of theorizing a guiding intuition in research that "nature is simple."   A representative, helpful example of such theorizing is the process of analysis and synthesis.   Natural phenomena are complex as a result of the interaction of certain factors.   Analysis requires close examination of those discrete factors, revealing the individual properties of them.   The process of weaving together the uncovered properties, i.e. synthesis, makes it possible to explain a natural phenomenon as a collection of factors.   Based on "Galilean methods" of this sort, scientists have discovered natural laws, such as universal gravitation, and they have provided principled explanations for a variety of phenomena.

From the 1950s to today, Noam Chomsky and his followers have demonstrated consistently that Galilean methods are also quite valid in the context of linguistic inquiry.   The term "linguistic inquiry" here implies Chomsky's linguistic research project that endeavors to explore the nature of the faculty of language (FL).   FL is the innate knowledge of language that the speaker of any language is supposed to possess and therefore can be considered a "mental organ" inherent to a particular area inside the human brain, which deserves investigation along with other biological organs.   According to Chomsky (1965: 59, 2005: 6, 9), exploration into FL is a task to identify three factors in the language design that intricately interact to determine I-languages attained (see also Kitahara (2011), Narita (2011)):

(1)   Three factors in the language design (from Chomsky (2013: 37)):

      a.   genetic endowment

      b.   external data

      c.   organism-independent factors, including principles of natural law, which play a crucial role in development as in evolution: e.g., the laws of physics that determine that cells divide into spheres rather than cubes; and for computational systems like language, principles of computational efficiency that may well be reducible to laws of nature

The first factor is genetic endowment, more specifically, Universal Grammar (UG), which is "apparently nearly uniform for the species" (Chomsky (2005: 6)).   UG is the theory of the initial state of FL, which is not directly tangible to us.   Something accessible is external data as "primary linguistic data," the second factor.   Such experience turns the initial state of FL into an attainable state that corresponds to an I-language, the theory of a particular language.   We can pick out the nature of UG in a bottom-up fashion by investigating linguistic phenomena generated by I-languages.   Once the content of UG is fixed, we can explore a set of consequences from the top down.   On the other hand, the third factor provides us with a different kind of "bottom-up" approach by serving as principles of natural law, which are not specific to FL. That is, the third factor principles narrow the range of possible components permitted within UG.

Based on the "nature is simple" view, Chomsky (2005: 6) regards the third factor principles as those of efficient computation in the context of computational systems such as language. A straightforward manifestation of this idea is characterized as the strong minimalist thesis (SMT), which takes FL to be a "perfect solution" to the performance systems (i.e. the conceptual-intentional system (CI) and the sensorimotor system (SM)) (Chomsky (2000, 2001, 2004, 2007a, 2008)):

(2)   The strong minimalist thesis (SMT):

FL is a perfect solution to the conditions imposed by CI and SM.

According to (2), FL is a system of minimal computation with no redundancy toward the performance systems, in which all linguistic phenomena find principled explanations.   SMT thus functions as an important "guideline" in exploring the theory of FL.

Bearing these considerations in mind, I would like to clarify our object, method, and goal in linguistic research.   The object to be examined is FL as a mental organ, whose initial state is UG and whose attainable states are I-languages fixed by experience.   The latter are easy to be tangible relatively, so we can initiate the fundamental inspection of FL through the observation of linguistic expressions generated by I-languages.   FL is, by its nature, a system that enables the "infinite use of finite means," given that it has a strong generative capacity despite it being resident in a particular area inside the human brain.   We thus need to elaborate a simple and explanatory theory of FL, along the lines of SMT.   To understand how we should construct such a theory, let me briefly sketch some of the significant paradigm shifts in the studies of phrase structure in Generative Grammar.

1.1.2.   Paradigm Shifts in the Theory of Phrase Structure

Researchers in Generative Grammar have often developed theories of FL through the studies of phrase structure.   The standard theory, presented by Chomsky (1965), captured such fundamental properties as discrete infinity, endocentricity, and ordering in phrase structure by developing a composite system of phrase structure rules and transformational rules (see Fukui (2001) and Narita (2011: chapter 1) for detailed outlines).   In an attempt to imbue such descriptively adequate formulations of FL with explanatory adequacy, the principles-and-parameters approach (Chomsky (1981, 1982, 1986a); see also Fukui (2006)) cultivated X′-theory (Chom-

sky (1970)).   Under this approach, discrete infinity is explained as recursive application of the X′-schema [$_{XP}$ ZP [$_{X'}$ X YP]], in which the featural properties of the head X are projected up to the categories X′ and XP, thereby yielding endocentricity.   An ordering property, which varies from language to language, is derived via the head-parameter: the head-first order [$_{XP}$ ZP [$_{X'}$ X YP]] (e.g. English) versus the head-last order [$_{XP}$ ZP [$_{X'}$ YP X]] (e.g. Japanese).   As a result, X′-theory, under the principles-and-parameters approach, not only brought about a breakthrough in solving the so-called "Plato's problem" (Chomsky (1986b)), but it also made it possible to address variation among I-languages.

Although X′-theory so refined seems sufficiently attractive, it still has a stipulative nature such as projection and thus demands a simpler apparatus.   The Minimalist Program (MP), advocated by Chomsky (1995a, b et seq.), accelerates this move under SMT.   This framework aims to construct a minimal theory of FL and ultimately to guide the constructed theory to "beyond explanatory adequacy" (Chomsky (2004, 2012: 18)), which we can also call "biological adequacy" (Narita (2010)).   As seen above, descriptive adequacy has sought to answer *what* FL is and explanatory adequacy to determine *how* FL enables language acquisition.   Beyond explanatory adequacy is an enterprise to explain the reason *why* FL has emerged and evolved as it has.   The close correlation of these three types of questions affords a clue for constructing the theory of FL (cf. Chomsky (2007b: 2)).   When we pursue beyond explanatory adequacy, it is desirable that FL only contains entities that are sufficient to achieve descriptive/explanatory adequacy, because the more complex FL is, the harder it becomes to explain its emergence and evolution.   In the best case, FL includes only one entity.

In effect, Chomsky proposes Merge as the only structure-building computational operation.   Merge is defined as producing a simple set (i.e. Merge (α, β) = {α, β}), which we may call a syntactic object (SO).   The rise of Merge over X′-schemata recaptures the aspect of discrete infinity in phrase structure as recursive application of Merge (i.e. Merge (γ, {α, β}) = {γ, {α,

β}}), a property specific to human beings (cf. Fujita (2009)). Merge does not entail the application of projection; rather, it purely ensures set formation, hence the labeling algorithm (LA) (Chomsky (2013: 43)):

(3)   The labeling algorithm (LA):

Suppose SO = {H, XP}, H a head and XP not a head. Then LA will select H as the label, and the usual procedures of interpretation at the interfaces can proceed.

Since Merge yields an SO as a set but does not name it for interpretation at the interfaces, it follows that (3) emerges as an independent computational algorithm. LA detects an SO's internal head under minimal search and selects the detected head as the label of the SO.

In this way, Chomsky's (2013) theory of phrase structure demands LA as a labeling process for interpreting SOs at the interfaces, which derives endocentricity independently of Merge. It is conceptually meaningful that Merge acquires independence in the system of grammar, given that Merge has solely eliminated a number of artificial materials lacking conceptual necessity such as D-structure and S-structure, including X′-schemata. Since Merge has no capacity of labeling SOs, some other mechanism has to fill that role. LA is the mechanism.

A simple set formed by Merge does not contain the information for ordering, so the rise of this operation eliminates even such concepts as complements and specifiers, leaving only heads detectable. We cannot then formulate the head-parameter because it regulates ordering between heads and complements. The issue of ordering is thus often attributed to the phonetic component in an MP framework (cf. Chomsky (2001: 37–38); see also note 6 of chapter 5 for an outlook for addressing the issue of ordering in this thesis).

The theory of FL based on Merge tells us that the combination of set formation and labeling is the substance of structure-building computation. SMT requires this combination to be

minimal. We are naturally led to explore the minimal theory of FL based on Merge. To achieve this ultimate goal, I will organize this thesis in the form described in the next section.

1.2. Organization

This thesis contains seven chapters. Chapter 2 will introduce some central ideas of the modern minimalist framework, such as Merge, Agree/Value, labeling, and Transfer, based primarily on Chomsky (2000 et seq.), Richards (2007), and Narita (2011). This introduction facilitates systematic exploration into FL in the subsequent chapters.

Based on the theoretical framework introduced in chapter 2, chapter 3 will propose a mechanism of valuation in Merge, which I refer to as the theory of (non)phasal valuation. According to Chomsky (2007a, 2008), there are two types of features in the theory of FL: the Agree feature (AF) and the edge feature (EF), both of which belong to lexical items (LIs), computational atoms stored in the Lexicon. The AF induces feature valuation based on Agree/Value as a property of a nonphase head LI such as T and V. In contrast, the EF, which any LI possesses, drives Merge. Under the assumption that both the AF and the EF are unvalued/uninterpretable features [uF], the valuation mechanism of the former is well-designed but the one of the latter lacks full consideration. The theory of (non)phasal valuation proposed in this chapter ensures that the EF undergoes valuation in Merge, but not Agree/Value, which enables SOs involving the valued EF to contribute to interpretation at the interfaces. Thus, this theory succeeds in locating the EF in the system of grammar.

In chapters 4 to 6, I will investigate the application of (non)phasal valuation. Chapter 4 will show that the mechanism of (non)phasal valuation allows us to derive the effects of labeling. According to Chomsky (2013), Merge creates an SO as a set but does not name the SO for interpretation at the interfaces, which means that a different mechanism guarantees the label of an SO. Although LA is a possible candidate for such a mechanism, it is just a stipulation. It is

6

preferable that the mechanism of labeling be derived from some other ingredient of the computational system. The theory of (non)phasal valuation accomplishes this purpose in a successful manner. In addition, (non)phasal valuation is considered a generalized version of the mechanism of clausal typing proposed by Cheng (1997) within the theory of phases. Under this view, the occurrence of phasehood-determining elements identifies the type of phases, which signify the domains of C and *v*, just as the occurrence of clausehood-determining elements identifies the type of clauses in the system of clausal typing. Thus, the domain of C derives the mechanism of clausal typing; the domain of *v* derives the mechanism of "predicate typing." Not only does the theory of (non)phasal valuation make an independent algorithm like LA unnecessary, but it also yields some significant implications for both clausal and predicate systems.

Chapter 5 will argue that vacuous movement phenomena follow from the interaction between the theory of (non)phasal valuation and the theory of feature inheritance. Traditionally, such phenomena have been taken to be regulated by the vacuous movement hypothesis (VMH), one of the principles constituting core grammar, under which a movement operation without an effect on PF output can be suspended (George (1980), Chomsky (1986a), Agbayani (2000, 2006), among others). By reducing the VMH to the theory of (non)phasal valuation that incorporates the mechanism of feature inheritance, I will demonstrate that this purported principle is unnecessary in the system of grammar. This reduction results in desirable consequences.

Chapter 6 will develop a novel analysis of parasitic gap constructions in which the adjunct clause with a parasitic gap functions as a restrictive relative clause that the mechanism of "afterthoughts" presented by Chomsky (2004) introduces separately from the main derivational workspace. In conjunction with the theory of (non)phasal valuation, the proposed analysis is shown to explain descriptive generalizations for parasitic gap phenomena discovered by a number of previous studies, including (i) anti-c-command effects, (ii) S-structure effects, (iii) the generalization of A′-movement versus A-movement, (iv) categorial restriction, and (v) an-

ti-reconstruction effects. In contrast to this analysis, Kasai's (2010) multiple dominance analysis derives parasitic gap constructions within a single derivational workspace approach by optimizing the theory of FL based on Merge. Both analyses are similar in their adherence to SMT in that they attempt to reduce Chomsky's (1986a) significant insight to the principles and operations of minimal syntax. The comparison between these two analyses guides us to some interesting consequences. In particular, the analysis proposed in this chapter, unlike Kasai's multiple dominance analysis, makes it possible to characterize a certain difference between the nature of core grammar and the surface aspect of grammar in concert with the analysis based on the parallelism condition, offered by Sakamoto (2011a, b).

Chapter 7 will conclude this thesis with an outlook for future investigation.

Chapter 2

Theoretical Backgrounds:

Structure-Building Computation in Minimal Syntax

2.1. Introduction

We saw in chapter 1 that exploration into the faculty of language (FL) is a task to identify three factors in the language design (Chomsky (1965: 59, 2005: 6, 9)) that intricately interact to determine I-languages attained (from Chomsky (2013: 37)):

(1)   Three factors in the language design:

        a.   genetic endowment

        b.   external data

        c.   organism-independent factors, including principles of natural law, which play a crucial role in development as in evolution: e.g., the laws of physics that determine that cells divide into spheres rather than cubes; and for computational systems like language, principles of computational efficiency that may well be reducible to laws of nature

A linguistic manifestation of the third factor principles is defined as the strong minimalist thesis (SMT), which takes FL to be a "perfect solution" to the performance systems (i.e. the conceptual-intentional system (CI) and the sensorimotor system (SM)) (Chomsky (2000, 2001, 2004, 2007a, 2008)):

(2)   The strong minimalist thesis (SMT):

FL is a perfect solution to the conditions imposed by CI and SM.

By behaving as an important "guideline" in exploring the theory of FL, SMT teaches us that FL is a system of minimal computation with no redundancy. Due to this instruction, we have achieved the theory of FL based on Merge, under which a combination of set formation and labeling is the substance of structure-building computation in minimal syntax. In what follows, I would like to consider how we should develop this theory by spending this chapter introducing some central ideas in the modern minimalist framework, based primarily on Chomsky (2000 et seq.), Richards (2007), and Narita (2011), thereby facilitating systematic exploration into FL in the subsequent chapters.

2.2.   Set Formation: Merge and the Edge Feature

SMT eliminates any redundancy in the system of grammar, requiring FL to be minimal computation. Such a consideration of computational efficiency guides us to the principle of Full Interpretation (FI):
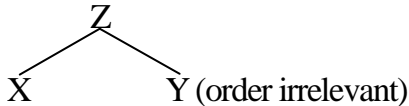
(3)   Full Interpretation (FI):

Every SO of SEM and PHON contributes to interpretation.

By applying Merge to lexical items (LIs), FL generates infinite pairs of SEM and PHON, i.e. syntactic objects (SOs), which FI instructs to receive proper interpretation. LIs bear a property referred to as the edge feature (EF), which I define in (4) based on Chomsky (2007a, 2008), Fukui (2008), Narita (2011: chapter 3):

(4)   a.   The EF is a feature that enables its bearer to be merged with some LI or SO.

         b.   The EF is undeletable throughout narrow syntax.

Application of Merge is unbounded insofar as there are LIs remaining in the derivational work-space, which is a place where the computational system implements structure-building computation. The EF therefore counts as an internal property of LIs, which can also have other properties, such as semantic and phonological features. The undeletability of the EF throughout narrow syntax ensures recursive application of Merge. Merge is a simple set-formation operation that takes objects X and Y and creates a new object Z:

(5)　　a.　　Merge $(X, Y) = \{X, Y\}$

　　　　b.
$$
\begin{array}{c}
Z \\
\diagup\diagdown \\
X \qquad Y
\end{array}
\text{(order irrelevant)}
$$

Let us provide explicit definitions for both LIs and SOs along the lines of (5). LIs are a bundle of features stored in the Lexicon, which serves as a computational atom. Merge yields an SO by combining an LI and some other LI or SO already constructed:

(6)　　Z is an SO iff Z is a set $\{X, Y\}$, where X and Y are either LIs or SOs already constructed, always excepting combination between SOs.

The merger of objects X and Y yields a new object Z. Z then corresponds to an SO, which comprises X as LIs or SOs and Y as LIs or SOs. Since LIs are the only bearers of the EF and hence SOs have no EF, SOs are not related to each other. SOs thus consist of either LIs plus LIs or LIs plus SOs already constructed (see also Narita (2011: chapter 3)).[1]

---

[1] Although our definition given in (6) views SOs as phrasal, Narita's (2011: 30) definition has no such entailment:

The relationship between X and Y can be either contiguous or noncontiguous:

(7)   a.   External Merge (EM):

$$X \qquad Y \qquad \longrightarrow \qquad \overset{\displaystyle Z}{\underset{X \qquad Y}{\diagup \diagdown}}$$
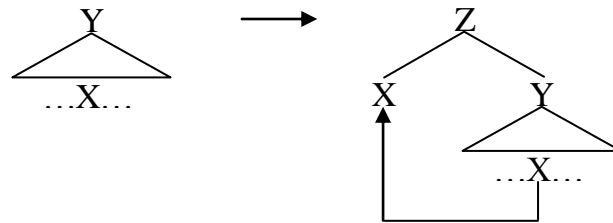
        b.   Internal Merge (IM):

The former is a case where neither X nor Y is part of the other, as in combining *read* and *LGB* to form SO = {X, Y}, which corresponds to *read LGB*.   We may call that case "external Merge" (EM).   The latter is produced by "internal Merge" (IM), which is the real nature of displacement phenomena such as *wh*-movement.   It is in such cases that one is part of the other, e.g., X is part of Y.   Then, the result is the same as in EM, with Merge forming SO = {X, Y}, but there are two copies of X in this case.   One is the original occurrence remaining in Y, and the other is the copied occurrence merged with Y.

Although IM produces copies, the two types of Merge are in essence identical in that they uniquely engage in structure-building computation at narrow syntax by combining two objects to define SO = {X, Y}.   Crucially, EM and IM would be expected to yield different effects at the interfaces (i.e. SEM and PHON) on the assumption that the means of FL are entirely ex-

---

(i)   α is a *syntactic object* (SO) iff
        a.   α is an LI, or
        b.   α is a set {β, γ}, where β and γ are SOs.

Under (i), SOs are not necessarily phrasal because they contains LIs by definition.   Narita (2011: chapter 3) thus states that phrasal/non-LI SOs do not bear the EF by defining non-LIs as phrases.

ploited by the performance systems (i.e. CI and SM). This reasoning is motivated in a suffi-

cient way, as Chomsky (2008: 140) states:

(8)   If the means of language are fully exploited by the interface systems, in accord with

        a reasonable interpretation of SMT, then we would expect the two types of Merge to

        have different effects at the interfaces. At the phonetic interface, they obviously do;

        IM yields the ubiquitous displacement phenomenon. At the semantic interface, the

        two types of Merge correlate well with the duality of semantics that has been studied

        within generative grammar for almost forty years, at first in terms of "deep and sur-

        face structure interpretation" (and of course with much earlier roots). To a large

        extent, EM yields generalized argument structure (θ-roles, the "cartographic" hierar-

        chies, and similar properties); and IM yields discourse-related properties such as old

        information and specificity, along with scopal effects.

As quoted in (8), Merge spectacularly integrates two distinctive aspects of the computational

system—structure building and movement—each of which D-structure and S-structure has been

considered to capture in the traditional sense. This integration is already eloquent proof of the

plausibility of the theory of FL based on Merge. Thus, both EM and IM, exploited by CI and

SM, are always available as a "cost-free" operation within the computational system.

2.3. "Detoxification" of Uninterpretability

2.3.1. Agree/Value

SOs formed by Merge have to be as simple as possible under SMT. In the best case,

only Merge should complete structure-building computation at narrow syntax. Nonetheless,

the system of FL contains an apparently "imperfect" aspect that Merge fails to express, which

we can characterize as "covariation of morphological or semantic features among multiple elements that possibly range over non-sister relations" (Narita (2011: 40)):

(9)   a.   There *seems* to be likely to be *a boy* in the garden.

       b.   There *seem* to be likely to be *boys* in the garden.

Examples (9a-b) display *there*-expletive constructions, in which the matrix verb shows long-distance agreement in number with the associate NP.   Given that Merge is the creator of a simple set, it is quite unlikely that Merge results in such featural covariation.   It is natural to consider that a distinct mechanism in the computational system is responsible for this effect.

Chomsky (1995a) introduces the notion of uninterpretability in his theoretical framework, according to which derivation crashes if it arrives at the interfaces, leaving uninterpretability in it. The source of uninterpretability is the existence of a certain feature on LIs.   Such an offending feature needs to be "detoxified" by undergoing Agree/Value under the Probe-Goal system (cf. Chomsky (2000) and many subsequent works):

(10)   Agree/Value:

        P > G   Agree/Value (P, G), where P is a probe and G is a matching goal, ">" is

                     a c-command relation: P c-commands G.

(11)   The Probe-Goal system:

        a.   Matching is non-distinctness.

        b.   G is the sister of P.

        c.   Locality reduces to "closest c-command."

        d.   P and G must be active.

Agree/Value under the Probe-Goal system relates two LIs via matching relations, as in (11a), with SMT minimizing this operation, as in (11b-d). An LI P with an unvalued/uninterpretable feature [uF] seeks a matching LI G with an innately valued/interpretable feature [vF]. [vF] of G is consequently copied into [uF] of P, hence the inactivity of P and G. We can define this process as feature valuation:

(12)   Feature valuation:

An LI P with [uF] seeks a matching LI G with [vF]. Then, [vF] of G is copied into [uF] of P.

Feature valuation detoxifies [uF] on LIs under minimal search. This mechanism explains the existence of featural covariation of the sort observed in (9). In that case, the LI P (i.e. T), with a set of unvalued/uninterpretable agreement features [uφ], undergoes feature valuation from the matching LI G (i.e. K(ase)), with a set of valued/interpretable agreement features [vφ], which induces long-distance agreement in number between the matrix verb and the associate NP.[2] The detoxification of uninterpretability enables the relevant derivation to converge at the interfaces, thereby observing FI.

2.3.2.   Labeling

The mechanism of labeling SOs could also have something to do with the detoxification system of uninterpretability. Recall that Merge does not entail application of projection, instead purely ensuring set formation, as in (5) (see section 1.1.2 in chapter 1). This situation leads Chomsky (2013: 43) to propose the labeling algorithm (LA):

---

[2] I ascribe this observation account to Narita's (2011) analysis, which I will sketch in section 2.6.

(13)   The labeling algorithm (LA):

Suppose SO = {H, XP}, H a head and XP not a head.   Then LA will select H as the

label, and the usual procedures of interpretation at the interfaces can proceed.

LA detects an SO's internal head under minimal search, as in Agree/Value, and it selects the de-

tected head as the label of the SO.   This mechanism allows every SO to receive interpretation at

the interfaces, thus satisfying the FI requirement.

Now, apart from LA, the assignment of labels to SOs entails the convergence of deriva-

tion at the interfaces.   In this sense, the licensing of labels evokes the detoxification of uninter-

pretability.   Nevertheless, we are not aware of a previous analysis that attributes labeling to

feature valuation or something analogous.   In chapter 3, I will propose a theory of valuation in

Merge—the theory of (non)phasal valuation—that permits the EF as [uF] of a sort to be detoxi-

fied.   This line of analysis is able to characterize a labeling process as an epiphenomenon of EF

detoxification, as indicated in chapter 4.

2.4.   Transfer: Mapping Structured Expressions onto SEM and PHON

We have thus far introduced computational operations pertaining to set formation and de-

toxification of uninterpretability: Merge (i.e. EM and IM), Agree/Value, and LA.   We have not,

however, mentioned another crucial aspect of the computational system.   Once syntactic deri-

vation proceeds up to some point, the computational system needs to hand a particular unit con-

sisting of certain SOs to SEM and PHON, each interfaced with CI and SM.   It is not until this

operation applies that this unit receives interpretation and fulfills the FI requirement.   We may

refer to that computational operation as Transfer:

(14)   Transfer (preliminary version):

Transfer maps SOs onto SEM and PHON.

Given minimal computation under SMT, any operation should be applied in a computationally efficient manner.   It goes without saying that the Transfer operation is no exception.   In recent years, Transfer has been proposed to work in concert with other operations such as Agree/Value and IM, thus yielding efficient computation.   In the following section, I would like to outline this line of approach initiated by Chomsky (2007a, 2008), Richards (2007), and Narita (2011: chapter 2).

2.5.   Simultaneous Application of Operations within Phases

2.5.1.   Feature Inheritance

Richards (2007) proposes, based on an original version of Chomsky (2008), that the mechanism of feature inheritance improves the computational system.   According to this proposal, the condition on efficient computation is that Agree/Value and Transfer of [uF] must take place simultaneously.   Within traditional frameworks (Chomsky (2000, 2001)), derivation cannot undergo efficient computation.   The nonphase head T is an inherent possessor of the Agree feature (AF), so feature valuation occurs prior to the introduction of the phase head C, which triggers the Transfer operation.[3, 4]   Put it more explicitly, Agree/Value applies before Transfer does.   This gap in operational application is a departure from SMT because Transfer

---

[3] In this thesis, I define the AF as a general term for value assignors/assignees based on Probe-Goal relations, such as [v/uφ], [Nom/uCase], and [Acc/uCase].   As necessary, I use these particular and general terms.

[4] Narita (2011) successfully defines the notion of phases in terms of convergence, which uncovers what transferred domains are.   In this subsection, I trivially consider CP and *v*P to be phases and TP and VP to be transferred domains for simplicity.   In section 2.6, I will review Narita's approach, thereby making clear the concepts of phases and transferred domains.

cannot immediately convey to the interfaces the AF that was assigned a value by Agree/Value. SMT thus requires that Agree/Value and Transfer of [uF] take place simultaneously. It is the theory of feature inheritance (Chomsky (2007a, 2008)) that implements this condition.

Feature inheritance forces nonphase heads to derivationally inherit the AF from phase heads. Such AF inheritance dictates that every operation within a phase apply in parallel, with a computationally efficient result:

(15)

CP
"Spec-C"
C[uAF][EF]
TP
TRANSFER →
"Spec-T"
T[uAF][EF]
vP
"Spec-v[vAF]"
v[uAF][EF]
VP
TRANSFER →
"Spec-V"
V[uAF][EF]   "Comp-V[vAF]"
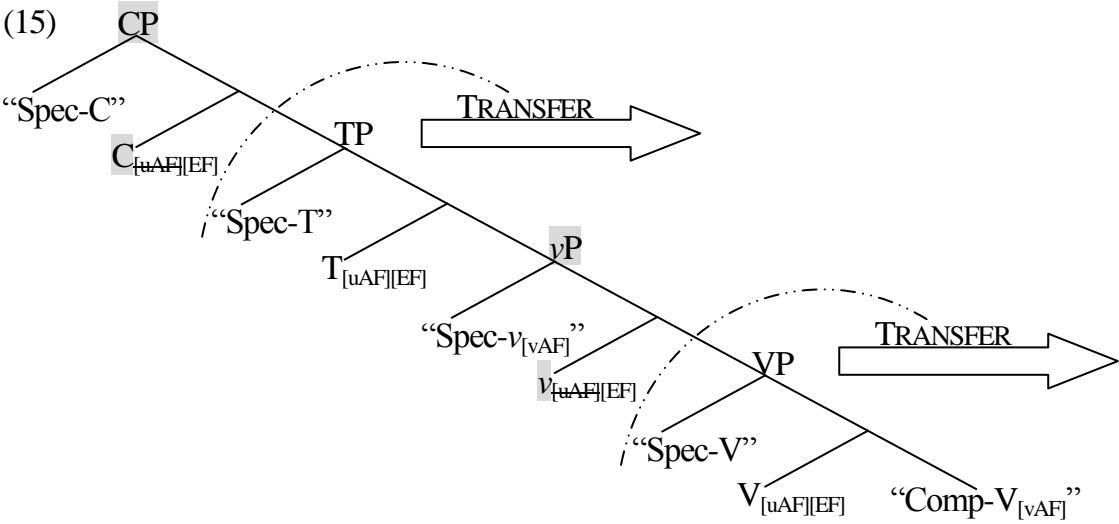
Feature inheritance transmits the unvalued/uninterpretable AF [uAF] from phase heads (i.e. C and v) to nonphase heads (i.e. T and V). Based on Agree/Value under the Probe-Goal system, the inherited [uAF] establishes feature valuation with the valued/interpretable AF [vAF] on elements such as "Spec-v" and "Comp-V" (cf. section 2.3.1). The EF drives IM of these elements into "Spec-C/T" and "Spec-v/V," respectively (cf. Chomsky (2008: 148–149)). Transfer sends off TP and VP to the interfaces. Unless any [uAF] remains, the derivation converges, observing FI. As expected, Agree/Value (and IM) applies simultaneously with Transfer at the phase level, thus yielding efficient computation. The mechanism of feature inheritance is therefore a crucial property with which the system of FL should be equipped.

2.5.2. Copy-Identification

Narita (2011: chapter 2) persuasively feeds operational simultaneity at the phase level with additional argumentation in terms of the mechanism of "copy-identification." We saw in section 2.2 that the system of FL makes available the two modes of Merge in (7), repeated here as (16).

(16) a. External Merge (EM):

        X   Y   ⟶

                                              Z
                                    X   Y

    b. Internal Merge (IM):
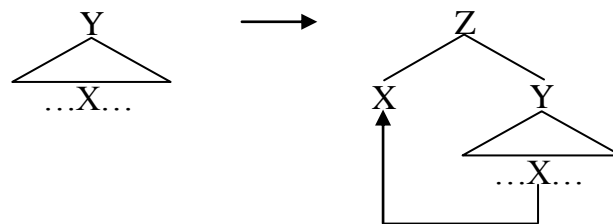
EM combines two independent objects X and Y; in contrast, IM links X internal to Y and Y by yielding two copies of X. The result is the same: a new object Z is formed. These two modes of Merge bring about different effects at the interfaces in such a way that they are exploited by CI and SM. (8) motivates this reasoning in a satisfactory manner.

To the extent that EM and IM create interpretive differences, they have to be distinguishable at the interfaces. In other words, the computational system must have "a way to distinguish copies created by IM from independently externally merged items with identical syntactic constitution" (Narita (2011: 37)). To grasp this situation, compare (16b) with the following schematized structure:

(17)   X   Y   ⟶   Z
                /‾‾\   /   \
              …X…   X   Y
                                              /‾‾\
                                            …X…
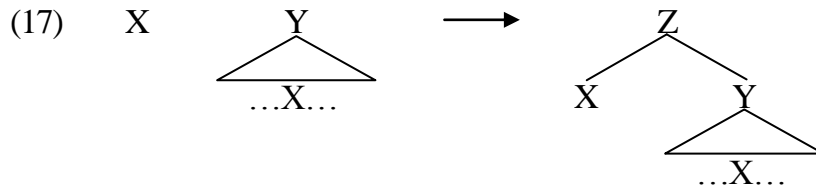
The internally merged item X in (16b) is identical to the externally merged item X in (17) with regard to syntactic constitution.   Thus, (16b) and (17) are representationally indistinguishable. In order for CI and SM to exploit the two types of Merge for interpretation, the computational system has to be able to distinguish between IM in (16b) and EM in (17).

The problem at issue falls under the mechanism of copy-identification.   We can find a crucial difference between EM and IM.   IM, unlike EM, involves copy formation.   There are two copies of X in (16b); (17) has two independent occurrences of X.   Since the representations of (16b) and (17) are exactly the same with regard to syntactic constitution, as stated above, only the applicational point of Merge (X, Y) can tell apart whether X is an independent item or not.   Narita (2011: 39) proposes that parallel application of IM and Transfer proves copy-identification.   If IM takes place before Transfer, neither interface can identify the copied items.   Where IM occurs after Transfer, the relevant derivation simply crashes at the interfaces in violation of FI.   Therefore, IM only applies simultaneously with Transfer.   EM does not, on the other hand, apply together with Transfer because it lacks copy formation and hence needs no copy-identification.

Recall here that Agree/Value under the Probe-Goal system involves copy formation in the sense that this operation enables [vF] on an LI to be reintroduced via feature valuation into derivation, as described in (12).   This fact means that the Agree/Value operation demands copy-identification, as in the case of IM.   We thus obtain the following result:

(18)   IM and Agree/Value apply simultaneously with Transfer.

In this way, Narita not only maintains a way for CI and SM to exploit EM and IM but also corroborates Richards's conclusion that Agree/Value and Transfer of [uF] take place in parallel. Based on these discussions, he also successfully defines phases and transferred domains in a derivational fashion, with the high hope of clarifying the rationale for an apparently "imperfect" aspect such as displacement phenomena and [uF] on LIs in the system of FL. In the subsequent section, we sketch these attempts.

2.6.   Implementation of Minimal Computation

Thus far, we have introduced, in section 2.2 to 2.4, the core computational operations (i.e. Merge, Agree/Value, LA, and Transfer) that any minimalist theory of FL should incorporate, and we have seen in section 2.5 that IM, Agree/Value, and Transfer must all apply simultaneously (see (18)).[5]   In this section, we would like to make sure how Merge, Agree/Value, and Transfer interact with each other, thereby making clear phases and transferred domains.

2.6.1.   Derivation of a Simple Sentence

Let us now examine the derivation of a simple sentence presented by Narita (2011):

---

[5] We have never discussed the interactions between LA and the other operations.   See section 2.7 for such a discussion.

(19)   … (Ø/that/if/whether) the boy will eat the apple.

(20)   a.

$K_{[uCase,\, u\varphi]}$ $D_{[v\varphi]}$ $N_{[v\varphi]}$

b.

$K_{[uCase,\, v\varphi]}$ $D_{[v\varphi]}$ $N_{[v\varphi]}$ TRANSFER

c. $K_{[uCase,\, v\varphi]}$
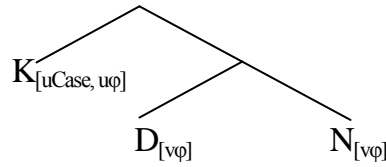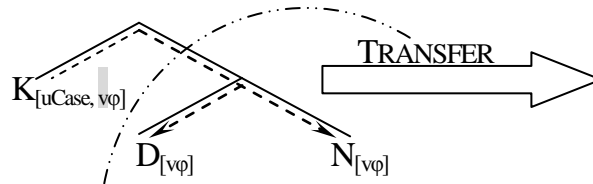
The assembling work of (19) starts with the object nominal phrase *the apple*, as illustrated in (20a), in which recursive application of EM creates the SO {K, {D (the), N (apple)}}.   Narita proposes that noun phrases are uniformly headed by a functional category with an unvalued Case-feature, which he represents as K(ase).   K contains the SO consisting of D and N.   D bears a valued/interpretable person feature and N number and gender features.   [vφ] expresses these features.   In addition to [uCase], K has a full set of unvalued/uninterpretable person, number, and gender features, which corresponds to [uφ].   EM of K introduces [uφ] in the derivational workspace, as in (20a).   The introduction of [uφ] triggers feature valuation (see (12)) based on Agree/Value under the Probe-Goal system (see (10) and (11)), which copies [vφ] of D and N into [uφ] of K.   Copy-identification guarantees simultaneous application of Agree/Value and Transfer, as depicted in (20b).   As a result, the object nominal phrase remains as a simplex LI K in the derivational workspace, as in (20c).   Bearing [uCase], K undergoes feature valuation in later computation.

