UNIVERSITY OF CALIFORNIA,

IRVINE

An Asymmetry in the Composition of Phrase Structure and its Consequences

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Social Science

by

Toru Ishii

Dissertation Committee:

Professor Naoki Fukui, Chair

Professor C.-T. James Huang

Professor Lisa Lai-Shen Cheng

1997

TABLE OF CONTENTS

CHAPTER 1
OVERVIEW

1.0 Introduction

This dissertation proposes a new system of the composition of phrase structure. Throughout this dissertation, I assume as its theoretical foundation the minimalist program (MP) proposed by Chomsky (1993) and further developed by Chomsky (1994, 1995, 1996). The discussion to follow presupposes that the reader has basic familiarity with the MP. In this chapter, I will first present a layout of the minimalist assumptions which are necessary for the understanding of the discussion of the following chapters. I will then outline the content of each chapter of this dissertation.

1.1 Basic Notions

The leading idea of the MP is that linguistic principles should be formulated only in terms of notions drawn from the domain of virtual conceptual necessity. This requires that the theory should only refer to notions indispensable for any theory of language. The design of language is therefore "economical" and language is surprisingly "perfect" in this sense. This section presents a set of basic notions of the MP which is necessary for the understanding of the following chapters.

1.1.1 Lexicon and Computational System

In generative grammar, language is taken to be part of the natural world (mind/brain). The human mind/brain provides the language faculty. The language faculty has at least two components, i.e., a cognitive system which stores information and a performance system which accesses that information and uses it in various ways. The cognitive component of the language faculty is called a language (or an I-language). The language is a generative procedure which consists of two components, i.e., a lexicon and a computational system. The lexicon specifies the items that enter into the computational system with their idiosyncratic properties. The computational system generates derivations and structural descriptions (SDs).

Since the language is embedded in two types of performance systems, i.e., articulatory-perceptual (A-P) and conceptual-intensional (C-I), it must provide instructions for these two performance systems to be usable at all. The language therefore generates linguistic expressions (structural descriptions (SDs)) consisting of two interface levels, i.e., PF and LF. PF (the A-P interface) interfaces with the articulatory-perceptual systems of cognition and provides instructions for articulating and perceiving. LF (the C-I interface), on the other hand, interfaces with the conceptual-intensional systems and provides instructions for interpreting, referring, reflecting, and other actions. These two interface levels are required by "bare output conditions" (BOCs), minimal requirements for language to be usable at all. BOCs ensure that at least some of the linguistic expressions generated by the language are "legible" to the external (performance) systems. They are "output conditions" in that they have to do with the properties of the linguistic expressions generated by the language. They are "bare" in that they are imposed from outside the language, not part of the computational system that generates linguistic expressions. Under the MP where everything is within the domain of virtual conceptual necessity, PF and LF, which are required by BOCs and thus conceptually indispensable, are taken to be the only linguistic levels in the language. This is in contrast with the Extended Standard Theory (EST) (see, among others, Chomsky (1972, 1975, 1977, 1980, 1981, 1982, 1986a, 1986b)), where an SD consists of D-structure, S-structure, PF, and LF .

The language, being a derivation-generating procedure, applies to a numeration (N) as its input. An N is a set of pairs (LI, *i*) where LI is a lexical item and *i* is its index, which is taken to be the number of times that LI is selected from the lexicon. Given the N, the language forms a sequence of syntactic objects ($\sigma_1, \sigma_2, ..., \sigma_n$), which are constructed from the N and syntactic objects already formed. The sequence only terminates if σ_n is a pair of PF and LF and every index in the N is reduced to zero. A sequence formed in this way is a derivation.

Each stage in a derivation is characterized as a set of syntactic objects $\Sigma = \{SO_1, SO_2, ..., SO_n\}$ (a set of phrase markers).[1] Operations apply to this set of syntactic objects Σ, yielding a new set of syntactic objects Σ'. In every theory of language, the following two operations are indispensable for constructing a derivation: Select and Merge. The operation Select selects a lexical item LI from an N reducing its index by 1 and introduces it into the derivation as a syntactic object. The syntactic objects formed by distinct applications of Select to an LI are distinguished. Two occurrences of an LI are marked distinct for the computational

[1]See Bobaljik (1995) for the alternative view that computation is the process of definition of (complex) terms and thus each stage in a derivation is characterized not as a set of phrase structures but as an unordered set of terms.

system if they are formed by distinct applications of Select accessing the same LI.[2] The operation Merge takes a pair of syntactic objects (SO_i, SO_j) and replaces them by a newly combined syntactic object SO_k. For instance, suppose that a derivation has reached a stage $\Sigma = \{\alpha, \beta, \delta_i, ..., \delta_n\}$. The application of Merge that forms γ from α and β converts Σ to $\Sigma' = \{\gamma, \delta_i, ..., \delta_n\}$. At any point in a derivation, we may apply Spell-Out, which strips away the elements relevant to PF. After Spell-Out, the computation continues, leading to LF. The parts of the computational system that are only relevant to PF are the PF component. The parts of the computational system that are only relevant to LF are the covert component. The parts of the computational system that are relevant to both PF and LF are the overt component.

1.1.2 Types of Lexical Features

Given these assumptions of the computational system, let us next look at types of lexical features each lexical item carries. First, lexical features are divided based on the accessibility by the computation system. Features accessible by the computational system are formal features. The other lexical features like phonetic and semantic features are not accessible by the computational system. Essentially following Chomsky (1995, 1996), we assume the following four types of formal features:

---

[2]This is a departure from the condition of inclusiveness, which is to be introduced later. As Chomsky (1995) argues, however, this departure seems to be indispensable and rooted in the nature of language.

(1) a. Categorial features

b. ϕ-features

c. Case features

d. Strong features

ϕ-features are bundles of person, number, and gender features. Strong features are the categorial features of functional elements which trigger overt category movement.[3]

Second, lexical features are imposed the following tripartite division by BOCs:

(2) a. Features which are interpreted at the A-P interface

b. Features which are interpreted at the C-I interface

c. Features which are not interpreted at either interface

Features which are interpreted at the A-P interface are phonetic features. Those which are interpreted at the C-I interface are semantic features. Formal features are never interpreted at the A-P interface. There are therefore no formal features which are also phonetic. There are, however, formal features which are also semantic. Such formal features are called

---

[3]It should be noted that our characterization of a strong feature is different from Chomsky's (1993, 1994, 1995, 1996). Under Chomsky's system, exactly like the interpretable/uninterpretable and intrinsic/optional distinctions, the strong/weak distinction is one of the characteristics which crossclassify formal features. Hence, there are formal features which are interpretable as well as strong like the Q-feature of C. Under our system, on the other hand, the expression "strong feature" is used just for sake of convenience to identify the categorial feature of a functional element which triggers overt category movement. Such a categorial feature does not represent the categorial status of the functional element itself but that of the moved category. Hence, it always counts as uninterpretable. T, for instance, has a so called "strong feature", an uninterpretable categorial feature D. This categorial feature triggers overt subject raising to the Spec of TP. We will argue in the next chapter that, strong features, being uninterpretable, must be checked off immediately in accordance with the Immediate Checking Principle (ICP) on uninterpretable formal features (UFFs), which is to be proposed below. So called "strong features" therefore trigger overt category movement. Hence, there is no theoretical notion of strong/weak feature or no strong/weak distinction in our system.

[+ Interpretable] formal features. Formal features which do not receive any interpretations at either of the interface levels and thus belong to (2c) are called [- Interpretable] formal features. For instance, categorial features and the ϕ-features of nouns are [+ Interpretable]. Case features, strong features, and the ϕ-features of verbs, on the other hand, are [- Interpretable]. Chomsky (1995) assumes that Q-features are [+ Interpretable] wherever they may appear. I claim contra Chomsky that the Q-feature of C is [- Interpretable]. Under our analysis, C is not the locus of the property of being interrogative/noninterrogative. Instead, a clause is interpreted as interrogative at LF when C or its specifier position is occupied by a Q-morpheme or *wh*-element, which has a [+ Interpretable] Q-feature. Otherwise, it is interpreted as noninterrogative.

Finally, formal features are classified into intrinsic and optional features. Intrinsic features are explicitly listed in the lexical entry or strictly determined by the listed features while optional features are added later. For instance, categorial features are all intrinsic. While the person and gender features of nouns are intrinsic, their number features are optional. The ϕ-features of verbs, on the other hand, are all optional. The Case features of nouns are optional while the Case features of verbs are intrinsic.

I claim that strong features like the D-feature of T and the Q-feature of C are all intrinsic. The strong D-feature of T is intrinsic. This is because overt subject raising to the Spec of TP takes place in non-finite as well as finite clauses. Let us assume the VP-internal subject hypothesis advocated by, among others, Fukui (1986), Kitagawa (1985, 1986), Koopman and Sportiche (1986), Kuroda (1988), and Pollock (1989). When T is finite, it is clear that overt subject raising takes place from the Spec of VP to the Spec of TP, which is triggered by the strong D-feature of a finite T.[4] As Chomsky (1995) argues, overt subject raising also takes place in nonfinite clauses:

(3) we are likely [t_3 to be asked [t_2 to [t_1 build airplanes]]]

In (3), the raising of *we* from t_1 to t_2 takes place in order to check the strong D-feature of the nonfinite T. Further raising from t_2 to t_3 is also triggered by the strong D-feature of the nonfinite T. Note that since the raising infinitive never assigns Case, it is impossible to claim that these instances of overt movement are triggered by Case features.[5] T therefore always has a strong D-feature whether it is finite or nonfinite. Hence, the strong D-feature of T is intrinsic.

Turning to the strong Q-feature of C, one might claim that it is optional because C may or may not have a strong Q-feature depending on whether it is interrogative or not. I rather argue that C is divided into two types, [+Q]-C and [-Q]-C. Between these two types of C, [+Q]-C has a strong Q-feature as its intrinsic feature while [-Q]-C never has a strong Q-feature. There are languages where these two types of C are realized as phonetically distinct lexical items. Let us look at the following Irish examples:[6]

---

[4]One might argue that overt subject raising in finite clauses is triggered not by the strong D-feature of T but by the Case feature of T. I will later argue, however, that overt raising only takes place when an interpretable feature is attracted. It then follows that the strong D-feature of T triggers overt subject raising, since it attracts the D-feature of the subject, which is an interpretable feature. Since Case features are always uninterpretable, on the other hand, they can never trigger overt raising.

[5]Also see note 4 above.

[6]*L* and *N* mark elements that induce lenition and nasalization respectively.

(4) a. cé **aL** deir siad **aL** chum an t-amhrán sin
who C say they C composed that song
'who do they say wrote that song'

b. deir siad **gurL** chum sé an t-amhrán sin
say they C composed he that song
'they say he wrote that song'

(McCloskey 1979:153)

In (4a), the two occurrences of complementizer *aL* are used. In (4b), on the other hand, the complementizer *gurL*, which is the past form of *goN*, is used. The *goN/aL* alternation is observed in (4).

McCloskey (1979, 1989) claims that *goN* is the complementizer which introduces a declarative clause. The complementizer *aL*, on the other hand, appears when its specifier position is occupied by a *wh*-element during a derivation. It introduces either an interrogative or noninterrogative clause depending on whether the *wh*-element stays in its specifier position at LF. In (4a), the *wh*-element *cé* 'who' first moves to the Spec of the embedded CP and then to the Spec of the matrix CP. Both the embedded and matrix C's are therefore realized as *aL*. In (4b), on the other hand, the embedded C is realized as *gurL*, the past form of *goN*, since there is no instance of wh-movement. The *goN/aL* alternation in Irish can be accounted for if we assume that *goN* is [-Q]-C while *aL* is [+Q]-C. Since the [+Q]-C *aL* has a strong Q-feature as its intrinsic feature, it always triggers movement of a *wh*-element into its specifier position. The [-Q]-C *goN* does not have a strong Q-feature and thus never triggers movement of a *wh*-element into its specifier position. Recall that we are assuming that the Q-feature of C is [- Interpretable].

A clause is interpreted as interrogative at LF when the Spec of its CP is occupied by a *wh*-element, which has a [+ Interpretable] Q-feature. Otherwise, it is interpreted as noninterrogative. Let us furthermore assume following Lasnik and Saito (1992) that the trace in the Spec of CP left by wh-movement is not interpreted as a *wh*-element. It then follows that the clause headed by [+Q]-C is not always interpreted as interrogative. The clause headed by [+Q]-C is interpreted as interrogative at LF only when its specifier position is occupied by a *wh*-element at LF. Otherwise, it is interpreted as noninterrogative. In (4), the embedded clause of (4a) is interpreted as noninterrogative although it is headed by the [+Q]-C *aL*. This is because its specifier position is not occupied by a *wh*-element at LF. On the other hand, the matrix clause of (4a), which is also headed by the [+Q]-C, is interpreted as interrogative, since its specifier position is occupied by the *wh*-element *cé* at LF.

I assume that [+Q]-C and [-Q]-C should be distinguished from each other even in languages like English where they are not phonetically distinct. [+Q]-C has a strong Q-feature as its intrinsic feature while [-Q]-C never has a strong Q-feature. For simplicity, the discussion to follow sticks to the expression that C has a strong Q-feature as its intrinsic feature unless any complication arises, though, to be precise, it is not C but [+Q]-C that has an intrinsic strong Q-feature.

1.1.3 The Theory of Phrase Structure

Chomsky (1995) claims that language meets the condition of inclusiveness. The condition of inclusiveness requires that any structure constructed by a computation from N to LF should only consist of elements present in the lexical items selected for the N. Phrase structure representations should therefore be "bare" in the sense that they exclude anything beyond lexical features and objects constructed from them. It then follows that no indices or bar levels in the sense of X-bar theory are allowed. Phrase structures are set-theoretic objects recursively constructed by Merge.[7] The syntactic objects Merge is applied to are of the following types:

(5) a. lexical items

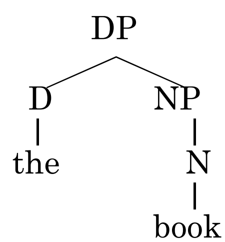
b. K = {γ, {α, β}}, where α, β are objects and γ is the label of K.

(Chomsky 1995:243)

Lexical items are complexes of features, which are listed in the lexicon. (5b) is the recursive step. α and β are constituents of K. γ is the label of K, which is the zero-level projection of the head of K.

Let us consider *the book* as an example. Under the standard X-bar theory, it was assigned the following structure:

(6)



Under the theory of "bare" phrase structure, on the other hand, *the book* is assigned the following structure:

---

[7]Lasnik and Kupin (1977) also argues that phrase markers should be defined in a set-theoretic way.

(7) {the, {the, book}}

the book

Structure (7) has *the* and *book* as its constituents. Of the two constituents of structure (7), its head is *the*. Structure (7) is therefore assigned *the* as its label.

The functioning elements in structures are called terms. The notion of term is defined as below:

(8) For any structure K,

a. K is a term of K.

b. If L is a term of K, then the members of the members of L are terms of K.

(Chomsky 1995:247)

In (7), for example, *{the, {the, book}}*, *the*, and *book* are terms. Terms correspond to nodes of the informal representations. For expository purposes, the discussion to follow sticks to the traditional way of representing phrase structures unless any complication arises.

1.1.4 Structural Relations

Under the minimalist conception of language, no structural relations are allowed to be invoked other than those required by BOCs like adjacency at PF and scope at LF and those induced in a natural way by the derivation itself like local relations to the head in its minimal domain. This would exclude EST notions like government, proper government, and binding relations which are internal to the derivation of linguistic expressions (SDs). The EST analyses which make use of these internal notions must therefore be reformulated under the MP.

1.1.5 Full Interpretation and Economy

The MP claims that each linguistic expression (SD) is "the optimal realization of the interface conditions, where "optimality" is determined by the economy conditions of UG" (Chomsky 1993: 4).

The condition of Full Interpretation (FI) requires that linguistic expressions, each a pair of PF and LF, must solely consist of legitimate objects that can receive an interpretation at the relevant interface level. A derivation converges at PF if the PF interface only consists of legitimate objects. Otherwise, it crashes at PF. A derivation converges at LF if the LF interface only consists of legitimate objects. Otherwise, it crashes at LF. A derivation converges if it converges at both interfaces. A linguistic expression must be a PF-LF pair formed by a convergent derivation.