Next, recursive application of EM assembles the *v*P-level structure {*v*, {V (eat), K (the apple)}}:

(21)	a.

$v_{[\text{Acc, u}\varphi]}$

V	$K_{[\text{uCase, v}\varphi]}$

b.

TRANSFER

$v_{[\text{Acc, v}\varphi]}$

V	$K_{[\text{Acc, v}\varphi]}$

c.	$v_{[\text{Acc, v}\varphi]}$
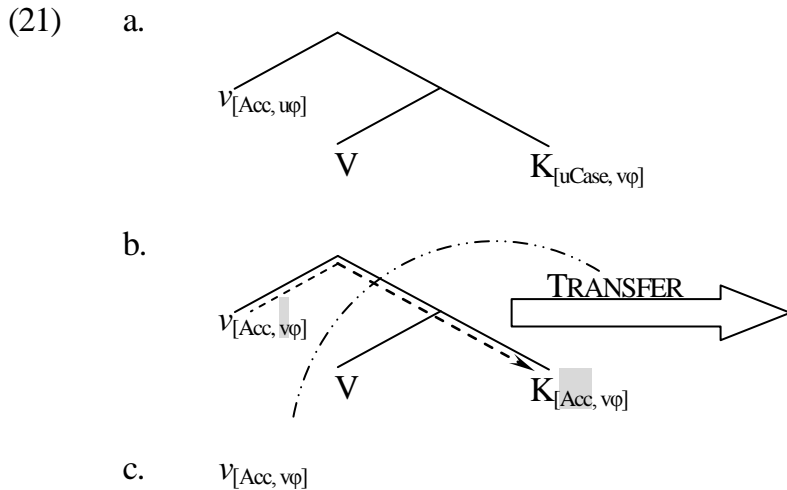
Following the standard analyses in the literature, Narita supposes that v bears a set of unvalued/uninterpretable agreement features [uφ] that is responsible for Accusative Case assignment to the object nominal.　EM of v in (21a) introduces [uφ] into the derivational workspace.　v with [uφ] probes the matching object K with [vφ] under minimal search, and [vφ] of K is consequently copied into [uφ] of v, with [uCase] of K concomitantly receiving an Accusative value. Based on copy-identification, Agree/Value and Transfer are applied in parallel, as depicted in (21b), with which only v stays as a simplex LI in the derivational workspace for the purpose of assigning its external θ-role to the subject nominal (see Narita (2011: section 5.3.4) for details).

Further, recursive application of EM yields the CP-level structure {C, {T (will), {K (the boy), v (eat the apple)}}}:

23

(22) a.

$C_{[Nom]}$ $T_{[u\phi]}$ $K_{[uCase, v\phi]}$ $v_{[Acc, v\phi]}$

b.

TRANSFER

$C_{[Nom]}$ $T_{[v\phi]}$ $K_{[Nom, v\phi]}$ $v_{[Acc, v\phi]}$ $K_{[Nom, v\phi]}$
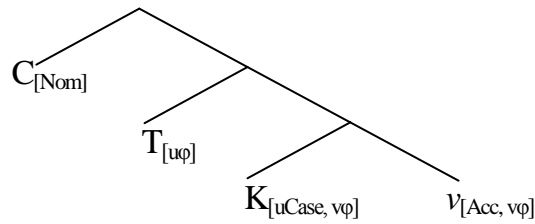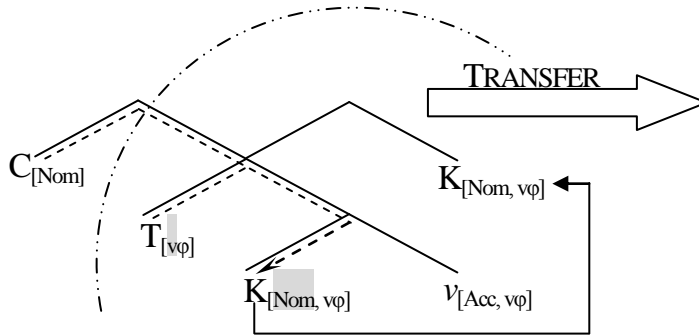
c. $C_{[Nom]}$

Under Narita's assumption, T bears a set of unvalued/uninterpretable agreement features [uϕ] that undergoes feature valuation from [vϕ] of the subject nominal, and C undertakes Nominative Case assignment to the subject nominal, which is introduced into the primary derivational workspace after becoming a simplex LI K in the same process as (20).   (22a) reflects these assumptions.   EM of K and *v* abides by the predicate-internal subject hypothesis (cf. Fukui and Speas (1986), Kitagawa (1986), Kuroda (1988), etc.), under which "the argument structure of a category X is fully 'saturated' strictly locally within the 'projection' of X" (Narita (2011: 55)). If we interpret this hypothesis within the context of Merge, then it holds that EM introduces all nominal arguments of a verbal category into the SO headed by *v*.   Keeping to this reinterpreted version of the predicate-internal subject hypothesis, the subject nominal originates in "Spec-*v*," namely, the "second complement" of *v*, using Narita's (2011: 207) terms.

The introduction of T and C by EM invokes a set of computational operations in a simultaneous fashion under the mechanism of copy-identification.   C and T probe the matching subject K and establish a feature-valuation relation, which copies [vϕ] of K into [uϕ] of T, with

[uCase] of K concomitantly receiving a Nominative value (note that v cannot probe K for the reason that the establishment of feature valuation in (21b) makes v inactive; cf. (11d)).[6]   It has generally been assumed that movement raises the subject nominal to "Spec-T," yielding the surface word order, traditionally known as EPP (Extended Projection Principle) effects, which the current theoretical framework reduces to the EF property (see (15)).   However, by the time IM raises the subject K to "Spec-T," EM has already linked C to the "T′-node" $\{T, \{K, v\}\}$ to form $\{C, \{T, \{K, v\}\}\}$.   IM of K to the node in question accordingly produces the "multi-rooted" structure illustrated in (22b), which corresponds to the two intersecting SOs:[7]

(23)   a. $\{C, \{T, \{K, v\}\}\}$

           b. $\{K, \{T, \{K, v\}\}\}$

The set of SOs in (23b), containing no [uF], undergoes the Transfer operation under copy-identification. $\{T, \{K, v\}\}$ of $\{C, \{T, \{K, v\}\}\}$ in (23a) is also transferred into the interfaces, overlapping with $\{K, \{T, \{K, v\}\}\}$ in (23b).   What remains after parallel application of Agree/Value, IM, and Transfer is a simplex LI C, as in (22c), to which operations in the next derivational stage apply if any.   If this item undergoes no further operations, the derivation in

[6] Narita considers the locus of Nominative/Accusative Case assignment to be C/v, but, for the purpose of his thesis, he refrains from adopting a proposal that the AF, such as [uφ], is originally installed on C/v and is derivationally inherited from C/v to T/V, the theory of feature inheritance (Chomsky (2007a, 2008), Richards (2007)), which I sketched in section 2.5.1.   Although I just outline Narita's analysis in unchanged form at this moment, consideration of minimal computation forces his analysis to incorporate the mechanism of feature inheritance (cf. section 2.5), thereby enabling Agree/Value in (22b) to take place simultaneously with IM and Transfer.   In chapter 5, I will show, based on some significant discussions made in chapters 3 and 4, that the theory of feature inheritance is absolutely imperative in capturing vacuous movement phenomena.

[7] Kitahara (2011) explicates the same analysis.   Chomsky (2007a, 2008), on the other hand, claims that "Spec-v-to-Spec-T raising" is one instance of "tucking-in" (Richards (2001)), which tampers with $\{C, \{T, \{K, v\}\}\}$ to define a new object $\{C, \{K, \{T, \{K, v\}\}\}\}$ in a countercyclic way.   I will show another strategy to circumvent this sort of countercyclic application of IM under the extended theory of feature inheritance proposed in chapter 5.

(20)-(22) is tantamount to constituting an undominated matrix clause.   Transfer therefore sends

off the whole CP to the interfaces, terminating the derivation with adherence to FI.

2.6.2.   Detection of Phase Cycles by Convergence

We saw in the last subsection that the interaction between EM and parallel application of

the subsequent operations Agree/Value, IM, and Transfer result in a set of SOs in a given cycle

and hands it to the interfaces.   The transferred domain is the largest convergent interior of KP in

(20), of vP in (21), and of CP in (22).   CP, vP, and KP, in contrast to nonphasal TP and VP, have

more or less been defined as "phases" (cf. Chomsky (2000 et seq.)), but its definition still re-

mains controversial, with its existence receiving empirical justification in cases such as topicali-

zation and (pseudo)clefting (cf. Narita (2011: 49)).   The notion of convergence, however, suc-

cessfully detects phase cycles in a derivational manner (cf. Narita (2011: 53)):

(24)   An SO Σ can be a phase only if the interior of Σ is convergent (i.e. containing no

[uF]).

(24) holds that the distribution of [uF] demarcates phase cycles.   If an LI with [uF] is intro-

duced into the derivational workspace, then an SO that contains the LI cannot be a phase unless

Agree/Value detoxifies [uF] on the LI.   Detoxification of [uF] on an LI renders an SO that con-

tains the LI convergent (i.e. valued/interpretable).   The convergent SO undergoes the Transfer

operation as the interior of a phase under the mechanism of copy-identification (cf. Narita (2011:

70)):

(25)   Transfer (ultimate version):

Applied to an SO Σ, Transfer maps the largest convergent term Σ' of Σ (the interior of

Σ) onto SEM and PHON, thereby identifying copies within Σ.[8]

This formulation derives the so-called phase impenetrability condition (PIC):

(26)   The phase impenetrability condition (PIC):

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 sends off the interior of a phase to the interfaces, eliminating it from the derivational

workspace.   Only the phase head and its edge are visible to computational operations in the

next higher phase (see Richards (2011) for issues surrounding the PIC).

Since its introduction into the theoretical framework of FL advocated by Chomsky

(1995a), the concept of uninterpretability has played an important role in the system of grammar.

The source of uninterpretability is characterized as [uF] on LIs, which Chomsky (2000: section

3.5) regards as a driving force for the ubiquitous displacement phenomenon.   At this early stage

of minimalist inquiry, the two examples of [uF] on LIs and the dislocation property were nothing

more than strikingly "imperfect" aspects of FL, in departure from SMT.   Generalizing EPP ef-

fects to the EF property, however, the theory of FL based on Merge succeeds in reducing the

dislocation property to IM, which constitutes one instance of Merge along with EM (cf. Chom-

sky (2004, 2007a, 2008)).   Accordingly, the dislocation property has become equivalent in

---

[8] The "term-of" relation is defined as follows (Narita (2011: 31)):

(i)   For any SO K,
        a.   K is a term of K;
        b.   If K is a term of L and K = {α, β}, then α and β are terms of L.

quality to the aspect of structure building, and its absence, rather than presence, has been characterized as an imperfection of the system of FL (cf. Chomsky (2004: 110)).

In this way, displacement has ceased to be an unexplained phenomenon far afield of SMT, but we are still exposed to the other "imperfect" example of [uF] on LIs. Narita's (2011) fresh perspective, however, counts even the paucity of [uF] as an imperfection in human language. That is, the local distribution of [uF] enables derivation to cyclically proceed in conjunction with the core computational operations, hence periodic reduction of computational load. We are consequently led to derivational formulation of phases and transferred domains. [uF] is therefore indispensable for implementation of minimal computation.

2.7. Summary

In this chapter, we introduced some central ideas of the modern minimalist framework, particularly based on Chomsky (2000 et seq.), Richards (2007), and Narita (2011). Within this theoretical model, derivation proceeds in the following way: (i) set formation, (ii) detoxification of uninterpretability, and (iii) mapping structured expressions onto the interfaces. Two modes of Merge—EM in (7a) and IM in (7b)—ensure procedure (i). EM yields argument structure at the vP-level and cartographic hierarchies at the CP-level. IM derives discourse-related properties. Agree/Value enables procedure (ii) in such a way that an LI P with [uF] seeks a matching LI G with [vF] and [vF] of G is subsequently copied into [uF] of P (see (12)). Transfer, which implements procedure (iii), hands structured expressions thus formed to SEM and PHON (see (25)), thereby satisfying the FI requirement (see (3)).

Procedures (i) and (iii) are uncontroversial for me. (ii) is also an uncontroversial procedure as far as the AF is concerned. However, if the EF is [uF] of a sort, then its detoxification process is still unclear. As noted, Chomsky (2007a, 2008) reduces the EPP feature to the EF, which any LI possesses. This reduction is highly significant in that it converts a driving force

for movement into that for Merge (i.e. EM and IM), which is the only general structure-building operation.   That said, I do not know a linguistic study that accurately locates a property termed the EF in the system of grammar.   Merge is the locus of structure-building computation in minimal syntax, and it is the EF that drives this operation.   Therefore, the appropriate location of the EF in the system of grammar calls for urgent attention.

In the next chapter, I will define the EF as [uF] of a sort and reveal its detoxification mechanism, which I designate as the theory of (non)phasal valuation.   (Non)phasal valuation is a process of assigning a certain value to the EF via Merge rather than Agree/Value, which makes it possible to capture labeling effects that LA guarantees as epiphenomena, as shown in chapter 4. By locating the EF in the system of grammar in this manner, I will explore the minimal theory of FL based on Merge.

Chapter 3

The Theory of (Non)phasal Valuation:

Locating the Edge Feature in the System of Grammar

3.1. Introduction

In the last chapter, we affirmed that the uninterpretability of derivation is attributable to the occurrence of an unvalued/uninterpretable feature [uF] on lexical items (LIs) and that [uF] is "detoxified" via feature valuation based on Agree/Value for interpretation at the interfaces. What [uF] entails in the generality is the (unvalued/uninterpretable) Agree feature (AF) on a certain LI.[1] A preminimalist framework often regarded the existence of a Case-feature as triggering movement of an element to the domain of a functional category such as Infl(ection). Put it differently, the AF served as a driving force for movement. Within a minimalist framework, it is not the AF but the edge feature (EF) (a generalized version of the EPP (Extended Projection Principle) feature) that creates movement. It is uncontroversial to some extent that Agree/Value implements the detoxification of the AF. On the other hand, it is unclear in the literature what detoxifies the EF.

In this chapter, I propose that the EF, an unvalued/uninterpretable property that any LI bears, is valued/detoxified through Merge, but not Agree/Value, and that EF detoxification determines the interpretation of any syntactic object (SO) involving the valued EF. This process of EF valuation based on Merge rendering SOs interpretable at the interfaces will be referred to as (non)phasal valuation. Crucially, the uninterpretability of derivation is detoxified not only by feature valuation based on Agree/Value but also by (non)phasal valuation based on Merge.

---

[1] As noted in section 2.5.1 in chapter 2, I define the AF as a general term for value assignors and assignees based on Probe-Goal relations, such as a set of valued/interpretable agreement features [vφ] and a set of unvalued/uninterpretable agreement features [uφ].

The existence of valuation in Merge enables us to locate the EF in the system of grammar in an appropriate manner.

This chapter is organized as follows. Section 3.2 presents the definition and mechanism of (non)phasal valuation, thereby revealing the status of the EF in the system of grammar. Section 3.3 gives concluding remarks.

3.2. Proposals

3.2.1. Detoxification of the Edge Feature as (Non)phasal Valuation

It has been argued in the literature that LIs constituting SOs bear a property called the EF. Based on Chomsky (2007a, 2008), Fukui (2008), Narita (2011: chapter 3), I characterize the EF as follows:

(1) a. The EF is a feature that enables its bearer to be merged with some LI or SO.

   b. The EF is undeletable throughout narrow syntax.

Application of Merge is unbounded insofar as there are LIs remaining in the derivational workspace. The EF therefore counts as an internal property of LIs. The undeletability of the EF throughout narrow syntax ensures recursive application of Merge, a property specific to human beings. According to Chomsky (2007a, 2008), the EF of a phase head triggers A′-movement, and the AF results in A-movement in conjunction with the EF of a nonphase head, based on the mechanism of feature inheritance (see section 2.5.1 in chapter 2).

Although the EF does not seem to involve feature valuation based on Agree/Value (see Kitahara (2011: 18)), there is a clear argument suggesting that the EF is associated with a certain kind of valuation. The EF is by definition undeletable (see (1b)) but is nonetheless obviously uninterpretable at the interfaces (cf. Fukui (2008: 15)). This means that the EF is [uF] of a sort

and should thus be valued/detoxified in some way.　Given that the EF triggers Merge, it is natural that Merge detoxifies the EF.　This reasoning leads us to put forth the hypothesis below:

(2)　　　The EF contributes to interpretation at the interfaces by receiving a certain type of value through Merge.

According to (2), the EF has something to do with valuation in Merge, just as the AF is associated with the process of feature valuation.　Under the strong minimalist thesis (SMT), the result of EF detoxification based on Merge must keep to Full Interpretation (FI):[2]

(3)　　　Full Interpretation (FI):

Every SO of SEM and PHON contributes to interpretation.

We saw in chapter 2 that the system of the faculty of language (FL) is based on minimal computation with no redundancy.　Any SO generated by FL must thus be interpretable with adherence to FI.　If the EF is unvalued/uninterpretable in nature, then it should be the case that FL has a means for detoxifying the EF thereby enabling SOs that contains LIs with the detoxified EF to contribute to interpretation at the interfaces.

　　　Then, what kind of contribution does the EF make at the interfaces?　As noted, there are two types of movement that the EF is supposed to produce: A′-movement and A-movement. However, is it a right definition that the EF of a (non)phase head triggers A/A′-movement? The answer is NO (if this definition is used to refer to the meaning of *the EF of a (non)phase head attracts an A/A′-moved element*).　The EF is an internal property only of LIs, with SOs

---

[2]　Fukui (2008: 15–16) argues that the EF should be eliminated at the timing of Transfer on the assumption that it contributes to no interpretation at the interfaces.

defined as follows (from section 2.2 in chapter 2):

(4)   Z is an SO iff Z is a set {X, Y}, where X and Y are either LIs or SOs already con-

            structed, always excepting combination between SOs.

The merger of objects X and Y yields a new object Z.   Z then corresponds to an SO, which

comprises X as LIs or SOs and Y as LIs or SOs.   Since LIs are the only bearers of the EF and

hence SOs have no EF, SOs are not related to each other.   SOs thus consist of either LIs plus

LIs or LIs plus SOs already constructed.[3]

    With the definition of SOs given in (4) in mind, let us consider a relevant case containing

A/A′-movement:

(5)
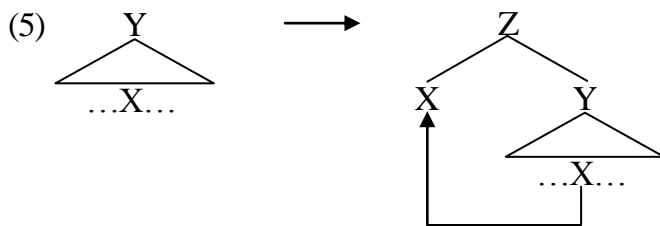
(5) is a case in which X undergoes A/A′-movement from within Y already constructed, an in-

stance of internal Merge (IM).   Y is an SO, but not an LI, and hence has no EF.   For the

A/A′-movement of X to be legitimate, the raising element X itself must be therefore an LI,

which is the only bearer of the EF.   Narita (2011: 89) presents (6) as a logical consequence of

any theory that adopts the EF.

(6)   Only LIs can undergo IM.

---

[3] Note that Narita (2011: 30) defines SOs either as LIs or as a set formed by Merge.   He thus states that
non-LI, namely, phrasal SOs have no EF.   See section 2.2 in chapter 2 for details.

Given (6), a nominal phrase KP, which equals to the SO {K, {D, N}}, must always become a simplex LI K in undergoing A/A′-movement, i.e. IM.   It is the Transfer operation that realizes this situation (see section 2.6.1 in chapter 2).   Within any theory that adopts the EF, Merge cannot relate two SOs to each other because LIs rather than SOs are the only possessors of the EF.   Transfer therefore converts at least one of them into a simplex LI, which is the bearer of the EF, thus licensing "XP-YP merger" (see section 4.5 in chapter 4 for related discussions).

If (6) is the right definition of IM, then we cannot accept the statement that the EF of a (non)phase head triggers A/A′-movement, because the trigger of A/A′-movement is the EF of an A/A′-moved element itself.   However, we can draw one important conclusion by exploiting the hypothesis presented in (2).   If we state with (2) that the EF of a (non)phase head triggers A/A′-movement, then it does not mean that the EF of a (non)phase head attracts an A/A′-moved element; rather, this statement indicates that the EF value of a (non)phase head determines whether a relevant movement has an A-property or an A′-property.   An A/A′-moved element is internally merged with the domain of a (non)phase head by utilizing its own EF.   Once IM takes place, the A/A′-moved element induces EF detoxification in the domain of the (non)phase head.   We may define this process as (non)phasal valuation, distinguishing feature valuation based on Agree/Value:

(7)   (Non)phasal valuation:

Merge links an LI to the domain of a (non)phase head and creates a new (non)phasal SO.   Then, the most prominent property of the LI for the relevant domain values the EF of the (non)phase head.   The valued EF identifies the interpretational property of the new (non)phasal SO at the interfaces.

According to (7), the EF, unlike the AF, does not undergo feature valuation based on

Agree/Value.   Merge associates an LI with the domain of a (non)phase head, when the most prominent property of the LI, typically a valued/interpretable feature [vF], enables EF valuation/detoxification.   The valued EF determines the interpretational status of the (non)phase, which corresponds to an SO that contains the (non)phase head LI with the valued EF.   This definition permits Merge to be accompanied with (non)phasal valuation, from which an A/A′-property emerges.

In this subsection, I defined the process of EF detoxification as (non)phasal valuation. (Non)phasal valuation is a process in which Merge values the EF of a (non)phase head and identifies the interpretational property of the (non)phasal SO.   This EF detoxification process consequently fulfills the FI requirement under SMT.   In the next subsection, I show the detailed mechanism of (non)phasal valuation.

3.2.2.   Mechanism of (Non)phasal Valuation

We showed in the last subsection that there is a valuation process that detoxifies [uF] via Merge in the system of grammar.   While feature valuation based on Agree/Value detoxifies the AF, (non)phasal valuation based on Merge detoxifies the EF.   Crucially, the detoxification of [uF] makes (non)phasal SOs involving the detoxified [uF] interpretable at the interfaces, thereby satisfying the FI requirement.   In this subsection, I would like to explore the mechanism of (non)phasal valuation.

Let us now consider the following schematized structure:

(8)   Ph(vF)

LI[vF]

PhH[uAF][EF]   Nph(vAF)

LI[vAF]

NphH[uAF][EF] ...
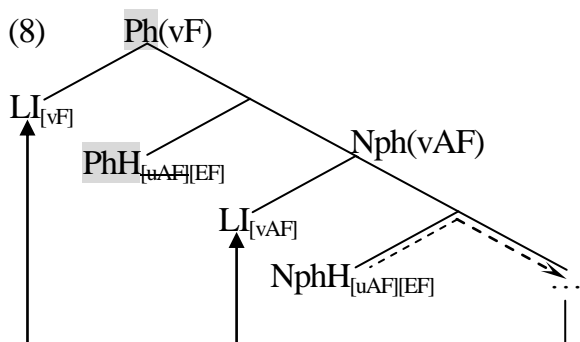
(8) organizes the set that consists of a phasal SO (Ph) and a nonphasal SO (Nph) in the tree diagram. As observed in section 2.5.1 of chapter 2, feature inheritance transmits the unvalued/uninterpretable AF [uAF] from a phase head (PhH) to a nonphase head (NphH), which enables Agree/Value and IM to apply simultaneously with Transfer at the phase level. Notice here that two LIs are internally merged into "Spec-PhH" and "Spec-NphH." These occurrences of the two LIs drive (non)phasal valuation.

Let us first examine the mechanism of nonphasal valuation, which is a less controversial process for the reason stated below. As illustrated in (8), nonphasal valuation is a process in which the most prominent property of an LI assigns [vAF] via Merge to a nonphasal SO by valuing the EF of the nonphase head. Why is then the most prominent property of the LI [vAF] in the case of nonphasal valuation? The reason is straightforward. (Non)phasal valuation is a process for interpreting SOs in essence. [vAF] is a valued/interpretable feature [vF], which contributes to interpretation at the interfaces and is thus a eligible value-assignor for the EF of a nonphase head. An LI that occupies "Spec-NphH" has already established feature valuation based on Agree/Value with its matching nonphase head LI, which renders [vAF] prominent. Therefore, [vAF] serves as a value-assignor for the EF of a nonphase head.

Let us turn to the mechanism of phasal valuation. This process is identical to nonphasal valuation in that Merge determines the interpretational status of the relevant SO, which in the case of phasal valuation corresponds to the SO whose head is a phase head. A phase head, un-

like a nonphase head, is not associated with feature valuation based on Agree/Value. Accordingly, [vAF] is not an appropriate value-assignor for the EF of a phase head even though an LI that occupies "Spec-PhH" bears [vAF]. What then constitutes a value-assignor for the EF of a phase head? As seen in chapter 2, the domain of a phase head has to do with discourse-related properties. Thus, a certain [vF] pertaining to such properties is defined as the most prominent property of the LI that occupies "Spec-PhH." As depicted in (8), the LI bearing such a [vF], occupying "Spec-PhH," values the EF of the phase head. Consequently, this EF detoxification process identifies the interpretational property of the relevant phasal SO along the lines of (7).[4]

There is a rubrical conceptual change in the analysis proposed here. (Non)phasal valuation can interpret every SO, phasal or nonphasal, at the interfaces. Under this analysis, (non)phasehood is underspecified at the initial stage of derivation but is specified via EF valuation by the mate of a (non)phase head in the course of derivation. In this respect, our analysis is in direct opposition to the standard assumption under which a (non)phase head is destined interpretationally and attracts its matching mate.

3.3. Concluding Remarks

I showed in this chapter that in addition to feature valuation based on Agree/Value, there exists a different type of feature detoxification, namely, (non)phasal valuation based on Merge. (Non)phasal valuation is a process in which Merge detoxifies the EF of a (non)phase head, thereby making the relevant (non)phasal SO interpretable at the interfaces. Chomsky (2007a, 2008) reduces the EPP property specific to a functional item to a more general property referred to as the EF that every LI shares, the existence of which guarantees unbounded application of

---

[4] FI dictates that every SO of SEM and PHON contributes to interpretation, as described in (3). This means that even SOs corresponding to "bar-level projections" have to be interpreted. Although I depict nothing about such SOs in (8), I consider that they receive the same interpretations as the whole (non)phasal SOs. In clear terms, the merger of "Spec" uniformly determines the interpretation of its relevant "projections" through EF detoxification.

Merge. However, the essential role that the EF plays has still been blurred. The theory of (non)phasal valuation formulated here allows us to locate the EF in the system of grammar. Under this theory, the EF is [uF] of a sort to which Merge rather than Agree/Value assigns a certain value. This detoxification mechanism identifies the interpretational status of an SO whose head bears the valued EF, thus meeting the FI requirement under SMT.

In the rest of this thesis, I will investigate the application of (non)phasal valuation, namely, what happens when FL absorbs the mechanism of (non)phasal valuation. It will be argued in chapter 4 that EF detoxification specifies clausehood in the CP-level structure and predicatehood in the vP-level structure. In this process, (non)phasal valuation is characterized as a generalized version of the mechanism of clausal typing proposed by Cheng (1997) within the theory of phases. This claim allows us to derive such effects as the labeling algorithm presented by Chomsky (2013) captures from the mechanism of (non)phasal valuation. I will also demonstrate that (non)phasal valuation derives the process of categorization proposed in the framework of Distributed Morphology, advocated by Halle and Marantz (1993). This demonstration means that (non)phasal valuation is a general mechanism that regulates not only the level of phrases but also the level of words.

Chapter 4

Functions of (Non)phasal Valuation

4.1. Introduction

We saw in chapter 1 that the Minimalist Program (MP) advocated by Chomsky (1995a, b), which explores "biological adequacy" of linguistic theory (Narita (2010)), has elaborated Merge as the simplest structure-building computational operation. Merge is defined as producing a simple set (i.e. Merge $(\alpha, \beta) = \{\alpha, \beta\}$), which we may call a syntactic object (SO). Once the rise of Merge over X′-schemata recaptures the aspect of discrete infinity as recursive application of Merge (i.e. Merge $(\gamma, \{\alpha, \beta\}) = \{\gamma, \{\alpha, \beta\}\}$), it favors labeling over projection. In other words, Merge does not entail application of projection, instead purely ensuring set formation, hence the labeling algorithm (LA) (Chomsky (2013: 43)):

(1)   The labeling algorithm (LA):

Suppose SO = {H, XP}, H a head and XP not a head. Then LA will select H as the label, and the usual procedures of interpretation at the interfaces can proceed.

Since Merge yields an SO as a set but does not name it for interpretation at the interfaces, it follows that (1) emerges as an independent computational algorithm. LA detects an SO's internal head under minimal search, as in Agree/Value, and selects the detected head as the label of the SO (see section 4.2.1).

In this way, Chomsky's (2013) theory of phrase structure demands LA as a labeling process for interpreting SOs at the interfaces, which derives endocentricity independently of Merge. It is conceptually significant that Merge acquires independence in the system of grammar, given that Merge has solely eliminated a lot of artificial materials lacking conceptual necessity such as

D-structure and S-structure, including X′-schemata. Since Merge has no capacity of labeling SOs, some other mechanism has to fill that role. Although LA is a possible candidate for such mechanism, it is just a stipulation. I hope to find a means by which a labeling process is not stipulated but is instead deducible from the interaction of core principles and operations in minimal syntax. To achieve this goal, I make the following claim: labeling follows as a natural manifestation of the mechanism of (non)phasal valuation, explored in chapter 3, and there is no need to posit an independent algorithm like LA.

(Non)phasal valuation can be regarded as a generalized version of the mechanism of clausal typing proposed by Cheng (1997) within the theory of phases (see section 4.2.2). Under the mechanism of clausal typing, the occurrence of clausehood-determining elements identifies the type of clauses. If we provide clausal typing with phase-theoretic interpretation, the result is that the occurrence of phasehood-determining elements identifies the type of phases, which signify the domains of C and v (see section 2.6.2 in chapter 2).[1] Then, the domains of C and v are relevant to deriving the mechanisms of clausal typing (see section 4.3) and "predicate typing" (see section 4.4) in cooperation with the theory of (non)phasal valuation, respectively. Not only does the theory of (non)phasal valuation make an independent algorithm like LA unnecessary, but it also yields some interesting implications for both clausal and predicate systems.

Our discussion proceeds as follows. Section 4.2 introduces the theory of LA and then reduces it to the theory of (non)phasal valuation. Sections 4.3 and 4.4 initiate the actual implementation of this theory in both clausal and predicate systems in English, respectively. Section 4.5 explores some consequences of the proposed analysis. Section 4.6 concludes this chapter.

---

[1] I do not distinguish between transitive/unergative verbs, i.e. v^* and passive/unaccusative verbs, i.e. v, as in Legate (2003). I argue, rather, that the distinction between verbs emerges simply from a different result of derivation. See section 4.4 for details.

4.2. Reduction of LA to (Non)phasal Valuation

4.2.1. The Role of Labeling in Minimalism

Before embarking on reduction of LA to (non)phasal valuation, I would like to explicate the role that labeling plays in an MP framework. SOs to be interpreted at the interfaces are generated by (recursive) application of Merge to lexical items (LIs) as a bundle of features stored in the Lexicon. Full Interpretation (FI) supervises SOs phase by phase. Derivation as a given set of SOs that undergo the Transfer operation, being interpretable, converges at the interfaces; otherwise, it would crash there (see section 2.6.2 in chapter 2). The source of uninterpretability, for one thing, is the existence of unvalued/uninterpretable features [uF] on LIs. Such offending features need to be detoxified by receiving certain values at narrow syntax under feature valuation based on Agree/Value (see section 2.3.1 in chapter 2). Uninterpretability can also come from the failure to label SOs. As mentioned in section 4.1, Merge is the creator of a simple set (e.g. Merge $(\alpha, \beta) = \{\alpha, \beta\}$) and does not name it. Some independent mechanism has to guarantee the label of such a set for interpretation at the interfaces. Chomsky (2013: 43) claims that there is a fixed LA that licenses SOs to permit them to be interpreted at the interfaces, operating at the phase level along with other operations (cf. (1)).

Let us here examine how LA operates on SOs. According to Chomsky (2013: 43), LA is just minimal search, as in Agree/Value and other operations, and it finds out the relevant information about an SO, which functions as the label of the SO. If SO = {H, XP}, then LA selects H as the label. This case is straightforward. The complicated case is SO = {XP, YP}, where neither is a head. In this case, minimal search is ambiguous because LA detects two heads, X of XP and Y of YP. We then have two strategies for disambiguation: (i) modify SOs

so that there is only one visible head or (ii) consider X and Y to be identical in that they provide the same label.[2]

To gain a better understanding of the circumstances at hand, let us consider some relevant constructions. Among those disambiguated by strategy (i) are predicate-internal subject constructions and copula constructions (Chomsky (2013: 44)):

(2)   a.   …T [$_\beta$ (EA) [$_v$ [V IA]]]

        b.   XP copula [$_\beta$ XP, YP]

As shown in (2), where EA stands for the external argument and IA for the internal argument, external Merge (EM) creates β, which has the form of SO = {XP, YP} and is not labeled by LA. Strategy (i) provides an ambiguous structure that EM produces with the raising strategy for labeling. If EA in (2a) moves out of its original position, β is labeled as v, with EA part of a discontinuous element invisible to LA. Where IA in (2a) becomes a subject material by undergoing a movement operation, β is still labeled as v in a similar vein. The same strategy is adapted to (2b). Based on Moro (2000), Chomsky takes copula structures to be of the form [copula-small clause], where the small clause has the form of SO = {XP, YP}. If internal Merge (IM) raises XP to subject position, β receives the label of YP. The reason is that the lower copy of XP is discontinuous and hence invisible to LA.[3]

---

[2] Narita's (2011: chapter 3) proposal adopted in chapter 3, under which Transfer converts either XP or YP into a simplex LI, can be a distinct strategy of generating a structure that licenses application of labeling. The analysis that we later present is compatible with Narita's analysis rather than Chomsky's analysis. See section 4.5 for related discussions.

[3] IM as well as EM can form an ambiguous structure to be disambiguated by strategy (i), which corresponds to the intermediate steps of successive-cyclic movement (Chomsky (2013: 44)). Here, however, we set aside such a case for simplicity. (i) is a strategy for getting rid of the ambiguity in labeling in the initial and intermediate stages of derivation.

Strategy (ii) makes viable the labeling of an ambiguous structure that IM yields (Chomsky (2013: 45)):

(3) a. $[_C C [_\alpha NP TP…]]$

  b. $[_\alpha NP CP…]$

Structures (3a) and (3b) indicate those of "Spec-*v*-to-Spec-T raising" and *wh*-movement, respectively. In both cases, IM of NP into TP/CP creates α, which has the form of SO = {XP, YP}. α is in the last stage of derivation, so the first strategy cannot be responsible for the labeling of α by raising either NP or TP/CP. Strategy (ii) leads α in (3) to proper application of labeling. NP in (3) shares a prominent feature with TP/CP: φ-features in (3a) and Q-features in (3b). Minimal search finds out those shared features and specifies them as the labels of α.

In sum, labeling is a process that makes SOs formed by Merge interpretable at the interfaces. As Chomsky (2013: 43) states, LA is a possible candidate for such a process. As just seen, however, consideration of cases with an ambiguous structure for labeling such as (2) and (3) complicates any analysis based on LA, which, to begin with, is just a stipulation, as stated in section 4.1. SMT should not tolerate any complication and stipulation within an MP framework. In section 4.2.2, I argue that LA is reducible to the simpler mechanism of (non)phasal valuation proposed in chapter 3.

4.2.2. Deriving Labeling Effects

In chapter 3, I proposed the detoxification process of the edge feature (EF) referred to as (non)phasal valuation. This process is crucially different from that of the Agree feature (AF) called feature valuation:

(4)

$$\text{Ph(vF)}$$

tree structure:
- $\text{LI}_{[vF]}$
- $\text{PhH}_{[uAF][EF]}$
- Nph(vAF)
  - $\text{LI}_{[vAF]}$
  - $\text{NphH}_{[uAF][EF]}$
  - $\dots$
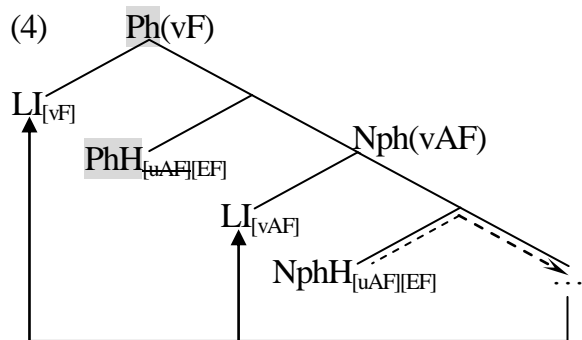
(Non)phasal valuation is a theory of valuation in Merge. The merger of an LI in the domain of a (non)phase head values the EF of the (non)phase head. The valued EF determines the interpretational status of the relevant (non)phasal SO. (Non)phasal valuation is in essence a process that makes SOs as (non)phases interpretable at the interfaces, which is just what LA attempts to accomplish. It then follows that we are reducing LA to (Non)phasal valuation.

It is noteworthy here that (non)phasal valuation successfully generalizes the clausal typing hypothesis (Cheng (1997: section 2.2)) within the theory of phases:

(5)   The clausal typing hypothesis:

Every clause needs to be typed. In the case of typing a *wh*-question, either a *wh*-particle in C^0 is used or else fronting of a *wh*-word to the Spec of C^0 is used, thereby typing a clause through C^0 by Spec-head agreement.

This hypothesis is a cross-linguistically valid generalization that any clause must be marked by a clausehood-determining element. For a *wh*-question, which Cheng addresses exclusively, the element is a *wh*-particle (e.g. Japanese) or a *wh*-word (e.g. English). If we place (5) under minimalism, the former and the latter can be cases of EM and IM, respectively. In any case, clausal typing now turns out to fall under our general formulation of labeling. Our analysis reduces Spec-head agreement in (5) to EF valuation. While Spec-head agreement identifies the

44

type of clauses, EF valuation determines the type of (non)phases.   Importantly, (non)phasal valuation is more general than clausal typing in that the former regulates the interpretational status of every SO, including clauses and predicates.   Under this analysis, clausal typing is characterized as part of a labeling process in the C-T domain.   In sections 4.3 and 4.4, I would like to show in what way (non)phasal valuation actually works as a labeling process in clausal and predicate systems.[4]

4.3.   Clausal Systems

In the preceding section, we reduced the theory of LA proposed by Chomsky (2013) to the theory of (non)phasal valuation, which not simply functions as a labeling process but also successfully formulates the intuition that clausal typing intends to capture within the theory of phases.   In this section, we handle the actual implementation of the proposed theory in the context of clausal systems in English.[5]

4.3.1.   EF Valuation in the C-T Domain

As shown in section 4.2.2, our approach regards labeling as a process of assigning a value to the EF of a (non)phase head via Merge.   We should thus specify the EF values of C and T. Let us begin with the discussion concerning phasal valuation for CP.   Phasal valuation for CP is a process where the merger of an LI assigns a certain value to the EF of C.   The domain of C has to do with the determination of the clause type, which hinges on the presence or absence of *wh*-operators.   If Merge associates a *wh*-operator with the domain of C, the relevant clause assumes *wh*-hood; otherwise, it partakes of non-*wh*-hood.   EF-value assignors in the domain of C

---

[4] See section 4.5.1 for an answer to the question of how the theory of (non)phasal valuation explains the cases in (2) and (3), which are complicated cases for labeling.

[5] In what follows, I will use *labeling* nearly synonymously with *EF valuation*/*detoxification* and thus *(non)phasal valuation*.

are therefore specified as (non-)*wh*-operators:

(6)   Phasal valuation in clausal systems:

Merge links a (non-)*wh*-operator to the domain of C and creates a phase CP.   Then, its (non-)*wh*-feature values the EF of C as [+WH] or [–WH].   The valued EF identifies the interpretational property of the phase CP at the interfaces.

According to (6), a phase CP is labeled as [+WH] when the merger of a *wh*-operator assigns the EF of C *wh*-hood; it is labeled as [–WH] if the merger of a non-*wh*-operator offers the EF of C non-*wh*-hood.

Let us turn to nonphasal valuation for TP.   This valuation is a less controversial process, in which Agree/Value and subsequent Merge of an LI values the EF of T.   A nonphase TP is created by IM that relates an LI that enters into a φ-agreement relation with T to the domain of T. Feature valuation based on Agree/Value thus makes prominent a set of valued/interpretable agreement features [vφ] on that LI.   The nonphase TP is consequently labeled as [vφ].[6]

The following tree diagram illustrates the result of (non)phasal valuation in the C-T domain:

(7)

```
        CP(±WH)
       /      \
 "Spec-C"      \
        C[uAF][EF]  TP(vφ)
              /      \
         "Spec-T"     \
           T[uAF][EF]  "Comp-T"
```

---
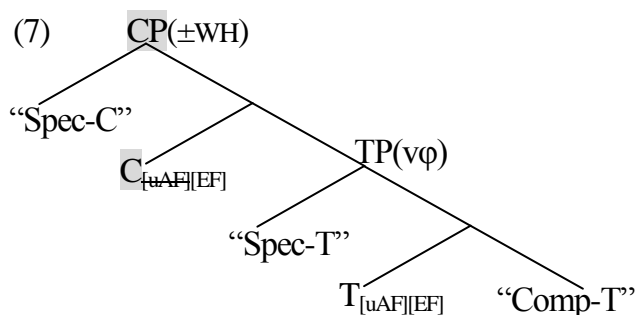
[6] The result of the labeling of TP found here is nearly the same as that found in the work of Chomsky (2013) (see section 4.2.1).

The mergers of "Spec-C" and "Spec-T," triggering (non)phasal valuation, label CP and TP as [±WH] and [vφ], respectively.   Our general formulation of labeling can interpret every SO, phasal or nonphasal, at the interfaces.   Under this approach, (non)phasehood in the C-T domain is underspecified at the initial stage of derivation but is specified via EF valuation by C's and T's mates in the course of derivation, as noted in general terms in chapter 3.   In the following subsection, I demonstrate that the proposed analysis fares well by presenting the derivations of both [+WH] and [−WH] constructions in English.

4.3.2.   Deriving Clauses

4.3.2.1.   [+WH] Constructions

Given phasal valuation in clausal systems formulated in (6), we may define [+WH] constructions as sentences containing a *wh*-operator.   Let us here consider the case of *wh*-questions:[7]

(8)   a.   Who did John see?

　　　b.   Who saw John?

(9)   a. $[_{\text{CP(+WH)}}$ Who $C_{[u\varphi][EF]}$ $[_{\text{TP(v}\varphi)}$ John $T_{[u\varphi][EF]}$ $[_{v\text{P}}$ <John$_{[v\varphi]}$> v $[_{\text{VP}}$ see <who>]]]]

　　　b. $[_{\text{CP(+WH)}}$ Who $C_{[u\varphi][EF]}$ $[_{\text{TP(v}\varphi)}$ who $T_{[u\varphi][EF]}$ $[_{v\text{P}}$ <who$_{[v\varphi]}$> v $[_{\text{VP}}$ saw John]]]]

In (9a), the derivation for the object *wh*-question in (8a), *who* and *John* bear a *wh*-feature and a φ-feature, respectively.   Merge links *who* to "Spec-C" and produces a new phasal SO.   Here,

---

[7] Here and below, I omit the labels of "bar-level projections" and of the intermediate steps of successive-cyclic movement and the exposition of feature inheritance for simplicity (see note 4 of chapter 3 for a way of addressing the labels of "bar-level projections" and section 2.5.1 in chapter 2 for the introduction of feature inheritance), and I refrain from investigating the cause of *do*-support, considering it a phonological phenomenon irrelevant to narrow syntax for now.   Moreover, the detailed analysis of predicate-internal derivation carries over into section 4.4; thus, for the moment, I adopt conventional notations such as *v*P and VP in the predicate domain.

the most prominent property of *who* for the relevant domain assigns the EF of C the value of [+WH]. The assigned value behaves as the label of the phasal SO. The SO thus labeled is interpreted as interrogative at the interfaces.[8] Retaining the services of Agree/Value, Merge relates *John* to "Spec-T" and creates a new nonphasal SO. In this case, the most prominent property of *John* values the EF of T as [vφ], with such a value identifying the interpretational property of the nonphasal SO at the interfaces. (Non)phasal valuation thus satisfies the requirement of FI.

In (9b), the derivation for the subject *wh*-question in (8b), *who* has both a *wh*-feature and a φ-feature, unlike in (9a). Merge raises one occurrence of *who* to "Spec-C" and Agree/Value plus Merge associates the other occurrence of *who* with "Spec-T," following the derivation of subject *wh*-questions proposed by Chomsky (2008: 149).[9] In this case, (non)phasal valuation labels the phasal and nonphasal SOs produced by the merger of two copies of *who* as [+WH] and [vφ], respectively. These SOs thus contribute to interpretation at the interfaces.

Let us turn to the derivation of yes-no questions, which can be viewed as a special case of *wh*-questions:

(10) a. Did Mary buy LGB?

b. $[_{CP(+WH)} \text{Op}_{y/n} \text{C}_{[u\phi][EF]} [_{TP(v\phi)} \text{Mary T}_{[u\phi][EF]} [_{vP} <\text{Mary}_{[v\phi]}> v [_{VP} \text{buy LGB}]]]]$

Yes-no questions are apparently different in nature from *wh*-questions, given that the former, un-

---

[8] A phasal SO that is labeled as [+WH] can be also interpreted as something other than interrogatives at the interfaces. Such difference in interpretation is considered to be attributed to the featural property of *wh*-phrases. The *wh*-phrase of interrogatives has a Q-feature (see e.g. Tsai (1994: section 2.1) for a relevant analysis), that of exclamatives has an E-feature (cf. Grimshaw (1979)), that of relatives has a feature pertaining to predication (see e.g. Takeda (1999: chapter 3) for a relevant discussion), among other examples. Nonetheless, specifying [+WH] as a label suffices for the syntactic computation to perform well.

[9] In chapter 5, I will refine the derivation of (non)phases by incorporating the extended mechanism of feature inheritance.

like the latter, have no visible operator.   It is argued in the literature, however, that both types of questions share the same property (see e.g. Lyons (1977: section 16.3)).   Although both yes-no questions and *wh*-questions introduce a variable in a proposition, the variable is "two-valued" in the former and "many-valued" in the latter.   The yes-no question *Is the door open?* can be answered with either *Yes* (which implies the proposition described in the statement *The door is open*) or *No* (which implies the proposition expressed by *The door is not open*) (cf. Lyons (1977: 757)).   In contrast, the *wh*-question *Who left the door open?* can be responded to with statements such as *John left the door open*, *That little boy left the door open*, or *Uncle Harry left the door open*, presupposing the proposition represented by *Someone left the door open*, in which the indefinite pronoun *someone* can be regarded as "a variable whose range of possible values depends upon the universe-of-discourse" (cf. Lyons (1977: 757–758)).   Both types of questions thus share the same property in the sense that they introduce a variable in a proposition that requires addressees to specify its value.   This shared property leads us to present the derivation of yes-no questions that has an operator in the domain of C, as in (10b).   In (10b), the merger of $Op_{y/n}$ into "Spec-C" labels the phasal SO as [+WH] via phasal valuation, and the merger of *Mary* into "Spec-T" labels the nonphasal SO as [vφ] via nonphasal valuation.   Both phasal and nonphasal SOs in the C-T domain consequently receive their proper interpretation at the interfaces.

4.3.2.2.   [–WH] Constructions

In contrast to [+WH] constructions, [–WH] constructions are characterized as sentences involving no *wh*-operator, as in (11a) and (11b), which are a topic and declarative sentence, respectively:

(11)   a.   John, Mary saw.

        b.   Mary saw John.

(12)   a. $[_{CP(-WH)}$ John C$_{[u\phi][EF]}$ $[_{TP(v\phi)}$ Mary T$_{[u\phi][EF]}$ $[_{vP}$ <Mary$_{[v\phi]}$> v $[_{VP}$ saw <John>]]]]

        b. $[_{CP(-WH)}$ Mary C$_{[u\phi][EF]}$ $[_{TP(v\phi)}$ Mary T$_{[u\phi][EF]}$ $[_{vP}$ <Mary$_{[v\phi]}$> v $[_{VP}$ saw John]]]]

(12a-b), the derivations for (11a-b), correspond essentially to derivations (9a-b). The difference between (9a) and (12a) is attributed to phasal valuation rather than to nonphasal valuation. In both (9a) and (12a), the C-T domain necessitates two different materials. The domain of T is the pivotal position of φ-agreement. In tandem with the AF [uφ] and the EF of T, the merger of *Mary*, a φ-feature bearer, into "Spec-T" gives the label of [vφ] to the nonphasal SO, hence nonphasal valuation comes along. With regard to phasal valuation, there is a difference in the presence or absence of *wh*-operators between (9a) and (12a). (12a), unlike in (9a), has no *wh*-operator and instead includes a topic feature bearer (i.e. *John*), which contributes to interpretation at the interfaces because a topic feature is in nature valued/interpretable and has to put a value into some [uF] under the requirement of FI. The merger of *John* into "Spec-C" assigns the value of [–WH] to the EF of C through the mechanism of phasal valuation. The value of [–WH] contributes via *John*'s topic feature to interpretation at the interfaces, hence the fulfillment of FI.

    The difference between (9b) and (12b) is also ascribed to phasal valuation but not to nonphasal valuation. In (12b), *Mary* is the possessor of a φ-feature but not that of a *wh*-feature. The merger of two occurrences of *Mary* with the C-T domain enables both phasal and nonphasal valuation, which labels the phasal and nonphasal SOs as [–WH] and [vφ], respectively. Derivation (12b) converges as a result without any FI violation.

    To wrap up section 4.3, we have tackled the actual implementation of the theory of (non)phasal valuation developed in chapter 3 in the context of clausal systems in English. In

the next section, we verify that (non)phasal valuation also works crucially on predicate systems.

4.4. Predicate Systems

Predicate systems as well as clausal systems are often supposed to comprise both phasal and nonphasal SOs, i.e. *v*P and VP (cf. Hale and Keyser (1993), Chomsky (1995a, 2000, 2001), Kratzer (1996), etc.). Not surprisingly, it is Merge that builds up such SOs in an MP framework. SOs formed by Merge, phasal or nonphasal, need their labels for interpretation at the interfaces. As shown in section 4.3, (non)phasal valuation is enough for guaranteeing the label of an SO. This mechanism of labeling is expected to apply to and play a crucial role in predicate systems in English as well. This section is devoted to showing that this is indeed the case.

4.4.1. EF Valuation in the *v*-V Domain

We showed in section 4.2.2 that the theory of (non)phasal valuation succeeds in abstracting and refining the mechanism of clausal typing within the theory of phases. Phases denote the domains of C and *v*. Naturally, the domain of *v* should serve to determine the type of predicates, just as the domain of C functions to identify the type of clauses. In this sense, phasal valuation derives the mechanism of predicate typing within predicate systems.[10]

In our approach, *v*'s phasehood is underspecified at the initial stage of derivation but is specified via EF valuation by *v*'s mate in the course of derivation. What is then *v*'s mate? *v*'s mate occupies "Spec-*v*," which purportedly involves the occurrence of a subject argument called EA. EA has a (valued/interpretable) agentive θ-feature and works as the subject of transitive or unergative verbs. It is thus natural to propose that the merger of EA into "Spec-*v*" values the EF of *v* via its agentive θ-feature as [TRANSITIVE/UNERGATIVE] and such a value behaves as the

---

[10] The term of predicate typing is used by Sakamoto (2012: 321–322) with a similar intention.

label of the phase vP. How, then, does IA play its own role? IA has a (valued/interpretable) thematic θ-feature and performs as the subject of unaccusative verbs. In the best case, IA labels a phase vP as [UNACCUSATIVE] by merging into "Spec-v" and by valuing the EF of v. Legate's (2003: 507–508) observation tells us about the plausibility of this scenario. She demonstrates, based on reconstruction effects, that successive-cyclic *wh*-movement leaves a copy in the VP-adjoined position of passive and unaccusative predicates. We can regard this movement into the intermediate position as the merger of IA into "Spec-v" for labeling.

Let us turn to nonphasal valuation for VP. It is reasonable to consider, based on Chomsky's (2008: 148–149) statement, that feature inheritance is also active in the v-V domain. A nonphase VP is thus created by IM that relates IA that enters into a φ-agreement relation with V to the domain of V. [vφ] on IA is then defined as a prominent EF-value assignor in the domain of V, as with the case of nonphasal valuation for TP in (7). As a result, the nonphase VP is labeled as [vφ].

Observe the following tree diagram, which depicts the result of (non)phasal valuation in the v-V domain:

(13)
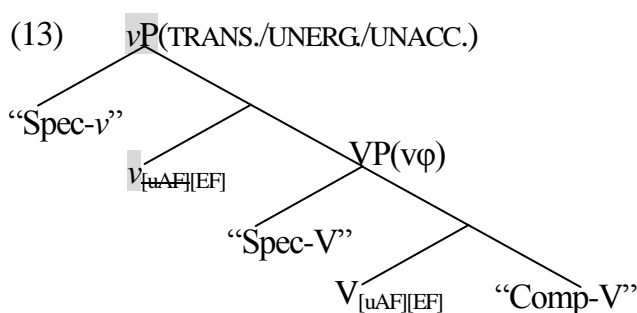
EM of EA into "Spec-v" labels a phase vP via its agentive θ-feature as [TRANSITIVE/UNERGATIVE]; IM of IA from "Comp-V" to "Spec-v" and from "Comp-V" to "Spec-V" labels a phase vP via its thematic θ-feature as [UNACCUSATIVE] and a nonphase VP as [vφ], re-

spectively.

Here, a question arises as to what warrants the information regarding argument/event structure that defines a particular structural relationship between predicates and arguments. I propose that the theory of (non)phasal valuation offers an explicit answer to this question along with the framework of Distributed Morphology (DM) advocated by Halle and Marantz (1993), in which every material, including words (i.e. LIs) and phrases (i.e. SOs), is generated by Merge at narrow syntax as the only derivational component, a perspective known as the single engine hypothesis (Arad (2003: 738)). Within this framework, the merger of category-neutral Roots and phasal categorizers derives LIs. A Root can be merged with any categorizer and can be any category in principle. For instance, "v" verbalizes, "n" nominalizes, and "a" adjectivizes a Root (note that I describe verbalizers as the unitalicized form v and (non)transitivizers as the italicized form v). Indeed, however, a Root cannot always be every category because it arguably involves a given semantic property compatible with a certain categorizer. Harley (2005) and Levinson (2007) argue that Roots are composed of at most three types of semantic properties, namely, events, individuals, and states (see also Marantz (2013: 159)), each of which would be related to verbness, nouniness, and adjectiveness. Under this idea, Roots that involve the semantic property of events are "verbal." Such Roots can be verbalized. In this sense, we may refer to Roots that can be verbs as "verbal Roots."

The merger of verbal Roots and v creates V, which constitutes a phonological and semantic unit and hence a phase. Such a verbal Root-v complex is thus defined as a phasal "SO," which undergoes the Transfer operation as a whole (cf. Marantz (2001, 2007)):[11]

---

[11] It is worthwhile to note here that our approach drawing on DM allows V (i.e. a "minimal head") to be a phase and VP (i.e. a "maximal projection") to be a nonphase. This is equivalent to stating that its syntactic status may shift when a head "projects up." This statement is consistent with the theory of labeling but is in contradiction to X′-theory. Under our analysis, phasal valuation and nonphasal valuation guarantee the label of V and the label of VP, respectively. This mechanism makes the former phasal and the latter nonphasal. X′-theory, on the other hand, demands the projecting head (see chapter 1). If

(14)

$$V \rightarrow_{\text{TRANSFER}} V_{[EF]}$$

V branches into √ROOT and v[EF]; TRANSFER converts it to V[EF].
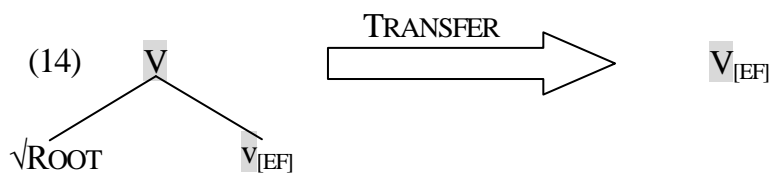
Transfer converts V as an SO into V as an LI, which is the only possessor of the EF (see chapter 3).[12]   Recall that (non)phasal valuation determines the interpretational property of SOs.   Since V is now a phasal "SO" comprising a verbal Root and v, it should receive interpretation at the interfaces in such a way that the most prominent property of the verbal Root determines the interpretational property of the phasal "SO" by valuing the EF of v.

Let us then search for the most prominent property of verbal Roots.   According to Harley (2005) and Levinson (2007), verbal Roots bear the semantic property of events, as mentioned above.   I would like to qualify this property as information about argument/event structure, which fixes the type of V as a verbal Root-v complex.   V takes IA as its argument.   Given that verbal Roots determine the type of V, the (in)ability to change the state of IA should be attributed to the property of verbal Roots.   This attributability leads us to define the most prominent property of verbal Roots as [±CHANGE].   V ensures the presence or absence of the changing state of IA, resting on whether its verbal Root is [+CHANGE] or [−CHANGE].   The internal property of IA identifies whether the change is spontaneous.   The interpretational property of an entire verbal SO {EA, {v, {V, IA}}} or {IA, {v, {V, IA}}} is determined by the inter-

minimal V is phasal, maximal VP is also required to be phasal.   However, we find it impossible to show that VP is a phase on both conceptual and empirical grounds (see e.g. Richards (2007)).   As far as DM is concerned, the computational system is likely to prefer labeling to projection (see also Chomsky (2013: 43), who concludes from a different perspective that "structures need not be endocentric").

[12] This mechanism derives the effect that the lexical integrity hypothesis (Lapointe (1980: 8)) produces, making the transferred domain invisible to further application of operations.   It should be noted here that the transferred domain in words is the whole phase as a Root-categorizer complex whereas that in phrases is the complement of phases (see chapter 2).   In the domain of phrases, unlike that of words, there is no need for SEM and PHON to interpret a phasal edge and complement simultaneously.   This distinction may explain the difference in the transferred domain between words and phrases.

action between the predicate type specified by phasal valuation and the meaning expressed by a constituent comprising V (essentially a verbal Root-v complex) and IA.

We have thus acquired phasal valuation in predicate systems:

(15)   Phasal valuation in predicate systems:

a.   Merge links EA or IA to the domain of *v* and creates a phase *v*P.   Then, its agentive or thematic θ-feature values the EF of *v* as [TRANSI-TIVE/UNERGATIVE] or [UNACCUSATIVE].   The valued EF identifies the interpretational property of the phase *v*P at the interfaces.

a.   Merge links a verbal Root to v and creates a phase V.   Then, its information about argument/event structure values the EF of v as [+CHANGE] or [–CHANGE].   The valued EF identifies the interpretational property of the phase V at the interfaces.

The merger of EA or IA into "Spec-*v*" specifies the type of predicate, and this specification must be consistent with the meaning represented by a constituent consisting of V and IA.   The proposed analysis successfully captures Marantz's (1981: chapter 2) insight into the introduction of logical subjects.   He corroborates that the choice of objects affects the semantic role of logical subjects but not vice versa.   This is tantamount to stating that VP takes a logical subject as its argument.   Our analysis can interpret this selectional requirement as follows.   The logical subject is a "Spec-*v*" occupier, which fixes the predicate type via phasal valuation as [TRANSI-TIVE/UNERGATIVE] or [UNACCUSATIVE].   If the fixed predicate type accords with the meaning expressed by a constituent comprising V and IA, it satisfies the selectional requirement; otherwise, it is in breach of the selectional requirement, with the relevant derivation not convergent at the interfaces.   This mechanism theorizes Marantz's insight into the introduction of logical

subjects.

Crucially, our general theory of labeling has indicated that the mate of a (non)phase head is a prior material in the computational system. Under this theory, the merger of Roots and (phasal) categorizers is not a category-assigning process but is rather a labeling process in which Roots fix the interpretational status of categorizers by providing them with the information regarding argument/event structure. This shift in perspective is the essence of our theory, as emphasized in chapter 3. The result is the elimination of the need to stipulate the merger of Roots and categorizers as the process of categorization, according to which Roots must be always licensed by categorizers for interpretation at the interfaces (cf. Embick and Marantz (2008: 6)). The merger of Roots and categorizers follows naturally from the theory of (non)phasal valuation.[13]

This subsection has extended our general formulation of labeling developed in the C-T domain to the v-V domain. It has been shown in this process that the stipulative mechanism of categorization in DM is reducible to the mechanism of labeling. In section 4.4.2, we confirm that the proposed analysis properly derives predicates in English.

4.4.2. Deriving Predicates

The analysis proposed in the preceding subsection enables us to classify predicates into two types: the predicate of transitives/unergatives typed by EM of EA into "Spec-v" and the predicate of unaccusatives typed by IM of IA from "Comp-V" into "Spec-v." In the next subsections, we consider the derivations of these two types of predicates.

---

[13] (Non)phasal valuation is expected to be a general labeling process for determining the interpretational status of every Root-categorizer complex (e.g. V, N, A) and its "projected" SO (e.g. VP/vP, NP/nP, AP/aP), but I leave clarification of the entire labeling process open for future research.

4.4.2.1. Transitives/Unergatives

As stated above, EM of EA into "Spec-*v*" characterizes the predicate of transitives/unergatives:

(16)　　a.　　The enemy destroyed the city.

　　　　b.　　[$_{vP(TRANSITIVE)}$ The enemy $v_{\text{[uφ][EF]}}$ [$_{VP(vφ)}$ the city destroy$_{[uφ][EF]}$ {$_{V(+CHANGE)}$ √DESTROY v$_{[EF]}$} <the city$_{[vφ]}$>]]

(17)　　a.　　John walked.

　　　　b.　　[$_{vP(UNERGATIVE)}$ John $v_{\text{[uφ][EF]}}$ [$_{VP(vφ)}$ PRO walk$_{[uφ][EF]}$ {$_{V(-CHANGE)}$ √WALK v$_{[EF]}$} <PRO$_{[vφ]}$>]]

In (16b), the predicate structure of (16a), EA *the enemy* is externally merged into "Spec-*v*." Phasal valuation enables its agentive θ-feature to label the phase *v*P as [TRANSITIVE]. The semantic interface verifies whether the labeled *v*P is compatible with the SO formed by the verb *destroy* and IA *the city*. The verb *destroy* is derived by the merger of √DESTROY and v (the curly bracket signifies the internal structure of V). In this case, √DESTROY values the EF of v as [+CHANGE] under the mechanism of phasal valuation. Subsequently, *destroy* and *the city* are externally merged (here and below, we disregard nonphasal valuation for VP for simplicity). The former guarantees the changing state of IA in which the intact state of *the city* is altered to the destructive state. The internal property of the latter in turn ensures that the destruction is nonspontaneous. That is, *the city* is in essence bound to be externally destroyed as its intrinsic nature if it incurs destruction. This state expressed by the SO consisting of *destroy* and *the city* conforms with the predicate type identified via phasal valuation by EM of EA *the enemy* into "Spec-*v*." Derivation (16b) consequently converges with adherence to FI.

In (17b), the predicate structure of (17a), EA *John* is externally merged into "Spec-*v*." Based on phasal valuation, *John*'s agentive θ-feature assigns the phase *v*P the label of [UNERGATIVE]. At the V-internal level, √WALK values the EF of v as [–CHANGE] and produces the verb *walk*. Suppose now that the act of *John*'s walking in (17a) denotes that *John* as an external force instructs himself to step forward slowly. Then, (17a) would have the predicate structure in (17b), in which EM associates *walk* with PRO. PRO is referentially empty and hence needed to be controlled by something. This situation represented by the SO including *walk* and PRO is consistent with the predicate type specified via phasal valuation by EM of EA *John* into "Spec-*v*." EA *John* that occupies "Spec-*v*" controls PRO within VP. The result is the FI observance.[14]

4.4.2.2. Unaccusatives

In contrast to transitive/unergative predicates, unaccusative predicates involve IM of IA into "Spec-*v*":

---

[14] The structure where the verb *walk* takes PRO as IA can be independently borne out from the fact that unergative predicates allow for the occurrence of reflexives (from Levin and Rappaport (1995: 35)):

  (i) a. Dora shouted herself hoarse.
     b. [$_{v\text{P(UNERGATIVE)}}$ Dora $v_{[u\varphi][EF]}$ [$_{\text{VP}(v\varphi)}$ herself shout$_{[u\varphi][EF]}$ {$_{\text{V(–CHANGE)}}$ √SHOUT v$_{[EF]}$} [$_{\text{SC}}$ <herself$_{[v\varphi]}$> hoarse]]]

We can analyze the predicate structure of (ia) as (ib), in which the verb *shout* takes a small clause as IA. Within the small clause, the reflexive *herself* establishes predication with *hoarse*. If this analysis is motivated, we could argue that unergative predicates generally admit the occurrence of IA in an abstract way but the invisible element is not realized without some special environment, such as in (ib), in which PRO is considered to be manifested as *herself* through the establishment of predication relation. We cannot state anything more promising than this claim at the moment, so we defer the in-depth analysis of the internal structure of unergative predicates to future investigation. Incidentally, we will deal with the issue of the labeling of the small clause in section 4.5.1.

(18) a. The rose blossomed.

   b. [$_{vP(\text{UNACCUSATIVE})}$ The rose $v_{\text{[u}\varphi\text{][EF]}}$ [$_{VP(v\varphi)}$ the rose blossom$_{\text{[u}\varphi\text{][EF]}}$ {$_{V(+\text{CHANGE})}$ √BLOSSOM $v_{\text{[EF]}}$} <the rose$_{\text{[}v\varphi\text{]}}$>]]

(18b) is the predicate structure of (18a), in which IM raises IA *the rose* to "Spec-*v*." In this case, its thematic θ-feature guarantees that the relevant phase *v*P receives the label of [UNACCUSATIVE] through phasal valuation. The verb *blossom* is derived in the form that √BLOSSOM values the EF of v as [+CHANGE]. Subsequently, *blossom* and *the rose* are externally merged. Because *the rose* comes into bloom through an internal force, it needs no external force (even if the spring of the internal force results from external factors, such as the sun, water, and/or plant food). This VP-internal information accords with the predicate type ensured via phasal valuation by IM of IA *the rose* into "Spec-*v*." Derivation (18b), observing FI, converges as a result.

4.5. Consequences: XP-YP Merger

Thus far, I have shown that the theory of (non)phasal valuation not only serves as a labeling process but also produces some desirable results both conceptually and empirically. In this section, by employing Narita's (2011) approach to Merge, I argue that our theory gives a more principled explanation for the cases of XP-YP merger kicked around in section 4.2.1 (i.e. (2) and (3)).

4.5.1. Chomsky's Cases

Chomsky's (2013) approach to labeling requires a (non)phase head to determine the label of its SO. Our approach is not inconsistent with this approach in this respect, but there is a crucial difference with respect to the locus of (non)phasehood. While Chomsky considers the property of a (non)phase head to be destined by default, our theory renders the same underspeci-

fied at the initial stage of derivation.   The former regards the property of a (non)phase head directly as the label of its SO, which is the theory of LA; the latter states that a certain LI merged with the domain of a (non)phase head identifies its property, which behaves as the label of its SO, the theory of (non)phasal valuation.

The cases of XP-YP merger, the complex labeling cases discussed earlier, guarantee the advantage of our analysis.   To ascertain the facts, let us consider how (non)phasal valuation operates on the cases in (2) and (3), repeated here as (19) and (20).

(19)　　a.　　　…T [$_\beta$ (EA) [$_v$ [V IA]]]

　　　　b.　　　XP copula [$_\beta$ XP, YP]

(20)　　a.　　　[$_C$ C [$_\alpha$ NP TP…]]

　　　　b.　　　[$_\alpha$ NP CP…]

According to Chomsky (2013: 43), the structure of XP-YP merger is labeled via two strategies by LA: (i) modify SOs so that there is only one visible head or (ii) consider X and Y to be identical in that they provide the same label (see section 4.2.1 for Chomsky's LA-based account of (19) and (20)).   This account not merely increases the complexity of any analysis based on LA but also faces a crucial problem.   The problem lies in the fact that LIs are the only possessors of the EF and thus SOs are not related to each other by Merge because they lack the EF (see chapter 3).   This fact means that there is no XP-YP merger in the computational system, which undermines Chomsky's analysis that presupposes XP-YP merger (see also note 16).