Legitimate objects at PF consist of elements that are interpreted in terms of articulatory and perceptual mechanisms in a language-invariant manner. Regarding LF, the following elements are legitimate, each a chain CH = (α_1, ..., α_n) (possibly, a one-membered chain), given the traditional A/A'-distinction:

(9) a. Arguments: each element is in an A-position.

b. Adjuncts: each element is in an A'-position.

c. Lexical elements: each element is in an X^0 position.

d. Predicates, possibly predicate chains if there is predicate raising, VP-movement in overt syntax, and other cases.

e. Operator-variable constructions, each a chain (α_1, α_2), where the operator α_1 is in an A'-position and the variable α_2 is in an A-position.

(adapted from Chomsky (1991a, 1993))

A linguistic expression must also meet the economy conditions. The economy conditions require that the derivation of a linguistic expression should be "optimal." In other words, it must be the "most economical" (the "least costly") derivation that forms the linguistic expression, where the "cost" of a derivation is defined by UG. Hence, "less economical" derivations are blocked by the "optimal" derivation even if they converge. A derivation selected by the economy conditions is an admissible derivation.

The language therefore generates the following three sets of derivations, i.e., the set D of derivations, a subset D_C of convergent derivations of D, and a subset D_A of admissible derivations of D. Chomsky (1993, 1995) claims that the economy conditions only compare convergent derivations. It then follows that D_A is a subset of D_C.

1.1.6 Attract/Move-F

Essentially following Chomsky (1995, 1996), we assume the operation Attract/Move-F (F a feature), which is a reinterpretation of the operation of movement.[8] According to the notion of Attract/Move-F, what is raised is not a category but a feature. Chomsky argues that what is raised should be just F unless it would result in a crashed derivation. In order to ensure this, he proposes the "no extra baggage" condition, which is one of the economy conditions:

(10) "No Extra Baggage" Condition

F carries along just enough material. z

(adapted from Chomsky 1995:262)

According to the "no extra baggage" condition, the derivation that raises just F should be chosen as "optimal" unless it would violate the FI.

In the overt component, however, a category, not just F, is raised to the target. In other words, a "generalized pied-piping" is always involved in the overt component. Chomsky (1995, 1996) argues that this displacement property follows from the "no extra baggage" condition and BOCs. Chomsky (1995) argues that if only F were raised to the target in the overt component, features of a single lexical item would be scattered. Only F would be in the checking domain of the target, but all the other features would remain in- situ. There is, however, a PF requirement that features of a single lexical item must be within a single X^0. A derivation with such scattered features violates the FI at PF and therefore crashes at that level. Hence, in the overt component, an "extra baggage" is required for PF-convergence; the whole category, not just F, is

[8]Chomsky (1995, 1996) totally eliminates the notion of Move, arguing that the traditional notion of movement should be reinterpreted as Attract-F. Under his view, the locus of the notion is totally shifted from the moved element to the target. I will argue in chapter 5, however, that the notion of Move is still needed to account for the distribution of *wh*-elements in-situ. The traditional operation of movement therefore should not be reinterpreted as Attract-F but as Attract/Move-F. See Takeda (1997) for further discussion of this subject.

raised to the target. Chomsky (1996), on the other hand, claims that a feature chain cannot be interpreted at the PF interface. This requires that in the overt component, the whole category should be raised, forming a category chain. If only F were raised to the target in the overt component, it would yield a feature chain at PF. Since a feature chain is an illegitimate object at PF, the derivation would crash at that level. Hence, an "extra baggage" is required in the overt component for PF-convergence.

Under either approach, overt category movement is forced by the BOCs which apply at the PF interface. In the covert component, on the other hand, the BOCs on the PF interface are irrelevant. According to the "no extra baggage" condition, therefore, only F rather than the whole category should be raised to the target.[9]

---

[9]When the feature F of a lexical item raises without pied-piping of the lexical item or any larger category, as always in the covert component, Chomsky (1995) assumes that F automatically carries along FF (LI), the set of formal features of the lexical item:

(i) Move F "carries along" FF[LI].

He argues that this much pied-piping takes place automatically and thus need not be required by any extraneous factors.

He presents the following binding and control facts as empirical evidence in favor of this view:

(ii) a. the DA [proved [the defendants to be guilty] during each other's trials]
b. *the DA [proved [that the defendants are guilty] during each other's trials]

(iii) a. there arrived three men (last night) [without PRO identifying themselves]
b. *I met three men (last night) [without PRO identifying themselves]

(Chomsky 1995:272-274)

In (iia), the Case and ϕ-features of *the defendants* are attracted by the matrix V. In (iiia), the Case and ϕ-features of *there* are attracted by the matrix T. He argues that these attracted features in (ii-iiia) carry along FF (LI). Since FF (LI) includes an A-position property, which has the ability to serve as a binder or controller, the contrast between (ii-iiia) and (ii-iiib) follows.

This empirical argument, however, is inconclusive. First, it is not clear which formal feature counts as having the A-position property. Second, under the MP, Although we follow Chomsky in claiming that the displacement property is forced by the "no extra baggage" condition and BOCs, I will later argue that overt category movement is forced by the BOCs on the LF interface not by those on the PF interface.

Following Chomsky (1993, 1995), we assume the copy theory of movement, where the trace left behind is the copy of the moved element. Under this approach, movement makes a term copied and introduced into the syntactic object a second time. Since we distinguish among distinct selections of a single lexical item from the lexicon, this is the only case where more than one terms can be identical in constitution. In other words, such pairs that consist of identical terms only arise through movement. I claim following Chomsky (1995) that although these terms are identical in constitution, they are distinguished from each other in terms of the context where they appear. For example, suppose that we have constructed (11a), where β is the head of K. From (11a), we derive (11b) by raising α, targeting K:

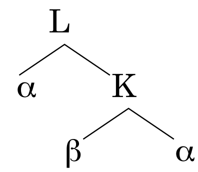
(11) a. K

β α

---

binding and control relations could be outside the domain of the computational system, as suggested by Chomsky (1993). Under such a view, binding and control theories do not regulate the relations between arguments at the LF interface. Rather, the LF interface only provides instructions for the interpretive version of binding and control theories. Then, it is conceivable that the attracted Case and ϕ-features in (ii-iiia) are sufficient to provide appropriate instructions for the establishment of binding and control relations.

Furthermore, the FF (LI) raising analysis has a conceptual problem. It does not explain why F always carries along FF (LI). Note that FF (LI), which is an "extra baggage," is not carried along for convergence. Such an "extra baggage" should be banned by the "no extra baggage" condition.

b. 

The operation that raises α introduces α a second time into the syntactic object. The element α appears twice in the syntactic object, in its initial and raised positions. Although these two terms are identical in constitution, they are positionally distinct. α in the initial position is identified as the pair <α, β>, where β is the co-constituent of the original α. α in the raised position, on the other hand, is identified as the pair <α, K>, where K is the co-constituent of the raised α. I claim that chain formation applies in the LF-component, forming a chain which consists of terms that are identical in constitution but positionally distinct. Hence, in (11), the chain CH = (α, α), or more precisely CH = (<α, K>, <α, β>), is formed.

1.2 Organization of the Dissertation

This dissertation proposes a new system of the composition of phrase structure in which there is an asymmetry with respect to merger. In the system proposed here, the terms required by uninterpretable formal features (UFFs) are merged cyclically whereas those not required by any UFFs are merged postcyclically. I will propose the Immediate Checking Principle (ICP) on UFFs and the Earliness Principle (EP) on Select, arguing that these two principles are conceptually attractive in that they contribute to the reduction of globality in the theory of language. It is shown that it is the need for these two principles to be satisfied which gives rise to the asymmetry in the composition of phrase structure. I will argue that the asymmetry receives strong support from a wide range of empirical facts. The empirical arguments of this dissertation constitute evidence in favor of the view that language is essentially derivational in character rather than in the representational mode. This dissertation also supports the language design that language is fundamentally global and thus its corresponding optimization problem belongs to Class NP, but there are language-specific computational devices which reduce its fundamental globality to local properties.

The organization of the rest of this dissertation is as follows. Chapter 2 presents conceptual arguments for our theory of phrase structure. I will introduce the ICP on UFFs and the EP on Select, both of which count as language-specific computational devices. The ICP requires that UFFs should be checked immediately when they become accessible to a computation. I will present conceptual arguments for the ICP through demonstrating that the ICP captures the D-structure and S-structure properties in the Extended Standard Theory (EST) in a local fashion and thus contributes to the reduction of globality in the theory of language. The EP states that lexical items must be selected from an N as early as possible. It is shown that the EP reduces fundamental globality induced by a condition on an N to local properties. I will argue that the ICP coupled with the EP guarantees the asymmetry in the composition of phrase structure.

The remaining chapters of this dissertation explore empirical justification for the asymmetry with the composition of phrase structure, which is required for the satisfaction of the ICP and the EP. I will argue that the asymmetry receives strong empirical support from a wide range of facts pertaining to movement constraints, scrambling in Japanese, the distribution of *wh*-elements in-situ, and reconstruction effects. It is also shown that these empirical facts lend support for the derivational view of the theory of language, since their analyses crucially make use of information which is available at an intermediate stage of a derivation but later "wiped-out" by an operation before the output representation.

Among these empirical facts, chapter 3 considers locality restrictions on feature-driven A'-movement, specifically the "domain barrier" effects, i.e., the Complex NP Constraint, the Adjunct Condition, the Subject Condition, and the non-bridge verb condition, and the ban against feature-driven extraction out of phrases which have undergone feature-driven A'-movement. I will argue that these locality conditions, which have not been given any principled account under the MP, can be accounted for by our theory of phrase structure. It is shown that our analysis diverges from all previous analyses in claiming that the locality conditions should not follow from restrictions on movement but from restrictions on merger.

In chapter 4, I will investigate scrambling in Japanese. It is shown that unlike English overt wh-movement and topicalization, Japanese scrambling does not obey any "domain barriers." I will argue that this asymmetry between the two types of movement with the "domain barrier" effects straightforwardly follows from our theory of phrase structure if we assume following Fukui (1993a) and Fukui and Saito (1996) that Japanese scrambling is not feature-driven. Since this asymmetry only follows from our locality theory but not from previous ones, it constitutes another empirical support in favor of our analysis. It is also pointed out that scrambling is not totally immune from any locality restrictions. I will argue that locality restrictions on scrambling should not be attributed to the "domain barriers" but to an A-over-A condition which applies in the PF-component.

Chapter 5 considers the distribution of *wh*-elements in-situ. It is pointed out that there are several asymmetries concerning the distribution of *wh*-elements in-situ which have not been given any principled account under the MP. First, unlike overt wh-movement, *wh*-arguments in-situ never exhibit any "domain barrier" effects. Second, like overt wh-movement and unlike *wh*-arguments in-situ, *wh*-adjuncts in-situ are constrained by the "domain barriers." Third, *wh*-arguments in-situ are constrained by the Wh-island Constraint in Japanese-type languages but not in English-type languages. I will propose that the Q-feature of a *wh*-element in-situ should undergo "overt" movement to an interrogative C in Japanese-type languages. It is shown that our theory of phrase structure coupled with "overt" Q-feature movement gives us a minimalist account of these asymmetries.

Chapter 6 investigates reconstruction effects with Condition C of the binding theory, variable binding, and the interpretation of *each other*. I will argue that reconstruction facts can be accounted for by our theory of phrase structure coupled with the assumption that binding relations are established at LF.

CHAPTER 2
CONCEPTUAL ARGUMENTS FOR AN ASYMMETRY IN THE COMPOSITION OF PHRASE STRUCTURE

2.0 Introduction

This chapter presents conceptual arguments for a theory of the composition of phrase structure which claims that there is an asymmetry with respect to merger. I will propose the Immediate Checking Principle (ICP) on uninterpretable formal features (UFFs) and the Earliness Principle (EP) on Select. The ICP on UFFs states that UFFs must be checked immediately when they become accessible to a computation. The EP on Select states that lexical items must be selected from a numeration N as early as possible. I will argue the ICP and the EP are language-specific computational devices which contribute to the reduction of globality in the theory of language. It is shown that it is the need for satisfaction of these two principles that gives rise to the asymmetry.

Section 1 considers problems of computational complexity in the theory of language. It is shown that globality, with its "look-ahead" or "look-back" properties that necessarily induce computational complexity, should be reduced to local properties. Section 2 presents Chomsky's (1993) attempt to eliminate D-structure and S-structure. Chomsky (1993) argues that D-structure and S-structure properties in the EST should be reformulated as conditions on the interface levels. It is shown that Chomsky's (1993) interface condition approach contributes to the elimination of these two theory-internal linguistic levels and thus counts as a step toward the goal of the MP. Section 3 argues against Chomsky's (1993) approach to the elimination of D-structure and S-structure. It is shown that Chomsky's (1993) approach raises a serious conceptual problem in that it sneaks in an element of globality into the theory of language. I will present Chomsky's (1995) derivational interpretation of strong features, which is intended to capture the displacement property, one of S-structure properties in the EST. It is shown that Chomsky's (1995) derivational constraint approach induces less globality than Chomsky's (1993) interface condition approach. Section 4 proposes the ICP on UFFs. The ICP reformulates D-structure and S-structure properties as constraints which apply throughout derivations. I will argue that our ICP approach is conceptually more desirable than Chomsky's (1993, 1995) in that the former captures D-structure and S-structure properties in a local fashion while the latter does not. Section 5 investigates consequences of the ICP for the theory of language. Section 6 discusses a consequence of the ICP concerning the composition of phrase structure. I propose the EP on Select, which reduces the fundamental globality induced by a condition on an N to local properties. It is shown that if we conform to the ICP coupled with the EP during derivations, the terms required by UFFs are merged cyclically whereas those not required by any UFFs are merged postcyclically. It then follows that arguments, which are required by UFFs, are merged cyclically. On the other hand, typical adjuncts, which are not required by any UFFs, are merged postcyclically. Section 7 makes concluding remarks.

2.1 Computational Complexity and the Theory of Language

Under the MP where the BOC-driven optimal design of language is assumed, each linguistic expression (SD) is the optimal realization of BOCs. The language therefore can be regarded as a procedure of finding a solution for the problem of optimization. The problem of optimization has been extensively studied in the field of theoretical computer science. Optimization problems can be classified in various ways. One classification, called computational complexity, is based on the amount of time, space, or other resources needed to solve a computational problem. The theory of computational complexity classifies computational problems into general complexity classes. Some computational problems are not only decidable but also feasible. Other computational problems, though decidable and thus computationally solvable in principle, are not solvable in practice because their solution requires an inordinate amount of time, memory, or other resources. In the development of generative grammar, we have come to the stage at which we can seriously ask which complexity class human language belongs to. In other words, the MP makes it possible to bring the problem of computational complexity in the theory of language to the research agenda for the first time in the history of generative grammar. This section considers the relation between the theory of computational complexity and human language. The discussion to follow, especially the implication of computational complexity for the theory of language, is largely based on Fukui (1996).

2.1.1 The Theory of Computational Complexity

This subsection overviews the theory of computational complexity. We must admit that our discussion here is quite brief and informal. To inquire further into the matter would lead us into that specialized area of computational complexity, and such a digression would undoubtedly obscure the outline of our argument. The reader should refer to works such as Hopcroft and Ullman (1979) and Johnson (1990) for more precise explanations.

The theory of computational complexity investigates time, memory, or other resources required for solving computational problems. Complexity theory classifies the problems which are decidable and thus computationally solvable in principle into general complexity classes. Among those complexity classes, those based on time are relevant to the present discussion. It has been claimed that when processing n data items, the solutions requiring no more than some polynomial involving n may be feasible. Those requiring more steps than can be described by any polynomial, on the other hand, are not feasible, though computationally solvable in principle.

If a problem has some solution where the number of steps to process n data items is no more than some polynomial involving n, the problem is defined to be in Class P. The problems in Class P are considered to have a feasible solution. In other words, they are computationally tractable. There are, however, problems whose solutions all require more than a polynomial amount of work. Those problems, which require an exponential or factorial amount of work or worse, are defined to be in Class NP. The problems in Class NP do not have any feasible solution. In other words, they are computationally intractable. The next subsection considers the relation between these two complexity classes and the theory of language.

2.1.2 Computational Complexity in the Theory of Language

Before turning to an examination of computational complexity in the MP, let us first define the notions of globality and locality in the theory of language.[1] Suppose that we come to a stage Σ in a derivation D where we have an option of applying an operation OP. Suppose further that the decision about whether OP applies to Σ is made by a condition C. Then the notions of globality and locality are defined as follows:

(1) a. C is local if it can determine whether to apply OP or not only on the basis of information available in Σ.

b. C is global if it cannot determine whether to apply OP or not only on the basis of information available in Σ.

According to the definitions in (1), the crucial difference between global and local conditions resides in the fact that while the former has the "look-ahead" or "look-back" property, the latter does not. In other words, global conditions require us to make reference to other entities than Σ in D. Some global conditions require us to make reference to Σ and Σ' that is in D or Σ' that is in another derivation D'. For example, the principle of Procrastinate proposed by Chomsky (1993) belongs to this type, since it refers to Σ and PF. Other global conditions require us to make reference to more than one derivations and inspect each of them as a whole. The shortest derivation requirement advocated by, among others, Chomsky (1991a, 1993), Epstein (1992), and Kitahara (1995, 1997), is an instance of

---

[1]For a discussion of this subject, see, among others, Chomsky (1995, 1996), Collins (1997), and Fukui (1996).

global conditions of this type, since it compares the number of steps involved in more than one derivations.

Returning to computational complexity in the MP, let us consider the relation between the global/local nature of conditions and computational complexity. We take two conditions proposed in the literature as examples, Greed and Suicidal Greed. Let us first consider the principle of Greed proposed by Chomsky (1994):

(2) Move raises α to a position β only if morphological properties of α itself would not otherwise be satisfied in the derivation.

(Chomsky 1994:14)

The principle of Greed (2) has the "look-ahead" property. Accordingly, it is global in nature. In order to decide whether to apply Move to α at a stage Σ in a derivation D, we have to know whether morphological properties of α would be satisfied at Σ' that is in another derivation D' where α does not move.

Let us next consider Suicidal Greed proposed by Chomsky (1996), a modification of the principle of Greed , which is part of the definition of Attract/Move:

(3) An uninterpretable formal feature (UFF) in the extended lexical item (ELI) seeks the closest matching feature F in its c-command domain and attaches it to ELI, UFF then erasing if the match is successful (where ELI is an object formed from a lexical item by attaching other features to it).

(adapted from Chomsky 1996:9)

Unlike the principle of Greed, the principle of Suicidal Greed is local in nature. Suicidal Greed enables us to decide whether to attract F at a stage Σ in a derivation D only on the basis of information available at Σ.

Considering these two principles from the viewpoint of the theory of computational complexity, let us first investigate the principle of Greed, which is a global condition. Its corresponding optimization problem is computationally intractable. To be specific, let us consider the following optimization problem. At a stage Σ in a derivation D, we have three elements, α, β, and γ, which have options of moving to X, Y, and Z, respectively. In order to solve this problem, 2^3 derivations must be inspected to see whether they converge or not. More generally, for *n* elements which have an option of movement, 2^n derivations must be reviewed. If the review of one derivation requires one step of work, this optimization problem requires 2^n steps of work. Since this optimization problem only has a solution which requires an exponential amount of work, it belongs to Class NP. Hence, it is computationally intractable.

Turning to the principle of Suicidal Greed, its corresponding optimization problem is computationally tractable. To be specific, let us consider the following situation. At a stage Σ in a derivation D, we come across a UFF. Suppose that there are *m* formal features within the c-command domain of the UFF. We scan all those formal features to see whether they match the UFF. Suppose that we have found *l* formal features which match the UFF within the c-command domain of the latter. Then, we have to inspect which formal feature is closest to the UFF. Let us assume following Chomsky (1995) that the notion of "closeness" is defined in terms of the notion of c-command. Since c-command is a transitive relation, for *l* features, we have to inspect the c-command relation between two features *l-1* times to decide which feature is closest to the UFF. If the inspection of a c-commanding feature and that of c-command relation between two elements each requires one step of work, this problem requires $m + (l - 1)$ steps of work. More generally, for n UFFs, this optimization problem requires $n\ (m + (l - 1))$ steps of work. Since this optimization problem has a solution where the number of steps to process n items is no more than a polynomial involving n, it belongs to Class P. Hence, it is computationally tractable. From the above discussion, we can conclude (4) concerning the relation between the global/local nature of conditions and computational complexity:

(4) Global conditions necessarily induce computational intractability while local conditions do not.

Before we leave this subsection, it is important to consider the relation between the problem of computational complexity and the design of language. Chomsky (1996) claims that the problem of computational complexity only arises in the theory of language as far as the following assumptions are supported:

(5) a. There is an empirical difference between "derivational" and "representational" interpretations of the recursive procedure constituting the I-language.

b. Language uses the derivational approach.

c. Considerations of computational complexity matter for a cognitive system (a "competence system," in the technical sense of this term).

(Chomsky 1996:10)

Let us first consider (5a) and (5b). A representational approach is the one which takes the recursive procedure to be nothing more than a convention for enumerating a set of linguistic expressions (SDs). Under the representational approach, rules are applied freely, but illegitimate outputs are excluded by output conditions as ill-formed. Note that under the representational approach where rules are allowed to apply optionally, we do not have to make the decision about whether to apply an operation during a derivation. If we make a wrong decision about the application of an operation during a derivation, its output representation is simply ruled out by an output condition. Hence, the problem of computational complexity never arises.[2]

A derivational approach, on the other hand, takes the recursive procedure literally, forming linguistic expressions (SDs) step-by-step by applying operations to features. At each step of a derivation, we have to make the decision about whether to apply an operation. Hence, the problem of computational complexity may arise given that a computation is sequential, not parallel. Recall that the MP which this thesis adopts as its theoretical foundation is derivational in nature. This is because the MP does not allow any superfluous step in a derivation. Operations are applied whenever necessary, not otherwise. At no point of a derivation, will there be an optional application of an operation. The problem of computational complexity therefore may arise in the MP.

Although it is typically possible to recode one approach in terms of the other, these two approaches are still empirically different. Chomsky (1995) argues that there is evidence which suggests that the derivational approach is on the right track.[3] Chomsky discusses the head movement constraint (HMC), which was first proposed by Travis (1984). The HMC states that head movement cannot pass over the closest c-commanding head, excluding examples like the following:

---

[2] See, among others, Brody (1995), Cinque (1990), Koster (1978a, 1986), and Rizzi (1986, 1990) for arguments in favor of the representational view.

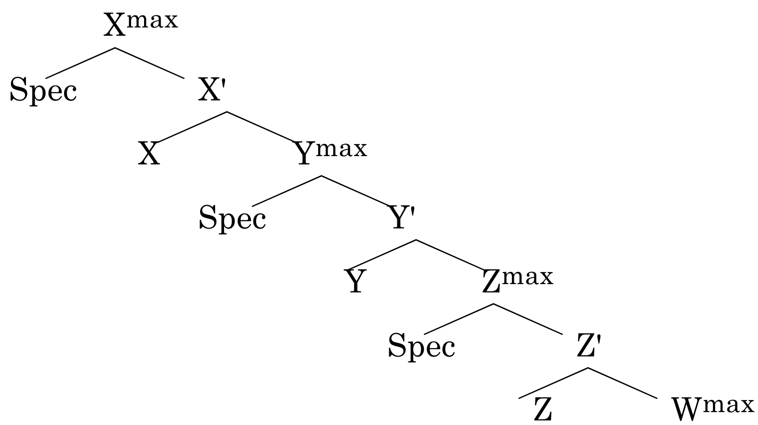
[3] See Chomsky and Lasnik (1993) for further discussion of this subject.

(6) *read John will *t* the book

In (6), *read* raises to the clause-initial position, crossing over *will*. Hence, (6) is excluded as illegitimate by the HMC. Chomsky argues that the HMC imposed on the step-by-step computation is not always reflected on the output representation.[4]

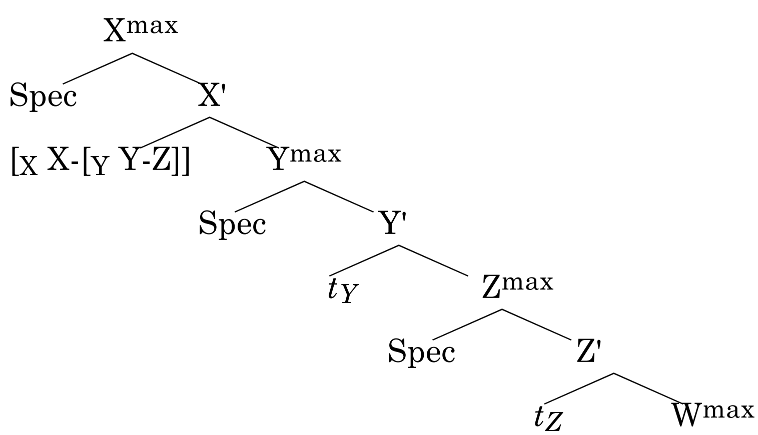
Let us consider the following structure as an example:

(7)



Suppose that Z adjoins to Y, leaving the trace of Z, and the amalgamated form [$_Y$ Y-Z] further adjoins to X, leaving the trace of Y:

(8)



---

[4]As argued by Lasnik (1994) and Robert (1994), however, Chomsky's (1995) argument for the derivational approach is inconclusive. In V-to-C constructions like (6), there is no feature of either V or C that is satisfied by the raising of V to C itself. Rather, since only a finite verb raises to C, Tense must be involved. Hence, it is possible to claim that the raising of *read* to C in (6) is excluded by the economy condition which bans a superfluous step in a derivation without any recourse to the HMC.

In the output representation (8), the chain (Z, t_Z) violates the HMC although each head movement obeys the HMC derivationally. In other words, the locality property of the adjunction of Z to Y is obscured in the output representation by being "wiped out" by the later operation. This dissertation agrees with Chomsky that there is an empirical difference between derivational and representational approaches and the former approach is on the right track, presenting further empirical evidence in support of the derivational interpretation of the recursive procedure.

Turning to (5c), suppose that considerations of computational complexity matter for fundamental aspects of language. It would then follow that the problem of computational complexity arises in the theory of language given the derivational interpretation of the recursive procedure. Hence, this view would require that global conditions, which necessarily induce computational intractability, should be reduced to local conditions, whose corresponding optimization problems may be feasible. Collins (1997) takes this view, making an attempt to reformulate global economy conditions as local economy conditions.

This dissertation, however, does not share the above view. I argue that considerations of computational complexity do not matter for fundamental aspects of language. Globality is one of the fundamental properties of language although it would necessarily induce "exponential blow-up" in construction and evaluation of derivations. As Chomsky (1991a, 1991b, 1993, 1994, 1995, 1996) argues, since there is no a priori reason to suppose that language is "usable" or "conducive to efficient use," language can be fundamentally computationally intractable due to its fundamental global properties. Its corresponding optimization problem therefore belongs to Class NP. Language, however, is usable in practice.

I argue following Chomsky that for purposes of normal life, usability of language is facilitated by what Chomsky calls "computational tricks," the biological counterparts of "heuristic algorithms," which enable us to obtain approximate solutions to the computationally intractable optimization problems. Such language-specific computational devices reduce fundamental globality to local properties, enabling language to be used in practice. To recapitulate, although language yields computationally intractable problems, they can often be overcome by "heuristic algorithms" ("computational tricks"). Only when such language-specific computational devices are available, expressions can be used. Unusable parts of language are simply not used. Hence, although considerations of computational complexity do not matter for fundamental aspects of language, they do matter for usable parts of language. For usable parts of the language, therefore, the problem of computational complexity arises given the derivational interpretation of the recursive procedure. Global conditions, which necessarily induce computational intractability, should be reduced to local "heuristic algorithms" ("computational tricks") for usable parts of language.

This dissertation presents empirical arguments for this view, proposing the ICP on UFFs and the EP on Select. I will argue that these principles serve as language-specific computational devices which reduce the globality induced by interface conditions on UFFs and numerations Ns to local properties, facilitating usability of language in practice. It is shown that these principles lead to an asymmetry in the composition of phrase structure, which receives strong empirical support from a wide range of facts.

The following sections are devoted to showing that the ICP on UFFs and the EP on Select serve as local "heuristic algorithms" ("computational tricks") for computationally intractable problems. I will first consider the ICP, investigating minimalist approaches to the elimination of the two theory-internal linguistic levels, i.e., D-structure and S-structure.

2.2 Chomsky's (1993) Interface Condition Approach

This section reviews Chomsky's (1993) attempt to reduce D-structure and S-structure properties in the EST to interface conditions. As mentioned in the previous chapter, under the MP where everything is within the domain of virtual conceptual necessity, an SD consists not of the traditional four levels, i.e. D-structure, S-structure, PF, and LF, but rather only of the latter two interface levels. It then follows that every principle which constrains a derivation applies either at the interface levels or at every step of a derivation where it is relevant. The principles and conditions which were assumed to apply at D-structure or S-structure in the EST now have to be captured in a different way. They have to be reformulated either as conditions on the interface levels or as constraints which apply throughout derivations. I call the former the interface condition approach and the latter the derivational constraint approach. Apart from the economy conditions, Chomsky (1993) pursues the interface condition approach, namely, reformulating the principles and conditions which were assumed to apply at D-structure and S-structure in the EST as conditions on the interface levels.

2.2.1 D-structure Properties

Let us first consider D-structure properties in the EST. Within the framework of the EST, D-structure is a representation of GF-θ, i.e., a pure representation of θ-structure. It is created by the operation SATISFY, which selects an array of items from the lexicon and presents it in a format satisfying the conditions of the X-bar theory. SATISFY is assumed to be an "all-at-once" operation in that all items that function at LF are drawn from the lexicon before a computation proceeds. Among the principles and conditions which were assumed to apply at D-structure is the θ-criterion, which states that each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument. The θ-criterion itself is a criterion of adequacy for LF. When coupled with the Projection Principle, however, the θ-criterion virtually applies at D-structure as well. This is because the Projection Principle requires that every syntactic representation, i.e., D-structure, S-structure, and LF, should be a projection of θ-structure.

Chomsky (1993) argues that there are conceptual and empirical problems with assuming the θ-criterion and the Projection Principle as D-structure conditions. Let us first look at their conceptual problem. Within the EST framework, the θ-criterion and the Projection Principle are required to apply at D-structure in order to ensure that D-structure has the basic properties of LF. At LF, these conditions are trivial, since if they are not met, the linguistic expression receives some deviant interpretation at that interface level. At D-structure, on the other hand, these conditions are only needed to make the EST picture coherent. If we abandon the EST picture, then these D-structure conditions lose their roles. Hence, they are dubious on conceptual grounds.

Among empirical problems with assuming the θ-criterion and the Projection Principle as D-structure conditions, let us look at the complex adjectival construction as an example:

(9) the man who is reading a book is easy to please

According to Chomsky (1981, 1986a, 1986b), (9) is assigned S-structure representation (10):

(10) [the man who is reading a book] is easy [***Op*** [PRO to please ***t***]]

In (10), *the man who is reading a book* is directly inserted in its surface position. The adjective *easy* takes as its complement a clause which has undergone an empty operator movement. As argued by, among others, Chomsky (1977, 1981, 1986a, 1986b), this empty operator movement analysis is supported by the bounding condition effects.

Under the empty operator movement analysis of the complex adjectival construction, we are led to assume that *the man who is reading a book* in (9) occupies a θ-position, since it is directly inserted in its surface position. There is, however, evidence to suggest that the subject position of the adjective *easy* is not a θ-position:

(11) it is easy to please the man who is reading a book

In (11), the expletive *it* appears in the subject position of the adjective *easy*, which suggests that the subject position of the adjective *easy* is a non-θ-position. Here we have a paradoxical situation unless we assume a dual lexical entry of the adjective *easy*. If we consider (11), where the expletive *it* appears in the subject position of the adjective *easy*, the subject position should be empty at D-structure in accordance with the θ-criterion and the Projection Principle. If we consider (10), where *the man who is reading a book*, being an argument, appears in the subject position of the adjective *easy*, the θ-criterion and the Projection Principle requires that the subject position should be filled at D-structure. This paradoxical situation suggests that assuming the θ-criterion and the Projection Principle as D-structure conditions should be on the wrong track.