Narita (2011: chapter 3) proposes an effective strategy by which any theory that adopts the EF licenses "XP-YP merger."   Because XP and YP are both SOs, they have no EF and are not linked by Merge.   The way in which such two SOs undergo Merge is that Transfer converts at least one of them into a simplex LI, which is the only bearer of the EF.   Under this proposal, the

structure of "XP-YP merger" is actually analyzed either as SO = {X, YP} or as SO = {Y, XP} (order irrelevant), from which subject and adjunct island effects are derived (see section 4.5.2 for detailed discussions).   If this analysis is on the right track, our theory has a simple explanation for the cases in (19) and (20):

(21)   CP(±WH)

   "Spec-C"

     C$_{[uAF][EF]}$ TP(vφ)

        "Spec-T"

          T$_{[uAF][EF]}$ *v*P(TRANS./UNERG./UNACC.)

            "Spec-*v*"

              v$_{[uAF][EF]}$ VP(vφ)

                "Spec-V"

                  V$_{[uAF][EF]}$ "Comp-V"
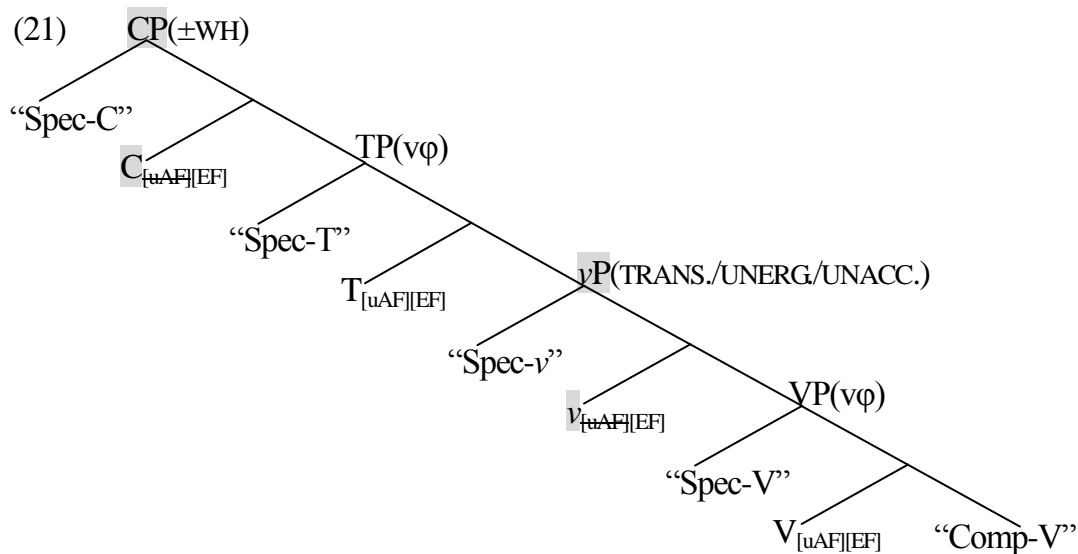
Let us start by explaining the predicate-internal construction in (19a).   Under the structure depicted in (21), the derivational stage that creates β in (19a), where EA/IA is externally/internally merged into "Spec-*v*," corresponds to the case of "XP-YP merger."   Narita's analysis provides an appropriate method of addressing this case.   Under his analysis, nominal phrases like EA/IA have the form of SO = {K(ase), {D, N}}, which Transfer converts into a simplex LI K as a phase head (see chapters 2 and 3 for details).   This mechanism allows for EM/IM of EA/IA, with EA/IA virtually being an LI with the EF via Transfer.   Phasal valuation then labels the relevant phase *v*P via EM/IM of EA/IA into "Spec-*v*" as [TRANSITIVE/UNERGATIVE] or [UNACCUSATIVE], the label of β in (19a) (see section 4.4.2).

The same account applies to the copula construction in (19b), where the property of YP identifies the label of β.   The (subject) XP in (19b), undergoing Transfer, becomes a simplex LI

K as a phase head, which is externally merged with YP. As a result, β in (19b) is defined as SO = {X, YP}, in which X is a phase head K having the EF. The property of YP then values the EF of X via phasal valuation, thus determining the label of β.

The cases of "Spec-*v*-to-Spec-T raising" and *wh*-movement in (20) receive essentially the same account. As depicted in (21), (Non)phasal valuation in the C-T domain labels the phasal and nonphasal SOs as [±WH] and [vφ], respectively. Based on Narita's analysis, the elements represented as NP, the subject element in (20a) and the subject or object *wh*-phrase in (20b), virtually obtain the status of a simplex LI K as a phase head by undergoing Transfer, and they are internally merged with the C-T domain:

(22)   a.   Mary saw John?   (= (11b))

         b.   [CP(–WH) Mary C[uφ][EF] [TP(vφ) Mary T[uφ][EF] [vP(TRANSITIVE) <Mary[vφ]> v[uφ][EF]
               [VP(vφ) John saw[uφ][EF] <John[vφ]>]]]]

(23)   a.   Who saw John?   (= (8b))

         b.   [CP(+WH) Who C[uφ][EF] [TP(vφ) who T[uφ][EF] [vP(TRANSITIVE) <who[vφ]> v[uφ][EF] [VP(vφ)
               John saw[uφ][EF] <John[vφ]>]]]]

(24)   a.   Who did John see?   (= (8a))

         b.   [CP(+WH) Who C[uφ][EF] [TP(vφ) John T[uφ][EF] [vP(TRANSITIVE) <John[vφ]> v[uφ][EF] [VP(vφ)
               who see[uφ][EF] <who[vφ]>]]]]

What characterizes "Spec-*v*-to-Spec-T raising" is nonphasal valuation for TP in (22b), in which IM of two copies of *Mary* labels CP and TP via (non)phasal valuation as [–WH] and [vφ], respectively. α in (20a) is accordingly defined as [vφ]. On the other hand, α in (20b) is characterized via phasal valuation for CP as [+WH] by a subject *wh*-phrase (see (23)) or by an object *wh*-phrase (see (24)).

As just seen, our analysis extending Narita's (2011) theory is more successful in that there is no stipulation—only a principled explanation—for the complex cases in (19) and (20). Only the principles and operations of minimal syntax, such as (non)phasal valuation and Transfer, are prerequisites. In section 4.5.2, we examine other cases of "XP-YP merger," including the merger of subject/adjunct materials and root phrases. Here too, our analysis is shown to easily address these cases without any difficulty in conjunction with Narita's approach, which would be abortive, aligning with Chomsky's LA-based analysis.

4.5.2. Island Cases

An interesting point of Narita's (2011) theory is that it derives subject/adjunct island effects. It is well known that subject/adjunct phrases resist subextraction from within, which Huang (1982) describes as CED (Condition on Extraction Domain). The relevant examples are cited from Narita (2011: 102, 107):

(25)   a.   ＊Which person$_i$ were [pictures of t_i] on sale?

          b.   ＊I know what$_i$ the man criticized Mary [after she said t_i].

The ungrammatical natures of (25a) and (25b) show that subject/adjunct phrases constitute islands for extraction in general. Narita explains this fact in an elegant manner. As already discussed, the merger of subject materials and root phrases is one instance of "XP-YP merger" (cf. (19a)), which is also true of the merger of adjunct materials and root phrases. Under Narita's analysis, Transfer converts subject/adjunct phrases into simplex LIs as phase heads. The transferred domains inside of subject/adjunct phrases are thus no longer visible to syntactic operations (see section 2.6.2 in chapter 2). These materials therefore resist subextraction from within.

If we situate Narita's analysis within the theory of LA, a potential problem for labeling arises. By contrast, our theory of labeling causes no damage to his analysis. To see if these assertions are the case, let us consider the merger of adjunct materials and root phrases in (25b). Suppose, for simplicity, that the adjunction in question targets at the phase vP, producing the structure of [$_\gamma$ vP CP], in which CP virtually behaves as a simplex LI C by undergoing the Transfer operation.[15] Then, LA predicts that the phase head C determines the label of γ. This prediction, however, is unnatural because this simplex LI is not an argument but an adjunct, which is merely a subsidiary object in the grammatical system. Labeling should be induced by something other than adjuncts (see Narita (2011: section 5.5) for a brief remark on his treatment of adjuncts).[16] Our approach to labeling, on the other hand, can properly handle such a case. The item at issue is a simplex LI C and hence bears the EF. Phasal valuation dictates that the property of vP, perhaps the label of [TRANSITIVE], must value the EF of C, thereby labeling γ in [$_\gamma$ vP CP]. The case of "XP-YP merger" in (25b) thus does not constitute any problem for labeling. Of course, the merger of subject materials and root phrases in (25a) also creates no challenge for labeling, as (22b) illustrates. Again, the theory of (non)phasal valuation has a more principled explanation than that offered by the theory of LA.[17]

In this section, by scrutinizing some cases of "XP-YP merger," we have shown that any analysis based on LA ends up in failure and that the theory of (non)phasal valuation allows for

---

[15] Lasnik and Saito (1992: section 4.1.2.2) demonstrate that adjunct clauses, such as the *after* clause in (25b), have a CP structure.

[16] Even if XP-YP merger is licensed in principle, there are doubts about whether LA can label γ in [$_\gamma$ vP CP]. Strategy (i) could not guarantee the label of γ, given that IM raises neither vP nor CP. As for strategy (ii), it is not obvious what feature the root vP shares with the adjunct CP.

[17] The theory of feature inheritance forces EM in the C-T domain of (22b) to generate the SO {C, {T, {EA, v}}} at first. If IM derives {EA, {C, {EA, {T, {EA, v}}}}} straightforwardly from {C, {T, {EA, v}}}, then the process would involve one countercyclic IM application of EA, as noted in note 7 of chapter 2. Setting aside this problem here, I will refine this sort of offending derivation under the analysis presented in chapter 5.

simple explanation, in conformance with SMT.

4.6. Concluding Remarks

In this chapter, we argued that the theory of LA proposed by Chomsky (2013), the process of giving a certain label to a simple set created by Merge, reduces to the theory of (non)phasal valuation explored in chapter 3.   Generalizing the mechanism of clausal typing proposed by Cheng (1997) within the theory of phases, the theory of (non)phasal valuation has introduced a distinctive perspective under which the mate of a (non)phase head is the locus of a labeling process in the computational system, in opposition to the standard MP view.   In doing so, we showed that the stipulative mechanism of categorization in DM follows naturally from our general formulation of labeling.   The proposed analysis brought about some interesting results. The case of "XP-YP merger" shows up our contribution, in which (non)phasal valuation, based on Narita's (2011) approach to Merge, leads to a simpler explanation than does any analysis based on LA.   In chapter 5, we will investigate further application of (non)phasal valuation by showing that the theory of (non)phasal valuation easily absorbs the mechanism of feature inheritance.

Chapter 5

Vacuous Movement Phenomena[*]

5.1. Introduction

A series of works in the literature (George (1980), Chomsky (1986a), Agbayani (2000, 2006), among others) has elaborated the vacuous movement hypothesis (VMH), which states that a movement operation without an effect on PF output (i.e. PHON) can be suspended (see also Abe and Hornstein (2010) and Mikami (2011)).   The VMH as presented in the seminal work of Chomsky (1986a) is formulated as follows:[1]

(1)   The vacuous movement hypothesis (VMH):

         Vacuous movement is not obligatory at S-structure.

This formulation indicates that vacuous movement can be delayed until LF (which is different in essence from SEM).   To put it differently, if the phonological effects of movement operations in overt syntax and in covert syntax are the same, movement preferably occurs in covert syntax. This manifests as a prominent difference between subject and object *wh*-constructions:

(2)   a.   Who saw John?

         b.   Who did John see?

---

[*] This chapter is a revised and extended version of Sakamoto (2012).

[1] This chapter does not consider Agbayani's (2000, 2006) version of the VMH.   His analysis is based on the theory of feature movement, which I do not adopt throughout this thesis.   See Shimada (2008) for convincing counterarguments to Agbayani's (2006) analysis.

In (2a), *wh*-movement to Spec-C is suspended at S-structure because covert *wh*-movement to Spec-C yields an equivalent effect on PF output. The situation differs in (2b), in which the *wh*-phrase moves from object position to Spec-C at S-structure, producing an effect on PF output. The VMH in (1) thus states that *wh*-subjects can remain in their original position at S-structure.

The plausibility of the VMH is bolstered by the observation that island effects weaken in *wh*-subject constructions:

(3)   a.   He is the man to whom$_i$ I wonder [who knew [which book to give t_i]].

        b.   He is the man to whom$_i$ I wonder [who John told [which book to give t_i]].

According to Chomsky (1986a: 50), (3a) is more acceptable than (3b).[2] The former includes relativization from within the island in which the *wh*-phrase functions as a subject. In contrast, the latter involves relativization outside of the island in which the *wh*-phrase functions as an object. If *wh*-subjects can stay in situ at S-structure, as guaranteed by (1), then embedded Spec-C in (3a) would serve as an escape hatch for successive-cyclic movement. On the other hand, embedded Spec-C in (3b) is occupied by the object *wh*-phrase *who*. This offending element interferes with the movement of *to whom*. This is why the contrast in acceptability is observed between (3a) and (3b).[3]

That vacuous movement of *wh*-subjects to Spec-C is suspended at S-structure would

---

[2] Because Chomsky (1986a) mentions his acceptability judgments for (3) only in the text, I do not assign any judgment marks to these sentences.

[3] Here, following Chomsky's (1986a) argument, I introduce the contrast in acceptability between (3a) and (3b) as deriving from one of the principles constituting core grammar, more specifically, the VMH. However, the analysis proposed later considers the contrast in (3) to be a manifestation of some sort of subject-object asymmetry and not to be directly related to core grammar (see section 5.2.2 for detailed discussion).

mean that it applies at LF.   The existence of this sort of LF-movement is independently evidenced by the ungrammaticality of (4).

(4)   *How$_i$ do you wonder [who fixed the car t_i]?

Chomsky (1986a: 49) attributes the ungrammaticality of this sentence to an ECP (Empty Category Principle) violation.   The LF-movement of *who* to embedded Spec-C, which satisfies semantic selection for *wonder* (cf. Grimshaw (1979)), eliminates the intermediate trace of *how*, and the proper government of *t* becomes impossible.   The result is an ECP violation.

The VMH is reminiscent of the principle of Procrastinate, which regards covert operations as less costly than overt operations.   If an operation does not need to be overt to meet some condition, the operation should apply at LF (cf. Chomsky (1995a: 69, 198)).   This view is not conceptually preferable because it demands comparison between the derivation with an overt operation and the derivation with an equivalent covert operation (cf. Collins (1997)).   In addition, movement operations reduce to Merge within the modern minimalist framework (Chomsky (2004, 2007a, 2008), among others).   Merge (i.e. external Merge (EM) and internal Merge (IM)) only applies at narrow syntax and yields pairs of SEM and PHON via Transfer (see sections 2.2 and 2.4 in chapter 2).   This means that there occurs no LF-movement in derivation. These minimalist tenets exclude the VMH, which premises the existence of LF-movement. The VMH must be reconsidered in the context of structure-building computation at narrow syntax.

Given these considerations, a specific question automatically arises as to what derives the VMH effect, in which subjects refuse vacuous *wh*-movement at overt syntax (i.e. narrow syntax) but behave as if they reside in Spec-C at covert syntax (i.e. SEM).   There are three logical possibilities for vacuous movement at narrow syntax: (i) vacuous movement is not permitted at all;

(ii) vacuous movement is optional, as stated in the VMH; or (iii) vacuous movement is regulated by some principle in a minimalist framework. This chapter shows that (iii), from which the VMH effect follows, is the case.

This chapter is organized as follows. Section 5.2 unveils the nature of vacuous movement phenomena in the form that the theory of (non)phasal valuation developed in chapter 3 incorporates the mechanism of feature inheritance. Section 5.3 discusses some consequences of the proposed analysis. Section 5.4 offers some concluding remarks.

5.2. Extension of the Mechanism of Feature Inheritance

We saw in the preceding chapters that the strong minimalist thesis (SMT) requires efficient computation with no redundancy. Given that (non)phasal valuation functions as a labeling process crucial for interpretation at the interfaces and hence is the locus of computation, this mechanism should be computationally efficient. Some studies based on a recent minimalist framework (Chomsky (2007a, 2008), Richards (2007)) have revealed that the mechanism of feature inheritance plays an important role in producing efficient computation. In this section, we argue that the theory of (non)phasal valuation can easily absorb the mechanism of feature inheritance, a conceptually favorable result. By pursuing the extensibility of the mechanism of feature inheritance under the theory of (non)phasal valuation, section 5.2.1 refines the derivations of clauses and predicates presented in chapter 4. Section 5.2.2 shows that vacuous movement phenomena emerge as a natural consequence of this refinement.

5.2.1. "Multiple Specs" Analysis: Amalgamation of Phase Heads with Nonphase Heads

In chapter 3, I constructed the theory of (non)phasal valuation in which the edge feature (EF) on any lexical item (LI) is detoxified by valuation in Merge. This theory enables the relevant syntactic object (SO) to contribute to interpretation at the interfaces, thus keeping to Full

Interpretation (FI).   Recall here that an LI can have the (unvalued/uninterpretable) Agree feature (AF).   Under the theory of feature inheritance, a nonphase head LI such as T and V receives the AF from a phase head LI such as C and *v* in a derivational fashion (see section 2.5.1 in chapter 2).   The inherited AF is detoxified by feature valuation based on Agree/Value, which makes the relevant SO interpretable at the interfaces, hence the FI observance.   In this manner, our present system allows every LI to be the locus of EF detoxification but only the nonphase head LI to be the locus of AF detoxification.   I find a problem here; that is, AF and EF detoxification does not operate in a computationally efficient manner, which is a departure from SMT.

To eliminate this problem, I propose that the computational system involves amalgamation of phase heads with nonphase heads:

(5)   CP(±WH)

             TP(vφ)

"Spec-C"

       "Spec-T"

           C–T$_{[AF][EF][EF(C)]}$ *v*P(TRANS./UNERG./UNACC.)

                      VP(vφ)

          "Spec-*v*"

               "Spec-V"

                 v–V$_{[AF][EF][EF(v)]}$ "Comp-V"
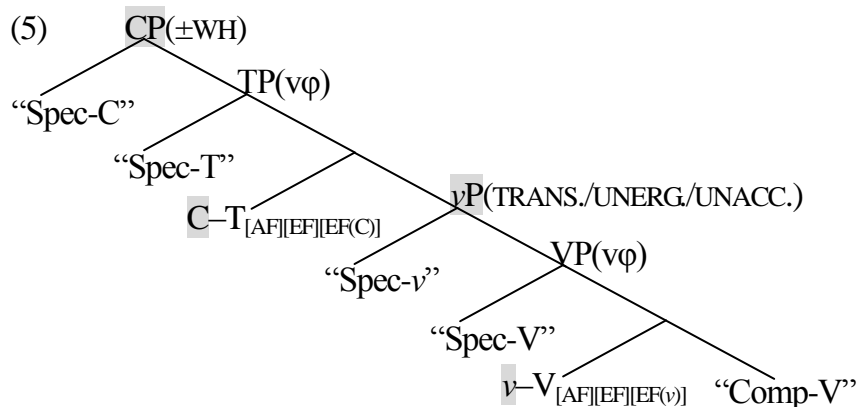
The structure schematized in (5) indicates that a phase head enters into the derivational workspace together with a nonphase head, with these two LIs amalgamated.[4]   Where the computational system tolerates the system of amalgamation (i.e. the C-T amalgam and the *v*-V amalgam), both feature valuation and (non)phasal valuation converge on one complex head, a computa-

---

[4] Tonoike (2013) develops a similar system according to which phase and nonphase heads constitute a "lexical complex" and only the former undergoes excorporation derivationally.   On the other hand, Kitada (2011) extends the mechanism of feature inheritance to the EF.   See also Shimada (2008) for a similar approach to C-T amalgamation.   He proposes a functional hybrid having the properties of both C and T, adopting the economy condition presented by Fukui and Takano (2000: 236).

tionally efficient result. This system in principle enables the syntactic amalgam to license "multiple Specs" merger by doubly employing the EF intrinsic to phase heads and the EF inherent to nonphase heads (note that I describe the EF from phase heads as EF(C) and EF(v), which are introduced only for expository purposes). Under this "dual EF system," "higher Specs" (i.e. "Spec-C" and "Spec-v") are phasal and hence A′-positions, and "lower Specs" (i.e. "Spec-T" and "Spec-V") are nonphasal and hence A-positions. The mechanism of (non)phasal valuation guarantees the labels of both phasal and nonphasal SOs (i.e. CP/vP and TP/VP) through the mergers of "higher Specs" and "lower Specs," respectively. The result is efficient computation in which (non)phasal valuation specifies every label within the C-T amalgam and within the v-V amalgam simultaneously with "multiple Specs" merger at the point of Transfer.[5]

An immediate consequence of the "multiple Specs" analysis proposed here is that it resolves the problem of countercyclicity with the derivations of clauses and predicates shown in chapter 4, namely, the countercyclic application of IM, including "Spec-v-to-Spec-T raising" after the merger of C and TP and "Comp-V-to-Spec-V raising" after the merger of v and VP (see also section 2.6.1 in chapter 2). According to our analysis, the process that produces "Spec-C" and "Spec-T" or "Spec-v" and "Spec-V" is simultaneous under both C-T amalgamation and v-V amalgamation. This operational simultaneity eliminates the existence of the countercyclic application of IM in question, hence efficient computation.[6]

Summarizing the points thus far, I proposed the "multiple Specs" analysis based on amal-

---

[5] It is generally assumed that Transfer maps nonphasal TP/VP onto the interfaces, leaving the phase head C/v and its edge "Spec-C/v" in the derivational workspace. The system of amalgamation depicted in (5), however, compels the phase head C/v to undergo the Transfer operation together with nonphasal TP/VP. This theoretical modification necessitates empirical justification, but I entrust this work to future research.

[6] Note in passing that the "multiple Specs" analysis based on amalgamation, sketched in (5), uniformly produces the OV order at narrow syntax. This statement suggests that the phonetic interface (i.e. PHON) is responsible for the surface VO order in English. What derives the difference in ordering among languages is a longstanding issue (cf. Kayne (1994), Fukui and Takano (1998); see also Takano (1996: chapter 2)). We leave it to future investigation to determine whether our suggestion can grapple with this perplexing issue.

gamation of phase heads with nonphase heads by extending the mechanism of feature inheritance within the theory of (non)phasal valuation. The proposed analysis purges the existence of the countercyclic application of IM, thereby refining the derivation of (non)phases. It is shown in the next subsection that mutational manifestation of the C-T amalgam is responsible for the VMH effect, an additional desirable result.

5.2.2. Deriving the VMH Effect

As indicated above, the interplay between (non)phasal valuation in clausal systems and feature inheritance derives the C-T amalgam with "multiple Specs":

(6) $[_{CP(\pm WH)}$ "Spec-C" $[_{TP(v\phi)}$ "Spec-T" $C{-}T_{[AF][EF][EF(C)]}{\ldots}]]$

Nonphasal valuation assigns the nonphasal SO (i.e. TP) the label of [vφ] as a result of Agree/Value plus subsequent Merge of an LI. Phasal valuation, on the other hand, labels the phasal SO (i.e. CP) as [±WH], as described in (7).

(7)   Phasal valuation in clausal systems:

Merge links a (non-)*wh*-operator to the domain of C and creates a phase CP. Then, its (non-)*wh*-feature values the EF of C as [+WH] or [–WH]. The valued EF identifies the interpretational property of the phase CP at the interfaces.

In this subsection, I investigate only the derivation of [+WH] constructions for the purpose of this chapter, based on the "multiple Specs" analysis in (6).

Let us now consider the derivation of the object *wh*-construction in (8a), the derivation of which, given in (8b), behaves as just a simple manifestation of (6).

(8)   a.   Who did John see?   (= (2b))

   b.   [$_{CP(+WH)}$ Who [$_{TP(vφ)}$ John C–T$_{[uφ][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <John$_{[vφ]}$> [$_{VP(vφ)}$ <who>

   v–V$_{[uφ][EF][EF(v)]}$ <who$_{[vφ]}$>]]]]

As shown in (8b), C-T amalgamation licenses "multiple Specs," in which the "higher Spec" and

the "lower Spec" are analyzed as "Spec-C" and "Spec-T," respectively.   Along with the "mul-

tiple Specs" merger, (non)phasal valuation labels the phasal SO as [+WH] and the nonphasal SO

as [vφ].   Such labels contribute to interpretation at the interfaces.[7]

   In contrast to (8a), the subject *wh*-construction in (9a) exhibits a special realization of (6).

(9)   a.   Who saw John?   (= (2a))

   b.   [$_{TP(+WH,\ vφ)}$ Who   C–T$_{[uφ][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <who$_{[vφ]}$>   [$_{VP(vφ)}$ John

   v–V$_{[uφ][EF][EF(v)]}$ <John$_{[vφ]}$>]]]

As depicted in (9b), the derivation for (9a), subject *wh*-constructions display a special case for

the "multiple Specs" analysis proposed here.   Notice that this derivation leaves the "higher

Spec" (i.e. "Spec-C") empty.   We know that a subject *wh*-phrase is composed of both a

wh-feature and a φ-feature.   The merger only of *who* with the C-T amalgam therefore drives

both phasal and nonphasal valuation, which labels the phasal and nonphasal SO as [+WH, vφ],

thus making the existence of the "higher Spec" (i.e. "Spec-C") redundant.   Given that (9b) has

only one "Spec" within the C-T amalgam, the semantic interface would interpret it as "Spec-T."[8]

---

[7] (8b) has the predicate-internal structure in which (non)phasal valuation labels the phasal SO via EM of *John* as [TRANSITIVE] and the nonphasal SO via IM of *who* as [vφ].   Below, we refrain for the sake of convenience from mentioning predicate-internal derivation.

[8] Unaccusatives seem to be equivalent to the predicate version of a special case for the "multiple Specs" analysis:

The derivation given here properly derives the VMH effect, in which subjects refuse vac-uous *wh*-movement at overt syntax (i.e. narrow syntax) but behave as if they reside in Spec-C at covert syntax (i.e. SEM). In (9b), the subject *wh*-phrase does not undergo overt vacuous *wh*-movement, namely, "Spec-T-to-Spec-C raising" because it creates only "Spec-T" based on C-T amalgamation. Nonetheless, this derivation produces the same effect as covert vacuous *wh*-movement because (non)phasal valuation labels the phasal and nonphasal SO as [+WH, vφ]. The label of [+WH] will be interpreted as interrogative at SEM. The VMH effect is conse-quently derived naturally from the theory of (non)phasal valuation that incorporates feature in-heritance.

It should be noted that the derivations of subject and object *wh*-constructions proposed here predict that (3a) as well as (3b) will exhibit *wh*-island effects uniformly:

(10)　　a.　　He is the man to whom I wonder [$_{TP(+WH, vφ)}$ who C–T$_{[uφ][EF][EF(C)]}$ [<who$_{[vφ]}$>

knew which book to give <to whom>]]

　　　　b.　　He is the man to whom I wonder [$_{CP(+WH)}$ who [$_{TP(vφ)}$ John C–T$_{[AF][EF][EF(C)]}$

[<John$_{[vφ]}$> told <who> which book to give <to whom>]]]

As illustrated in (10a-b), both structures are parallel in involving C-T amalgamation. The dif-ference is attributed to whether or not the C-T amalgam establishes "multiple Specs." The *wh*-phrase *who* in (10a), being a subject element bearing the dual nature in question, forms only "Spec-T." In contrast, *who* in (10b) is an object element, which does not enter into a

---

(i)　　a.　　The rose blossomed.
　　　b.　　[$_{VP(UNACCUSATIVE, vφ)}$ The rose *v*–V$_{[uφ][EF][EF(v)]}$ <the rose$_{[vφ]}$>]

In (ib), the structure for (ia), IA *the rose,* bearing both a thematic θ-feature and a φ-feature, fulfills all fea-tural requirements in the *v*-V amalgam by creating only one "Spec," thus rendering the "higher Spec" (i.e. Spec-*v)* redundant. Unaccusative predicates thus constitute a special case for the *v*-V amalgam, just as subject *wh*-constructions behave as a special case for the C-T amalgam.

φ-agreement relation with the AF [uφ] on T. It is *John* that establishes such a relation. The C-T amalgam in (10b) thus produces "multiple Specs." Although there is a difference in the presence or absence of "multiple Specs" between (10a) and (10b), both cases have a *wh*-island as the complement of *wonder* because phasal valuation guarantees the label of [+WH] in any case. Given that the structures in (10) both involve the process of relativization from within their complement clauses that function as *wh*-islands, the sentences in (3) would be uniformly ruled out.[9]

The view laid out here is different from Chomsky's. Recall that Chomsky (1986a) ascribes the subtle contrast in acceptability between (3a) and (3b) to the VMH as one of the principles constituting core grammar (see section 5.1). Our system, however, equally excludes both sentences for the reason stated above and counts the subtle contrast in (3) just mentioned as a manifestation of some sort of subject-object asymmetry. In fact, some previous studies (Cheng (1997), Richards (2001), Agbayani (2006), etc.) consider there to be no grammatical difference between extraction from the island where the *wh*-phrase functions as a subject (cf. (3a)) and extraction from the island where the *wh*-phrase functions as an object (cf. (3b)). Our present conclusion conforms with this perspective (see section 5.3 for the account of phenomena that are supposed to display the VMH effect in the true sense).[10]

---

[9] Here, I set aside an issue as to what exactly induces a *wh*-island effect because consideration of the issue digresses from the main topic of this chapter.

[10] The following grammatical contrast provided by Chomsky (1986a: 48–49) should also be treated in our system:

(i) a. What$_i$ do you wonder [who saw t_i]?
     b. *How$_i$ do you wonder [who fixed the car t_i]?   (= (4))

Sentence (ia), which is parallel to (3a) in that it contains extraction from within the island headed by a subject *wh*-phrase, displays a weaker *wh*-island violation than (ib). Within a preminimalist framework, the contrast in (i) is ascribed to the ECP. In (ib), the trace of *how* is not properly governed for the reason already stated in section 5.1. The trace in (ia), on the other hand, is properly governed because it is an argument of the embedded verb.

75

To sum up this section, I have captured the VMH effect in such a way that the theory of (non)phasal valuation absorbs the mechanism of feature inheritance thereby deriving the "multiple Specs" analysis. This means that the VMH is no longer operative in the system of grammar. I show in section 5.3 that the proposed analysis also gives a principled explanation for other phenomena that exhibit the VMH effect.

5.3.   Consequences

5.3.1.   Licensing Movement with No Effect on PHON

5.3.1.1.   The Reconciliation between the PISH and the VMH

It has been proposed in the literature (Fukui and Speas (1986), Kitagawa (1986), Kuroda (1988), among others) that subject materials originate in predicate-internal positions. This proposal has widely been accepted as the predicate-internal subject hypothesis (PISH). Under any system that postulates the PISH, subject materials are forced to undergo "Spec-*v*-to-Spec-T raising" for the purpose of ensuring a certain type of valuation, which corresponds to labeling/EF detoxification within the theory of (non)phasal valuation. The existence of the VMH would be problematic for such a system since this alleged principle incorrectly rules out "Spec-*v*-to-Spec-T raising" due to its string-vacuous nature. While previous studies have formulated the VMH by focusing on "Spec-T-to-Spec-C raising," they have not taken seriously the conflict between the PISH and the VMH. The analysis proposed in the last section, rendering the VMH inoperative in the system of grammar, reconciles this conflict straightforwardly. The inoperativity of the VMH does not prohibit us from adopting the PISH, more precisely, the

However, we can no longer utilize the ECP as a tool for explaining the contrast in (i) because it does not get along with minimalist tenets. It is expected from the present system that sentences (ia, b) would be equivalently excluded as a violation of minimality, as with the situation in (3), and that the contrast in acceptability between them should be attributed to a difference in referentiality between the relevant *wh*-phrases, no matter what general principle the ECP reduces to in a minimalist framework. I leave a full account of the contrast in (i) for future research.

"Spec-*v*-to-Spec-T raising" approach to subject materials.   In other words, our system allows of movement with no effect on PHON.   Differently from such a principle as the VMH, the occurrence of movement, i.e. IM, crucially depends on the system of amalgamation.[11]

This discussion over the PISH confirms explicitly that our system is superior to any system in which the VMH is operative in that only the former is consistent with the "Spec-*v*-to-Spec-T raising" approach to subject materials.[12]

5.3.1.2.   The Ban on Vacuous Application of Scrambling

Within any system where the VMH does not work, another significant issue emerges as to what excludes "vacuous application of scrambling" in Japanese (Hoji (1985), Takano (1996)). To understand the significance of this issue, let us observe a scopal effect in scrambling (from Hoji (1985: 342)):

---

[11] Note that the absence of the VMH does not also prevent us from employing any analysis that premises movement of phonologically-null elements such as null operators.   We will explore such an analysis in chapter 6.

[12] One *English Linguistics* reviewer points out, based on Takahashi's (1994) arguments that subject materials can stay in situ, that the argument here does not fall into place.   However, an aspect of the matter under consideration lies in the point that, when subject materials move from "Spec-*v*" to "Spec-T" under the PISH, such movement has a string-vacuous nature.   This type of movement follows unproblematically from our system, but not from any system in which the VMH is operative.   The account of the phenomenon noted by the reviewer might be ascribed to the copy theory of movement (Chomsky (1995a)), which allows the lower copy of a moved element to be interpreted and pronounced at the interfaces (cf. Nunes (2004), Mikami (2010), etc.), although I yield substantial pursuit of this possibility to another occasion.

(11) a. Dareka-ga   daremo-o   semeta.

someone-Nom   everyone-Acc criticized

"Someone criticized everyone." (unambiguous)

b. Daremo-o   dareka-ga   <daremo-o> semeta.

everyone-Acc   someone-Nom   criticized

"Everyone, someone criticized." (ambiguous)

(11a) exemplifies the basic word order in a Japanese transitive sentence.   This sentence is sco-

pally unambiguous, with the quantified subject only taking wide scope over the quantified object.

In (11b), the quantified object is scrambled over the quantified subject.   In this case, scopal am-

biguity is observed between the quantified subject and the quantified object.

Under the assumption that the scopal ambiguity in question results from a c-command

relation established between the relevant quantified elements, the observed effect of scrambling

allows us to expect that the two quantified phrases in (11a) exhibit scopal ambiguity if they un-

dergo multiple scrambling in a string-vacuous fashion, as shown in (12).

(12)   [Dareka-ga   [daremo-o   [<dareka-ga> [<daremo-o> semeta]]]]

(12) illustrates that multiple scrambling of the two quantified phrases produces the same word

order as in (11a).   (12), unlike (11a), should be amenable to ambiguous scope interpretation

because scrambling is applied to the quantified object, as observed in (11b).   However, such

scopal ambiguity is not obtained in (12).   Based on facts of this sort, Hoji (1985: 352) reaches

the following conclusion:

(13)   A syntactic adjunction operation cannot apply if it does not change the order of the overt lexical string.