These conceptual and empirical problems have led Chomsky (1993) to give up assuming the θ-criterion and the Projection Principle as D-structure conditions. The Projection Principle is totally eliminated while the θ-criterion is reformulated as a condition which only applies at LF. Hence, SATISFY does not have the property of "all-at-once" any more. Instead, a lexical item is selected from the lexicon freely at any point as a computation proceeds.

2.2.2 S-structure Properties

Let us turn to S-structure properties in the EST. Among principles and conditions which were assumed to apply at S-structure in the EST, let us consider structural Case assignment, which was instrumental in explaining the displacement property regarding NP-movement.

It was assumed in the EST that structural Case should be assigned at S-structure. The arguments for assuming structural Case assignment as a S-structure property were as follows. In some languages, Case is morphologically realized while in others it is not. It had been assumed, however, that Case should be assigned in a uniform way whether it is morphologically realized or not. Hence, Case features must appear at PF. In Chomsky (1981, 1986b), the Case Filter, which applies in the PF-component, was intended to capture this fact:

(12) *NP if NP has phonetic content and has no abstract Case.

(adapted from Chomsky 1981:49)

Even in languages where Case is not morphologically realized, Case was assumed to be assigned at S-structure though it is abstract.

Furthermore, it was claimed that Case features must be "visible" at LF (see, among others, Aoun (1985a) and Chomsky (1981, 1986b)). Aoun and Chomsky propose the visibility condition on θ-role assignment, which informally states that only those elements that are assigned Case features are "visible" and thus able to retain θ-roles at LF. They present some empirical arguments for the visibility condition. Among the arguments is the Case requirement of variables. Let us look at the following example:

(13) ***who** does it seem [***t*** to see Mary]

If a variable appears in the non-Case-marked position as in (13), the result is deviant. The Case Filter (12), being a PF requirement, cannot accommodate this fact, since variables do not have any phonetic content and thus should be exempted from the Case Filter. Under the visibility condition approach, on the other hand, variables as in (13) are not assigned any Case. Hence, they cannot retain any θ-role at LF. This violates the θ-criterion. We can correctly predict that examples like (13) are deviant.

One might argue, however, that the Case Filter (12) suffices to rule out examples like (13). One could argue that what needs Case is not the variable, but the *wh*-phrase. The *wh*-phrase must inherit a Case feature from the variable it binds. In (13), since the variable is not assigned any Case feature, *who* does not inherit any Case feature from the variable. Hence, the *wh*-phrase *who* violates the Case Filter (12). There are, however, cases which even the Case Filter (12) coupled with the Case inheritance mechanism cannot accommodate. Let us consider the following example:

(14) *the man [that [you tried [***t*** to win]]]

Let us assume following, among others, Chomsky (1977) that wh-movement is involved in the formation of relative clauses. In (14), the *wh*-phrase originates in the subject position of *win*. It is raised to the Spec of CP and then deleted there. Since the *wh*-phrase is deleted in (14), it is not submitted to the Case Filter (12). Hence, the Case Filter with the inheritance mechanism cannot provide any basis for excluding examples like (14). The visibility condition, on the other hand, explains why examples like (14) are deviant. This is because since the variable in (14) is not assigned any Case feature, it is "invisible" at LF and thus unable to retain any θ-role at LF; this violates the θ-criterion. These arguments suggest that Case features should appear both at PF and LF. Hence, Case features must be present at the time when derivations reach S-structure.

Chomsky (1993) argues, however, that once we assume the checking theory, the above arguments for assuming structural Case assignment as a S-structure property collapse. Structural Case assignment, which explains the displacement property regarding NP-movement, can be reformulated as conditions on the interface levels. Under Chomsky's (1993) checking theory, lexical items are fully inflected when selected from the lexicon. Fully inflected lexical items take the following form, where α is the morphological complex [R-$Infl_1$, ..., $Infl_n$], R is a root and $Infl_i$ is an inflectional feature:

(15) LI = (α, $Infl_1$, ..., $Infl_n$)

$Infl_i$ in (15) is an inflectional feature, which is removed from the LI when it is checked. There is assumed to be a strong/weak distinction regarding features. Strong features are illegitimate objects at PF. If they do not enter into checking relations before Spell-Out and thus remain at PF, the derivation crashes. Weak features, though legitimate at PF, are illegitimate at LF. They must be checked off before the derivation reaches LF; otherwise the derivation crashes.

Let us look at how the checking theory works, taking (16) as an example:

(16) John saw Mary.

(16) is assigned the following structure at the time of Spell-Out:

(17) $[_{AGRsP}$ **John** $[[$**T** $AGRs]_{AGRs}$ $[_{TP}$ ***t*** $[_{AGRoP}$ Spec $[AGRo$ $[_{VP}$ ***t*** [saw Mary]]]]]]]

John has a Nominative Case feature and ϕ-features. *Mary* has an Accusative Case feature and ϕ-features. The verb *saw* has an Accusative Case feature, a Tense feature, and ϕ-features. AGR and T, being functional heads, have V-features as well as N-features. V-features are those which check the inflectional features of verbs like Tense features and ϕ-features. N-features, on the other hand, are those which check the inflectional features of nouns like Case features and ϕ-features. Let us first consider the Nominative Case assignment. Since the N-feature of T, i.e., the Nominative Case feature, is strong in English, T overtly raises to AGRs in order to have its Case feature checked off through entering into a checking relation with a noun. This is because it was assumed that formal features can only be checked through the mediation of AGR.

Then, *John* overtly raises to the Spec of AGRsP to enter into a checking relation with the amalgamated form [$_{\text{AGRs}}$ T-AGRs]. If these operations do not apply overtly, the strong Nominative Case feature of T remains at PF and thus the derivation crashes.

Turning to the Accusative Case assignment, the N-features of V, including an Accusative Case feature, are weak in English. V does not have to raise overtly to AGRo to have its Accusative Case feature checked off. The principle of Procrastinate, which prefers covert operations to overt operations, requires that V should raise to AGRo covertly, not overtly. Then, *Mary* raises covertly to the Spec of AGRoP to enter into a checking relation with the amalgamated form [$_{\text{AGRo}}$ *saw*-AGRo], resulting in the following structure:

(18) [$_{\text{AGRsP}}$ John$_i$ [[T$_j$-AGRs]$_{\text{AGRs}}$ [$_{\text{TP}}$ t_j [$_{\text{AGRoP}}$ **Mary**

[[**saw**-AGRo]$_{\text{AGRo}}$ [$_{\text{VP}}$ t_i [***t*** ***t***]]]]]]]

The amalgamated form [$_{\text{AGRo}}$ *saw*-AGRo] further adjoins to T and then to AGRs in order to check the V-features of AGRs. If these operations do not apply, the features remain at LF and thus the derivation crashes. Therefore, if we assume the checking theory, structural Case assignment, which was assumed to apply at S-structure in the EST, can be reformulated as conditions at the interface levels.

To summarize Chomsky's (1993) approach to the elimination of D-structure and S-structure, the principles and conditions which were assumed to apply at either of these levels within the EST are reformulated as conditions on the interface levels. The θ-criterion, which was a D-structure condition within the EST, is reformulated as a condition at LF. Structural Case assignment, which was a S-structure property within the EST that subsumes the displacement property regarding NP-movement, is reformulated as conditions at PF and LF. Such an interface condition approach to the elimination of D-structure and S-structure contributes to the elimination of these two theory-internal linguistic levels and thus counts as a step toward the goal of the MP. In the next section, I will first present Chomsky's (1995) derivational interpretation of strong features, which is intended to reformulate the displacement property, one of S-structure properties in the EST, as a condition which applies through derivations. It is shown that Chomsky's (1995) derivational constraint approach should be preferred over Chomsky's (1993) interface condition approach on conceptual grounds.

2.3 Chomsky's (1995) Derivational Constraint Approach

The last section has considered Chomsky's (1993) interface condition approach to the elimination of D-structure and S-structure and pointed out that it contributes to the elimination of the two theory-internal linguistic levels. In this section, however, I will first point out that such an interface condition approach raises a serious conceptual problem in that it sneaks in an element of globality in the theory of language. I will then present Chomsky's (1995) derivational interpretation of strong features, which is intended to capture the displacement property, one of S-structure properties in the EST. It is shown that Chomsky's (1995) derivational constraint approach is more desirable than Chomsky's (1993) interface condition approach in that the former reduces computational burden.

Under Chomsky's (1993) approach to the elimination of D-structure and S-structure, the properties of these two levels are reduced to conditions on the interface levels. Although Chomsky's (1993) interface condition approach counts as a step toward the goal of the MP in that it eliminates the two theory-internal linguistic levels, it raises a serious conceptual problem. Let us consider structural Case assignment as an example. Under Chomsky's (1993) checking theory, we should apply the raising operation for checking a Case feature only if the derivation would otherwise reach the interface levels with an illegitimate object. Hence, at the stage of the derivation where we have an option of raising DP to check its Case feature, we cannot decide whether to apply the raising operation or not only on the basis of information available at that stage. We must look ahead to see whether the application/non-application of the raising operation would yield the interface levels only with legitimate objects. In other words, global considerations are needed for the decision about whether to apply the raising operation or not. As argued in section 1, globality, which necessarily induces computational intractability, should be reduced to local properties.

It should be noted that we are arguing against globality not against interface conditions. Interface conditions can be formulated not to trigger any operations during derivations but to simply rule out illegitimate interface representations. As will be argued later, the binding theory, the uniformity condition on chains, and the bans against vacuous quantification and free variables could be formulated as such local interface conditions. Such local interface conditions are not conceptually problematic, since they do not induce any globality.

Reformulating the previous D-structure and S-structure principles and conditions as constraints which apply throughout derivations greatly contributes to the reduction of globality in the theory of language. This is because such constraints can be formulated in such a way that they only need local considerations, but not global considerations. Chomsky's (1995) treatment of strong features goes along this line. As mentioned in the previous chapter, strong features are those which trigger overt category movement. Recall that in Chomsky (1993), strong features are characterized as illegitimate objects at PF and thus required to be checked off before Spell-Out for PF-convergence. Hence, they trigger overt category movement. As I have argued, however, this analysis needs very global considerations in the sense that when we come to a stage of a derivation where we have an option of checking a strong feature, we cannot decide whether to check it only on the basis of information available at the stage. This is because we must look ahead to see whether the checking/non-checking of the strong feature would yield PF consisting only of legitimate objects. Chomsky (1995), on the other hand, defines strong features as those that derivations "cannot tolerate" in the sense stated in (19):

(19) Suppose that a derivation D has formed a structure containing α with a strong feature F. Then, D is canceled if α is in a category not headed by α.

(adapted from Chomsky 1995:234)

This condition gives us a less global interpretation of strong features.

Before considering how (19) works, let us clarify two points on which Chomsky (1995) departs from Chomsky (1993). First, among the functional categories, Chomsky (1995) eliminates AGR, which exists only for theory-internal reasons. The Case features and ϕ-features of T and Verb are now assumed to enter into checking relations with those of nominals without any mediation of AGR. Second, Chomsky (1995) assumes that a strong feature always calls for a certain category in its checking domain. Hence, the parametric variations among languages regarding overt category movement are attributed to the existence/nonexistence of strong categorial features. Overt subject raising to the Spec of TP, for instance, is triggered not by the strong Nominative Case feature of T as in Chomsky (1993) but by the strong D-feature of T.

For an illustration of Chomsky's (1995) derivational interpretation of strong features, let us consider overt subject raising to the Spec of TP, taking the embedded clause in (20) as an example:

(20) Bill believes that John saw Mary.

Under Chomsky's (1995) theory, the embedded clause in (20) is analyzed as follows:

(21) [$_{v\mathrm{P}}$ John [v [$_{\mathrm{VP}}$ saw Mary]]]

The v is a light verb which is assumed in the transitive and unergative constructions but not in the unaccusative construction. The subject *John* appears in the Spec of vP where it receives an external θ-role in the v-VP configuration. Then, we come to the following structure:

(22) [T$_{[\mathrm{D}]}$ [$_{v\mathrm{P}}$ John [v [$_{\mathrm{VP}}$ saw Mary]]]]

Recall that under Chomsky's (1995) theory, what triggers overt subject raising is the existence of the strong D-feature of T. In English, where subjects are raised overtly, T is assumed to have a strong D-feature.

There are logically two possible continuations of (22); we either move *John* to the Spec of TP in order to check the D-feature of T or merge C with (22). Between these two possibilities, we must choose the former option, yielding the following structure:

(23) [$_{\mathrm{TP}}$ **John** [T [$_{v\mathrm{P}}$ ***t*** [v [$_{\mathrm{VP}}$ saw Mary]]]]]

This is because if we merged C with (22) as in (24), then the derivation would be canceled:

(24) [C [$\mathrm{T}_{[\mathrm{D}]}$ [$_{v\mathrm{P}}$ John [v [$_{\mathrm{VP}}$ saw Mary]]]]]

In (24), the strong D-feature of T is contained in the category which is not headed by T; this violates (19). Hence, (19), which is a condition which applies throughout derivations, triggers overt subject raising in English without recourse to any interface conditions. More generally, the displacement property, an S-structure property in the EST, can be reduced to the derivational constraint (19).

Chomsky's (1995) derivational constraint approach to strong features is conceptually more attractive than Chomsky's (1993) interface condition approach. This is because the former needs less global considerations than the latter. In order to decide whether to apply a checking operation of a strong feature, Chomsky's (1995) derivational constraint approach does not have to look at the interface levels, but only the next stage of the derivation. If the inspection of the next stage tells us that the non-checking of the strong feature would result in a structure which violates (19), we should apply the operation to check the strong feature at the present stage.

2.4 A Strictly Derivational Constraint Approach

2.4.1 Reinterpretation of Strong Features

The previous section has shown that Chomsky's (1995) interpretation of strong features (19) is conceptually more attractive than Chomsky's (1993) interpretation. It was pointed out that the former needs less global considerations than the latter. I will argue, however, that Chomsky's (1995) interpretation, though more plausible than Chomsky's (1993), still has conceptual and empirical problems.

Conceptually, Chomsky's (1995) interpretation of strong features (19) cannot reduce globality to local properties. This is because it has to look ahead to inspect the next stage in a derivation to make the decision about whether to apply an operation. To be specific, suppose that we come to a stage Σ_i in a derivation D where we have an option of applying an operation OP. Suppose further that the inspection of the next stage Σ_{i+1} tells us that the non-application of OP at Σ_i would yield a structure which violates the derivational constraint (19) at Σ_{i+1}. Then, we know that it is necessary to apply OP at Σ_i. Hence, Chomsky's (1995) interpretation of strong features (19), which has the "look-ahead" property, is still global, since it cannot decide whether to apply OP only on the basis of information available at Σ_i. Accordingly, the corresponding optimization problem is computationally intractable. To be specific, suppose that we come across n strong features at Σ_i in D. In order to decide whether to apply operations to check these features, it is necessary to make reference to Σ_{i+1} in D and the derivations which share a sequence of syntactic objects up to Σ_i with D. Overall, 2^n derivations must be reviewed to see whether they violate (19). If the review of one derivation requires one step of work, this optimization problem requires 2^n steps of work (apart from the work required to search matching formal features within the c-command domain of the strong features). Since this optimization problem only has a solution which at least requires an exponential amount of work, it belongs to Class NP. Hence, it is computationally intractable. We can see from the above discussion that Chomsky's (1995) derivational constraint approach, though it needs less global considerations than Chomsky's (1993), is still global. Its corresponding optimization problem is computationally intractable. Such global conditions should be reduced to local properties.[5]

Chomsky's (1995) interpretation of strong features (19) also has an empirical problem. It cannot trigger root overt movement like overt wh-movement in the matrix clause. Let us consider (25) as an example:

(25) **what** did you read ***t***

During the derivation of (25), we come to the stage where the strong Q-feature of C can be checked off:

(26) [$C_{[Q]}$ [you read what]]

In order to derive (25), we have to move the *wh*-phrase *what* to the Spec of CP and check the strong Q-feature of C. Derivational constraint (19),

---

[5]Note in passing that exactly like Chomsky's (1995) interpretation of strong features, the local economy conditions proposed by Collins (1997) are also global, contrary to what Collins himself claims. Let us consider his Last Resort condition:

(i) Last Resort
An operation OP involving α may apply only if some property of α is satisfied.
(Collins 1997:9)

It should be noted that in order to decide whether an operation OP involving α satisfies some property of α, it is necessary to make reference to the next stage in a derivation. The Last Resort condition (i) cannot decide whether to apply OP only on the basis of information available at the present stage. Hence, it is a global condition. Accordingly, its corresponding optimization problem is computationally intractable.

however, cannot trigger this overt wh-movement. Let us consider why (19) does not work in such a case. If the *wh*-phrase *what* does not move to the Spec of CP, then the strong Q-feature remains. According to (19), however, this derivation is not canceled. This is because the CP is the root clause and thus never contained in another category. (19) would thus claim that the strong Q-feature in (26) does not have to be checked at this stage by the application of overt wh-movement. According to the principle of Procrastinate, which prefers covert operations to overt operations, *what* should move in the covert component rather than in the overt component. Hence, there is no way to derive (25).

In order to solve these conceptual and empirical problems, I propose the following constraint on strong features while pursuing the derivational constraint approach:

(27) Strong features must be checked immediately when they become accessible to a computation.

(27) requires that strong features should be checked and deleted immediately when they become accessible to a computation; otherwise, the derivation is canceled. I call this approach the strictly derivational constraint approach to strong features.

Let us look at how the strictly derivational constraint approach works, considering the overt subject raising to the Spec of TP as an example. I will take the embedded clause in (20) (repeated here as (28)) as an example:

(28) Bill believes that John saw Mary.

Let us assume that the embedded clause in (28) is analyzed as follows:[6]

(29) [$_{VP}$ John [see Mary]]

The subject *John*, which is base-generated in the Spec of VP, receives an external θ-role in that position. Then, we come to the following structure where T with a strong D-feature is introduced:

(30) [T$_{[D]}$ [$_{VP}$ John [see Mary]]]

Recall that under Chomsky's (1995) approach, there are logically two possible continuations of (30); we either move *John* to the Spec of TP in order to check the D-feature of T or merge C with structure (30). Between these two possibilities, we choose movement of *John* to the Spec of TP, since merger of C with (30) would violate derivational constraint (19) at the next stage and thus make the derivation canceled. Under the strictly derivational constraint approach, on the other hand, there exists only one option available at this stage, i.e., the raising of *John* to the Spec of TP in order to check off the strong D-feature of T. This is because (27) requires that the checking operation of the strong D-feature of T should be applied prior to any other operations. This correctly yields the following structure:

(31) [$_{TP}$ **John** [T [$_{VP}$ ***t*** [see Mary]]]]

It should be noted that the strictly derivational constraint approach to strong features can solve the conceptual and empirical problems which Chomsky's (1995) derivational constraint approach faces. Let us first consider the conceptual problem. Unlike Chomsky's (1995) approach, the

---

[6] I do not assume Chomsky's (1995) light verb analysis of the transitive and unergative constructions. It should be pointed out, however, that the arguments to follow are valid even under Chomsky's (1995) light verb analysis.

strictly derivational constraint approach is local, since at a stage Σ of a derivation D, it can decide whether to apply an operation OP only on the basis of information available at Σ. Let us look at (30) again as an example. Under the strictly derivational constraint approach, at stage (30) where the strong D-feature of T is introduced into the derivation, the information available at this stage tells us that we should apply the raising operation to check the strong D-feature. Unlike Chomsky's (1995) derivational constraint approach, the strictly derivational constraint approach does not has to look ahead to inspect the next stage to make the decision about whether to apply the raising operation. Accordingly, the corresponding optimization problem is computationally tractable. To be specific, suppose that we come across *n* strong features at Σ in D. Suppose further that there are *m* formal features within the c-command domain of those strong features. For each strong feature, we scan all those c-commanded formal features to see whether it matches them. If the inspection of one c-commanded feature requires one step of work, this problem requires *nm* steps of work. Since this optimization problem has a solution where the number of steps to process *n* items is no more than some polynomial involving *n*, it belongs to Class P. Hence, it is computationally tractable.

Turning to the empirical problem, let us consider (25) (repeated here as (32)) again as an example:

(32) **what** did you read ***t***

During the derivation of (32), we come to the stage where C is introduced into the derivation:

(33) [$C_{[Q]}$ [you read what]]

C has a strong Q-feature. The strictly derivational constraint approach requires that this strong feature should be checked immediately by the raising of the *wh*-phrase *what* to the Spec of CP, correctly deriving (32). Hence, unlike Chomsky's (1995) approach, the strictly derivational constraint approach can correctly trigger root overt movement.

In this subsection, I have proposed the strict derivational constraint approach to strong features and argued that it should be preferred over Chomsky's (1995) approach on both conceptual and empirical grounds. It is shown that the displacement property, an S-structure property in the EST, can be reduced to the strictly derivational constraint. In the next subsection, I will argue that other D-structure and S-structure properties in the EST can also be subsumed under the strictly derivational constraint.

2.4.2 The Immediate Checking Principle

2.4.2.1 Strong Features

In the previous subsection, I have proposed the strictly derivational constraint approach to strong features, which claims that strong features must be checked immediately when they become accessible to a computation. It was shown that the displacement property, an S-structure property in the EST, can be reduced to the strictly derivational constraint (27). I have argued that the strictly derivational constraint approach to the elimination of this S-structure property is more attractive than the previous approaches on both conceptual and empirical grounds. The question to be addressed at this point is why strong features are subject to the strictly derivational constraint. I argue that the notion [+/-Interpretable] in Chomsky's (1995) sense plays a crucial role in characterizing features which are subject to the strictly derivational constraint. I argue that the uninterpretable characteristic of strong features makes them subject to the strictly derivational constraint. In other words, strong features, because they are [- Interpretable] and thus illegitimate at LF, must be checked immediately when they become accessible to a computation.[7]

It is then natural to claim that the strictly derivational constraint not only applies to strong features but also the other uninterpretable formal features (UFFs). I propose the ICP on UFFs, arguing that all UFFs are subject to the strictly derivational constraint:[8]

(34) The Immediate Checking Principle (ICP)

Uninterpretable formal features (UFFs) must be checked immediately when they become accessible to a computation.

The ICP requires that UFFs must be checked and deleted (made invisible at LF) immediately when they become accessible to a computation; otherwise the derivation is canceled.[9] According to the ICP, UFFs appear in derivations just to be deleted. I will argue that the ICP gains support from the fact that it subsumes other D-structure and S-structure properties than the displacement property in the EST. I will first consider two S-structure properties, i.e., structural Case assignment and

---

[7]Recall that our characterization of a strong feature is different from Chomsky's (1993, 1994, 1995, 1996). There is no theoretical notion of strong feature in our system. The expression "strong feature" is used just for sake of presentation to identify the uninterpretable categorial feature of a functional element which triggers overt category movement. According to our system, therefore, the Q-feature of C, which triggers overt wh-movement, is uninterpretable. This is in contrast with Chomsky's system where the strong Q-feature of C is assumed to be interpretable.

[8]Conceptual plausibility of the ICP, especially its status in the theory of language, will be discussed in detail in section 2.5.1.

[9]I will later argue that there is no operation of checking. Accordingly, the ICP will be reformulated without recourse to the operation of checking.

agreement relations. I will then consider selectional restrictions, which were D-structure properties in the EST.

Before turning to the ICP approach to the elimination of D-structure and S-structure, let us explicate the operation of checking. Following Chomsky (1995), let us assume the following formulation of the checking operation:

(35) A checked feature is deleted when possible.

(Chomsky 1995:280)

Deleted features are invisible at LF, but still accessible to a computation. (35) claims that a checked feature is deleted only when it would not contradict the principle of recoverability of deletion, which states that unrecoverable items may not be deleted. Interpretable features, which receive interpretation at LF, cannot be deleted when they are checked. This is because if checked interpretable features are deleted, it would violate the principle of recoverbility of deletion. Uninterpretable features, on the other hand, are deleted when they are checked. Since uninterpretable features do not contribute to any content at the LF interface, their deletion does not violate the principle of recoverbility of deletion.

I also assume with Chomsky (1995) that checking relations can be established either by Attract/Move or Merge between a head H and an element in its neighborhood. Although Chomsky assumes that checking relations are only established between H and an element in its Spec position, I claim that they are established between H and an element in its Spec or complement position.

2.4.2.2 Structural Case Assignment

Within the framework of the EST, structural Case assignment was assumed to apply at S-structure. As I have presented in section 2.2.2, the arguments for the S-structure property of structural Case assignment were based on the fact that Case features appear at both PF and LF and thus must be present at the time when derivations reach S-structure. I have presented Chomsky's (1993) checking theory of Case, according to which the S-structure property of structural Case assignment can be reduced to the interface conditions. It was shown that it counts as a step toward the goal of the MP in that it contributes to the elimination of S-structure. As I have argued above, however, Chomsky's (1993) approach raises a conceptual problem, since it would sneak in an element of globality in the theory of language. In order to solve the globality problem, I argue that structural Case assignment should be subsumed under the ICP. Under the ICP approach, where structural Case assignment is reduced to a local condition, the problem of globality does not arise.

Recall that since Case features are uninterpretable no matter where they may appear, they are always subject to the ICP. The ICP then requires that when Case features become accessible to a computation, they should immediately enter into checking relations either by Attract/Move or Merge. After Case features enter into checking relations, they are deleted and made invisible at LF, since they are uninterpretable.

Let us consider (36) as an example:

(36) John saw Mary

We first select the verb *see* from the numeration N:

(37) $\text{see}_{[\text{ACC}]}$

In (37), the Accusative Case feature is represented as ACC. Recall that transitive verbs like *see* have Accusative Case features as their intrinsic features. The ICP requires that the Accusative Case feature of *see* should be checked immediately. It is checked by selecting *Mary*, which has a Case feature as its optional feature, and merging *see* with *Mary*, as shown below:[10]

(38) $[_{V^{max}}$ see Mary$]$

Note that if the Case feature of *Mary* does not coincide with that of *see*, the derivation is canceled due to a violation of the ICP.

One might say that this derivation violates the ICP, since Select *Mary* intervenes between Select *see*, which makes its uninterpretable Accusative Case feature accessible to the computation, and merger of *see* with *Mary*, which checks the uninterpretable feature. I argue that the ICP only regulates operations which manipulate terms in phrase structures, i.e., Merge and Attract/Move. The operation Select selects a lexical item from an N and introduces it into a derivation. It does not manipulate terms. Select is therefore immune from the ICP. Hence, in (38), although we apply Select *Mary* before merger of *see* with *Mary*, there is no violation of the ICP.

Another important point to note concerns the introduction of optional features. Chomsky (1995) claims that they are added

[10]Essentially following Chomsky (1995) and Muysken (1982), we define the notion of maximal projection derivationally. In (38), for instance, the dominating node is assigned the categorial status of V^{max}, since it is the top node of the V projection at this stage. If it further projects up, its categorial status will change to an intermediate projection of the V rather than remain as a maximal projection of the V.

arbitrarily as a lexical item enters an N. I depart from Chomsky, proposing that optional features are added at any point of a derivation:[11]

(39) Optional features are added arbitrarily at any point of a derivation.

In (38), recall that the Case feature of *Mary* is optional. Hence, *Mary* does not have any Case feature when it is selected from the N. We add an Accusative Case feature to *Mary* after it is selected from the N. We then merge *see* with *Mary* and check their Accusative Case features, constructing (38). Note that the operation of adding an optional feature to a lexical item is immune from the ICP, since it does not manipulate terms in phrase structures.

After having constructed *see Mary*, we merge *John* and *see Mary*:

(40) [$_{V^{max}}$ John [see Mary]]

---

[11]The question that arises here is whether optional features are still elements associated with an N. I claim that lexical items have slots for its optional features (if it has ones) which are to be filled by specific features during a derivation. For example, a noun has a slot for its optional Case feature. The slot is to be filled by a specific Case feature during the derivation. Under this view, lexical items in an N only have slots for its optional features but not the features themselves. I claim that although optional features are not part of lexical items within an N, they are in an N as independent elements ("floating" in a sense) and later added to lexical items at any stage of a derivation in order to fill their slots. Note that it is plausible to assume that optional features exist as independent elements in an N, since, to be precise, lexical items themselves are collections of features. Alternatively, it might be possible to say that optional features are not included in an N. This view, however, would increase reference sets and hence induce computability problems. It is also against the idea that an N is what specifies an input for a derivation.

Chomsky (1995) considers two possible ways of introducing optional features. The one is to add optional features to a lexical item when an N is formed. The other is to add optional features when a lexical item is selected from an N and introduced into a derivation. He claims that both of these approaches are compatible with the MP. Between these two approaches, an N is the place where a lexical item and its optional features get together only in the former but not in the latter. Our view therefore can be regarded as a further extension of the latter approach . I will argue in chapter 5 that this interpretation of optional features is also instrumental in accounting for the distribution of *wh*-elements in-situ in languages like Japanese. I am indebted to Lisa Cheng (personal communication) for bringing this subject to my attention.

Note that the Case feature of *John*, being optional, has not been added at this point of the derivation yet.

The next step is to select the finite T. It has a Nominative Case feature, which is to be checked by an element in its specifier position. Apart from the Case feature, the finite T also has a D-feature and V-feature. As will be discussed in detail in section 2.4.2.4, these features are both uninterpretable. The V-feature is checked by merger of T with a projection of V. The D-feature is checked by the raising of D^{max} to the Spec of T^{max}. I argue that these formal features, i.e., the Nominative Case feature, the D-feature, and the V-feature, are not just listed in an unordered fashion in the lexical entry of the finite T. There is a hierarchical structure among these formal features within the lexical entry, which ensures a specific ordering among the applications of their checking operations. The finite T has the following hierarchical structure concerning its formal features, where NOM, D, and V represent the Nominative Case feature, the D-feature, and the V-feature, respectively:

(41) $T_{[[\text{NOM, D}]\ \text{V}]}$

Let us assume that only the structurally highest feature is accessible to a computation. Embedded features, which are "covered" by the higher features, are not accessible to a computation. In (41), the computation is only accessible to the V-feature, which is structurally highest, but not to the embedded Nominative Case feature and D-feature. It is only after the V-feature is eliminated by the checking operation that the Nominative Case feature and the D-feature become accessible to the computation. This ensures that the V-feature is checked by merger of T with its complement while the Nominative Case feature and the D-feature are checked by the raising of D^{max} to the Spec of T^{max}. Putting it in traditional terms, the feature hierarchy (41) tells us that the finite T takes a projection of V as its internal argument and a D^{max} with a Nominative Case feature as its external argument.[12]

The hierarchical structure among formal features like (41) corresponds to the traditional argument structure, which is advocated by, among others, di Sciullo and Williams (1987), Grimshaw (1990), Levin and Rappaport (1986), Marantz (1984), Williams (1981, 1994), and Zubizarreta (1987), though the hierarchical ordering between external and internal arguments advocated here is opposite of what is assumed in their works. It should be noted that such a hierarchical structure within a lexical entry is needed anyway, since it is necessary to make a distinction between external and internal arguments. Since our feature hierarchy and the traditional argument structure are both meant to express the same fact, it is more desirable if they can be collapsed into one. I claim that the traditional argument structure should be replaced by our feature hierarchy. Feature-wise, an internal argument is higher than an external argument. Because of the ICP, the external argument appears in a structurally higher position than the internal argument in phrase structures.[13]

---

[12]Alternatively, it is conceivable that the notion of closeness comes into play. It has been assumed that the notion of closeness only plays a role when a triggering feature searches for its matching feature in its c-commanding domain within a single phrase structure. Let us assume contra this standard view that the notion of closeness is also relevant when a triggering feature searches for its matching feature within a lexical entry. The structurally higher within the feature hierarchy of a lexical entry a feature is, it is "closer" to the triggering feature. It then follows that at each stage in a derivation, it is only the "closest" feature within a lexical entry which can enter into a checking relation.

[13]I am indebted to Jim Huang (personal communication) for bringing this subject to my attention.

Returning to the derivation of (36), when we select the finite T, its V-feature becomes accessible to the computation. Note that the Nominative Case feature and the D-feature are "covered" by the structurally higher V-feature and thus not accessible to the computation at this stage. The ICP requires that the V-feature of the finite T should be checked immediately by merger of the finite T with the V^{max} *John see Mary*. The V- feature, being uninterpretable, is deleted and made invisible at LF.