Relating (13) to the VMH as proposed by George (1980), Hoji states that vacuous application of scrambling is not allowed in Japanese (see Hoji (1985: section 3.5) for details).

Hoji's (1985) conclusion is in conflict with the proposed analysis, since our analysis does not prevent vacuous application of movement (cf. section 5.3.1.1) and scrambling is generally grouped as a kind of movement (cf. Hoji (1985), Saito (1985)).   There seem to me to be two possible strategies to resolve this conflict.   One strategy analyzes scrambling as optional movement.   This analysis characterizes scrambling as requiring no grammatical factor that makes it obligatory, as proposed by Fukui (1993).   If scrambling is indeed an instance of optional movement, then it differs in nature from what we discussed in the last subsection, namely, vacuous movement of subject materials from "Spec-v" to "Spec-T," which obligatorily occurs following the mechanism of (non)phasal valuation.   It is impossible, however, that we adopt Fukui's (1993) approach to optional movement straightforwardly because it crucially hinges on the head-parameter, which is not formulable within the theory of FL based on Merge (see section 1.1.2 in chapter 1).   It is difficult, further, to regard scrambling as optional movement since it is often defined as a kind of movement, which the modern minimalist framework reduces to IM, an unconstrained operation.   It is unclear in what way we can formulate whether IM is obligatory or optional.   We cannot therefore attribute the conflict between our analysis and Hoji's (1985) analysis to the difference in movement types.

It is highly promising that we pursue the possibility of the other strategy, which reduces the ban on vacuous application of scrambling to the condition on the syntax-phonology interface referred to as linearization preservation (cf. Fox and Pesetsky (2003), Ko (2007), Takano (2010), etc.):

79

(14)   Linearization preservation (from Fox and Pesetsky (2003: 2)):

The linear ordering of syntactic units is affected by Merge [EM] and Move [IM] *within* a Spell-out domain, but is fixed once and for all at the end of each Spell-out domain.   (brackets added)

(14) holds "that linear order is determined cyclically and that linear order established at the end of a given Spell-Out domain must be preserved at the end of each later Spell-Out domain," which "refers to a syntactic constituent relevant to determining the linear order of syntactic elements" (Takano (2010: 98)).   On the assumption that linearization preservation yields efficient computation, we can understand this condition as a straightforward embodiment of SMT, which properly rules in (11a-b) and out (12) with no recourse to an independent principle such as the VMH.

Setting aside the question of what exactly scrambling is, let us consider for present purposes the derivations of (11b) and (12):

(15)   a. $[_\alpha \text{ Y X } <\text{Y}> \text{ V}]$ (cf. (11b))

b. $\text{X} [_\alpha \text{ Y } <\text{X}> <\text{Y}> \text{ V}]$ (cf. (12))

(15a) and (15b) denote the schematic derivations of (11b) and (12) based on Takano (2010), respectively.   Suppose that CP and *v*P are Spell-Out domains, as defined by Ko (2007), and that α in (15) is CP.   Then in (15a), at the end of the Spell-Out domain α, the order between X and Y is fixed in the form that Y precedes X, with X *dareka-ga*, Y *daremo-o*, and V *semeta*.   Linearization preservation dictates that the fixed order Y-X(-V) be preserved at later Spell-Out domains. Derivation (15b) is excluded because multiple scrambling disrupts this order, creating the altered order X-Y(-V), in violation of linearization preservation.   Note that (11a) is properly linearized

in such a way that X precedes Y, involving no application of scrambling. We can therefore eliminate the conflict between our analysis and Hoji's (1985) analysis, banning vacuous application of scrambling in (12) without appealing to the VMH but with adherence to linearization preservation.

We saw in this section that there are two types of IM with no effect on PHON: (i) vacuous movement of subject materials from "Spec-v" to "Spec-T" (see section 5.3.1.1) and (ii) vacuous application of scrambling (see section 5.3.1.2). Reducing the VMH effect to the theory of (non)phasal valuation incorporating feature inheritance, the analysis proposed in section 5.2 does not prohibit these types of IM from appearing because of the inoperativity of the VMH in the system of grammar. The presence of (i) displays a straightforward advantage for our analysis because any analysis based on the VMH incorrectly does away with such movement to be requisite for labeling/EF detoxification. The tolerance of (ii), on the other hand, can be problematic for our analysis due to the lack of evidence proving its existence. The independent, well-established condition called linearization preservation, however, felicitously excludes the possibility of (ii). We can thus maintain the present analysis with the benefit of (i).

5.3.2.   Superiority Effects

The analysis explored in section 5.2 also covers superiority effects (cf. Chomsky (1973, 1981, 1995a, 2008), Lasnik and Saito (1992), etc.):

(16)   a.   Who saw what?

           b.   *What did who see?

The grammatical asymmetry between these sentences illustrates that in multiple *wh*-constructions, object materials cannot occupy sentence-initial position by crossing subject

materials.   Early minimalism accounted for the observed effect under the minimal link condition, according to which an attractor such as C raises a closer element to its domain if there are more than one candidate for attraction (see Chomsky (1995a: section 4.5.5) for more details). Recall, however, that the mechanism of (non)phasal valuation does not rest on a relation between attractors and attractees on the occasion of movement; rather, a driving force for movement is the EF of a raising element itself (see chapter 3).   Not adopting the system based on attraction, the theory of (non)phasal valuation needs to seek another principled explanation for superiority effects.

Let us now suppose that (16a) corresponds to the well-formed derivation in (17a) and (16b) to the ill-formed derivation in (17b).

(17)   a. $[_{TP(+WH, \ v\varphi)}$ Who   C–T$_{[u\varphi][EF][EF(C)]}$ $[_{vP(TRANSITIVE)}$ $<$who$_{[v\varphi]}>$ $[_{VP(v\varphi)}$ what

v–V$_{[u\varphi][EF][EF(v)]}$ $<$what$_{[v\varphi]}>$]]]

b.   *$[_{CP}$ What $[_{TP}$ who C–T$_{[u\varphi][EF][EF(C)]}$ $[_{vP(TRANSITIVE)}$ $<$who$_{[v\varphi]}>$ $[_{VP(v\varphi)}$ $<$what$>$

v–V$_{[u\varphi][EF][EF(v)]}$ $<$what$_{[v\varphi]}>$]]]]

(17a) and (17b) differ in that only the latter involves object *wh*-movement.   Our system does not allow for this movement insofar as subject *wh*-movement occurs.   As stated earlier, subject *wh*-phrases, unlike object ones, bear both a *wh*-feature and a φ-feature.   This dual property enables a subject *wh*-phrase to solely label the phasal and nonphasal SO in the C-T amalgam as [+WH, Vφ], thereby making the existence of the "higher Spec" (i.e. "Spec-C") redundant.

Note, however, that our theory is always regulated by the principle of FI, under which every SO contributes to interpretation at the interfaces.   Object *wh*-phrases as well as subject *wh*-phrases bear a valued/interpretable *wh*-feature, which has to assign a value to [uF] under the

requirement of FI.[13]   IM of both *what* and *who* into the C-T amalgam, however, leads to a failure in labeling/EF detoxification.   More specifically, two value-assignors with a feature that is equative in quality compete against one another for one value-assignee, the EF of C.   This situation makes the derivation of (17b) illegitimate.[14]   Why then does (17a) converge despite the object *wh*-phrase not being an apparent member of the C-T amalgam?   The key to solving this question is the unselective binding approach to quantification in the sense of Tsai (1994: 58). In conjunction with his approach, our system enables the in-situ *wh*-phrase in (17a) to have its quantificational force.   In (17a), EF(C) becomes "active" by receiving the value of [+WH] based on phasal valuation.   The activated EF(C) unselectively binds the in-situ *wh*-phrase. This in-situ *wh*-phrase accordingly has its quantificational force and behaves as a member of the C-T amalgam.   The derivation of (17a) thus observes FI.

The proposed analysis captures the fact that only one element must undergo *wh*-movement in English multiple *wh*-constructions (Watanabe (1992, 2001)):

(18)   a.   *I wonder who what bought.

        b.   *I wonder what who bought.

Examples (18a-b) show that the complement of *wonder* licenses no multiple *wh*-movement. The account of (17b) is equally applicable to the ill-formedness of these embedded multiple *wh*-constructions.   Here too, (17a) includes the only way that makes this sort of illegitimate

---

[13] Nominal phrases like *wh*-phrases are an SO headed by K, which Transfer converts into a simplex LI K (see chapters 2 and 3).   If a valued/interpretable *wh*-feature on K contributes to interpretation by valuing some [uF], it can be said that the whole SO headed by K fulfills the FI requirement.

[14] It turns out from consideration of multiple *wh*-constructions in German, where no superiority effect is observed, that the language allows for a structure like (17b).   Based on this fact, Chomsky (2008) claims that English is also allowed to have a derivation for generating the sentence in (16b).   However, I believe that English cannot have such a derivation for the reason discussed here.   The origin of the lack of superiority effects in German is unclear, so I leave this issue open for future work.

structure licit.   These facts thus confirm that the analysis presented here is on the right track.

5.3.3.   Semantic Selection

Furukawa and Fukuda (2009) point out a paradigm of semantic selection for verbs that challenges any analysis based on LF-movement:

(19)   I wonder [who saw what].   (Chomsky (1986a: 52))

(20)   a.   *Who wondered [John saw what]?

          b.   Who wondered [what John saw]?   (Furukawa and Fukuda (2009: 271))

The verb *wonder* requires an interrogative complement at the level of semantic representation (see Grimshaw (1979)).   If the verb *wonder* in (19) could satisfy semantic selection via LF-movement of the subject *wh*-phrase to embedded Spec-C, the grammatical contrast in (20) would remain a mystery.   Specifically, it is unclear why sentence (20a) fails to obtain a grammatical status, with the object *wh*-phrase moving to embedded Spec-C at LF.

Our system gives a principled explanation for the paradigm:

(21)   I wonder [$_{TP(+WH, \, v\varphi)}$ who C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <who$_{[v\varphi]}$> [$_{VP(v\varphi)}$ what

          v–V$_{[u\varphi][EF][EF(v)]}$ <what$_{[v\varphi]}$>]]]

(22)   a.   *Who wondered [$_{TP(-WH, \, v\varphi)}$ John C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <John$_{[v\varphi]}$> [$_{VP(v\varphi)}$

                what v–V$_{[u\varphi][EF][EF(v)]}$ <what$_{[v\varphi]}$>]]]

          b.   Who wondered [$_{CP(+WH)}$ what [$_{TP(v\varphi)}$ John C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$

                <John$_{[v\varphi]}$> [$_{VP(v\varphi)}$ <what> v–V$_{[u\varphi][EF][EF(v)]}$ <what$_{[v\varphi]}$>]]]]

The embedded clause of (21) corresponds to a subject *wh*-construction.   Of particular im-

portance is the derivation of the C-T amalgam. Recall that the C-T amalgam of a subject *wh*-construction derives the VMH effect by forming only the "lower Spec" (i.e. "Spec-T"), which solely labels the phasal and nonphasal SO as [+WH, vφ] based on the mechanism of (non)phasal valuation. The labeled SO will be interpreted as interrogative at SEM. The verb *wonder* consequently fulfills semantic selection by taking its proper complement.

This account of semantic selection readily captures the difference in grammaticality between (22a) and (22b). In (22a), *John* rather than *what* occupies embedded "Spec-T." Because *John* has a φ-feature but not a *wh*-feature, (non)phasal valuation labels TP as [−WH, vφ], which is interpreted as declarative at SEM. As a result, the verb *wonder* fails to take its proper complement, and, in addition, *what*, bearing a *wh*-feature, does not value any [uF], thus yielding an FI violation. By contrast, the verb *wonder* can take its proper complement in (22b), which is parallel to an object *wh*-construction. In this structure, *what* but not *John* occupies embedded "Spec-C." Phasal valuation feeds CP with the label of [+WH], which is interpreted as interrogative at SEM, with *John* with a φ-feature independently labeling TP as [vφ]. This mechanism of labeling/EF detoxification satisfies semantic selection for *wonder* and thus the requirement of FI.

Under the VMH, the verb *wonder* in (21) needs to fulfill semantic selection via LF-movement of *who* to embedded Spec-C because the overt counterpart at narrow syntax is suspended. If this is the case, the grammatical contrast in (22) would not fall into place for the reason mentioned above. In this way, the discussion here has further ensured the validity of our system.

5.3.4. Anaphor Binding in Embedded Topicalization

Next, we consider the grammatical asymmetry of anaphor binding in embedded topicalization between (23a) and (23b), which is given by Lasnik and Saito (1992) and later discussed

by Agbayani (2000, 2006) in the context of vacuous movement.

(23)   a.   John$_i$ thinks that himself$_i$, Mary likes t_i.

          b.   *John$_i$ thinks that himself$_i$, t_i likes Mary.

In (23a), the anaphor *himself*, which is topicalized from object position in the embedded clause, can take the matrix subject *John* as its antecedent. On the other hand, in (23b), which involves potential embedded subject topicalization corresponding to embedded object topicalization, there is no anaphor binding relation between *John* and *himself*.[15]

Here, we adopt the hypothesis that Condition A is an "anywhere" condition (Belletti and Rizzi (1988)), under which any level of representations (i.e. D-structure and S-structure) can meet the binding requirements for anaphoric elements. If this hypothesis is reinterpreted within the phase-based system postulated throughout this thesis, it then follows that anaphoric elements can be licensed phase by phase at narrow syntax. Condition A therefore abides by the phase impenetrability condition (PIC) (see section 2.6.2 in chapter 2).

Further, we propose that the verb heading an irrealis complement, such as *think*, *believe*, and *expect*, make defective the AF on the embedded C-T amalgam in the case of ECM (Exceptional Case-Marking) and the AF on the matrix *v*-V amalgam in the case of non-ECM:

---

[15] Sentence (i) from Lasnik and Saito (1992: 198), though somewhat marginal, illustrates that embedded anaphoric subjects can undergo topicalization in principle.

(i)   ?? John$_i$ thinks that himself$_i$, Mary said t_i won the race.

In this sentence, the embedded anaphoric subject, occupying the topicalized position, is permitted to be bound by its matrix antecedent. Thus, the ungrammaticality of (23b) does not mean that its embedded anaphoric subject cannot be subject to topicalization.

(24) a. John believes her to be happy. (ECM)

  b. John believes that she is happy. (non-ECM)

(25) a. [$_{TP(-WH, v\varphi)}$ John C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <John$_{[v\varphi]}$> [$_{VP(v\varphi)}$ her

  v–V$_{[u\varphi][EF][EF(v)]}$ [$_{TP(-WH)}$ (to) <her$_{[v\varphi]}$> $\boxed{\text{C–T}_{[EF][EF(C)]}}$ [$_{VP(UNACCUSATIVE, v\varphi)}$ <her$_{[v\varphi]}$>

  v–V$_{[u\varphi][EF][EF(v)]}$ [$_{SC}$ <her$_{[v\varphi]}$> happy]]]]]]

  b. [$_{TP(-WH, v\varphi)}$ John C–T$_{[u\varphi][EF][EF(C)]}$ [$_{VP(TRANSITIVE)}$ <John$_{[v\varphi]}$> $\boxed{v\text{–V}_{[EF][EF(v)]}}$ [$_{TP(-WH, v\varphi)}$

  (that) she C–T$_{[u\varphi][EF][EF(C)]}$ [$_{VP(UNACCUSATIVE, v\varphi)}$ <she$_{[v\varphi]}$> v–V$_{[u\varphi][EF][EF(v)]}$ [$_{SC}$

  <she$_{[v\varphi]}$> happy]]]]]]

The parts placed in squares express the defectiveness of the relevant amalgams, namely, the lack of the AF. The standard analysis based on the theory of feature inheritance views the ECM complement as TP. The lack of C entails that of AF inheritance, which in turn shows that there is no φ-agreement and thus Case-valuation. Instead, the v-V relation in the matrix clause, being responsible for φ-agreement and Accusative valuation, triggers subject-to-object raising, i.e. movement from embedded "Spec-T" to matrix "Spec-V" (cf. Chomsky (2008: 148)). The principle of minimal computation, however, expects clausal structure uniquely to be C-T-v-V (cf. Richards (2007)), regardless of ECM or non-ECM constructions. And the only characterizations of ECM and non-ECM constructions are the parts placed in squares in (25), respectively. That is, we are presenting the following generalization:

(26) The generalization of AF defectiveness:
  The verb heading an irrealis complement makes defective the AF on a locally related phase head.

According to (26), either of the parts placed in squares in C-T-v-V-C-T-v-V (i.e. either matrix v or embedded C) is defective with respect to the AF. It remains unclear in what way AF defectiveness comes about. That said, it suffices for us to describe the derivational difference between ECM and non-ECM constructions.

Let us now consider (25a), which signifies the derivation of the ECM construction in (24a). (26) states that the matrix verb *believe* renders either the AF on the matrix v or the AF on the embedded C inert. As illustrated in (25a), ECM constructions correspond to the latter case, so the embedded C-T amalgam is defective with respect to the AF. The element *her*, which has been raised from the original position via embedded "Spec-V" to embedded "Spec-T" on the assumption that it is base-generated within the small clause, undergoes IM for the purpose of labeling matrix VP. In contrast, the corresponding element *she* in (25b), the derivation of the non-ECM construction in (24b), stays in embedded "Spec-T" due to the inertness of the AF on the matrix v-V amalgam. The ECM complement, receiving the label of [–WH] but not that of [$v\varphi$], functions as a nonfinite clause. The non-ECM complement, on the other hand, behaves as a finite clause because it gains the label of [–WH, $v\varphi$].[16]

The generalization given in (26) is never an ad hoc one, being in its nature descriptive. Based on the fact that the irrealis infinitive of a control verb, such as *expect*, *hope*, and *want*, is susceptible to heavy NP shift from within, Hirai (2004) demonstrates that it is a weak CP phase ($C^{w}P$ in his term). If the theory of (non)phasal valuation sets in this analysis, it can be said that irrealis control infinitives are headed by defective C-T amalgams, as with the case of the derivation of an ECM construction in (25a). It would not be surprising if there is a defective v-V amalgam, given that the existence of a defective C-T amalgam is confirmable to a certain ex-

---

[16] The current analysis considers *to* in (25a) and *that* in (25b) as a phonetic realizations of the labels of the relevant complement clauses.

tent.[17]   Though necessary to receive a principled explanation, (26) is indeed not an adhoc generalization.

Let us now consider how the proposed analysis explains the grammatical asymmetry of anaphor binding in (23):

(27)   a.   [$_{\text{TP(–WH, v}\varphi\text{)}}$ John C–T$_{[u\varphi][EF][EF(C)]}$ [$_{\text{VP(TRANSITIVE)}}$ <John$_{[v\varphi]}$> $\boxed{v\text{–V}_{[EF][EF(v)]}}$ [$_{\text{CP(–WH)}}$

(that) **himself** [$_{\text{TP(v}\varphi\text{)}}$ Mary C–T$_{[u\varphi][EF][EF(C)]}$ [$_{v\text{P(TRANSITIVE)}}$ <Mary$_{[v\varphi]}$> [$_{\text{VP(v}\varphi\text{)}}$

<himself> v–V$_{[u\varphi][EF][EF(C)]}$ <himself$_{[v\varphi]}$>]]]]]]

b.   *[$_{\text{TP(–WH, v}\varphi\text{)}}$ John C–T$_{[u\varphi][EF][EF(C)]}$ [$_{\text{VP(TRANSITIVE)}}$ <John$_{[v\varphi]}$> $\boxed{v\text{–V}_{[EF][EF(v)]}}$ [$_{\text{TP(–WH, v}\varphi\text{)}}$

(that) **himself**   C–T$_{[u\varphi][EF][EF(C)]}$ [$_{v\text{P(TRANSITIVE)}}$ <himself$_{[v\varphi]}$> [$_{\text{VP(v}\varphi\text{)}}$ Mary

v–V$_{[u\varphi][EF][EF(v)]}$ <Mary$_{[v\varphi]}$>]]]]]

According to the analysis proposed above, derivations (27a-b) have a defective matrix v-V amalgam, with the verb *think* with an irrealis complement taking its non-ECM complement. AF defectiveness means that the relevant v-V amalgam has nothing to do with feature valuation and thus nonphasal valuation, although phasal valuation takes place based on EM of the external argument *John*.   Feature valuation entails the application of Agree/Value and nonphasal valuation the one of Agree/Value plus IM.   The matrix v-V amalgam in (27), which is defective with respect to the AF, involves no application of Agree/Value and IM.   Recall that both Agree/Value

---

[17]   In contrast to the irrealis infinitive of a control verb, Hirai (2004) defines the realis infinitive of a control verb, such as *bother*, *claim*, and *decline*, as strong CP phases (C*P in his term) due to the fact that it oppose against heavy NP shift from within.   If this definition is correct, our analysis regards realis control infinitives as headed by complete C-T amalgams, including the AF.   This view would incorrectly indicate, however, that realis control infinitives correspond to *that*-complements, as in (25b), and that realis control infinitives allow for the occurrence of overt subject elements rather than PRO.   For the full, uniform account of the paradigm of (non-)ECM complements and the paradigm of (ir)realis control infinitives, we would need to investigate (i) the lexical property of the matrix verb and (ii) the licensing condition for the occurrence of PRO/overt subject elements.   Such considerations, however, defeat our main purpose, so I leave these issues open for future investigation.

and IM, requiring copy formation, fall under the mechanism of copy-identification, which forces these copy-formation operations to apply simultaneously with Transfer (see section 2.5.2 in chapter 2 for details). Conversely, the absence of copy-formation operations indicates that there is no application of Transfer. Therefore, until transmission of the matrix TP into the interfaces, the matrix VP in (27), undergoing no Transfer operation, suspends a PIC effect, with the shaded domain representing where Transfer applies.

Since Condition A applies phase by phase at narrow syntax, as assumed above, the domain of its application is in the shaded domain in (27). Because the anaphor *himself* in (27a), a topic-feature bearer, occupies the "higher Spec" of the embedded C-T amalgam for the purpose of phasal valuation, it evades the lower transferred domains. This anaphoric element is allowed to be bound by its antecedent *John* in the highest transferred domain, the matrix VP undergoing the Transfer operation together with this domain for the reason stated above. Thus, in (23a), the coreferential reading obtains between *John* and *himself*.

In contrast, sentence (23b) can establish no such coreferential reading because the anaphor *himself* cannot evade the lower transferred domain, as illustrated in (27b). The embedded clause of (27b) is parallel to the derivation of subject *wh*-constructions, in which subject *wh*-phrases, which are both a *wh*-feature and a φ-feature, form only the "lower Spec" (i.e. "Spec-T") in the C-T amalgam. Similarly, the anaphoric element in (27b), bearing both a topic-feature and a φ-feature, occupies "Spec-T," thereby driving (non)phasal valuation. This situation keeps *himself* in the lower transferred domain. By the time its antecedent *John* emerges, this anaphoric element is not in the derivational workspace, hence no anaphor binding relation is established between *John* and *himself*.

Under the system that posits the VMH, it is not easy to capture the asymmetric behavior of anaphor binding because the VMH does not prevent the anaphoric subject from undergoing LF-movement, which might circumvent a Condition A violation in (23b). This situation does

not actually occur, however.　　Again, it has been confirmed that our system is more plausible.

5.3.5.　Across-the-Board Movement Phenomena

　　　　Our ensuing consideration turns to across-the-board (ATB) movement phenomena of *wh*-phrases.　　It has been argued in the literature that ATB movement is feasible under structural parallelism between conjuncts (cf. Ross (1967), George (1980), Goodall (1987), Bošković and Franks (2000), etc.).　　To see how ATB movement phenomena are derived, I present the following example from George (1980: section 5.3):

(28)　　They removed the prisoner, who(m) the judge has sentenced and (who(m)) the warden will execute.

Each conjunct in (28) has a *wh*-phrase that undergoes movement to the parallel position, "Spec-C" in general.　　In this case, the *wh*-phrase in the second conjunct can be elided, as the parenthesis shows.　　Such situations are known as ATB movement phenomena.　　Our theory can interpret the condition imposed on ATB movement as the parallelism in labeling between the largest SOs:

(29)　　They removed the prisoner, $[_{CP(+WH)}$ who(m) $[_{TP(v\phi)}$ the judge C–T$_{[u\phi][EF][EF(C)]}$ $[_{vP(TRANSITIVE)}$ <the judge$_{[v\phi]}$> $[_{VP(v\phi)}$ <who(m)> *v*–V$_{[u\phi][EF][EF(v)]}$ <who(m)$_{[v\phi]}$>]]]] and $[_{CP(+WH)}$ (who(m)) $[_{TP(v\phi)}$ the warden C–T$_{[u\phi][EF][EF(C)]}$ $[_{vP(TRANSITIVE)}$ <the warden$_{[v\phi]}$> $[_{VP(v\phi)}$ <who(m)> *v*–V$_{[u\phi][EF][EF(v)]}$ <who(m)$_{[v\phi]}$>]]]]

As shown in (29), phasal valuation labels both of the largest SOs via object *wh*-movement as [+WH].　　This parallelism in labeling between the largest SOs licenses ATB movement.

Consideration of other examples supports the proposed analysis.   Again, the relevant data are cited from George (1980: section 5.3):

(30)　　a.　　They removed the prisoner, who has lost his appeal and *(who(m)) the warden will execute.

　　　　b.　　They removed the prisoner, who(m) the judge has sentenced and *(who) will now appeal.

Sentence (30a) includes the *wh*-subject in the first conjunct and the *wh*-object in the second conjunct.   In (30b), the situation is the other way around.   As our system predicts, these sentences license no ATB movement:

(31)　　a.　　They removed the prisoner, [$_{TP(+WH, v\varphi)}$ who C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <who$_{[v\varphi]}$> [$_{VP(v\varphi)}$ his appeal *v*–V$_{[u\varphi][EF][EF(v)]}$ <his appeal$_{[v\varphi]}$>]]] and [$_{CP(+WH)}$ *(who(m)) [$_{TP(v\varphi)}$ the warden C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <the warden$_{[v\varphi]}$> [$_{VP(v\varphi)}$ <who(m)> *v*–V$_{[u\varphi][EF][EF(v)]}$ <who(m)$_{[v\varphi]}$>]]]]

　　　　b.　　They removed the prisoner, [$_{CP(+WH)}$ who(m) [$_{TP(v\varphi)}$ the judge C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <the judge$_{[v\varphi]}$> [$_{VP(v\varphi)}$ <who(m)> *v*–V$_{[u\varphi][EF][EF(v)]}$ <who(m)$_{[v\varphi]}$>]]]] and [$_{TP(+WH, v\varphi)}$ *(who) C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(UNERGATIVE)}$ <who$_{[v\varphi]}$> [$_{VP(v\varphi)}$ PRO *v*–V$_{[u\varphi][EF][EF(v)]}$ <PRO$_{[v\varphi]}$>]]]

As depicted in (31a), the *wh*-phrase in the first conjunct is raised via (non)phasal valuation to "Spec-T"; the *wh*-phrase in the second conjunct is raised via phasal valuation to "Spec-C."   As just mentioned, the situation in (31b) is reversed.   In both cases, there is no parallelism in labeling between the largest SOs.   These derivations thus fail to undergo ATB movement.

Recall again that the VMH permits subject *wh*-phrases to undergo LF-movement to Spec-C. If subject *wh*-phrases may be raised to Spec-C at LF, each conjunct in (30a) and (30b) would have a parallel structure in which the *wh*-phrase occupies Spec-C. This situation incorrectly predicts that the derivations of (30a) and (30b) are allowed to undergo ATB movement. Our system, on the other hand, does not have any difficulty in predicting the applicability of ATB movement.

5.4. Concluding Remarks

In this chapter, I inspected additional application of the theory of (non)phasal valuation presented in chapter 3, which works as a labeling/EF detoxification process in both clausal and predicate systems, as shown in chapter 4. According to this inspection, EF detoxification yields efficient computation by going along with AF detoxification, based on the theory of (non)phasal valuation that incorporates the extended mechanism of feature inheritance. One important result is that this refined theory eliminates the countercyclic application of IM by offering the "multiple Specs" analysis under amalgamation of phase heads with nonphase heads. This theory was also shown to produce some other significant implications and consequences. Of particular significance is that the VMH no longer works as an independent principle because its effect follows as a special case of the C-T amalgam. This means that there is no reason to ban movement with no effect on PHON. In effect, we demonstrated that this is the case. In other words, the phenomena that the VMH has captured traditionally are reducible to just the principles and operations of minimal syntax.

In chapter 6, I will explore another application of the system of amalgamation proposed here. This exploration indicates that the theory of (non)phasal valuation is responsible for the curious phenomena that parasitic gap constructions display, in conjunction with an independent demonstration that the adjunct clause with a parasitic gap behaves as a restrictive relative clause

that the mechanism of "afterthoughts" offered by Chomsky (2004) introduces separately from the main derivational workspace.

Chapter 6

Parasitic Gap Constructions

and Their Implication for the Derivational Workspace

6.1. Introduction

Parasitic gap constructions have attracted much attention of linguistic research in Generative Grammar over several decades (see e.g. Ross (1967), Taraldsen (1981), Chomsky (1982, 1986a), Engdahl (1983, 1984), Kearney (1983), Contreras (1984, 1993), Kiss (1985), Frampton (1990), Brody (1995), Kim and Lyle (1996), Kennedy (1997), Karimi (1999), Shimada (1999), Nissenbaum (2000), Nunes (2001, 2004), Kasai (2007, 2010)).   Though apparent peripheral phenomena in the sense that native speakers rarely encounter them, parasitic gap constructions deserve inspection in that they show up the nature of core grammar.   With careful observation of parasitic gap phenomena, Chomsky (1982) improves the system of government and binding and Chomsky (1986a) elaborates the theoretical framework of barriers.   Some subsequent inquiries based on a minimalist framework have attempted to reduce the properties of parasitic gap constructions discovered by a multitude of works to principles and operations indispensable to the theory of the faculty of language (FL).   A representative study along this line of ambitious attempts is the work of Kasai (2010), which offers the multiple dominance analysis of parasitic gap constructions.   This analysis derives parasitic gaps under the single derivational workspace approach by optimizing two modes of Merge, external Merge (EM) and internal Merge (IM).

In contrast to this approach, the theory of (non)phasal valuation developed in this thesis makes it possible to put forward the multiple derivational workspace approach to parasitic gap constructions, based on the mechanism of "afterthoughts" as presented by Chomsky (2004). Under this approach, the adjunct clause including a parasitic gap serves as a restrictive relative clause that the mechanism of afterthoughts derives separately from the main derivational work-

space.   Both approaches share the same perspective in that they endeavor to reduce Chomsky's

(1986a) insight to the principles and operations of minimal syntax.   In addition, these two ap-

proaches are capable of explaining well-known descriptive generalizations for parasitic gap

phenomena, including (i) anti-c-command effects, (ii) S-structure effects, (iii) A′-movement

versus A-movement, (iv) categorial restriction, and (v) anti-reconstruction effects.   Both ap-

proaches are thus worth exploring.   By comparing Kasai's single derivational workspace ap-

proach with our multiple derivational workspace approach, this chapter will show that the latter

successfully characterizes a certain difference between the nature of core grammar and the sur-

face aspect of grammar in concert with the parallelism condition on parasitic gap constructions

proposed by Sakamoto (2011a, b).   The proposed analysis enjoys an advantage in this respect.

The organization of this chapter is as follows.   After introducing core questions that any

approach to parasitic gap constructions needs to answer, section 6.2 reviews Chomsky (1986a)

and Kasai (2010), which are representative approaches adopting preminimalist and minimalist

frameworks, respectively.   Section 6.3 seeks the possibility of an alternative approach to para-

sitic gap constructions, based on the mechanism of afterthoughts proposed by Chomsky (2004).

Section 6.4 compares the proposed analysis with Kasai's multiple dominance analysis, thereby

emphasizing our advantage.   Section 6.5 concludes this chapter.

6.2.   Previous Studies: From Preminimalism to Minimalism

6.2.1.   Introducing Core Questions

Before outlining previous works, it is important to clarify what parasitic gap constructions

are and what core questions are that encourage essential understanding of parasitic gap construc-

tions if we answer them.   In order to do this, let us observe the following examples (from Kasai

(2010: 236)):

(1)   a.   Which paper$_i$ did you file t_i [without reading pg_i]?

   b.   *Which book$_i$ did you review this paper [without reading t_i]?

(1a) exhibits a typical parasitic gap sentence, in which there are two gaps represented as t in the main clause and as pg in the adjunct clause. Both gaps are interpretationally associated with the single operator in the main clause, as the subscript expresses. It is hard to conceive, however, that movement raises this operator from the position of pg in the adjunct clause to sentence-initial position, as exemplified in (1b). The ungrammaticality of (1b) indicates that adjunct clauses generally resist subextraction from within. With this in mind, sentence (1a) would have an ungrammatical status if extraction generates a gap in the adjunct clause. The *wh*-phrase in (1a) is thus linked derivationally not to the second gap pg but to the first gap t. Owing to the fact that the existence of a gap in the adjunct clause is crucially dependent on the occurrence of a gap in the main clause, the former and the latter are referred to as a parasitic gap and a real gap, respectively.

The contrast in (1) clarifies core questions that any approach to parasitic gap constructions has to answer: (i) what derives a parasitic gap? and (ii) what establishes a coreferential relation between a parasitic gap and its seeming operator? In what follows, we see what sorts of answers previous works find to these two contiguous questions. We first gain a quick overview of the analysis based on null operator movement and chain composition initiated by Chomsky (1986a) and then we grab the gist of Kasai (2010), understandable as a minimalist reduction of Chomsky's seminal work.[1]

---

[1] It is known in the literature that parasitic gaps can also appear in the subject phrase and that such subject-internal parasitic gaps behave differently from adjunct-internal parasitic gaps under consideration (see e.g. Johnson (1985), Shimada (1991), Kurogi (2007)). This chapter thus confines the target of analysis to adjunct-internal parasitic gaps.

6.2.2.   Chomsky (1986a): Null Operator Movement and Chain Composition

Chomsky (1986a) proposes that null operator movement derives a parasitic gap and that the algorithm of chain composition guarantees coreferentiality between the derived gap and its seeming operator.   The procedure is illustrated below:

(2) $[_{CP}$ Which paper did you file t $[_{PP}$ without $[_{CP}$ Op [reading pg]]]]

(3)   a.   C = (*which paper*, *t*)

         b.   C' = (*Op*, *pg*)

(4)   If C = $(\alpha_1,\ldots,\alpha_n)$ is the chain of the real gap, and C' = $(\beta_1,\ldots,\beta_n)$ is the chain of the parasitic gap, then the "composed chain" (C, C') = $(\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_n)$ is the chain associated with the parasitic gap construction and yields its interpretation.

(5)   (C, C') = (*which paper*, *t*, *Op*, *pg*)

As shown in (2), the representation of (1a), null operator movement raises the invisible operator *Op* from the position of *pg* to embedded Spec-C; similarly, regular *wh*-movement raises the visible operator *which paper* from the position of *t* to matrix Spec-C.   This is tantamount to stating that representation (2) has the two chains in (3).   These independent chains are composed under the algorithm of chain composition in (4), thereby yielding the composed chain in (5).   The formed chain has *which paper* as its head and *pg* as its tail, with *t* and *Op* behaving as intermediate traces.   In this way, a coreferential relation obtains between the parasitic gap and its literal operator.[2]

As just seen, Chomsky (1986a) separates the derivation of the adjunct clause including a parasitic gap from the derivation of the main clause by adopting null operator movement.   Null

---

[2] See e.g. Frampton (1990), Kim and Lyle (1996), Karimi (1999), Nissenbaum (2000), and Sakamoto (2011a, b) for this line of approach based on Chomsky (1986a).

operator movement to embedded Spec-C crosses no island, as with the case of regular *wh*-movement observed in the main clause. (1a) thus displays no island effect. Chain composition, on the other hand, serves as a strategy for integrating the two separated derivational workspaces. Derivationally being composed of two independent workspaces, parasitic gap constructions behave as if they organize only one workspace with respect to interpretation. In the next subsection, we review the work of Kasai (2010), which succeeds in capturing Chomsky's insight in the theory of FL based on Merge on the premise that parasitic gap constructions are derived within the single derivational workspace.

6.2.3. Kasai (2010): A Single Derivational Workspace Approach

We saw in chapter 2 that there are EM and IM in the theory of FL and that both modes of Merge are always available as a "cost-free" operation in the sense that the performance systems are free to exploit them. The strong minimalist thesis (SMT) requires all linguistic phenomena to be reduced to such principled computational apparatus, including principles as well as operations. We can count the work of Kasai (2010) as one ambitious enterprise along this line of approach.[3] In this subsection, I would like to not only review his work in detail but also consider its theoretical implication.

6.2.3.1. Multiple Dominance

Kasai (2010) explicates a multiple dominance approach to parasitic gap constructions, assuming that Universal Grammar (UG) allows nodes to have multiple mothers, rooted in the preminimalist work of McCawley (1982) and followed by such minimalist works as Wilder (1999), Citko (2005), Riemsdijk (2006), Gracanin-Yuksek (2007), Hiraiwa and Bodomo (2008),

---

[3] Similar attempts are made by Nunes (2001, 2004), but I do not outline his work because Kasai (2010) critically reviews it.

and Bachrach and Katzir (2009).　This approach is highly significant in deriving the properties of parasitic gap constructions by optimizing the two modes of Merge of EM and IM.

　　Let us now examine the derivation of parasitic gap constructions proposed by Kasai (2010):

(6)　　　N = {{$_A$ PRO, *v*, reading, Asp, which, paper} {$_B$ C, T} {$_C$ you, *v*, file, Asp, without} {$_D$ C$_Q$ did}}

Kasai adopts Nunes and Uriagereka's (2000: 40) assumption that "prepositions that select clausal complements must belong to the subordinating array, and not to the array associated with the complement clause."　Under this assumption, the derivation of (1a) proceeds along the lines of Numeration shown in (6), in which the subarray C involves *without*.　A and B stand for the subarrays with which the adjunct clause is assembled, and C and D for the subarrays with which the matrix clause is built up.　The computational system first enters into assembling of the adjunct clause:

(7)　　　[$_{CP}$ which paper$_i$ C [$_{TP}$ PRO T [$_{vP}$ which paper$_i$ [$_{vP}$ PRO *v* [$_{AspP}$ which paper$_i$ Asp reading which paper$_i$]]]]]

EM creates the *v*P-level structure where *which paper* and PRO are generated as the object of *read* and as the external argument, respectively.　IM raises the object *wh*-phrase to "Spec-Asp" and then to the edge of *v* in a successive-cyclic way (see Borer (1994) and Hiraiwa (2005) for the functional category Asp(ect), which Kasai considers to be responsible for Accusative Case

assignment).[4]   After EM yields the CP-level structure, IM relates PRO to "Spec-T" and *which paper* to "Spec-C."

Further, EM combines the matrix verb *file* and the object *wh*-phrase linked to embedded "Spec-C," as in (8).

(8)　　VP　　　　　CP
　file　　which paper$_i$
　　　　　　　　reading which paper$_i$
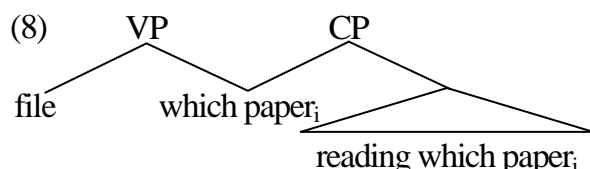
This is an example of the multiple dominance structure in question, in which VP and CP dominate the one node.   IM additionally raises *which paper* to matrix "Spec-Asp," yielding the structure in (9).

(9)　　AspP
which paper$_i$
　　　　Asp　　　VP　　　　CP
　　　　　　file　　which paper$_i$
　　　　　　　　　　　　reading which paper$_i$
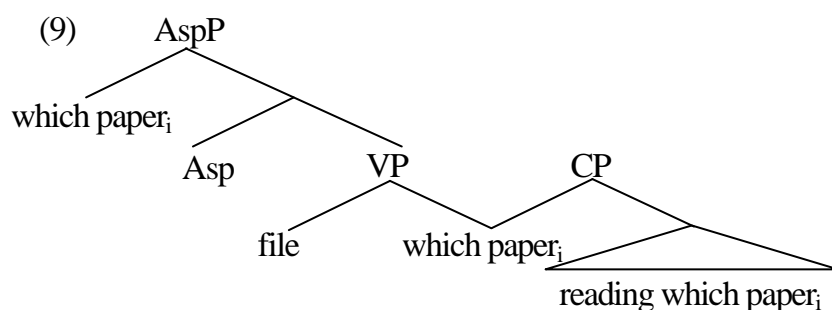
Given that an Asp head is responsible for Accusative Case assignment, the object *wh*-phrase would receive Case once in the adjunct clause and once again in the matrix clause.   Kasai argues that the activity condition revised in the form of (10), originally proposed by Chomsky (1986b: 94), makes multiple Case assignment possible:

---

[4] See Johnson (1991), Koizumi (1993), and Runner (1995) for an argument confirming the existence of overt object shift in English.

(10)   If a DP α, which has already received a theta-role θ_1 and a Case value c_1, receives

another theta-role θ_2, then α needs to be assigned another Case value c_2.

Agree/Value enables a set of unvalued/uninterpretable agreement features [$u\varphi$] on a lexical item to establish feature valuation with a set of valued/interpretable agreement features [$v\varphi$] on its matching lexical item, with Case valuation occurring concomitantly (see sections 2.3.1 and 2.6.1 in chapter 2).   The establishment of feature valuation renders the relevant features inactive in a normal situation, the activity condition (cf. Chomsky (2000, 2001)).   According to Kasai, however, a nominal phrase is permitted to receive multiple Case if it is subject to multiple θ-role assignment, as in (10).[5]   It thus follows that the object *wh*-phrase in (9), obtaining two θ-roles from the embedded and matrix verbs, receives Accusative Case first in embedded "Spec-Asp" and then in matrix "Spec-Asp."

The derivation heads for the last stage:

---

[5] Kasai believes that English can have the system of multiple Case assignment as a possible option of UG, based on cross-linguistic studies such as Moore (1998) and Ura (1998).

(11)

CP
- which paper$_i$
- C
  - TP
    - you
    - T
      - vP
        - which paper$_i$
        - you
          - vP
            - v
            - AspP
              - AspP
                - which paper$_i$
                - Asp
                - VP
                  - file
                  - which paper$_i$
              - PP
                - without
                - CP
                  - C
                  - reading which paper$_i$
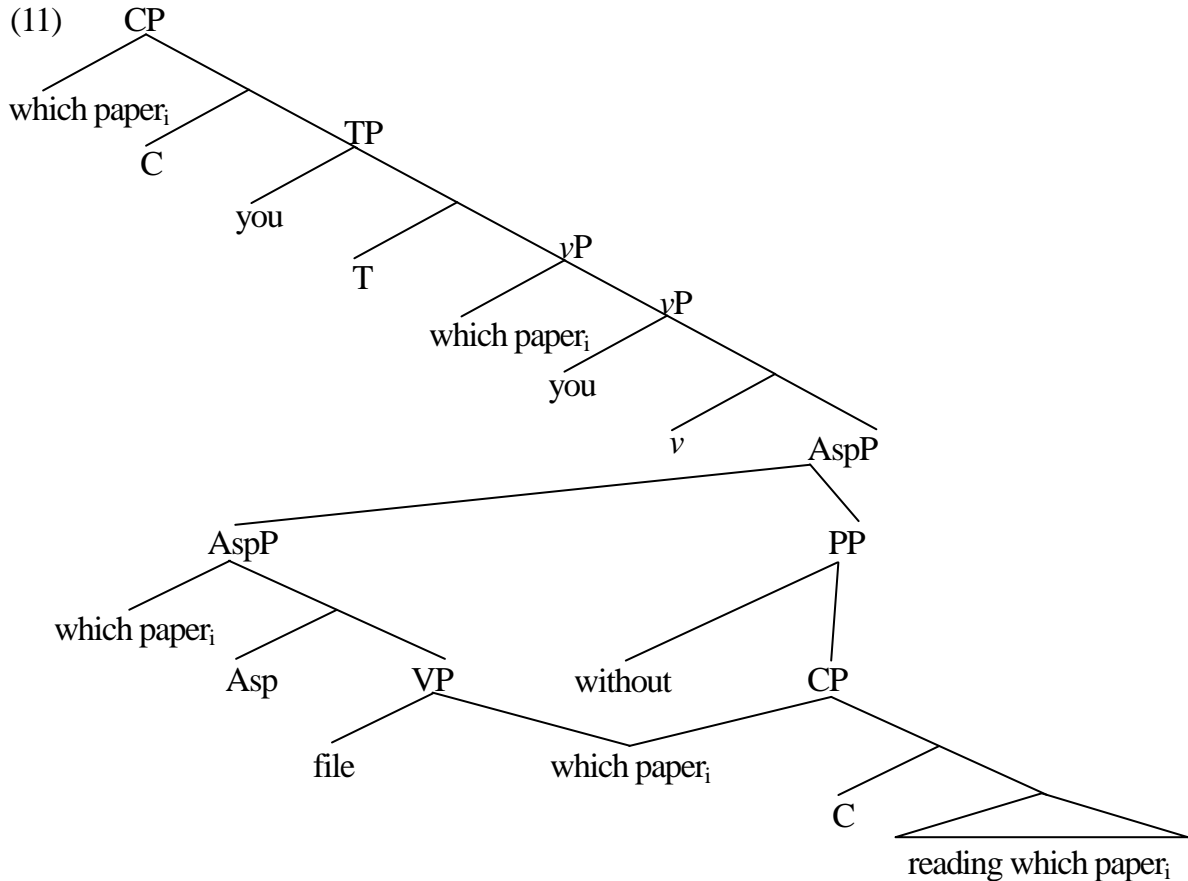
As depicted in (11), the merger of the embedded CP and the preposition *without* produces the adjunct PP, which is externally merged with the matrix AspP.[6]   After EM forms the matrix CP-level structure, IM raises the object *wh*-phrase via the edge of *v* as a landing site of succes-

---

[6] Kasai supports the plausibility of the structure in which an object material (i.e. "Spec-Asp") occupies a higher position than an adjunct PP (from Runner (1995: 26)).

(i)   a.   * I called him$_i$ without dialing John's$_i$ number.
             b.   * I saw the bastard$_i$ before John$_i$ left.
             c.   ? The DA accused the defendants during each other's trial.
             d.   I saw none of the children after any of the parents left.

The ungrammaticality of (ia-b) shows the violation of Condition C; the grammaticality of (ic-d) indicates the observance of Condition A and the licensing of the negative polarity item, respectively.   These (un)grammatical patterns imply that object position c-commands into the adjunct clause, with c-command defined as follows:

(ii)   X c-commands Y iff X and Y are categories and X excludes Y and for every category Z that dominates X, (a) Z dominates Y or (b) one of the segments of Z dominates Y.

Under this definition, the object *wh*-phrase in (11) properly c-commands into the adjunct PP.

sive-cyclic movement to matrix "Spec-C." Crucially, *wh*-movement "crossing" the adjunct clause in (11) induces no island effect because such movement takes place before the adjunct PP is merged with the matrix clause, provided that adjunct clauses do not function as islands until they are merged with some structure (cf. Nunes and Uriagereka (2000), Nunes (2001, 2004)):

(12)   *Which book$_i$ did you borrow t_i [after leaving the bookstore [without finding pg_i]]?

The ungrammatical status of (12) means that island effects emerge if the adjunct clause including a parasitic gap is embedded within another adjunct clause. Given the present analysis, the existence of an island effect in (12) comes from the reality that the adjunct clause including a parasitic gap already behaves as an island by merging with the structure in the derivational stage of applying *wh*-movement.

In sum, we can understand Kasai's (2010) multiple dominance approach as a minimalist revision of Chomsky's (1986a) analysis. Recall that we have two core questions to be answered in addressing parasitic gap constructions: (i) what derives a parasitic gap? and (ii) what establishes a coreferential relation between a parasitic gap and its seeming operator? Chomsky (1986a) regards null operator movement and chain composition as the answers to (i) and (ii), respectively. Null operator movement separates between the derivational workspaces of the main clause and the adjunct clause including a parasitic gap, hence no island effect; chain composition articulates them with respect to interpretation by bringing together the real gap chain of the main clause and the parasitic gap chain of the adjunct clause, hence the coreferentiality between a parasitic gap and its seeming operator. Kasai successfully captures this insight along his single derivational workspace approach, under which full utilization of Merge, i.e. EM and IM, derives the same effect as that of the interaction between null operator movement and chain composition. His analysis has thus arrived at principled explanation in conformity with SMT.

104

In the next subsection, we see that the properties of parasitic gap constructions follow from this analysis.

6.2.3.2. Deriving the Properties of Parasitic Gap Constructions

A number of works have thus far formulated the licensing conditions of parasitic gaps in the following form: (i) anti-c-command effects, (ii) S-structure effects, (iii) A′-movement versus A-movement, (iv) categorial restriction, and (v) anti-reconstruction effects. In this subsection, we survey how Kasai's multiple dominance approach derives the effects obtained by these conditions.

Engdahl (1983: 22) presents the anti-c-command condition in (13) to capture such an instance as (14):

(13)   The anti-c-command condition:

A parasitic gap may not be c-commanded by the real gap.

(14)   *I wonder which man$_i$ t_i called you before you met pg_i.

Sentence (14) has a complement clause headed by the subject *wh*-phrase that acts as an operator of the real gap. We find there to be a c-command relation between the real gap t and the parasitic gap pg, when coreferential interpretation is blocked. What generalizes such a licensing pattern is the anti-c-command condition in (13), which Chomsky (1986a) reduces to Condition C. Let us consider how the multiple dominance analysis derives the effect of this condition:

(15)

vP   CP

which man$_i$

v AspP   you met which man$_i$
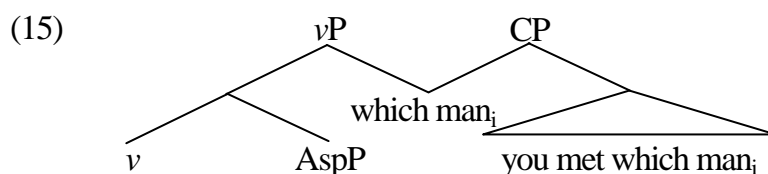
As mentioned above, the sentence under consideration contains the matrix subject *wh*-phrase that acts as an operator of the real gap.   The multiple dominance analysis allows the same element to work as an object *wh*-phrase in the embedded clause.   As depicted in (15), IM associates the object *wh*-phrase with embedded "Spec-C," and EM introduces the same element as the subject *wh*-phrase into matrix "Spec-*v*."   Note that the analysis presented in the last subsection supposes that the adjunction site of an adjunct clause is the edge of matrix Asp, which is lower than object position in the matrix clause.   The derivation already assembles the matrix *v*P-level structure at the point of applying multiple dominance.   The adjunct clause cannot thus adjoin to matrix AspP, from which anti-c-command effects follow.

Let us turn to S-structure effects:

(16)   *Who filed which report$_i$ without reading *pg*$_i$?

Example (16) involves the in-situ object *wh*-phrase that is intended to be coreferential with the parasitic gap, which undergoes no overt movement.   Such environments do not license parasitic gaps, as originally observed by Engdahl (1983: 14).   The preminimalist framework thus considered that S-structure movement, rather than LF-movement, licenses parasitic gaps, which we may call S-structure effects.[7]

Kasai (2010) captures S-structure effects in the following manner:

---

[7] Shimada (1999) challenges S-structure licensing of parasitic gaps in a minimalist framework, in which there is no linguistic level but the semantic and phonetic interfaces.   In effect, he demonstrates with elliptical sentences including subject-internal parasitic gaps that LF licensing of parasitic gaps is available in no uncertain manner.   See also Kim and Lyle (1996: section 5.1) and Kennedy (1997) for related discussions.

(17)

```
           AspP
          /    \
which report_i   \
              Asp  VP
                  /    \
               file    which report_i
                            \
                             CP
                            /  \
                           /    \
                    without reading which report_i
```
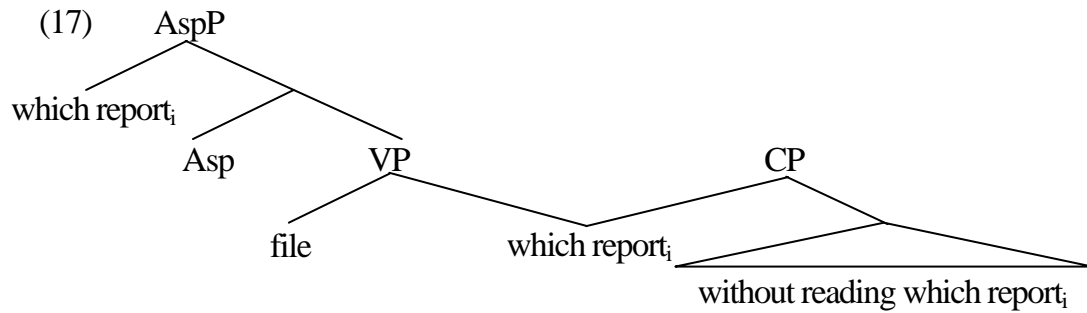
As illustrated in (17), IM relates *which report* to embedded "Spec-C," and EM introduces the same element as an object of the matrix verb *file*, thus yielding a multiple dominance structure. The in-situ object *wh*-phrase is further raised to matrix "Spec-Asp" and stays at that position without undergoing movement to matrix "Spec-C." Given that the in-situ object *wh*-phrase has a feature that activates *wh*-movement ([–wh] in Kasai's (2010: 250) term) and that this feature is not checked at an intermediate step of successive-cyclic movement (Bošković (2007)), derivation (17) crashes at the interfaces, leaving the [–wh] feature on the in-situ object *wh*-phrase unchecked.

Kasai also simulates a derivation in which the in-situ object *wh*-phrase has no activation feature for *wh*-movement:

(18)

```
        CP
       /  \
      C    TP
          /   \
        DP     \
       AspP  T  vP
      /  \      /  \
which report_i  DP   \
        Asp   VP  v   AspP
             /  \    /   \
          file  which report_i
                      Asp   VP
                           /   \
                      reading  which report_i
```

107
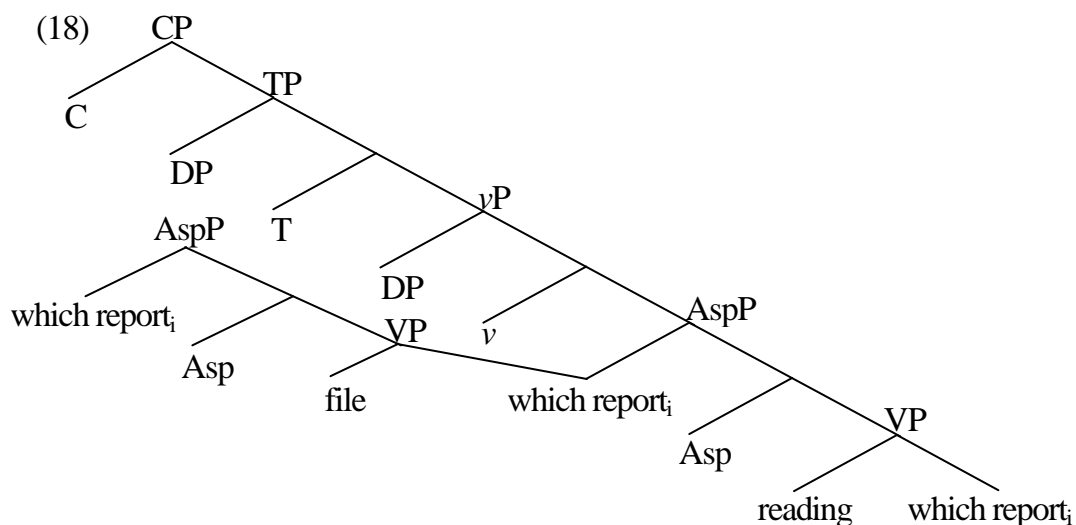
As shown in (18), the in-situ object *wh*-phrase, bearing no activation feature for *wh*-movement, is only raised to embedded "Spec-Asp." After the computational system assembles the embedded CP-level structure, multiple dominance tries to enter into EM of the matrix verb and the in-situ object *wh*-phrase. Transfer, however, has already handed the embedded AspP to the interfaces on the assumption that *v*P and CP are phases. The phase impenetrability condition prevents any operation from operating on the in-situ object *wh*-phrase within the embedded AspP. Derivation (18) thus fails.

The same account applies to the generalization of A′-movement versus A-movement, attributed to Engdahl (1983: 13):

(19)   *The book$_i$ was filed t_i without my reading *pg*$_i$ first.

The ungrammaticality of (19) shows that A-movement, unlike A′-movement (cf. (1a)), licenses no parasitic gaps. This generalization is explained along the lines of (18). The element *the book* in (19) is introduced via EM as an object in the adjunct clause. IM raises *the book* to embedded "Spec-Asp." Because *the book* has no activation feature such as [–wh], it remains in that position. By the time multiple dominance applies, Transfer renders this object material invisible for any operation in the matrix clause. Sentence (19) thus fails to allow for the occurrence of a parasitic gap.

The generalization of categorial restriction also follows naturally from the multiple dominance analysis:

(20)   a.   *[How sick]$_i$ did John say he felt t_i before getting *pg*$_i$?

b.   *[How long]$_i$ did John drink t_i before lecturing *pg*$_i$?

c.   *This is a topic [about which]$_i$ you should think t_i before talking *pg*$_i$.

It is reported in the literature (cf. Emonds (1985), Cinque (1990)) that parasitic gaps are restricted to the nominal category.   As shown in (20), the environment where adjective, adverbial, and prepositional categories undergo *wh*-movement is not able to license parasitic gaps.   The multiple dominance analysis assigns (20a) the following structure:

(21)

```
              vP
            /    \
          DP      \
                 AspP
               /      \
              v        \
           AspP          PP
          /    \        /   \
       AspP    VP   before   CP
       /    \    \          /  \
    feel  how sick_i    getting how sick_i
```
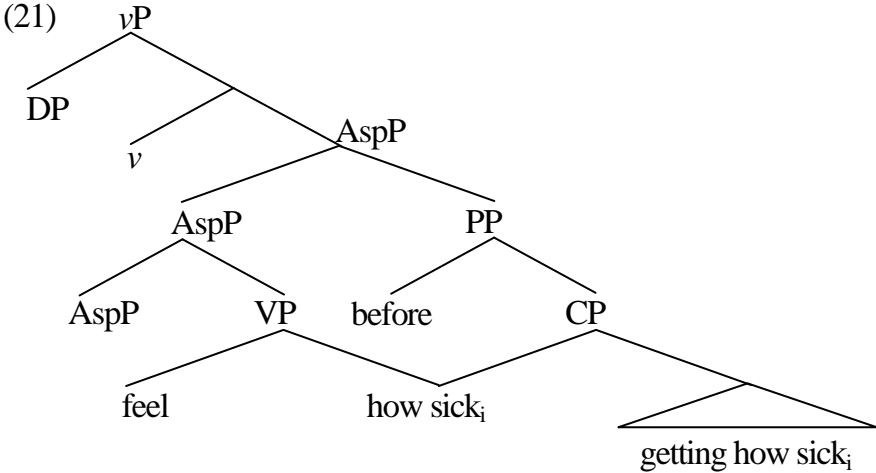
As depicted in (21), EM introduces *how sick* as a complement of the embedded verb *get* and IM links this element to embedded "Spec-C" in a successive-cyclic fashion.   At this stage of derivation, EM combines the matrix verb *feel* and *how sick* to form a multiple dominance structure. The element *how sick* bears an adjective category, so it requires no Case, which keeps this element in the VP-internal position.   Recall that the merger of the adjunct PP with the matrix clause makes the former an island.   Because *how sick* is not raised to matrix "Spec-Asp," it cannot escape from the adjunct island, unlike in the case of (11).   If *wh*-movement raises *how sick* via the edge of *v* to matrix "Spec-C," then that movement would cross an island, yielding ungrammaticality.   Categorial restriction therefore follows.

Finally, Kasai (2010) argues that the multiple dominance approach also derives anti-reconstruction effects, originally observed by Kearney (1983):

(22)   a.   [Which books about himself$_j$]$_i$ did John$_j$ file t_i before Mary read pg_i?

          b.   *[Which books about herself$_j$]$_i$ did John file t_i before Mary$_j$ read pg_i?

In (22a), a coreferential relation obtains between *John* and *himself*.   This fact means that *which books about himself* undergoes reconstruction into the real gap t but not into the parasitic gap pg, thereby satisfying Condition A.   If *which books about herself* in (22b) can undergo reconstruction into the parasitic gap, *Mary* should establish a coreferential relation with *herself*, contrary to fact.   The contrast in (22) thus indicates that parasitic gaps resist reconstruction, which we may refer to as anti-reconstruction effects.

Chomsky's (1986a) analysis easily explains this contrast, having no possibility of reconstructing *wh*-phrases into parasitic gaps, because parasitic gaps are traces generated by null operator movement in the adjunct clause, not ones created by visible *wh*-movement in the main clause.   Kasai's (2010) analysis, on the other hand, cannot derive this contrast straightforwardly. Under his analysis, *wh*-movement directly generates parasitic gaps by utilizing multiple dominance.   This analysis itself does not prohibit the possibility of reconstructing *wh*-phrases into parasitic gaps as well as real gaps at all.   What then captures the contrast in (22)?   Kasai explains this grammatical contrast by employing the theory of wholesale late merger proposed by Takahashi (2006) and Takahashi and Hulsey (2009), which enables the restrictor of a determiner to be introduced into derivation in a countercyclic manner:

(23)

VP   CP

file   which [books about…]

C

TP

Mary

T

*v*P

Mary

*v*P

v

which

AspP

Asp

which

VP

read   which
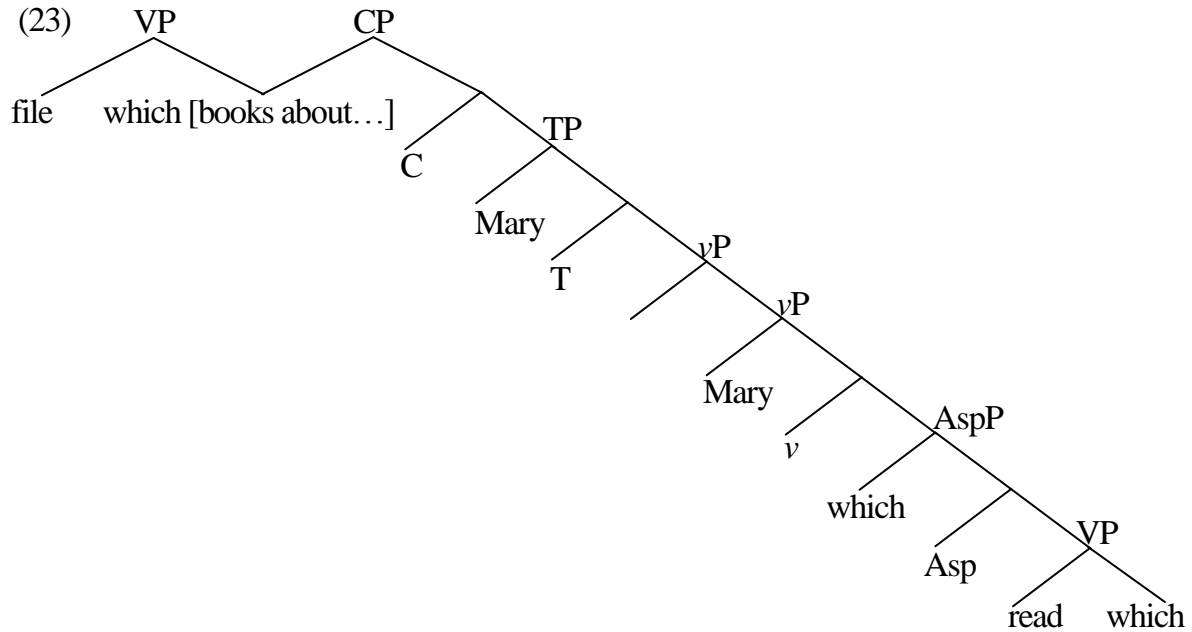
As illustrated in (23), the *wh*-determiner is base-generated via EM as an object of the verb *read* in the adjunct clause and undergoes successive-cyclic movement without being followed by its restrictor NP, on the assumption that determiners can discharge a Case-assigning property on heads such as Asp.   At the point where *which* is raised to embedded "Spec-C," wholesale late merger associates *which* with its restrictor NP.   After that, EM combines the matrix verb and the whole DP consisting of the *wh*-determiner and its restrictor NP, yielding a multiple dominance structure.   The restrictor NP of *which* receives Accusative Case from the matrix Asp. This analysis properly captures the contrast in (22) because the c-command domain of *Mary* does not involve the nominal phrase with *himself/herself*.   Anti-reconstruction effects therefore fall into place.

Given, however, that late merger is assumed to be an optional operation, Kasai states that there are two possible unwelcome derivations of (22b) to be considered.   One is the derivation by which the embedded verb *read* is merged with the whole DP *which books about herself* in the first place.   In this case, we expect (22b) to be grammatical, observing Condition A, contrary to fact.   According to Kasai, consideration of (global) economy rules out this undesired derivation,

in which *books about herself* is assigned Case twice in the adjunct clause and in the matrix clause.   In contrast, *books about herself* in (23) receives Case only in the matrix clause.   The former is less economical than the latter in that derivation (23) contains fewer Case assignments.

The other is the derivation whereby wholesale late merger applies when IM relates *which* to the edge of *v* in the adjunct clause:

(24)   [$_{TP}$ Mary T [$_{vP}$ which [books about herself] [$_{vP}$ Mary *v* [$_{AspP}$ which Asp [$_{VP}$ read which]]]]]]

In this derivation, *Mary* c-commands *books about herself*, so the observance of Condition A should ensure a coreferential relation between *Mary* and *herself*, contrary to fact.   Note that *books about herself* is outside of the Case licenser Asp, which means that it receives Case only in the matrix clause, as with the case of (23).   Economy consideration of the sort mentioned above cannot thus exclude the derivation in (24).   Instead, Kasai argues that locality is responsible for the illicitness of (24).   *Which* in (24) is inactive in the sense of the activity condition proposed by Chomsky (2000, 2001) because its Case-feature is valued via a φ-agreement relation with Asp in the adjunct clause.   *Books about herself*, by contrast, is not Case-valued, undergoing wholesale late merger at the edge of *v*, and hence the entire *wh*-phrase is still active.   Because this active material lies between T and *Mary*, it is regarded as an intervener for φ-agreement, which renders the derivation in (24) illegitimate.

In this subsection, we have seen that Kasai's (2010) multiple dominance analysis receives empirical justification, deriving the effects that the licensing conditions of parasitic gaps produce, including (i) anti-c-command effects, (ii) S-structure effects, (iii) A′-movement versus A-movement, (iv) categorial restriction, and (v) anti-reconstruction effects.   Crucially, his analysis revises the work of Chomsky (1986a) under a single derivational workspace approach fol-

lowing minimalist tenets, which allows the computational system to derive parasitic gaps via multiple dominance by integrating the derivations of the main clause and the adjunct clause. Given that Chomsky's (1986a) analysis originally separates between the derivational workspaces of the main clause and the adjunct clause including a parasitic gap, it is natural to expect there to be a multiple derivational workspace approach following minimalist tenets. In section 6.3, we endeavor to demonstrate the existence of such an approach.

6.3.   A Multiple Derivational Workspace Approach

The previous two subsections illustrated Chomsky's preminimalist analysis and Kasai's minimalist analysis of parasitic gap constructions. A crucial difference between these analyses lies in whether a workspace for computation is multiple or not. Chomsky's chain composition analysis utilizes a multiple workspace, while Kasai develops the analysis based on a single workplace. This section shows that a theoretical framework involving the mechanism of (non)phasal valuation is to present an analysis of parasitic gap constructions in terms of a multiple workplace. Specifically, I propose a multiple derivational workspace approach under which the adjunct clause including a parasitic gap serves as a restrictive relative clause that the mechanism of afterthoughts proposed by Chomsky (2004) derives separately from the main derivational workspace. Section 6.3.1 introduces the mechanism of afterthoughts, based on which section 6.3.2 provides the derivation of parasitic gap constructions and explains the descriptive generalizations for parasitic gap constructions seen in section 6.2.3.2.

6.3.1.   Afterthoughts: Eliminating QR and Late Merger

Fox (2002) argues with Fox and Nissenbaum (1999) that antecedent-contained deletion (ACD) is derived by a composite operation, rightward quantifier raising (QR) followed by late merger of adjuncts (cf. Lebeaux (1988)). Pointing out two theoretical problems with this anal-

ysis, Chomsky (2004) proposes that the mechanism of afterthoughts derives ACD.