Note that at this stage of the derivation, the Nominative Case feature and D-feature of the finite T become accessible to the computation. Recall that both of these features are uninterpretable and thus subject to the ICP. The question then arises as to how to check off more than one UFFs which become accessible to the computation at the same time without violating the ICP. I claim following Chomsky (1993, 1995) that when a certain operation checks one feature, it can also check other features simultaneously as "free riders." Hence, both of these UFFs can be checked off in conformity with the ICP. At this point, we first add an optional Nominative Case feature to *John*. Then, we check the Nominative Case feature and D-feature of T through the raising of *John* to the Spec of T^{max} in accordance with the ICP. The resultant structure is as follows, given the copy theory of Attract/Move:

(42) [$_{T^{max}}$ John [T [$_{V^{max}}$ John [see Mary]]]]

As illustrated above, we can construct (36) in conformity with the ICP. Hence, structural Case assignment, which was assumed to apply at S-structure in the EST, can be subsumed under the ICP.

2.4.2.3 Agreement Relations

Apart from structural Case assignment, agreement relations also counted as S-structure properties within the EST. The arguments for assuming agreement relations as S-structure properties were as follows. In some languages, agreement is morphologically realized while in others it is not. It has been claimed, however, that agreement features, called ϕ-features, should occur in a uniform way whether they are morphologically realized or not. Then, ϕ-features appear at PF. Furthermore, the ϕ-features of nouns receive interpretations at LF. Then, they also appear at LF. Hence, agreement was assumed to occur at S-structure before derivations bifurcate into PF and LF.

Chomsky (1993) argues that once we assume the checking theory, agreement relations can be reduced to conditions at the interface levels. Considering the verb-noun agreement as an example, it was assumed that both verbs and nouns are morphologically fully inflected before the checking operations for the verb-noun agreement take place. The agreement morphology is always realized at PF wherever the checking operations for the agreement take place. Hence, the realizations of agreement morphology at PF do not necessarily count as evidence in support of the view that agreement takes place at S-structure.

Concerning the LF interpretability of the ϕ-features of nouns, ϕ-features have already been assigned to both nouns and verbs within the lexicon. The ϕ-features of nouns therefore appear at LF wherever the checking operations for the verb-noun agreement take place. Hence, the LF interpretability of the ϕ-features of nouns does not constitute evidence for the S-structure property of agreement relations, either.

As I have pointed out repeatedly, Chomsky's checking theory, although it contributes to the elimination of the theory-internal linguistic levels, raises the problem of globality. I argue that agreement relations should also be subsumed under the ICP. Recall that while the ϕ-features of nouns are interpretable, those of verbs are uninterpretable. Hence, only the ϕ-features of verbs, but not those of nouns, are subject to the ICP.

Let us consider how the ICP approach to agreement relations works, taking (36) (repeated here as (43)) as an example. For expository purposes, the discussion of this subsection only explicates how ϕ-features are checked, ignoring Case features:

(43) John saw Mary

We first select the verb *see*. Since the ϕ-features of verbs are optional, *see* does not have any ϕ-features when selected. After selecting *see*, therefore, we add ϕ-features to it. The ϕ-features are hierarchically structured as follows:

(44) $\text{see}_{[[3SM]\ 3SF]}$

In (44), 3rd person, singular, male, and female features are represented as 3, S, M, and F, respectively. Since the ϕ-features of verbs are uninterpretable, they must be checked immediately when they become accessible to the computation. Since the ϕ-features of *see* are accessible to the computation at this point, they are checked by selecting *Mary*, which has the matching ϕ-features as its intrinsic property, and combining *see* with *Mary*. The resultant structure is as below:

(45) $[_{V^{max}}\ \text{see}_{[3SM]}\ \text{Mary}_{[3SF]}]$

Note that if the ϕ-features of *see* do not coincide with those of *Mary*, the derivation is canceled due to a violation of the ICP. While the ϕ-features of *see*, being uninterpretable, are deleted when checked, those of *Mary*, being interpretable, remain intact.

At this stage of the derivation, the remaining ϕ-features of *see* become accessible to the computation. Since they are uninterpretable, the ICP requires that they should be checked immediately. They are checked by selecting *John*, which has those ϕ-features as its intrinsic property, and combining *John* with *see Mary*:

(46) $[_{V^{max}}$ $John_{[3SM]}$ [see $Mary_{[3SF]}$]]

The ϕ-features of *see*, being uninterpretable, are deleted when checked. Those of *John*, on the other hand, remain intact and receive interpretations at LF. In this way, agreement relations can be subsumed under the ICP.[14]

---

[14]In our system, the ϕ-features of verbs are deleted and made invisible at LF through checking operations. One might wonder how verbs are interpreted if their ϕ-features are deleted especially in polysynthetic languages (see, among others, Baker (1988, 1996)). It should be noted, however, our system is only claiming that the ϕ-features of verbs, being formal features, are deleted. Apart from formal features, verbs also have pure semantic features. Chomsky (1995) claims that formal features typically have their purely semantic correlates which reflect semantic properties (accusative Case and transitivity, for example). Hence, the nonexistence of the ϕ-features of verbs at LF does not necessarily mean that verbs do not have any features that receive interpretations at LF regarding their agreement. I am indebted to Lisa Cheng (personal communication) for bringing this subject to my attention.

2.4.2.4 Selectional Restrictions

It has been assumed by, among others, Abney (1987), Chomsky (1965, 1981, 1986b), and Fukui (1986) that lexical items can be classified into two types, i.e., thematic elements like Noun, Verb, Adjective, and Preposition and functional elements like Complementizer, Tense, and Determiner. Among several differences between these two types of elements is the way they impose restrictions on their arguments, i.e., the items which appear in their specifier and complement positions.[15]

Thematic items choose the thematic types of arguments they take. For instance, the verb *believe* in (47), being a two-place predicate, assigns the θ-role of Agent to its first argument *John* and the θ-role of Theme to its second argument *that Bill saw Mary*:

(47) John believes that Bill saw Mary.

It should be noted that exactly what kinds of θ-roles each thematic item assigns is irrelevant to the following discussion. We just use them as identifying the arguments of thematic items.

Functional elements, on the other hand, do not assign any θ-role, but only choose the categorial status of their complement. Another salient property of functional elements is that those belonging to the same category share their categorial selection property. All the functional elements belonging to C select T^{max} as their complement. Similarly, all the functional elements belonging to T and D select V^{max} and N^{max}, respectively, as their complements. In the following, I will use the notion selectional restriction as a cover term for the θ-role assignment properties of thematic items and the categorial selection properties of functional

[15]See Grimshaw (1979) and Pesetsky (1982) for further discussion of this subject.

items. It is important to note that the present definition of the notion selectional restriction differs from that of Chomsky (1965). The latter specifies the restrictions which verbs impose on the semantic features of their arguments like [+/- Human] and [+/- Abstract].

It was assumed in the EST framework that these selectional restriction properties of functional and thematic items should be satisfied at D-structure. This is because it was assumed that the base component, including the lexicon and the phrase structure component in the EST, should generate D-structure, respecting the selectional restriction properties of functional and thematic items. As shown in section 2.2.1, the θ-criterion coupled with the Projection Principle used to ensure that D-structure is a representation of GF-θ. Hence, the selectional restrictions of thematic items were assumed to be satisfied at D-structure. The selectional restrictions of functional items were also required to be satisfied at D-structure by the Projection Principle.[16] Under the MP, however, selectional restrictions, which were stated at the level of D-structure in the EST, must be reformulated either as conditions on the interface levels or constraints which apply throughout derivations.

Chomsky (1993, 1995) pursues the former possibility, arguing that the selectional restrictions of functional and thematic items should be satisfied at LF. As mentioned above, he totally eliminates the Projection Principle and claims that lexical items are selected from the lexicon freely at any point of a derivation. The θ-criterion, which was responsible for

---

[16] In the earlier stage of the EST where the phrase structure component is still posited, the selectional restriction properties of functional items are specified in terms of phrase structure rules. See, among others, Higginbotham (1983, 1985), Speas (1984) and Stowell (1981) for arguments in favor of the elimination of the phrase structure component.

stating the selectional restrictions of thematic items in the EST, is reformulated as a condition at LF. As I have extensively argued above, however, reformulating the properties of the theory-internal linguistic levels as conditions on the interface levels would raise a conceptual problem in that it needs global considerations. For this reason, I would rather pursue the latter possibility, i.e., reformulating selectional restrictions as constraints which apply throughout derivations, arguing that selectional restrictions should be subsumed under the ICP.

Let us consider how the ICP approach to selectional restrictions works. For expository purposes, the discussion of this subsection restricts itself to explicating how selectional restriction features can be checked in accordance with the ICP, ignoring Case and agreement features unless they become relevant to the discussion. Functional elements have categorial features to be checked through merger with their complements. For instance, C has the categorial feature [T], which is to be checked by its merger with a projection of T. T has the categorial feature [V], which is to be checked by its merger with a projection of V. D has the categorial feature of [N], which is to be checked by its merger with a projection of N. These categorial features of functional items are uninterpretable, since they do not specify the categorial status of the functional items themselves and thus do not receive any interpretations at LF. The categorial features of the complements, on the other hand, are interpretable, since the features specify the categorial status of the complements themselves.

Turning to thematic items, let us assume contra Chomsky (1995) that thematic features are formal features and thus accessible to a computation. Thematic items have thematic features, which are to be checked by merger with their argument(s). For instance, the verb *see* has thematic features AGENT and THEME. Its lexical entry also specifies that the former thematic feature is checked by merger of the verb with an element in its specifier position while the latter is checked by merger of the verb with an element in its complement position. The verb *see* has the following structure concerning its thematic features, where AGENT and THEME features are represented as A and T, respectively:

(48) [[A] T]

The thematic features of thematic items are uninterpretable, since thematic items themselves do not receive any interpretations regarding thematic types like Agent and Theme at LF. The thematic features of the specifiers and complements, on the other hand, are interpretable, since they receive interpretations regarding their thematic types at LF.

Let us consider the embedded clause of (47) (repeated here as (49)) as an example:

(49) John believes that Bill saw Mary

We first select the verb *see*, which has uninterpretable thematic features AGENT and THEME:[17]

---

[17]This thematic feature hierarchy is exactly the opposite of the thematic hierarchy advocated by, among others, Foley and van Valin (1984), Jackendoff (1972), Nishigauchi (1984), Randoll (1988), Rappaport and Levin (1988), Schwartz (1988), and Wilkins (1988). It might be possible to claim that the thematic hierarchy is defined by our thematic feature hierarchy. Feature-wise, for instance, Theme is structurally higher than Agent. Because of the ICP, however, we derive the thematic hierarchy as a phrase structure generalization. As far as thematic items are concerned, it is possible to achieve the same hierarchical effect without recourse to any feature hierarchy. Let us assume the lexical decomposition approach advocated by, among others, Chomsky (1995), Hale and Keyser (1993), and Huang (1994). Then, each thematic item is only associated with one thematic feature. Regarding the selectional restrictions of functional elements, however, we still need a feature hierarchy to achieve the hierarchical effects in phrase structures. I am indebted to Jim Huang (personal communication) for bringing this subject to my attention.

(50) $\text{see}_{[[A]\ T]}$

When the verb *see* is selected from the N, the ICP requires that THEME should be checked immediately. It should be noted that AGENT, being embedded in the structure of thematic features and thus not accessible to the computation, cannot be checked at this point of the derivation. Since the thematic features of nominals are optional, we add THEME feature to *Mary*. The THEME feature of *see* is checked by selecting *Mary* and combining *see* with *Mary*, in accordance with the ICP:

(51) $[_{V^{max}}\ \text{see}_{[A]}\ \text{Mary}_{[T]}]$

The THEME feature of *see*, being uninterpretable, is deleted when checked. The THEME feature of *Mary*, on the other hand, remains, since it is interpretable. At this point of the derivation, the AGENT feature of *see* becomes accessible to the computation. The ICP requires that it should be checked immediately. We select *Bill* from the N and add an AGENT feature to it, since its thematic feature is optional. We then merge *Bill* with *see Mary*, resulting in the following structure:

(52) $[_{V^{max}}\ \text{Bill}_{[A]}\ [\text{see Mary}_{[T]}]]$

While the AGENT feature of *see*, being uninterpretable, is deleted, that of *Bill*, being interpretable, remains intact. It should be noted that if the thematic types of *see* do not coincide with those of its arguments, the derivation is canceled due to a violation of the ICP.

The next step is to select the finite T from the N. The finite T, being a functional item, has a V-feature as its selectional restriction property as well as a D-feature as its strong feature. These categorial features are hierarchically structured within the lexical entry of the finite T:

(53) $T_{[[D]\ V]}$

This ensures that the finite T takes a projection of V as its internal argument and a projection of D as its external argument. Since both of these categorial features are uninterpretable, the ICP requires that they should be checked immediately when they become accessible to the computation. When T is selected, only its V-feature is accessible to the computation. Hence, the V-feature must be checked immediately by merger of T with V^{max}. After the V-feature of T undergoes deletion through the checking operation, the D-feature of T becomes accessible to the computation. The D-feature must be checked immediately by the raising of *Bill* to the Spec of T^{max} at this point of the derivation in accordance with the ICP. The D-feature, being uninterpretable, is deleted when checked. The resultant structure is as follows:

(54) $[_{T^{max}}$ $Bill_{[A]}$ [T $[_{V^{max}}$ $Bill_{[A]}$ [see $Mary_{[T]}$]]]]

Finally, we select the complementizer *that* from the N. The complementizer *that*, being a functional item, has the categorial feature [T] as its selectional restriction property. This categorial feature is uninterpretable and thus must be checked immediately by merger of *that* with T^{max}. The categorial feature [T] of *that*, being uninterpretable, is deleted when checked:

(55) $[_{C^{max}}$ that $[_{T^{max}}$ $Bill_{[A]}$ [T $[_{V^{max}}$ $Bill_{[A]}$ [see $Mary_{[T]}$]]]]]

It was shown that the selectional restriction properties of lexical items, which were assumed to be D-structure properties in the EST, can be subsumed under the ICP. Unlike Chomsky's (1995) approach, the former D-structure properties are now reduced to the constraint which applies throughout derivations without recourse to any interface conditions.

To summarize section 2.4.2, I have shown that the displacement property, structural Case assignment, agreement relations, and selectional restrictions, which were assumed to be D-structure or S-structure properties in the EST, can be subsumed under the ICP. Left untouched among the properties of the theory-internal linguistic levels in the EST is the binding theory, which I will discuss in the next section.

2.5 Consequences of the Immediate Checking Principle

This section considers theoretical consequences of the ICP. Section 2.5.1 investigates interpretation of the ICP in the theory of language. Section 2.5.2 claims that if we assume the ICP, we can virtually eliminate the notion of LF-convergence. Section 2.5.3 shows that the ICP is a language-specific computational device which greatly contributes to the reduction of globality in the theory of language.

2.5.1 Interpretation of the Immediate Checking Principle

The above discussion has shown that the ICP enables us to capture the D-structure and S-structure properties in the EST. Especially, since the ICP accommodates the displacement property, it subsumes the principle of Last Resort advocated by Chomsky (1991a) and Chomsky and Lasnik (1993) (or its modified principles, Greed proposed by Chomsky (1993, 1995) and Suicidal Greed proposed by Chomsky (1996)), which was intended to capture that property. Chomsky (1995, 1996) argues that Last Resort should not count as an economy condition, since it is inviolable. It should rather be part of the definition of Attract/Move-F, as can been seen in Chomsky's (1995) definition:[18]

(56) Attract/Move-F

F raises to target K only if F enters into a checking relation with a sublabel of K (where a sublabel of K is a feature of the zero-level projection of the head H(K) of K).