To understand the background of these studies, let us begin by reviewing Fox (2002), who proposes to revise an account of ACD that draws upon QR shown by Larson and May (1990) under the copy theory of movement. In order to see how this proposal works, consider the following:
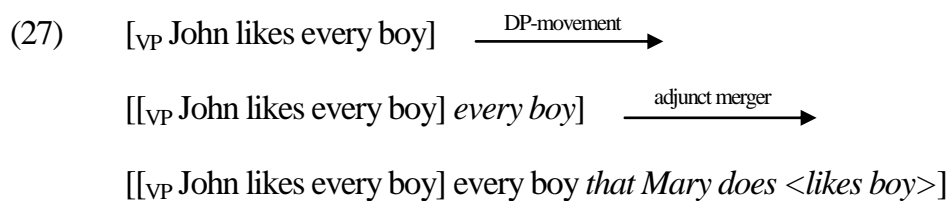
(25) a. John likes every boy Mary does.

  b. John [$_{VP}$ likes every boy Mary does [$_{VP}$ e]].

  c. John [$_{VP}$ likes every boy Mary does [$_{VP}$ <likes every boy Mary does [$_{VP}$ e]>]].

(26) a. [every boy Mary does [$_{VP}$ e]] [$_{TP}$ John [$_{VP}$ likes t]].

  b. [every boy Mary does [$_{VP}$ <likes t>]] [$_{TP}$ John [$_{VP}$ likes t]].

The ACD sentence in (25a), unlike typical VP-ellipsis sentences, has a structure in which the elided VP is contained in the antecedent VP. This structure is depicted in (25b). Given that an elided VP undergoes its interpretation via copying an antecedent VP onto the ellipsis site at LF (cf. Fiengo and May (1994)), the elided VP in (25b), denoted with e, should be interpreted by being reconstructed in the form of (25c). However, the structure so created still includes the elided VP, so it fails to obtain its appropriate interpretation, which is conventionally termed the infinite regress problem.

Traditionally, QR has been supposed to resolve the infinite regress problem along the lines of the structures illustrated in (26). In (26a), the nominal expression that contains the elided VP moves via QR outside the antecedent VP, which allows the elided VP to be properly interpreted, as shown in (26b). The application of QR converts the nominal expression that contains the elided VP into a simplex trace. LF-copying attaches the antecedent VP consisting of the verb and the trace with no ellipsis site onto the elided VP. The formed representation receives inter-

pretation without inducing the infinite regress problem.

　　Fox (2002) attempts to feed this account of ACD with a minimalist revision in terms of the copy theory of movement.　This theory requires a QRed element to leave behind an identical copy, not a trace.　This requirement jeopardizes anew the QR-based account of ACD. Bearing this in mind, Fox proposes a novel QR-based account of ACD by employing the derivation based on rightward QR followed by late merger of adjuncts, as originally proposed by Fox and Nissenbaum (1999) for the purpose of deriving the extraposition of adjuncts:

(27)　　[$_{VP}$ John likes every boy]　$\xrightarrow{\text{DP-movement}}$

　　　　[[$_{VP}$ John likes every boy] *every boy*]　$\xrightarrow{\text{adjunct merger}}$

　　　　[[$_{VP}$ John likes every boy] every boy *that Mary does <likes boy>*]

In (27), after the structure VP is assembled, the DP *every boy* undergoes rightward QR, which is followed by the late merger of the relative clause *that Mary does* and the deletion of the upper copy of *every boy* (see Kayne (1976), Cinque (1981–82), and Sauerland (1998) for the detailed structures of relatives).　The current analysis of ACD correctly resolves the infinite regress problem, with the ellipsis site outside the antecedent VP.

　　Although Fox's (2002) analysis is attractive in that it results in some further consequences, Chomsky (2004) points out that there are two theoretical problems: (i) late merger, which is countercyclic, should be replaced by a cyclic adjunction operation, and (ii) it is unclear why QR applies to the right though it is a movement operation displaying no ordering property.[8]　Also, Chomsky states that these theoretical problems disappear by proposing the mechanism of after-

---

[8]　We can give two additional problems: (iii) any movement, including QR, should be motivated by some feature (cf. Hornstein (1999), Kitahara (1996)), and (iv) it is not obvious why the upper copy is deleted and the lower copy is pronounced with respect to the QRed DP.

thoughts to derive ACD:

(28)   John likes every boy (that is, more accurately...) every boy Mary does <likes *t*>.

In (28), after the matrix clause *John likes every boy* is built up, it is combined with the after-thought for the object element *every boy Mary does*. *Every boy* as part of this afterthought is destressed and thus deleted at the phonetic interface. As just seen, the derivation in (28) involves neither late merger nor QR, so it is able to straightforwardly overcome the theoretical problems noted above.

The analysis of ACD based on afterthoughts is reminiscent of that of parasitic gap constructions based on null operator movement and chain composition proposed by Chomsky (1986a). Specifically, dissociating the derivation of the adjunct clause including an ellipsis site from the derivation of the main clause, the mechanism of afterthoughts makes it possible to identify the elided VP with the antecedent VP with respect to interpretation. This analysis produces the same effect as that of the interplay between null operator movement and chain composition. Null operator movement separates between the derivational workspaces of the main clause and the adjunct clause including a parasitic gap; chain composition articulates them with respect to interpretation by bringing together the real gap chain of the main clause and the parasitic gap chain of the adjunct clause. Crucially, the afterthought approach is more principled than Chomsky's (1986a) approach because the former involves not an independent algorithm such as chain composition but rather a general mechanism of identification based on relativization. In the next subsection, we demonstrate an afterthought approach to parasitic gap constructions under which the adjunct clause including a parasitic gap serves as a restrictive relative clause that the mechanism of afterthoughts derives separately from the main derivational workspace.

6.3.2.   Parasitic Gap Constructions as Restrictive Relative Clauses

We saw in section 6.3.1 that afterthoughts function to "share" one element with the main derivational workspace by utilizing the mechanism of identification by relativization and that this analysis is justified in that it eliminates the theoretical problems surrounding ACD.[9]   In this subsection, we extend this analysis to parasitic gap constructions.

6.3.2.1.   Answering Core Questions

In section 6.2.1, we introduced two core questions that lead us to a deeper understanding of parasitic gap phenomena if we answer them: (i) what derives a parasitic gap? and (ii) what establishes a coreferential relation between a parasitic gap and its seeming operator?   In this subsection, we answer these contiguous questions by showing an afterthought approach to parasitic gap constructions.

Let us now consider the derivation of the parasitic gap sentence exemplified in (1a), which is repeated here as (29).

(29)   Which paper did you file without reading?

Recall here that chapter 3 acquired a general theory of labeling/edge feature (EF) detoxification referred to as (non)phasal valuation and that chapter 5 refined this theory by incorporating the system of amalgamation based on feature inheritance.   Under these analyses, the afterthought approach provides (29) with the following structures:

---

[9] Chomsky (2004) also shows that the afterthought analysis applies to "extraposition from NP."

(30)   a.   [$_{CP(+WH)}$ Which paper [$_{TP(v\varphi)}$ you C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <you$_{[v\varphi]}$> [$_{VP(v\varphi)}$

                    <which paper> v–V$_{[u\varphi][EF][EF(v)]}$ <which paper$_{[v\varphi]}$>]]]]

    b.   [[~~$_{KP}$ which paper~~] [$_{CP(+WH)}$ (without) Op [$_{TP(v\varphi)}$ PRO C–T$_{[u\varphi][EF][EF(C)]}$

                    [$_{vP(TRANSITIVE)}$ <PRO$_{[v\varphi]}$> [$_{VP(v\varphi)}$ <Op> v–V$_{[u\varphi][EF][EF(v)]}$ <Op$_{[v\varphi]}$>]]]]]]

(30a) and (30b) signify the derivations of the main clause and the afterthought phrase, respectively. The former corresponds to the derivation of an object *wh*-question. At the *v*P-level structure, EM of *you* into "Spec-*v*" labels *v*P as [TRANSITIVE], and IM of *which paper* into "Spec-V" labels VP as [vφ]. At the CP-level structure, IM of *which paper* into "Spec-C" assigns the label of [+WH] to CP, and IM of *you* into "Spec-T" assigns the label of [vφ] to TP. The labeled syntactic objects (SOs) satisfy the requirement of Full Interpretation (FI) by undergoing cyclic Transfer.

The derivation proceeds to introduce an afterthought phrase corresponding to the adjunct clause including a parasitic gap. As illustrated in (30b), I assume *which paper without reading* to be the afterthought phrase in question. Within this afterthought phrase, *without reading* functions as a restrictive relative clause predicated of the external relative head noun *which article*, an element copied from the main derivational workspace in (30a), which is finally omitted at the phonetic interface. The predication relation here is thus implemented by a "matching" relative (see e.g. Lees (1960, 1961), Chomsky (1965)) rather than a "raising" relative (see e.g. Vergnaud (1974), Kayne (1994)). The former involves the adjunction of a relative CP to its external relative head noun; the latter is a relative clause whose external relative head noun is in a derivational relation with a clause-internal argument position. In (30b), which involves a matching relative, IM raises the null operator *Op* to "Spec-C" in a successive-cyclic fashion. The merger of *Op* into "Spec-C" labels CP via EF detoxification as [+WH] on the assumption that *Op* is an operator that bears a valued/interpretable *wh*-feature. The whole relative CP la-

beled by *Op* enters into a predication relation with the external relative head noun *which paper* copied from the main derivational workspace. *Op* is identified with the external relative head noun *which paper* as a result.[10]

Notice that the proposed analysis succeeds in deriving and interpreting parasitic gaps. IM of *Op* into "Spec-C" creates a gap, more precisely, a copy in the afterthought phrase. This copy is parasitic in the sense that its semantic content is saturated by some element copied from the main derivational workspace. We can now answer the two core questions noted above: (i) what derives a parasitic gap? and (ii) what establishes a coreferential relation between a parasitic gap and its seeming operator? The answer to (i) is IM of *Op* into "Spec-C" in relativization. The answer to (ii) is the saturation of the semantic content of *Op* based on predication.

Importantly, the following contrast from Lasnik and Uriagereka (1988: 75) strongly suggests that the multiple derivational approach to parasitic gap constructions proposed here is on the right track:

(31) a. ＊Without reading, which report did you file?

   b. After visiting Bill, who did you file?

---

[10] Given that the semantic content of *Op* is saturated through predication, it is unlikely that this element has a set of valued/interpretable agreement features [vφ] at least at the point of constructing the *v*P-level structure. If this is the case, nonphasal valuation for VP would go wrong, because it is not until the computational system assembles the CP-level structure that *Op* gains [vφ], which labels VP. Cyclic Transfer might map this unlabeled VP onto the interfaces by the time predication obtains, which renders the derivation nonconvergent at the interfaces. Recall, however, that convergence (i.e. the state where unvalued/uninterpretable features [uF] are detoxified by feature valuation and (non)phasal valuation) detects a phase and that the largest convergent SO of the detected phase undergoes Transfer (see section 2.6.2 in chapter 2). The lack of the semantic content of *Op* should not induce valuation/convergence, which means that a derivation like (30) containing *Op* in object position undergoes no Transfer operation at the *v*P-level structure. After *Op* acquires its semantic content at the CP-level structure, every operation within the entire CP, including feature valuation and (non)phasal valuation, is considered to take place simultaneously with Transfer, which leads the relevant derivation to convergence at the interfaces. In addition, the absence of Transfer at the *v*P-level structure would easily guarantee an coreferential relation between the external relative head noun and the object element.

The ungrammatical status of (31a) means that if the adjunct clause including a parasitic gap is preposed to sentence-initial position, the relevant sentence fails to license the parasitic gap. The observed effect follows naturally from the proposed approach, according to which the adjunct clause including a parasitic gap works as a restrictive matching relative clause that modifies a certain material copied from the main derivational workspace. It is thus impossible for the computational system to introduce the adjunct clause including a parasitic gap prior to the main clause. As is clear from the grammatical contrast in (31), the adjunct clause including a parasitic gap is in essence different from a normal adjunct clause. Lasnik and Uriagereka (1988: 75) state, based on this grammatical contrast, that "parasitic gaps might exhibit some kind of left-right asymmetry." Our approach is capable of reducing this sort of left-right asymmetry to the mechanism of afterthoughts in a principled manner.

In this subsection, we showed, based on the mechanism of afterthoughts proposed by Chomsky (2004), that the adjunct clause including a parasitic gap serves as a restrictive matching relative clause that is derived separately from the main derivational workspace, which gives explicit answers to core questions about parasitic gap phenomena. In the following subsection, we offer some arguments for the "parasitic relative" analysis proposed here.

6.3.2.2. Supporting Arguments

6.3.2.2.1. The Reanalysis from Prepositions to Complementizers

According to the parasitic relative analysis presented in the preceding subsection, the categorial status of the adjunct clause including a parasitic gap is CP. We noted in the last chapter that *that* in the non-ECM (Exceptional Case-Marking) complement can be analyzed as a phonetic realization of the label of CP (see note 16 of chapter 5). In light of these analyses, it is expected that such elements as *without*, *before*, and *after* heading a parasitic relative are also phonetic manifestations of the label of CP, which we can define as virtual complementizers.

Huang (1982: 83–86) demonstrates that prepositional elements can have an ability of introducing clauses:

(32)   For John to come would be difficult.

(33)   a.   Zheijian shi   gen [ta lai   bu lai]   meiyou guanxi.

                         this   matter with he come not come no   relation

                         "This matter has nothing to do with whether he is coming or not."

             b.   Ta ba   [Lisi Jiehun] bu dang yihui shi.

                         he BA   marry   not treat one   matter

                         "He does not take it serious that Lisi is getting married."

(32) and (33) indicate that prepositional elements (i.e. *for* in (32), *gen* in (33a), and *ba* in (33b)) can function to head clausal complements in English and Chinese, respectively. These facts suggest that a parasitic relative has the status of CP in which a prepositional element like *without*, *before*, and *after* behaves as a complementizer.[11]

    We can elicit the same conclusion from Stowell's (1982) observations:

(34)   a.   Jenny remembered [PRO to bring the wine].

             b.   Jenny remembered [PRO bringing the wine].

(35)   a.   Jim tried [PRO to lock the door].

             b.   Jim tried [PRO locking the door].

---

[11] See Dubinsky and Williams (1995) for a possible counterargument to the analysis made here. They analyze a temporal preposition (e.g. *after*, *before*, *while*) as the head C and a nontemporal preposition (e.g. *without*, *despite*, *about*) as the head P.

The (a)-examples and (b)-examples contain the infinitival complements and the gerundive complements, respectively. Whereas the former are only understood as [unrealized] with respect to tense interpretation (cf. Bresnan (1972)), the latter can have diverse tense interpretations such as [past], [present], and [unrealized] according to the semantic properties of their matrix verbs. Stowell explains this contrast by assuming that there is a COMP position (i.e. Spec-C) in the infinitival complement but not in the gerundive complement, based on den Besten's (1978) statement that tense operators occur in the D-structure COMP position, which in passing we can regard as the conceptual emergence of the mechanism of feature inheritance. The lack of COMP position entails that of a tense operator. The infinitival complement, having an independent tense operator, is always unambiguous with respect to tense interpretation. In contrast, the gerundive complement cannot bear a unique tense interpretation in the absence of its own tense operator; rather, it must be interpreted according to the semantic property of its matrix verb.

Returning to the analysis of a parasitic relative, we find that its prepositional element, such as *without*, *before*, and *after*, forces a unique tense interpretation on the heading gerundive complement. This finding suggests that the preposition-like element of a parasitic relative is a virtual complementizer, which in turn shows that the adjunct clause including a parasitic gap has the status of CP. The present analysis is also evidenced by the following examples from Chomsky (1986a: 55–56):

(36)   a.   This is the man John interviewed *t* [before expecting us to ask you [which job to give to *pg*]].

          b.   Who did they expect us to ask you [which job to give to *t*]?

Sentence (36a) illustrates that if a parasitic gap is deeply embedded, it degrades acceptability. This grammatical pattern is parallel to that of (36b), in which *wh*-movement out of the *wh*-island

deteriorates acceptability.[12]   These facts support the view that the adjunct clause including a

parasitic gap is a restrictive relative clause CP and that there occurs a relativization operation that

raises an empty operator to "Spec-C" in the parasitic relative.[13]

6.3.2.2.2.   Categorial Restriction and Anti-Reconstruction Effects

We saw in section 6.2.3.2 that parasitic gaps are restricted to the nominal category.   The

relevant examples are repeated here as (37).

(37)　　a.　　*[How sick]$_i$ did John say he felt t_i before getting pg_i?

　　　　　b.　　*[How long]$_i$ did John drink t_i before lecturing pg_i?

　　　　　c.　　*This is a topic [about which]$_i$ you should think t_i before talking pg_i.

As shown in (37), parasitic gaps are not licensed in environments where adjective, adverbial, and

prepositional categories undergo *wh*-movement.   Our analysis easily captures these licensing

patterns.   Let us take (37a) as an instance:

---

[12] Note that Chomsky (1986a: 55–56) does not give any judgment marks to the sentences in (38).   He
mentions the deterioration of acceptability in question in the text.

[13] Interestingly, there are cases where adverbial *before*, *when*, and *after* work as adnominal clauses by
heading clausal complements:

(i)　　a.　　the interval before she spoke was appreciate, and that was against the rules of the
　　　　　　　game.　　　　　　　　　　　　　　　　　　　　　　　　　　　(Haan (1989: 106))
　　　　b.　　…he remembered Ron's expression when he had seen her kissing Dean,…
　　　　c.　　Just take it as a challenge, because some of them are very hard to get, but the satisfac-
　　　　　　　tion after you complete the stage is a very valuable prize.　　(Shizawa (2011: 229))

The existence of these expressions might also bolster up the analysis developed here if adverbial *before*,
when, and *after* are regarded as complementizers in the sense noted above.

(38) *[[$_{AP}$ how sick] [$_{CP(+WH)}$ (before) Op [$_{TP(v\phi)}$ PRO C–T$_{[u\phi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <Op$_{[v\phi]}$>

   [$_{VP(v\phi)}$ <Op> v–V$_{[u\phi][EF][EF(v)]}$ <Op$_{[v\phi]}$>]]]]]]

(38) denotes the derivation of the afterthought phrase in (37a).   Under the proposed analysis, the adjunct clause including a parasitic gap functions as a restrictive relative clause that modifies a *wh*-phrase copied from the main derivational workspace.   As is well known, it is only a nominal phrase that a restrictive relative clause can modify.   The *wh*-phrase to be modified in (37a) belongs to the adjective category, as indicated in (38).   The ungrammaticality of (37a) thus follows naturally.

   The plausibility of our analysis also stems from the anti-reconstruction effects observed in (22), which is reproduced here as (39).

(39)   a.   [Which books about himself$_j$]$_i$ did John$_j$ file t_i before Mary read pg_i?

   b.   *[Which books about herself$_j$]$_i$ did John file t_i before Mary$_j$ read pg_i?

This grammatical contrast means that (39a) observes, and (39b) violates, Condition A, which in turn shows that parasitic gaps refuse, and real gaps accept, reconstruction.   Our analysis provides the sentences in (39a) and (39b) with the derivations in (40) and (41).

(40) a. [$_{CP(+WH)}$ [Which books about himself] [$_{TP(v\phi)}$ John C–T$_{[u\phi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$

<John$_{[v\phi]}$> v [$_{VP(v\phi)}$ <[which books about himself]> v–V$_{[u\phi][EF][EF(v)]}$ <[which

books about himself$_{[v\phi]}$]>]]]]

b. [[$_{KP}$ Which books about himself] [$_{CP(+WH)}$ (before) Op [$_{TP(v\phi)}$ Mary

C–T$_{[u\phi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <Mary$_{[v\phi]}$> [$_{VP(v\phi)}$ <Op> v–V$_{[u\phi][EF][EF(v)]}$

<Op$_{[v\phi]}$>]]]]]

(41) a. *[$_{CP(+WH)}$ [Which books about herself] [$_{TP(v\phi)}$ John C–T$_{[u\phi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$

<John$_{[v\phi]}$> v [$_{VP(v\phi)}$ <[which books about herself]> v–V$_{[u\phi][EF][EF(v)]}$ <[which

books about herself$_{[v\phi]}$]>]]]]

b. *[[$_{KP}$ Which books about herself] [$_{CP(+WH)}$ (before) Op [$_{TP(v\phi)}$ Mary

C–T$_{[u\phi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <Mary$_{[v\phi]}$> [$_{VP(v\phi)}$ <Op> v–V$_{[u\phi][EF][EF(v)]}$

<Op$_{[v\phi]}$>]]]]]

The (a)-derivations and the (b)-derivations are of the main clauses and the afterthought phrases, respectively. In (40a), a c-command relation holds between *John* and *himself*, hence the observance of Condition A (see also section 4.3.4 in chapter 5 for the definition of Condition A in this thesis). Under the afterthought approach, *which books about himself* in the main derivational workspace is reintroduced into the different derivational workspace in (40b) and is modified by the matching parasitic relative. Because the referent of *himself* has already been fixed in (40a), it needs not to establish a binding relation in (40b). Derivations (40a) and (40b) therefore fall into place without any discrepancy.

Derivations (41a) and (41b), on the other hand, pose a Condition A violation. Our system separates the derivation of the main clause and the derivation of the adjunct clause including a parasitic gap. Each derivation has to seek convergence in its own way. *Herself* in (41a) cannot find its referent in the derivational workspace to which it belongs. Because the deriva-

tion crashes at this point, the computational system cannot embark on the derivation of the adjunct clause including a parasitic gap. If at all, *herself* in (41b) would not discover its referent because the derivational workspace to which it belongs involves a matching relative, which in nature disallows reconstruction. Anti-reconstruction effects therefore follow.

Section 6.3.2.2 verified the validity of the parasitic relative analysis proposed in section 6.3.2.1 by guaranteeing the categorial status of a parasitic relative as CP and by explaining categorial restriction and anti-reconstruction effects. However, we are still left with unexplained phenomena: anti-c-command effects, S-structure effects, and A′-movement versus A-movement. In the next subsection, I show that the theory of (non)phasal valuation makes it possible to account for these phenomena as well.

6.3.2.3. The Condition on Predication as Part of the Mechanism of Afterthoughts

In this subsection, we propose that the licensing of parasitic gaps is sensitive to a certain condition on predication as part of the mechanism of afterthoughts. The proposed analysis is shown to capture descriptive generalizations such as anti-c-command effects, S-structure effects, and A′-movement versus A-movement.

6.3.2.3.1. Anti-C-Command Effects

According to the anti-c-command condition given in (13), parasitic gaps refuse to be c-commanded by real gaps. The relevant sentence is repeated here as (42).

(42)   *I wonder which man$_i$ t_i called you before you met pg_i.

The real gap in (42), which is the trace produced by movement of *which man*, c-commands the parasitic gap in the adjunct clause. Such environments do not admit the occurrence of parasitic

gaps, as the ungrammatical status of (42) indicates. Our approach assigns (42) the following derivations along with the theory of (non)phasal valuation:

(43)   a.   I wonder $[_{\text{TP(+WH, v}\varphi)}$ which man C–T$_{[u\varphi][EF][EF(C)]}$ $[_{v\text{P(TRANSITIVE)}}$ <which man$_{[v\varphi]}$>

        $[_{\text{VP(v}\varphi)}$ you v–V$_{[u\varphi][EF][EF(v)]}$ <you$_{[v\varphi]}$>]]]

    b. $[[_{\text{KP}}$ ~~which man~~] $[_{\text{CP(+WH)}}$ (before) Op $[_{\text{TP(v}\varphi)}$ you C–T$_{[u\varphi][EF][EF(C)]}$ $[_{v\text{P(TRANSITIVE)}}$

        <you$_{[v\varphi]}$> $[_{\text{VP(v}\varphi)}$ <Op> v–V$_{[u\varphi][EF][EF(v)]}$ <Op$_{[v\varphi]}$>]]]]]

(43a) and (43b) signify the derivations of the main clause and the afterthought phrase, respectively. Recall here that the theory of (non)phasal valuation involves the system of amalgamation. This system enables the syntactic amalgam to license "multiple Specs" by doubly utilizing the phasal EF and the nonphasal EF. An object *wh*-construction exhibits a straightforward manifestation of the "multiple Specs" analysis. The derivation of the parasitic relative given in (43b) is parallel to that of this construction. The C-T amalgam in (43b) has "multiple Specs," with *Op* and *you* internally merging into "Spec-C" and "Spec-T," respectively. As a result, (non)phasal valuation feeds CP and TP with the labels of [+WH] and [vφ], respectively (here and below, we disregard (non)phasal valuation for *v*P and VP for simplicity). The derivation of the main clause, in contrast, displays a special realization of the "multiple Specs" analysis because the main clause corresponds to a subject *wh*-construction. A subject *wh*-phrase, unlike an object *wh*-phrase, bears both a *wh*-feature and a φ-feature. Accordingly, in (43a), the solitary merger of *which man* with the C-T amalgam labels the phasal and nonphasal SO as [+WH, vφ]. Because the C-T amalgam in (43a) leaves the "higher Spec" empty, the semantic interface interprets the relevant "Spec" as "Spec-T."

    Suppose now that the licensing of parasitic gaps is susceptible to a certain condition on predication as part of the mechanism of afterthoughts:

(44)   The null operator *Op* heading a parasitic relative modifier and the (*wh-*)modifiee heading a main clause license a nondistinctive label.

The condition in (44) requires *Op* and the (*wh-*)modifiee to license a nondistinctive label in derivation.   If derivation satisfies this condition, predication as part of the mechanism of afterthoughts holds between a parasitic relative and a main clause, thus observing FI.

Returning to derivations (43a) and (43b), we find that *Op* and the (*wh-*)modifiee provide a distinctive label: [+WH] versus [+WH, $v\varphi$].   The result is a violation of (44), which disallows the appearance of the parasitic gap in (42).   The anti-c-command condition thus follows without any problem.   Of course, the derivation of a typical parasitic gap sentence of the sort described in (29) fulfills this condition because *Op* and the (*wh-*)modifiee both license the label of [+WH].[14]

6.3.2.3.2.   S-Structure Effects

The same account holds true of S-structure effects.   The relevant example is repeated here as (45), the derivations of which are given in (46).

(45)   *Who filed which report$_i$ without reading *pg*$_i$?

(46)   a.   [$_{TP(+WH, v\varphi)}$ Who C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <who$_{[v\varphi]}$> [$_{VP(v\varphi)}$ which report v–V$_{[u\varphi][EF][EF(v)]}$ <which report$_{[v\varphi]}$>]]]

        b.   [[$_{KP}$ which report] [$_{CP(+WH)}$ (without) Op [$_{TP(v\varphi)}$ PRO C–T$_{[u\varphi][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$ <PRO$_{[v\varphi]}$> [$_{VP(v\varphi)}$ <Op> v–V$_{[u\varphi][EF][EF(v)]}$ <Op$_{[v\varphi]}$>]]]]]

---

[14] The formulation of (44) suggests that the labeling by the lower copy, which is often taken to be invisible computationally, does not pertain to the establishment of predication as part of the mechanism of afterthoughts.

As illustrated in (46a) and (46b), *Op* licenses the label of [+WH] but the *wh*-modifiee the label of [vφ] (see section 4.3.2 in chapter 5 for the treatment of in-situ *wh*-phrases in this thesis).　This distinctiveness in labeling deranges the establishment of predication as part of the mechanism of afterthoughts, from which S-structure effects are derived.

6.3.2.3.3.　A′-Movement versus A-Movement

　　　Finally, we deal with the generalization of A′-movement versus A-movement observed in (19), which is repeated here as (47).

(47)　　*The book$_i$ was filed t_i without my reading *pg*$_i$ first.

As indicated in (47), it is argued in the literature that A-movement, being a kind of S-structure movement like A′-movement, is not compatible with a parasitic gap.　Our analysis explains this incompatibility in the following way:

(48)　　a.　　[$_{TP(-WH,\ vφ)}$ The book C–T$_{[uφ][EF][EF(C)]}$ [$_{VP(UNACCUSATIVE,\ vφ)}$ <the book$_{[vφ]}$>

　　　　　　　　v–V$_{[uφ][EF][EF(v)]}$ <the book$_{[vφ]}$>]]

　　　b.　　[[~~$_{KP}$ the book~~] [$_{CP(+WH)}$ (without) Op [$_{TP(vφ)}$ my C–T$_{[uφ][EF][EF(C)]}$ [$_{vP(TRANSITIVE)}$

　　　　　　　　<my$_{[vφ]}$> [$_{VP(vφ)}$ <Op> *v*–V$_{[uφ][EF][EF(v)]}$ <Op$_{[vφ]}$> first]]]]]

Derivations (48a) and (48b) show that *Op* licenses the label of [+WH] whereas the modifiee *the book* licenses the label of [–WH, vφ].　Here too, (44) is violated; thus, predication as part of the mechanism of afterthoughts does not obtain between *Op* and *the book*.　Hence, the inconsistency of A-movement with a parasitic gap.

　　　Summarizing section 6.3, we argued that the adjunct clause including a parasitic gap be-

haves as a restrictive matching relative clause that the mechanism of afterthoughts introduces separately from the main derivational workspace and that a relevant predication falls under the condition formulated in (44). This analysis based on the multiple derivational workspace approach differs from Kasai's (2010) multiple dominance analysis sketched in section 6.2.3, which adheres rigidly to the single derivational workspace approach. Section 6.4 compares these analyses, thereby exploring some predictive differences between the single and multiple derivational workspace approaches.

6.4. Consequences: Some Predictive Differences between the Single and Multiple Derivational Workspace Approaches

Thus far, we have seen that there are two analyses of parasitic gap constructions: the multiple dominance analysis (see section 6.2.3) and the afterthought analysis (see section 6.3.2). These are like-minded analyses in the sense that they endeavor to reduce the significant insight of Chomsky (1986a) to the principles and operations of minimal syntax. At the same time, these two analyses are vested in an entirely separate line of inquiry in that the multiple dominance analysis initiates the single derivational workspace approach while the afterthought analysis develops the multiple derivational workspace approach, respectively. Keeping to SMT, both analyses are worth pursuing. We dedicate this section to investigating some differences in prediction between the single and multiple derivational workspace approaches.

6.4.1. The Problem of Overgeneration

Let us begin with the problem of overgeneration. In contrast to the afterthought analysis, the multiple dominance analysis, which adopts the single derivational workspace approach, is always exposed to this problem. It is not obvious at which point multiple dominance applies in a precise sense. If multiple dominance creates a plurality of derivations, the computational

system must compare them with respect to economical efficiency. In effect, Kasai's account of anti-reconstruction effects necessitates the consideration of global economy in such a way that it works in tandem with the theory of wholesale late merger. Collins (1997) argues that it is desirable to construct the computational system based on local economy, under which the decision about whether an operation may apply to an SO is made only on the basis of information available in the SO. Local economy forbids the computational system from measuring the derivation with an operation against another derivation without the same operation on the assumption that an identical Numeration produces both derivations, which corresponds to the consideration of global economy. In chapter 5, we successfully reduced the global nature that the vacuous movement hypothesis exhibits to the local nature that feature inheritance displays, which suggests that local economy is superior to global economy. In order for the multiple dominance analysis to work out well, it needs to surmount the problem of overgeneration and thus global economy.

6.4.2. The Anti-C-Command Condition Redux

In this subsection, we discuss the issue surrounding the anti-c-command condition described in (13), which is reproduced here as (49).

(49)   The anti-c-command condition:

A parasitic gap may not be c-commanded by the real gap.

It is noted in the literature (cf. Engdahl (1984), Kiss (1985), Brody (1995), Sakamoto (2011a, b)) that (49) encounters some counterevidence. First, apart from the two analyses seen above, we examine counterexamples to this condition discovered by previous studies. Subsequently, we consider implications of such counterexamples for the single and multiple derivational work-

space approaches.

6.4.2.1. Examining Counterexamples

In contradiction to the anti-c-command condition, it is known that there are cases where parasitic gaps are licensed though they are c-commanded by real gaps:

(50)　　a.　　Which man did the police warn *t* [that they would arrest *pg*]?

　　　　b.　　*The police warned him$_i$ [that they would arrest John$_i$].

　　　　c.　　The police warned everybody$_i$ [that they would arrest him$_i$].

(Kiss (1985: 42, 45))

(51)　　?Who did Bill believe *t* visited you [without you having invited *pg*]?

(Brody (1995: 83))

In (50a), the real gap occupying matrix object position c-commands the parasitic gap within the clausal complement.　　This c-command relation is clear from (50b, c).　　In (50b), coreferential interpretation is not established between *him* and *John*.　　In (50c), the pronoun *him* can be interpreted as a variable bound by the operator *everybody*.　　These interpretational patterns indicate that matrix object position c-commands into the clausal complement.　　Similarly, the real gap in (51) c-commands the parasitic gap, with the adjunct clause adjoined to the embedded verbal domain.　　Despite the fact that the parasitic gaps are c-commanded by the real gaps, sentences (50a) and (51) still remain grammatical.　　It thus follows from the observation in (50) and (51) that the anti-c-command condition is not qualified as a licensing condition of parasitic gaps.

Given these circumstances, Sakamoto (2011a, b) argues, based on Chomsky's (1986a) analysis, that the parallelism condition described in (52) is a true licensing condition of parasitic gap phenomena.

(52)   The parallelism condition:

Composed chains can only consist of parallel chains.

This condition holds that parasitic gaps are licensed via the parallelism in chain types between a real gap chain and a parasitic gap chain.[15]   (50a) and (51) have the following composed chains, respectively:

(53)   a. $(C, C') = (which\ man_{CP}, t, Op_{CP}, pg)$

b. $(C, C') = (who_{CP}, t, Op_{CP}, pg)$

As shown in (53a) and (53b), the real gap chain and the parasitic gap chain are parallel because there occur chain formations to Spec-C in both chains, thus satisfying the condition in (52). The parasitic gaps in (50a) and (51) are licensed as a result.

The parallelism condition also predicts the following grammatical contrasts ((54) and (55) are cited from Sakamoto (2011a: 202) and Engdahl (1983: 20–21), respectively):

(54)   a.   ? Who did Bill believe *t* visited John while him having refused to meet *pg*?

b.   * Who did Bill believe *t′* was visited *t* by John while him having refused to meet *pg*?

(55)   a.   Which Caesar did Brutus imply *t* was no good while ostensibly praising *pg*?

b.   * Which articles did you say *t′* got filed *t* by John without him reading *pg*?

---

[15] Kim and Lyle (1996) and Karimi (1999) propose a similar condition on parasitic gap constructions. These works, in common with Sakamoto (2011a, b), are based on Chomsky's (1986a) analysis (see also Sakamoto (2012: section 3.5) for related discussions.

The (b)-examples, unlike the (a)-examples, disallow the occurrence of parasitic gaps.　The differences in chain types explain these grammatical contrasts:

(56)　　a.　　　$(C, C') = (who_{CP}, t, Op_{CP}, pg)$

　　　　b.　　*$(C, C') = (who_{CP}, t_{TP}, t, Op_{CP}, pg)$

(57)　　a.　　　$(C, C') = (which\ Caesar_{CP}, t, Op_{CP}, pg)$

　　　　b.　　*$(C, C') = (which\ articles_{CP}, t_{TP}, t, Op_{CP}, pg)$

The (a)-chains consist of parallel chains, as with (53), so they constitute no problem for (52). The (b)-chains, on the other hand, comprise different types of chains because the real gap chains contain A-chains, hence a violation of (52).　The grammatical contrasts in (54) and (55) thus follow.[16]

　　The account here expects a parasitic gap sentence to be grammatical if an A-chain intervenes both in a real gap chain and in a parasitic gap chain.　In fact, this prediction is borne out ((58a) and (58b) are cited from Sakamoto (2011a: 204)):

---

[16] The grammatical status of (54a) can be problematic for the parallelism condition under the predicate-internal subject hypothesis, which allows *who* in (54a) to undergo A-movement in the embedded clause.　If this is the case, (54a) should have the same composed chain as (56b), an unwelcome result. To find a way out of this offending situation, Sakamoto (2011a: note 8) suggests that the parallelism condition can disregard the chain created by vacuous movement.　According to this analysis, (54a) produces the composed chain $(C, C') = (who_{CP}, t_{TP}, t, Op_{CP}, pg)$, but the parallelism condition disregards t_{TP} for its vacuous nature.　It then follows that the virtual chain that the parallelism condition evaluates is $(C, C') = (who_{CP}, t, Op_{CP}, pg)$, which consists of parallel chains.　Sentence (54a) thus tolerates the occurrence of a parasitic gap.

　　It should be noted, however, that vacuous movement no longer constitutes an unwelcome operation under our system, as discussed in chapter 5.　The theory of phases also poses a potential problem for the analysis based on the parallelism condition, forcing a *wh*-phrase to undergo successive-cyclic movement, as Sakamoto (2011a: note 8) states.　It is not clear whether the parallelism condition can disregard the chains created by successive-cyclic movement.　In any case, we need to provide the parallelism condition with a more precise characterization if we pursue the analysis based on the parallelism condition.

(58)　　a.　　?Which candidate do you think *t′* was hired *t* without believing *pg′* to have

been fired *pg* before?

b.　　(C, C') = (*which candidate*$_{CP}$, *t*$_{TP}$, *t*, *Op*$_{CP}$, *pg*$_{TP}$, *pg*)

As illustrated in (58a), there is an occurrence of A-movement that steps between A′-movement

to Spec-C both in the main clause and in the adjunct clause, which yields the composed chain in

(58b).　It is obvious that this formed chain is parallel with respect to chain types, thus observing

the parallelism condition in (52), which renders (58a) grammatical.

The parallelism condition is also responsible for anti-c-command effects, S-structure ef-

fects, and the generalization of A′-movement versus A-movement in a similar manner (see Sa-

kamoto (2011a, b, 2012) for details) but is silent on categorial restriction and anti-reconstruction

effects.　In other words, this condition ends as a descriptive generalization without attaining the

level of a deeper explanation, although it is more general than the anti-c-command condition.

Our ultimate objective is to attribute all the properties of parasitic gap constructions to the nature

of core grammar in accordance with SMT.　As it turns out, however, our analysis, which works

to complete this objective, seems to be left with certain unexplained phenomena that the paral-

lelism condition can capture exclusively, as noted in the next subsection.　We can speculate

from this result that the parallelism condition characterizes not the nature of core grammar but

rather a different aspect in the system of grammar; that is, the parallelism requirement contrib-

utes to reducing processing load.　In section 6.4.2.2, we show that this speculation is on the

right track by reinterpreting the counterexamples to the anti-c-command condition observed

above.

6.4.2.2.　Reinterpreting Counterexamples

Let us start by pointing out that (51) and (54a) are problematic for the multiple dominance

analysis.   To understand why these examples raise a problem for this analysis, recall that Kasai (2010) derives anti-c-command effects as follows:

(59)   *I wonder which man$_i$ t_i called you before you met pg_i.

(60)

<div style="text-align:center">

     vP   CP

     which man$_i$

 v AspP you met which man$_i$

</div>
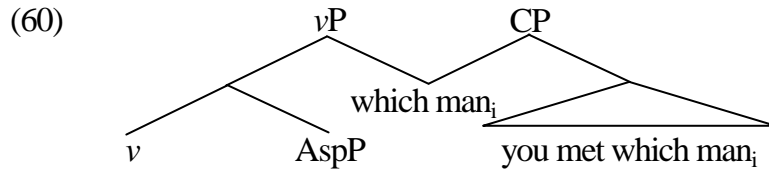
According to his assumption, the adjunction site of an adjunct clause is the edge of matrix Asp, which is lower than object position in the matrix clause.   As depicted in (60), however, the derivation already assembles the matrix vP-level structure at the point of applying multiple dominance.   The adjunct clause cannot thus adjoin to matrix AspP, from which anti-c-command effects are derived.

   With this in mind, consider examples (51) and (54a).   These examples have subject *wh*-phrases in the main clause, as with the case of the ungrammatical example in (59).   This means that (51) and (54a) are derived in a similar way to (60).   In plain words, the derivation already assembles the embedded vP-level structure in the main clause at the point of applying multiple dominance.   Although the adjunct clause in (51) and (54a) has to adjoin to embedded AspP in the main clause, it is impossible for it to do so.   The reason is that the derivation proceeds to the embedded vP-level structure in the main clause, which is structurally higher than embedded AspP in the main clause.   Accordingly, (51) and (54a) should not converge, but they are actually permitted to emerge as grammatical sentences.   The multiple dominance analysis is thus able to capture anti-c-command effects but fails to explain certain counterexamples like (51) and (54a).

   In contrast, the afterthought analysis makes a different prediction.   For simplicity, all rel-

evant examples given in the last subsection are reproduced:

(61)   Which man did the police warn *t* [that they would arrest *pg*]?   (= (50a))

(62)   ?Who did Bill believe *t* visited you [without you having invited *pg*]?   (= (51))

(63)   a.   ?Who did Bill believe *t* visited John while him having refused to meet *pg*?

        b.   *Who did Bill believe *t′* was visited *t* by John while him having refused to

              meet *pg*?   (= (54a-b))

(64)   a.   Which Caesar did Brutus imply *t* was no good while ostensibly praising *pg*?

        b.   *Which articles did you say *t′* got filed *t* by John without him reading *pg*?

                                                                             (= (55a-b))

(65)   a.   ?Which candidate do you think *t′* was hired *t* without believing *pg′* to have

              been fired *pg* before?   (= (58))

We presented (61) and (62) as counterexamples to the anti-c-command condition.   Because we

discussed (63), (64), and (65) to confirm the validity of the parallelism condition, it is only the

sentence in (63a) that behaves as a counterexample to the anti-c-command condition in (63)-(65).

Apart from these two conditions, the afterthought analysis can make a correct prediction for all

examples but (63b) and (64b).   Recall that the afterthought analysis imposes the condition in

(44), repeated here as (66), on predication as part of the mechanism of afterthoughts.

(66)   The null operator *Op* heading a parasitic relative modifier and the (*wh*-)modifiee

        heading a main clause license a nondistinctive label.

Based on (66), all the sentences in (61)-(65) should license a parasitic gap because *Op* and the

wh-modifiee provide a nondistinctive label, namely, [+WH] by undergoing IM into "Spec-C."

In effect, however, (63b) and (64b) fail to qualify as parasitic gap sentences.

What then derives the difference between (63b) and (64b) on one hand and the other grammatical sentences on the other hand? Putting together our analysis and the analysis based on the parallelism condition, we can find out one possible answer. As noted, the afterthought analysis intends to show up the nature of core grammar with adherence to SMT and did so; the parallelism condition is a descriptive generalization that does not guide us to the level of a deeper explanation. It is natural to consider that the latter imposes a parallelism requirement at a more surface level. It is unlikely, however, that the parallelism requirement exists in such a form as in (52) for the reason stated in note 16. Rather, the parallelism condition should be formulated to eliminate the voice mismatch between the main clause and the adjunct clause including a parasitic gap. The facts observed in (63)-(65) suggest that the disappearance of voice mismatch ameliorates acceptability by reducing processing load, as Merchant (2013) makes a similar observation with clausal ellipses such as sluicing. We can therefore safely conclude that the afterthought analysis and the analysis based on the parallelism condition characterize the nature of core grammar and the surface aspect of grammar, respectively.

To sum up section 6.4, we inspected some differences in prediction between the single and multiple derivational workspace approaches to parasitic gap constructions, each of which corresponds to the multiple dominance analysis and the afterthought analysis. Through this inspection, we found that the afterthought analysis, unlike the multiple dominance analysis, leads us to some interesting consequences.

6.5. Concluding Remarks

In this chapter, we developed a multiple derivational workspace approach to parasitic gap constructions. This approach analyzes the adjunct clause including a parasitic gap as a restrictive matching relative clause that the mechanism of afterthoughts proposed by Chomsky (2004)

derives separately from the main derivational workspace (see section 6.3.2). By contrast, Kasai's (2010) multiple dominance analysis derives parasitic gap constructions within a single derivational workspace by optimizing EM and IM (see section 6.2.3). In this way, both analyses differ in quality but share the same perspective in that they endeavor to reduce the significant insight of Chomsky (1986a) to the principles and operations of minimal syntax. Abiding by SMT, both the afterthought analysis and the multiple dominance analysis are worth pursuing. By comparing these two analyses in detail, we found some differences in prediction. Of particular significance is that the afterthought analysis enables us to discriminate between the nature of core grammar and the surface aspect of grammar in conjunction with the analysis based on the parallelism condition (see section 6.4). However, there is an unclear point at the present stage; that is, it is not obvious what derives the condition on predication as part of the mechanism of afterthoughts (see section 6.3.2.3). To answer this question, we need to refine the system of predication in concert with the theory of labeling/EF detoxification proposed in this thesis. Also, it is favorable to explore the nature of related constructions such as *tough*-constructions that are generally supposed to involve null operator movement. Through these investigations, it is expected that we can characterize the theory of FL based on Merge in a more explicit way.

Chapter 7

Conclusion

In this thesis, I have investigated a mechanism of valuation in Merge in the modern minimalist framework. What valuation evokes in general is feature valuation based on Agree/Value, a process of "detoxifying" an unvalued/uninterpretable property called the Agree feature (AF). It has often been argued in a preminimalist framework that the existence of the AF, such as a Case-feature, is responsible for the dislocation property of the faculty of language (FL). Within a minimalist framework, however, it is not the AF but the edge feature (EF) that drives movement operations. Chomsky (2007a, 2008) reduces the EPP (Extended Projection Principle) property specific to a functional item to a more general property referred to as the EF that every lexical item (LI) shares, but its essential role remains unclear. In this thesis, I proposed that the EF, an unvalued/uninterpretable property that any LI bears, is valued/detoxified through Merge, but not Agree/Value, and that EF detoxification determines the interpretation of any syntactic object (SO) involving the valued EF, which is a process that I referred to as (non)phasal valuation. Crucially, the uninterpretability of derivation is detoxified not only by feature valuation based on Agree/Value but also by (non)phasal valuation based on Merge. The existence of valuation in Merge enabled us to locate the EF in the system of grammar in an appropriate manner.

I also showed that the theory of (non)phasal valuation derives the effects of labeling. According to Chomsky (2013), labeling plays a crucial role in the theory of FL based on Merge. Merge yields an SO as a simple set but does not name the SO because its application does not entail that of projection. The computational system thus necessitates the mechanism of labeling independent of Merge for the purpose of interpreting SOs at the interfaces. Chomsky proposes that the labeling algorithm (LA) is in charge of the interpretation of SOs. Pointing out

that LA is just stipulation, I reduced LA to the mechanism of (non)phasal valuation. Importantly, the theory of (non)phasal valuation offered a distinctive perspective in which the property of a (non)phase head is not underspecified at the initial stage of derivation but is rather specified via the merger of its mate in the course of derivation, in direct opposition to the standard theory. This opposite perspective succeeded in eliminating the stipulative mechanism of categorization adopted by Distributed Morphology.

By incorporating the theory of feature inheritance into the theory of (non)phasal valuation, we acquired a system of amalgamation of phase heads with nonphase heads. This system made it possible to not only eliminate the countercyclic application of internal Merge but also derive a wide range of phenomena surrounding vacuous movement. The system of amalgamation also helped to explain the properties of parasitic gap constructions in conjunction with the afterthought analysis based on Chomsky (2004), under which the adjunct clause with a parasitic gap functions as a restrictive relative clause that the mechanism of afterthoughts introduces independently of the main derivational workspace.

Given that (non)phasal valuation fixes the interpretational status of every SO, including clauses, phrases, and words, the field for linguistic inquiry based on this theory is unbounded without regard to languages. Although there are many remaining issues to be addressed in future investigation, I hope that this thesis contributes to further understanding of FL.

References

Abe, Jun and Norbert Hornstein (2010) "'Lasnik-Effects' and String-Vacuous ATB Movement," ms., Tohoku Gakuin University and University of Maryland. [Published in *Ways of Structure Building*, ed. by Myriam Uribe-Etxebarria and Vidal Valmala, 2012, 169–205, Oxford University Press, Oxford.]

Agbayani, Brian (2000) "*Wh*-Subjects in English and the Vacuous Movement Hypothesis," *Linguistic Inquiry* 31, 703–713.

Agbayani, Brian (2006) "Pied-Piping, Feature Movement, and *Wh*-Subjects," Wh-*Movement: Moving on*, ed. by Lisa L.-S. Cheng and Norbert Corver, 71–93, MIT Press, Cambridge, MA.

Arad, Maya (2003) "Locality Constraints on the Interpretation of Roots: The Case of Hebrew Denominal Verbs," *Natural Language and Linguistic Theory* 21, 737–778.

Bachrach, Asaf and Roni Katzir (2009) "Right-Node Raising and Delayed Spell-Out," *Inter-Phases: Phase-Theoretic Investigations of Linguistic Interfaces*, ed. by Kleanthes K. Grohmann, 283–316, Oxford University Press, Oxford.

Belletti, Adriana and Luigi Rizzi (1988) "Psych-Verbs and θ-Theory," *Natural Language and Linguistic Theory* 6, 291–352.

den Besten, Hans (1978) "Auxiliary Deletions and the Interplay between Local Deletive Rules and Filters," paper presented at the 1978 GLOW Colloquium.

Borer, Hagit (1994) "The Projection of Arguments," *Functional Projections: University of Massachusetts Occasional Papers* 17, ed. by Elena Benedicto and Jeff Runner, 19–48, GLSA.

Bošković, Željko (2007) "On the Locality and Motivation of Move and Agree: An Even More Minimal Theory," *Linguistic Inquiry* 38, 589–644.

Bošković, Željko and Steven Franks (2000) "Across-the-Board Movement and LF," *Syntax* 3, 107–128.

Bresnan, Joan (1972) *Theory of Complementation in English Syntax*, Doctoral dissertation, MIT.

Brody, Michael (1995) *Lexico-Logical Form: A Radically Minimalist Theory*, MIT Press, Cambridge, MA.

Cheng, Lisa L.-S. (1997) *On the Typology of* Wh-*Questions*, Garland, New York.

Chomsky, Noam (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

Chomsky, Noam (1970) "Remarks on Nominalization," *Readings in English Transformational Grammar*, ed. by Roderick A. Jacobs and Peter S. Rosenbaum, 184–221, Ginn and Company, Waltham, 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: The Pisa Lectures*, Foris, Dordrecht.

Chomsky, Noam (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge, MA.

Chomsky, Noam (1986a) *Barriers*, MIT Press, Cambridge, MA.

Chomsky, Noam (1986b) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.

Chomsky, Noam (1995a) *The Minimalist Program*, MIT Press, Cambridge, MA.

Chomsky, Noam (1995b) "Bare Phrase Structure," *Government and Binding Theory and the Minimalist Program*, ed. by Gert Webelhuth, 383–439, Blackwell, Oxford.

Chomsky, Noam (2000) "Minimalist Inquiries: The Framework," *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michael and

Juan Uriagereka, 89–155, MIT Press, Cambridge, MA.

Chomsky, Noam (2001) "Derivation by Phase," *Ken Hale: A Life in Language*, ed. by Michael Kenstwicz, 1–52, MIT Press, Cambridge, MA.

Chomsky, Noam (2004) "Beyond Explanatory Adequacy," *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*, 104–131, Oxford University Press, Oxford.

Chomsky, Noam (2005) "Three Factors in Language Design," *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam (2007a) "Approaching UG from Below," *Interfaces + Recursion = Language?: Chomsky's Minimalism and the View from Syntax-Semantics*, ed. by Uli Sauerland and Hans M. Gärtner, 1–29, Mouton de Gruyter, Berlin.

Chomsky, Noam (2007b) "Biolinguistic Explorations: Design, Development, Evolution" *International Journal of Philosophical Studies* 15, 1–22.

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 L. Zubizarreta, 133–166, MIT Press, Cambridge, MA.

Chomsky, Noam (2012) "Introduction," *Gengokisoronshu* (Foundations of Biolinguistics: Selected Writings), ed. by Naoki Fukui, 17–26, Iwanami Shoten, Tokyo.

Chomsky, Noam (2013) "Problems of Projection," *Lingua* 130, 33–49.

Cinque, Guglielmo (1981–82) "On the Theory of Relative Clauses and Markedness," *The Linguistic Review* 1, 247–294.

Cinque, Guglielmo (1990) *Types of Ā-Dependencies*, MIT Press, Cambridge, MA.

Citko, Barbara (2005) "On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge," *Linguistic Inquiry* 36, 475–496.

Collins, Chris (1997) *Local Economy*, MIT Press, Cambridge, MA.

Contreras, Heles (1984) "A Note on Parasitic Gaps," *Linguistic Inquiry* 15, 698–701.

Contreras, Heles (1993) "On Null Operator Structures," *Natural Language and Linguistic The-*

ory 11, 1–30.

Dubinsky, Stanley and Kemp Williams (1995) "Recategorization of Prepositions as Complementizers: The Case of Temporal Prepositions in English," *Linguistic Inquiry* 26, 125–137.

Embick, David and Alec Marantz (2008) "Architecture and Blocking," *Linguistic Inquiry* 39, 1–53.

Emonds, Joseph (1985) *A Unified Theory of Syntactic Categories*, Foris, Dordrecht.

Engdahl, Elisabet (1983) "Parasitic Gaps," *Linguistics and Philosophy* 6, 5–34. [Reprinted in *Parasitic Gaps*, ed. by Peter W. Culicover and Paul M. Postal, 2001, 69–98, MIT Press, Cambridge, MA.]

Engdahl, Elisabet (1984) "Why Some Empty Subjects Don't License Parasitic Gaps," *WCCFL* 3, 91–104.

Fiengo, Robert and Robert May (1994) *Indices and Identity*, MIT Press, Cambridge, MA.

Fox, Danny (2002) "Antecedent-Contained Deletion and the Copy Theory of Movement," *Linguistic Inquiry* 33, 63–96.

Fox, Danny and Johnathan Nissenbaum (1999) "Extraposition and Scope: A Case for Overt QR," *WCCFL* 18, 132–144.

Fox, Danny and David Pesetsky (2003) "Cyclic Linearization and the Typology of Movement," Lecture Notes, Linguistic Society of America Summer Institute, Michigan State University, East Lansing.

Frampton, John (1990) "Parasitic Gaps and the Theory of *Wh*-Chains," *Linguistic Inquiry* 21, 49–77.

Fujita, Koji (2009) "A Prospect for Evolutionary Adequacy: Merge and the Evolution and Development of Human Language," *Biolinguistics* 3, 128–153.

Fukui, Naoki (1993) "Parameters and Optionality," *Linguistic Inquiry* 24, 399–420. [Reprinted

in *Theoretical Comparative Syntax: Studies in Macroparameters*, ed. by Carlos P. Otero, 2006, 69–86, Routledge, New York.]

Fukui, Naoki (2001) "Phrase Structure," *The Handbook of Contemporary Syntactic Theory*, ed. by Mark Baltin and Chris Collins, 374–406, Blackwell, Oxford. [Reprinted in *Theoretical Comparative Syntax: Studies in Macroparameters*, ed. by Carlos P. Otero, 2006, 258–288, Routledge, New York.]

Fukui, Naoki (2006) *Theoretical Comparative Syntax: Studies in Macroparameters*, ed. by Carlos P. Otero, Routledge, New York.

Fukui, Naoki (2008) "Gengo no Kihonenzan o Meguru Oboegaki (A Note concerning Basic Operations in Language)," *Gengokenkyu no Genzai: Keishiki to Imi no Intafeisu* (The State of the Art in Linguistic Research: The Syntax-Semantics Interface), ed. by Yoshiaki Kaneko, Daiko Takahashi, Yoshiki Ogawa, Akira Kikuchi and Etsuro Shima, 1–21, Kaitakusha, Tokyo.

Fukui, Naoki and Margaret Speas (1986) "Specifiers and Projections," *MIT Working Papers in Linguistics* 8, 128–172. [Reprinted in *Theoretical Comparative Syntax: Studies in Macroparameters*, ed. by Carlos P. Otero, 2006, 9–37, Routledge, New York.]

Fukui, Naoki and Yuji Takano (1998) "Symmetry in Syntax: Merge and Demerge," *Journal of East Asian Linguistics* 7, 27–86. [Reprinted in *Theoretical Comparative Syntax: Studies in Macroparameters*, ed. by Carlos P. Otero, 2006, 132–178, Routledge, New York.]

Fukui, Naoki and Yuji Takano (2000) "Nominal Structure: An Extension of the Symmetry Principle," *The Derivation of VO and OV*, ed. by Peter Svenonius, 219–254, John Benjamins, Amsterdam. [Reprinted in *Theoretical Comparative Syntax: Studies in Macroparameters*, ed. by Carlos P. Otero, 2006, 229–257, Routledge, New York.]

Furukawa, Takeshi and Minoru Fukuda (2009) "Kuidokasetsu ni tsuite (On the Vacuous Movement Hypothesis)," *Bulletin of Miyazaki Municipal University Faculty of Humani-*

ties 16, 265–279.

George, Leland M. (1980) *Analogical Generalization in Natural Language Syntax*, Doctoral dissertation, MIT.

Goodall, Grant (1987) *Parallel Structures in Syntax: Coordination, Causatives and Restructuring*, Cambridge University Press, Cambridge.

Gracanin-Yuksek, Martina (2007) *About Sharing*, Doctoral dissertation, MIT.

Grimshaw, Jane B. (1979) "Complement Selection and the Lexicon," *Linguistic Inquiry* 10, 279–326.

Haan, Pieter de (1989) *Postmodifying Clauses in the English Noun Phrase: A Corpus-Based Study*, Rodopi, Amsterdam.

Hale, Kenneth L. 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 L. Hale and Samuel J. Keyser, 53–109, MIT Press, Cambridge, MA.

Halle, Morris and Alec Marantz (1993) "Distributed Morphology and the Pieces of Inflection," *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Kenneth L. Hale and Samuel J. Keyser, 111–176, MIT Press, Cambridge, MA.

Harley, Heidi (2005) "How Do Verbs Get Their Names?: Denominal Verbs, Manner Incorporation, and the Ontology of Verb Roots in English," *The Syntax of Aspect*, ed. by Nomi Erteschik-Shir and Tova Rapoport, 42–64, Oxford University Press, Oxford.

Hirai, Daisuke (2004) "Control Infinitives and Two Types of CP Phrases," *English Linguistics* 21, 241–264.

Hiraiwa, Ken (2005) *Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture*, Doctoral dissertation, MIT.

Hiraiwa, Ken and Adams Bodomo (2008) "Object Sharing as Symmetric Sharing: Predicate

Clefting and Serial Verbs in Dagaare," *Natural Language and Linguistic Theory* 26, 795–832.

Hoji, Hajime (1985) *Logical Form Constraints and Configurational Structure in Japanese*, Doctoral dissertation, University of Washington.

Hornstein, Norbert (1999) "Minimalism and Quantifier Raising," *Working Minimalism*, ed. by Samuel D. Epstein and Norbert Hornstein, 45–75, MIT Press, Cambridge, MA.

Huang, C.-T. James (1982) *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT.

Johnson, Kyle (1985) *The Case for Movement*, Doctoral dissertation, MIT.

Johnson, Kyle (1991) "Object Positions," *Natural Language and Linguistic Theory* 9, 577–636.

Karimi, Simin (1999) "A Note on Parasitic Gaps and Specificity," *Linguistic Inquiry* 30, 704–713.

Kasai, Hironobu (2007) *Multiple Dominance in Syntax*, Doctoral dissertation, Harvard University.

Kasai, Hironobu (2010) "Parasitic Gaps under Multiple Dominance," *English Linguistics* 27, 235–269.

Kayne, Richard S. (1976) "French Relative *Que*," *Current Studies in Romance Linguistics*, ed. by Marta Luján and Fritz Hensey, 255–299, Georgetown University Press, Washington, D.C.

Kayne, Richard S. (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, MA.

Kearney, Kevin (1983) "Governing Categories," ms., University of Connecticut, Storrs.

Kennedy, Christopher (1997) "VP-Deletion and 'Nonparasitic' Gaps," *Linguistic Inquiry* 28, 697–707.

Kim, Soowon and James Lyle (1996) "Parasitic Gaps, Multiple Questions, and VP Ellipsis," *WCCFL* 14, 287–301.

Kiss, Katalin É. (1985) "Parasitic Chains," *The Linguistic Review* 5, 41–74. [Reprinted in *Parasitic Gaps*, ed. by Peter W. Culicover and Paul M. Postal, 2001, 99–124, MIT Press, Cambridge, MA.]

Kitada, Shin-Ichi (2011) "C-to-T Inheritance of Edge Features," *Studies in English Literature* 52, 77–101.

Kitagawa, Yoshihisa (1986) *Subjects in Japanese and English*, Doctoral dissertation, University of Massachusetts, Amherst.

Kitahara, Hisatsugu (1996) "Raising Quantifiers without Quantifier Raising," *Minimal Ideas*, ed. by Werner Abraham, Samuel D. Epstein, Hoskuldor Thrainsson and C. Jan-Wouter Zwart, 189–198, John Benjamins, Philadelphia.

Kitahara, Hisatsugu (2011) "Relations in Minimalism," *English Linguistics* 28, 1–22.

Ko, Heejeong (2007) "Asymmetries in Scrambling and Cyclic Linearization," *Linguistic Inquiry* 38, 49–83.

Koizumi, Masatoshi (1993) "Object Agreement Phrases and the Split VP Hypothesis," *MIT Working Papers in Linguistics* 18, 99–148.

Kratzer, Angelika (1996) "Severing the External Argument from Its Verb," *Phrase Structure and the Lexicon*, ed. by Johan Rooryck and Laurie Zaring, 109–137, Kluwer, Dordrecht.

Kuroda, S.-Y. (1988) "Whether We Agree or Not: A Comparative Syntax of English and Japanese," *Linguisticae Investigationes* 12, 1–47.

Kurogi, Takayoshi (2007) "Kiseikusho Kobun ni okeru Ninka no Chigai (On the Licensing Disparities of Parasitic Gaps)," *JELS* 24, 91–100.

Lapointe, Steven (1980) *A Theory of Grammatical Agreement*, Doctoral dissertation, University of Massachusetts, Amherst.

Larson, Richard and Robert May (1990) "Antecedent Containment or Vacuous Movement: Reply to Baltin," *Linguistic Inquiry* 21, 103–122.

Lasnik, Howard and Mamoru Saito (1992) *Move α: Conditions on Its Application and Output*, MIT Press, Cambridge, MA.

Lasnik, Howard and Juan Uriagereka (1988) *A Course in GB Syntax*, MIT Press, Cambridge, MA.

Lebeaux, David (1988) *Language Acquisition and the Form of the Grammar*, Doctoral dissertation, University of Massachusetts, Amherst.

Lees, Robert B. (1960) *The Grammar of English Nominalization*, Mouton, The Hague.

Lees, Robert B. (1961) "The Constituent Structure of Noun Phrases," *American Speech* 36, 159–168.

Legate, Julie A. (2003) "Some Interface Properties of the Phase," *Linguistic Inquiry* 34, 506–516.

Levin, Beth and M. Rappaport Hovav (1995) *Unaccusativity: At the Syntax-Lexical Semantics Interface*, MIT Press, Cambridge, MA.

Levinson, Lisa (2007) *The Roots of Verbs*, Doctoral dissertation, New York University.

Lyons, John (1977) *Semantics: Volume 2*, Cambridge University Press, Cambridge.

Marantz, Alec (1981) *On the Nature of Grammatical Relations*, Doctoral dissertation, MIT.

Marantz, Alec (2001) "Words and Things," Handout, MIT.

Marantz, Alec (2007) "Phases and Words," *Phases in the Theory of Grammar*, ed. by Sook-Hee Choe, 199–222, Dong In, Seoul.

Marantz, Alec (2013) "Verbal Argument Structure: Events and Participants," *Lingua* 130 152–168.

McCawley, James (1982) "Parentheticals and Discontinuous Constituent Structure," *Linguistic Inquiry* 13, 91–106.

Merchant, Jason (2013) "Voice and Ellipsis," *Linguistic Inquiry* 44, 77–108.

Mikami, Suguru (2010) "The Locative Inversion Construction: Topicalization and the Pro-

nounciation of the Lower Copy," *English Linguistics* 27, 297–328.

Mikami, Suguru (2011) *Parametric Variations in A-Movement between Subject-Prominent Languages and Focus-Prominent Languages*, Doctoral dissertation, University of Tsukuba.

Moore, John (1998) "Turkish Copy-Raising and A-Chain Locality," *Natural Language and Linguistic Theory* 16, 149–189.

Moro, Andrea (2000) *Dynamic Antisymmetry*, MIT Press, Cambridge, MA.

Narita, Hiroki (2010) "The Tension between Explanatory and Biological Adequacy," *Lingua* 120, 1313–1323.

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

Nissenbaum, Johnathan (2000) *Investigations of Covert Phrase Movement*, Doctoral dissertation, MIT.

Nunes, Jairo (2001) "Sideward Movement," *Linguistic Inquiry* 32, 303–344.

Nunes, Jairo (2004) *Linearization of Chains and Sideward Movement*, MIT Press, Cambridge, MA.

Nunes, Jairo and Juan Uriagereka (2000) "Cyclicity and Extraction Domains," *Syntax* 3, 20–43.

Richards, Marc D. (2007) "On Feature Inheritance: An Argument from the Phase Impenetrability Condition," *Linguistic Inquiry* 38, 563–572.

Richards, Marc D. (2011) "Deriving the Edge: What's in a Phase?" *Syntax* 14, 74–95.

Richards, Norvin (2001) *Movement in Language: Interactions and Architectures*, Oxford University Press, Oxford.

van Riemsdijk, Henk (2006) "Grafts Follow from Merge," *Phases of Interpretation*, ed. by Mara Frascarelli, 17–44, Mounton de Gruyter, Berlin.

Ross, John R. (1967) *Constraints on Variables in Syntax*, Doctoral dissertation, MIT.

Runner, Jeffrey (1995) *Noun Phrase Licensing and Interpretation*, Doctoral dissertation, Uni-

versity of Massachusetts, Amherst.

Saito, Mamoru (1985) *Some Asymmetries in Japanese and Their Theoretical Implications*, Doctoral dissertation, MIT.

Sakamoto, Akihiko (2011a) "The Licensing Condition on Parasitic Gaps in English," *Tsukuba English Studies* 29, 193–209.

Sakamoto, Akihiko (2011b) "Eigo Kiseikusho no Ninka Joken ni tsuite (The Licensing Condition on Parasitic Gaps in English)," *JELS* 28, 136–142.

Sakamoto, Akihiko (2012) "Feature Inheritance and Vacuous Movement," *English Linguistics* 29, 316–343.

Sauerland, Urlich (1998) *On the Making and Meaning of Chains*, Doctoral dissertation, MIT.

Shimada, Masaharu (1991) "A Note on Subject-internal Parasitic Gaps," *Tsukuba English Studies* 10, 275–292.

Shimada, Masaharu (1999) "VP-Ellipsis and Parasitic Gaps," *English Linguistics* 16, 145–151.

Shimada, Masaharu (2008) "*Wh*-Movement and Linguistic Theory," (Review Article: *On* Wh-*Movement: Moving on*, by Lisa L.-S. Cheng and Norbert Corver, MIT Press, Cambridge, MA, 2006,) *English Linguistics* 25, 519–540.

Shizawa, Takashi (2011) "Adnominal Conditionals and Their Licensing Conditions," *Tsukuba English Studies* 29, 211–230.

Stowell, Timothy A. (1982) "The Tense of Infinitives," *Linguistic Inquiry* 13, 561–570.

Takahashi, Daiko (1994) *Minimality of Movement*, Doctoral dissertation, University of Connecticut, Storrs.

Takahashi, Shoichi (2006) *Decompositionality and Identity*, Doctoral dissertation, MIT.

Takahashi, Shoichi and Sarah Hulsey (2009) "Wholesale Late Merger: Beyond the A/A-Bar Distinction," *Linguistic Inquiry* 40, 387–426.

Takano, Yuji (1996) *Movement and Parametric Variation in Syntax*, Doctoral dissertation, Uni-

versity of California, Irvine.

Takano, Yuji (2010) "Scrambling and Control," *Linguistic Inquiry* 41, 83–110.

Takeda, Kazue (1999) *Multiple Headed Structures*, Doctoral dissertation, University of California, Irvine.

Taraldsen, Knut T. (1981) "The Theoretical Interpretation of a Class of Marked Extractions," *Theory of Markedness in Generative Grammar, Proceedings of the 1979 GLOW Conference*, 475–516, Scuora Normale Speriore, Pisa.

Tonoike, Shigeo (2013) "Elimination of Inheritance, Counter-Cyclicity and EPP," paper presented at 2013 International Conference on English Linguistics, Korea University and Korea Military Academy.

Tsai, W.-T. Dylan (1994) *On Economizing the Theory of A'-Dependencies*, Doctoral dissertation, MIT.

Ura, Hiroyuki (1998) "Checking, Economy, and Copy-Raising in Igbo," *Linguistic Analysis* 28, 67–88.

Vergnaud, Jean-Roger (1974) *French Relative Clauses*, Doctoral dissertation, MIT.

Watanabe, Akira (1992) "Subjacency and S-Structure Movement of *Wh*-in-situ," *Journal of East Asian Linguistics* 1, 255–291.

Watanabe, Akira (2001) "*Wh*-in-situ Languages," *The Handbook of Contemporary Syntactic Theory*, ed. by Mark Baltin and Chris Collins, 203–225, Blackwell, Oxford.

Wilder, Chris (1999) "Right Node Raising and the LCA," *WCCFL* 18, 586–598.