(adapted from Chomsky (1995))

Our analysis claims, on the contrary, that the ICP, which subsumes Last Resort, is an independent principle rather than part of the definition of Attract/Move. This is because, as argued above, the ICP constrains not only Attract/Move but also Merge. If the ICP were part of the definition of Attract/Move, then it would also be part of the definition of Merge. This would be theoretically undesirable, however, since it is redundant to incorporate the ICP into the definitions of the two distinct operations. Concerning Chomsky's argument for incorporating Last Resort as part of the definition of an operation, we should recall that the ICP is inviolable. There is therefore no need to incorporate the ICP into definition of an operation in order to derive the inviolability of the ICP.[19]

---

[18]This definition differs from Chomsky's (1995) original one in the following respects. First, Chomsky (1995) totally eliminates the notion of Move, arguing that the traditional notion of movement should be reinterpreted as Attract-F. Under this view, the locus of the notion is completely shifted from the moved element to the target. We are assuming, on the other hand, that the notion of Move is still needed. The traditional operation of movement is reinterpreted as Attract/Move-F rather than Attract-F. Second, Chomsky incorporates the Minimal Link Condition (MLC) into the definition of Attract/Move-F. I argue contra Chomsky that the MLC is also an independent principle. As will be discussed soon, our analysis claims that the last resort condition on Attract/Move should be subsumed by the ICP, which is an independent principle. Hence, the MLC, which only regulates the operations satisfying the ICP, should also be an independent principle under our analysis.

[19]Collins (1997) also argues that his economy conditions, Last Resort and Minimality, should not be part of the definition of Attract/Move.

Given that the ICP is an independent principle, a question then arises as to interpretation of the ICP in the theory of language. I argue that the ICP is a local "heuristic algorithm" ("computational trick") which gives us an approximate solution to a computationally intractable problem induced by the global interface condition on UFFs. Recall that UFFs are those which are illegitimate at LF. According to the BOC-driven optimal design of language assumed by the MP, we have an interface condition which requires that UFFs, being illegitimate at LF, should not remain at that level. If they remain at LF, the derivation crashes at that level. As argued above, however, the postulation of the interface condition on UFFs alone necessarily leads to computational intractability. This is because triggering a checking operation of a UFF in terms of the interface condition would necessarily need global considerations. It would then follow that the part of language involving UFFs, whose corresponding optimization problem is computationally intractable, is not usable in practice, contrary to fact. Hence, there should exist a local "heuristic algorithm" ("computational trick") which gives us an approximate solution to this computationally intractable problem, making that part of the language usable in practice.

I argue that it is the ICP which plays this role. First, the ICP brings about an approximate solution to this computationally intractable problem. Suppose that we have an option of applying a checking operation of a UFF at a certain stage of a derivation. In most cases, the ICP gives us the same answer to this computational problem as the interface condition on UFFs. In other words, the ICP only requires us to apply the checking operation when the interface condition on UFFs would also require us to do so. For example, when we have an option of applying overt subject raising to check a strong D-feature of T during a derivation, the ICP requires us to apply that operation, and so would the interface condition on UFFs. Hence, in such cases, the ICP brings about a perfect solution to the computationally intractable problem induced by the interface condition on UFFs.

There are, however, cases where the ICP does not give us the same answer to a computational problem as the interface condition on UFFs. For example, suppose that we select V, which has a UFF that can be checked by its merger with D^{max}, which is in the N. Suppose further that we also have a so called VP-adverb in the N. Here, we have two options. We can select D^{max} and merge it with V. Alternatively, we can select the adverb and merge it with V. Recall that while merger of V with D^{max} is triggered by the UFF of V, merger of V with the adverb is not triggered by any UFF. The interface condition on UFFs would claim that we may apply either of these operations at this stage as long as the UFF of V eventually enters into a checking relation before LF. The ICP, on the other hand, only allows us to apply merger of V with D^{max} at this stage, preventing us from applying merger of V with the adverb. This is because the ICP requires us to apply the checking operation of the UFF of V before any other operations that manipulate terms. In cases like this, the ICP does not gives us the same answer to a computational problem as the interface condition on UFFs. Hence, the solution which the ICP brings about to the computationally intractable problem induced by the interface condition on UFFs is only approximate.

Furthermore, the ICP is local. This is because when we have an option of applying a checking operation of a UFF at a certain stage of a derivation, it can decide whether to apply the checking operation only on the basis of information available at that stage.

In conclusion, the interface condition on UFFs, which is required by BOCs, fundamentally induces globality and thus its corresponding optimization problem is computationally intractable. The ICP, being a local "heuristic algorithm" ("computational trick") for this computationally intractable problem, reduces computational burden and facilitates usability of language in practice. Hence, it is conceptually plausible to claim that UFFs are subject to the ICP. It should also be noted that the ICP would make a derivation canceled before it reaches LF which would otherwise crash due to a violation of the interface condition on UFFs. If we conform to the ICP throughout derivations, UFFs never remain at LF and the interface condition on UFFs is always satisfied. Hence, the interface condition on UFFs virtually plays no role once we assume the ICP.

It is important to point out that the ICP approach supports the language design that considerations of computational complexity do not matter for fundamental aspects of language. Under this view, globality is one of the fundamental properties of language. Its corresponding optimization problem is fundamentally intractable. There are, however, language-specific computational devices, "heuristic algorithms" ("computational tricks"), which reduce its fundamental globality to local properties, facilitating its usability in practice. The ICP is one such language-specific computational device. Hence, if our ICP approach is on the right track, it constitutes evidence against the language design assumed by, among others, Collins (1997) that considerations of computational complexity matter for fundamental aspects of language and no globality is allowed.

It might be possible to claim that the ICP is also compatible with the language design that considerations of computational complexity matter for fundamental aspects of language. Under this design of language, the ICP would either count as the defining property of UFFs or as a local economy condition in the sense of Collins (1997). We are reluctant to take this view, however, since we would lose an explanation of why UFFs are subject to the ICP. Recall that we are claiming that UFFs are subject to the ICP because the latter gives us an approximate solution to the computationally intractable problem induced by the BOC-driven interface condition on UFFs. Without having the ICP based on the interface condition on UFFs, there would be no reason why only UFFs, but not the other features, must be checked and deleted immediately.

To summarize, the ICP, being an independent principle, functions as a local "heuristic algorithm" ("computational trick") for the computationally intractable problem induced by the interface condition on UFFs. This explains why only UFFs, but not the other features, must be checked and deleted immediately when they become accessible to a computation.

2.5.2 Elimination of the Notion of LF-Convergence

Another theoretical consequence of the ICP is that it enables us to virtually eliminate the notion of LF-convergence. Recall that under Chomsky's (1993, 1994, 1995) framework, a derivation converges when its interface levels only consist of legitimate objects and thus satisfy BOCs. Otherwise, the derivation crashes. UFFs like strong features, Case features, and the ϕ-features of verbs, are illegitimate objects at LF. If they are checked and deleted (made invisible at LF) during a derivation, the derivation converges at LF. If they remain unchecked and undeleted, the derivation crashes at LF. Hence, the notion of LF-convergence is needed.

The ICP requires, on the other hand, that UFFs should be checked and deleted (made invisible at LF) immediately when they become accessible to a computation. Otherwise, the derivation is canceled. Thus, it never happens that a derivation reaches LF with UFFs being left unchecked and undeleted. The ICP makes it mandatory that UFFs are checked and deleted (made invisible at LF) before a derivation reaches that interface level. The interface condition on UFFs is therefore always satisfied. In other words, a derivation which reaches LF is always convergent as far as UFFs are concerned. Hence, the ICP virtually eliminates the notion of LF-convergence regarding UFFs.

There are, however, other conditions which are assumed to apply at LF, i.e., the binding theory, the uniformity condition on chains, and the condition on the operator-variable construction, which consists of the bans against vacuous quantification and free variables.[20] One might say that we still need the notion of LF-convergence for these principles. It is not clear, however, whether violations of these principles would lead derivations to crash at LF. It might be possible to claim that the binding theory, the uniformity condition on chains, and the condition on the operator-variable construction purely count as principles of interpretation on the LF-interface level. They never play any role in making

---

[20]See Fukui (1993b) for the alternative view that the uniformity condition on chains should apply derivationally.

derivations (computations) converge or crash. Their violations would not make derivations crash, but only make their interpretations deviant.[21]

There is another principle where the notion of LF-convergence plays a crucial role, that is, the economy conditions like the Minimal Link Condition (MLC) and the principle of Procrastinate. Recall that Chomsky (1993, 1994, 1995) assumes that the economy conditions only compare convergent derivations. Among convergent derivations, the economy conditions choose the most "economical" one, blocking all the others.[22] Several attempts have been made to eliminate the economy conditions and find alternative treatments. Chomsky (1995) pursues the possibility of incorporating the MLC as part of the definition of Attract-F. Collins (1997) suggests that the principle of Procrastinate should be subsumed under the modified definitions of strong/weak features. As far as we can do away with the economy conditions, we do not need the notion of LF-convergence for the purpose of the economy conditions either.

If the above line of reasoning is correct, we can virtually eliminate the notion of LF-convergence through nullifying the effects of the interface condition on UFFs in terms of the ICP. I must admit, however, that the discussion of this subsection is only sketchy and much still remains to be done.

---

[21]Among the conditions of the binding theory, Condition A might make reference to formal features and thus play a role in making derivations converge or crash. See, for example, Chomsky (1986b, 1993) for this view of Condition A of the binding theory, where reflexives and reciprocals are assumed to undergo movement at LF. See Battistella (1989), Cole, Hermon, and Sung (1990), Huang and Tang (1991), Katada (1991), Lebeaux (1983), and Pica (1987) for further arguments in favor of the LF-movement analysis of anaphors.

[22]See Ura (1995) for the different view that economy conditions compare not only convergent but also non-convergent derivations.

2.5.3 Reduction of Globality to Local Properties

The MP requires us to reformulate the D-structure and S-structure properties in the EST either as interface conditions or as derivational constraints. I have investigated the interface condition approach, arguing that it is conceptually problematic, though empirically adequate in capturing the D-structure and S-structure properties in the EST. This is because the interface condition approach has the "look-ahead" property and thus necessarily induces globality. Pursuing the derivational constraint approach, I have proposed the ICP, which is conceptually desirable in that it greatly contributes to reducing globality to local properties. It should be noted that we are arguing against globality not against interface conditions. Interface conditions do not induce any globality problems as far as they are local in nature. This is a good place to consider whether other interface conditions are local and therefore conceptually desirable.

Under our analysis, the remaining interface conditions are the binding theory, the uniformity condition on chains, and the condition on the operator-variable construction. It is conceivable that these interface conditions are all local in nature. Recall that global interface conditions like Chomsky's (1993) interface condition on UFFs require us to look at the interface levels in order to decide whether to apply an operation at an intermediate stage of a derivation. Global interface conditions, with this "look-ahead" property, necessarily induce computational intractability and thus raise conceptual problems. The binding theory, the uniformity condition on chains, and the condition on the operator-variable construction, on the other hand, can plausibly be interpreted as local interface conditions which only make reference to LF. Since these interface conditions have nothing to do with formal features, they should never trigger the application of an operation during a computation given the minimalist assumption that only formal features are accessible to a computation.[23] It follows that these interface conditions may not trigger any operation during a computation in order to avoid their violations at the LF interface level. The LF-representations which violate any of these interface conditions are simply ruled out. Since these interface conditions as such formulated only make reference to a single stage in a derivation, i.e., the LF interface level, they are local in nature.[24]

As a result, we can largely dispense with globality in the theory of language. What remains global is the economy conditions which crucially rely on the notion of convergence. As mentioned in the previous subsection, however, the economy conditions may possibly be eliminated. As far as we can do away with the economy conditions, we can totally eliminate globality in the theory of language.

---

[23]As mentioned in note 21, Condition A of the binding theory might involve formal features.

[24]Tsai (1994) claims that the condition on the operator-variable construction triggers overt category movement. His analysis, however, is conceptually problematic. First, it has the "look-ahead" property and thus necessarily induces globality. Second, it is against the minimalist view that only formal features are accessible to a computation. His analysis would require a computation to make reference to other entities than formal features, since the condition on the operator-variable construction does not involve any formal features.

2.6 Adjuncts and the Immediate Checking Principle

Before closing this chapter, let us consider a consequence of the ICP for the theory of phrase structure. I will propose the Earliness Principle (EP) on Select, a local "heuristic algorithm" ("computational trick") for a global interface condition on a numeration N, arguing that the ICP coupled with the EP gives rise to an asymmetry concerning the composition of phrase structure. I will argue that while the terms required by UFFs are merged cyclically, those not required by UFFs are merged postcyclically. It then follows that arguments, which are required by UFFs, are merged cyclically. On the other hand, typical adjuncts, which are not required by any UFFs, are "hooked-up" to the skeletal argument structure postcyclically. It is shown that postcyclic merger of adjuncts gives us a derivational way of capturing the argument/adjunct distinction, which is compatible with the minimalist spirit and thus theoretically desirable. I will then investigate Lebeaux's (1988, 1991) theory of phrase structure, which also claims that the argument/adjunct distinction should be made derivationally. It is shown that under Lebeaux's theory, adjuncts are <u>allowed</u> to be merged postcyclically, which is in contrast to our theory where adjuncts are <u>required</u> to be merged postcyclically.

2.6.1 Postcyclic Merger of Adjuncts

The ICP requires that UFFs should be checked immediately when they become accessible to a computation. In other words, when a UFF becomes accessible to a computation, we have to apply its checking operation before any other operations that manipulate terms. Hence, if we have two possible continuations at any stage of a derivation, one is triggered by an UFF and the other is not, then the ICP requires that we should always choose the former option. It then follows that typical adjuncts, whose merger is not triggered by any UFF, are forced to be merged postcyclically.

Let us consider in detail why this generalization follows from the ICP, taking (57) as an example:

(57) Bill said that John saw Suzy after he met Mary.

We will consider how to construct the structure of (57) under the interpretation where the adjunct *after he met Mary* modifies the embedded clause. (57) can be divided into two parts; what we call the main structure, i.e., *Bill said that John saw Suzy*, and what we call the adjunct, i.e., *after he met Mary*.

Let us first look at how to construct the adjunct, i.e., *after he met Mary*. We first select the verb *meet*, whose lexical entry is as follows:

(58) $\text{meet}_{[[\text{A}]\ \text{T, ACC}]}$

Recall that since the ϕ-features of verbs are optional, *see* does not have any ϕ-features when selected. After selecting *see*, we add ϕ-features to it, resulting in the following hierarchical structure among its features:

(59) $\text{meet}_{[[\text{A, 3SM}]\ \text{T, ACC, 3SF}]}$

At this point of the derivation, the THEME, Accusative Case, and [3rd person, singular, female] features of *meet* are accessible to the computation. Since these features are uninterpretable, the ICP requires that they should be checked immediately by selecting *Mary* and combining *meet* with *Mary*. Note that *Mary* has [3rd person, singular, female] as its intrinsic feature. Its THEME and Accusative Case features, being optional, are added after it is selected:

(60) $[_{\text{V}^{\text{max}}}\ \text{meet}_{[\text{A, 3SM}]}\ \text{Mary}_{[\text{T, 3SF}]}]$

The Accusative Case features of *meet* and *Mary*, being uninterpretable, are deleted when checked. While the THEME and ϕ-features of *Mary*, being interpretable, remain when checked, those of *meet*, being uninterpretable, are deleted.

At this stage of the derivation, the AGENT and [3rd person, singular, male] features of *meet* become accessible to the computation. These features, being uninterpretable, must be checked immediately by selecting *he* and combining *he* with *meet Mary* in accordance with the ICP. Note that *he* has the ϕ-features as its intrinsic property. Its AGENT feature, being optional, is added after it is selected:

(61) $[_{V^{max}}$ $he_{[A,\ 3SM]}$ [meet $Mary_{[T,\ 3SF]}$]]

While the AGENT and [3rd person, singular, male] features of *he*, being interpretable, remain when checked, those of *meet*, being uninterpretable, are deleted.

When we come to the stage of the derivation where (61) is constructed, the next step is to select the finite T. Recall that the finite T has the following hierarchical structure of features:

(62) $T_{[[D,\ NOM]\ V]}$

Since the V-feature of T, being uninterpretable, is accessible to the computation at this stage, the ICP requires that it must be checked immediately by merger of T with V^{max}, as shown below:

(63) $[_{T^{max}}$ $T_{[D,\ NOM]}$ $[_{V^{max}}$ $he_{[A,\ 3SM]}$ [meet $Mary_{[T,\ 3SF]}$]]]

At this point, the Nominative Case feature and D-feature of T, being uninterpretable, become accessible to the computation. They are checked by the raising of *he* to the Spec of T^{max}, conforming to the ICP:

(64) $[_{T^{max}}$ $he_{[A,\ 3SM]}$ $[_{T'}$ $[_{V^{max}}$ $he_{[A,\ 3SM]}$ [meet $Mary_{[T,\ 3SF]}$]]]]

Note that T can never be selected at the earlier stage of the derivation. If it were selected before V^{max} is constructed, its UFFs could not be checked immediately; this would lead to a violation of the ICP.

The next step must be to select the P *after*. The ICP requires that the selectional restriction feature of *after*, which states that it takes T^{max} as its complement, should be checked immediately by combining *after* with the T^{max} (64):

(65) $[_{P^{max}}$ after $[_{T^{max}}$ $he_{[A,\ 3SM]}$ $[_{T'}$ $[_{V^{max}}$ $he_{[A,\ 3SM]}$ [meet $Mary_{[T,\ 3SF]}$]]]]]

Note again that the P *after* cannot be selected at the earlier stage of the derivation due to the ICP. As shown above, we can construct the adjunct *after he met Mary* through checking the UFFs of the selected lexical items, conforming to the ICP.

Let us next look at how to construct the main structure, i.e., *Bill said that John saw Suzy*. In its derivation, we come to the stage where structure (66) is constructed. Although we do not look at how to construct this structure in detail, we can construct it through checking the UFFs of the selected items, conforming to the ICP:

(66) $[_{T^{max}}$ $John_{[A,\ 3SM]}$ $[_{T'}$ $[_{V^{max}}$ $John_{[A,\ 3SM]}$ [see $Suzy_{[T,\ 3SF]}$]]]]

Let us suppose that the adjunct clause (65), i.e., *after he met Mary*, should be adjoined to the T^{max} (66) for its proper interpretation at LF. Then, there are two possible continuations at this stage of the derivation: (i) Selection of C and merger of C with the T^{max} (66), or (ii) Merger of the adjunct clause *after he met Mary* with the T^{max} (66). I propose the

Earliness Principle (EP) on Select (67) and argue that it is selection of C that applies at this stage of derivation:[25]

(67) Earliness Principle (EP) on Select

Lexical items must be selected from a numeration (N) as early as possible.

The EP requires that lexical items should be selected from an N at the earliest possible stage of a derivation unless it would lead a derivation to be canceled due to a violation of the ICP.

In the present case, since merger of the adjunct clause with the main structure is not triggered by any UFF and thus not required by the ICP to be applied immediately, selection of C should apply because of the EP. When C is selected, the ICP requires that the next step should be to combine C with T^{max} in order to check the selectional restriction feature of C, as shown below:[26]

---

[25]Note that the EP on Select differs from Pesetsky's (1989) earliness condition, which states that filters should be satisfied as early as possible on the hierarchy of levels S-structure > LF. Pesetsky's earliness condition essentially requires that overt operations should be preferred over covert operations. The EP on Select, on the other hand, makes a decision among choices at each step of a derivation within the overt component.

[26]One might wonder what ensures that we select C, but not some other element, at this stage. Note that all selected lexical items are immediately subject to Merge except the first selected lexical item in a derivation. Ignoring the first selected item, therefore, we have a condition which states that whenever Select takes place, it must be immediately followed by Merge. It then follows that we are forced to select C at this stage. This is because if we selected any other lexical item than C, the selected lexical item could not be immediately merged with any structure. Hence, no other lexical item than C may be selected at this stage of the derivation.

One might argue that the EP is not local and its corresponding computational problem is computationally intractable, since we have to look at the next stage of a derivation to decide which lexical item is to be selected. We have to look ahead to see whether the selected item may be immediately merged with any structure. I argue, however, that it does not induce any computational intractability. Its corresponding problem has a solution which needs only a polynomial amount of work. Suppose that the EP requires us to select a lexical item from an N at a stage Σ of a derivation. Suppose further that there are *n* lexical items in N whose index is not reduced to zero at Σ. In order to decide which lexical item is to be selected, *n* lexical items each must be inspected to see whether it may be immediately merged with any structure if selected.

(68) [$_{C}$max that [$_{T}$max John$_{[A,\ 3SM]}$ [T [$_{V}$max John$_{[A,\ 3SM]}$ [see Suzy$_{[T,\ 3SF]}$]]]]]

The derivation proceeds further, checking the UFFs of the selected items, resulting in the following structure:

(69) [$_{T}$max Bill$_{[A,\ 3SM]}$ [T [$_{V}$max Bill$_{[A,\ 3SM]}$ [say [$_{C}$max that [$_{T}$max John$_{[A,\ 3SM]}$ [T [$_{V}$maxJohn$_{[A,\ 3SM]}$ [see Suzy$_{[T,\ 3SF]}$]]]]]]]]

It is important to point out that until this final stage of the derivation, the adjunct (65) has not been allowed to be merged with the main structure due to the ICP and the EP. At this final stage of derivation, we can combine the adjunct (65) with the main structure (69) by adjoining the former to the embedded T^{max} of the latter.

Like the ICP, the EP also receives strong conceptual support. Recall that a numeration N consists of lexical items and their index. Chomsky (1993, 1994, 1995) proposes a condition on N, which states that if the index is not reduced to zero at the interface levels, the derivation crashes. As argued by Collins (1997), however, this condition is global in nature, since it requires us to look at the interface levels in order to decide whether to apply Select at a certain stage of a derivation. Its corresponding optimization problem is therefore computationally intractable. It would then follow that this part of the language is unusable in practice, contrary to fact. Hence, there should exist a local "heuristic algorithm" ("computational trick") which brings about an

---

If the review of one lexical item requires one step of work, this problem requires *n* steps of work. Since this problem has a solution which only requires a polynomial amount of work, it belongs to Class P. Hence, it is computationally tractable.

approximate solution to this computationally intractable optimization problem. I argue that the EP serves for this purpose.

Let us first consider the local property of the EP. Note first that the EP is overridden by the ICP. Recall that the EP requires that we should exhaust an N as early as possible unless it would lead a derivation to be canceled. If the ICP requires some other operation to apply at a certain stage of a derivation in order to check a UFF, the selection of a lexical item cannot be applied at that stage. This is because if we applied the select operation, the derivation would be canceled due to a violation of the ICP. It then follows from the EP together with the ICP that at each step of a derivation, a lexical item should be selected from an N unless there exists any UFF accessible to a computation. To be specific, suppose that we come to a stage Σ of a derivation where we have an option of applying Select. If there is a UFF which is accessible to the computation at Σ, we should not apply Select but rather apply an operation to check the UFF. If there is no UFF which is accessible to the computation at Σ, we should apply Select. Hence, the EP is local in that when we come to Σ where we have an option of applying Select, it can determine whether to apply Select only on the basis of information available at Σ.

The EP also brings about an approximate solution to the computationally intractable optimization problem induced by the interface condition on N. In most cases, the EP gives us the same answer to a computational problem as the interface condition on N. In other words, the EP only requires us to apply Select when the interface condition on N would also require us to do so. For example, suppose that we have constructed T^{max}. Here, we have an option of selecting C from the N. The EP requires us to apply that operation, and so would the interface condition on N. Hence, in such cases, the EP brings about a perfect solution to the computationally intractable problem induced by the interface condition on N. There are, however, cases where the EP does not give us the same answer to a computational problem with the interface condition on N. For example, suppose that as in the derivation of (57) discussed above, we come to a stage Σ where there are two independent syntactic objects, T^{max} and an adjunct which modifies the T^{max}. We have two options at Σ. We can select C and merge it with T^{max}. Alternatively, we can merge the adjunct with T^{max}. The interface condition on N allows us to apply either of these operations at Σ as long as C is eventually selected before the derivation reaches the interface levels. The EP, on the other hand, only allows us to select C at Σ, excluding merger of the adjunct with T^{max}. In such cases, the EP does not give us the same answer to a computational problem with the interface condition on N. Hence, the solution which the EP brings about to the computationally intractable problem induced by the interface condition on N is only approximate.

In conclusion, the EP on Select receives conceptual support from the fact that it serves as a local "heuristic algorithm" ("computational trick") which gives us an approximate solution to the computationally intractable problem induced by the condition on N. It should also be noted that if we conform to the EP during a derivation, it never happens that the derivation reaches the interface levels with any lexical items being left unselected from N. The condition on N is therefore always satisfied and virtually plays no role once we assume the EP.

To summarize, if we conform to the ICP and the EP during derivations, adjuncts, whose merger is not triggered by any UFF, are required to be merged postcyclically.

2.6.2 A Derivational Notion of Adjuncts

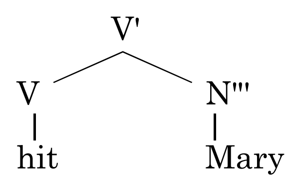
It has been assumed in the pre-minimalist period (see, among others, Chomsky (1972) and Jackendoff (1977)) that the argument/adjunct distinction is made representationally. Given the X-bar theory, while arguments are attached under X'-level, adjuncts are attached under higher-bar levels. Let us consider the following:

(70) a. John hit Mary

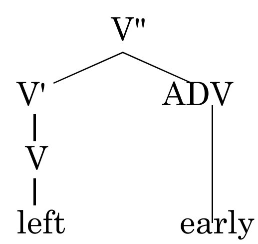
b. John left early

While *Mary* in (70a) is the argument of the verb *hit*, *early* in (70b) is an adjunct. Under Jackendoff's (1977) X-bar theory where the uniform three-level hypothesis is adopted, for instance, (70a) and (70b) are represented as in (71a) and (71b), respectively, with the irrelevant parts being ignored:

(71) a.



b.



While *Mary* in (70a), being an argument, is attached under V', *early* in (70b), being a restrictive modifier and thus an adjunct, is attached under V''.

Such a representational argument/adjunct distinction, however, is no longer available in the MP. Recall that the MP requires that phrase structures should be "bare." Crucially, neither non-branching nodes like V' in (71b) nor bar-levels in the sense of the X-bar theory are allowed any more. Hence, we need an alternative way of making the argument/adjunct distinction which is compatible with the minimalist spirit.

The ICP coupled with the EP gives us a minimalist way of capturing the argument/adjunct distinction. Our theory claims that arguments are merged cyclically whereas adjuncts are merged postcyclically. Arguments and adjuncts are therefore distinguished by means of derivational terms instead of representational terms. Our derivational notion of adjuncts is theoretically desirable. This is because it does not rely on the notion of non-branching nodes or bar-levels and thus can be accommodated under the bare phrase structure in accordance with the minimalist spirit. I will argue in the following chapters that apart from this conceptual support, there is strong empirical support in favor of this derivational notion of adjuncts.

2.6.3 Lebeaux's (1988, 1991) Analysis

Lebeaux (1988, 1991) also argues for a derivational distinction between arguments and adjuncts. He proposes a heterogeneous licensing of phrase structures within the framework of the EST, arguing that elements in adjunct-of relations are licensed in a different way from those in argument-of relations. Following Chomsky (1981), Lebeaux assumes that the Projection Principle holds at all levels of representation. Then, elements in argument-of relations must be present at all levels of representation, crucially at D-structure. Elements in adjunct-of relations, not being subject to the Projection Principle, may not be present at D-structure. They may be added by the operation called Adjoin-α in the course of derivations.

Our ICP analysis and Lebeaux's Adjoin-α analysis agree on the view that there is an argument/adjunct asymmetry with respect to merger. They differ, however, as to obligatoriness/optionality of postcyclic merger of adjuncts. Our theory claims that adjuncts <u>must</u> be merged after argument-of relations are established. Lebeaux's theory, on the other hand, claims that adjuncts <u>may</u> be merged after argument-of relations are established. In the following chapters, I will argue that our theory should be preferred over Lebeaux's theory, presenting various empirical facts which only follow from our theory but not from Lebeaux's theory.[27]

2.6.4 Strict Cyclicity and Postcyclic Merger of Adjuncts

It has been proposed by, among others, Chomsky (1993), Collins (1997), Kitahara (1995, 1997), and Watanabe (1995) that merger should be subject to strict cyclicity. Since our analysis allows countercyclic merger of adjuncts, this is a good place to consider empirical evidence which has been presented in support of cyclic application of merger. It is

---

[27]Fukui (1986) also argues for the argument/adjunct asymmetry with respect to structure building within the EST framework. As a consequence of his theory of phrase structure, he claims that non-thematic elements appear after every thematic element has been introduced within a clause. Our analysis agrees with Fukui's in claiming that non-thematic elements are forced to be introduced into derivations after thematic elements. In this respect, we can say that our analysis counts as a minimalist extension of Fukui's analysis. Our analysis, however, differs from Fukui's in that the latter only regulates the order of merger between arguments and adjuncts within a clause. In complex sentences, Fukui's analysis allows the embedded adjuncts to be merged earlier than the matrix arguments whereas our analysis does not.

shown that empirical arguments for cyclic merger are only based on merger of arguments but not on merger of adjuncts. Hence, they can be accommodated under our analysis, where only adjuncts, but not arguments, are allowed to be merged countercyclically. I will also argue that cyclic merger of arguments straightforwardly follows from the ICP coupled with the EP without recourse to any extra device.

Let us consider empirical arguments which have been presented in favor of cyclic merger. First, as pointed out by Collins (1997), under the AGRo or small *v* theory of clausal structure, if merge were allowed to apply countercyclically, then examples like (72) would be acceptable:

(72) *Mary John hit
'John hit Mary'

Suppose that *Mary*, which has a Nominative Case feature, is generated in the complement position of the verb *hit*, yielding *hit Mary*. Suppose further that before *John* is inserted in the Spec of VP, *hit Mary* is merged with AGRo or *v* and the resultant structure is merged with T. Then, since *John* is not present, the raising of *Mary* to the Spec of TP does not violate the MLC. After the raising of *Mary* to the Spec of TP, *John*, which has an Accusative Case feature, is inserted in the Spec of VP countercyclically. Its Accusative Case feature moves to AGRo or *v* to be checked off in the covert component. Hence, if we allowed countercyclic merger, examples like (72) would be acceptable, contrary to fact.

Second, the relativized minimality effects would not be explained by the MLC if countercyclic merger were allowed. The MLC is intended to subsume the relativized minimality effects, i.e., superraising (73a), the Wh-island Constraint (73b), and the Head Movement Constraint (73c):

(73) a. ***John** seems that it is certain ***t*** to be here

b. ***how$_j$** did John wonder **what$_i$** Mary fixed ***t$_i$ t$_j$***

c. ***fix** John can ***t*** the car

If countercyclic merger were allowed, we would lose an MLC account of the relativized minimality effects. In (73a), it would be possible to raise *John* directly to the matrix Spec of TP and later insert *it* in the embedded Spec of TP countercyclically. In (73b), we could raise *how* directly to the matrix Spec of CP. Then, we could countercyclically insert the embedded C, which would attract *what*. In (73c), it would be possible to raise *fix* directly to C and then insert *can* countercyclically. Hence, if we allowed countercyclic merger, examples like (73) would be generated without violating the MLC.

Third, if we allowed countercyclic merger, examples like (74) could be generated without violating the Subject Condition:

(74) *who$_i$ was [pictures of ***t$_i$***]$_j$ taken t_j by Bill

If countercyclic merger were allowed, it would be possible to move *who* to the Spec of CP first and then move the remaining NP *pictures of t* to the Spec of TP countercyclically. In this derivation, there would be no violation of the Subject Condition. Examples like (74) would be acceptable, contrary to fact.

It should be noted, however, that these arguments only provide empirical support for cyclic merger of arguments but not cyclic merger of adjuncts. No empirical arguments have ever been presented in favor of cyclic merger of adjuncts. Recall that under our theory, arguments are required to be merged cyclically, though adjuncts, which are required to be merged postcyclically, may be merged countercyclically. The empirical arguments which have been presented in favor of strict cyclicity therefore are compatible with our theory of phrase structure. Since our theory of phrase structure ensures cyclic merger of arguments, it can subsume the principles or conditions which are intended to derive strict cyclicity as far as the overt component is concerned.[28]

2.7 Concluding Remarks

To recapitulate this chapter, I have first introduced Chomsky's (1993) interface condition approach to the elimination of D-structure and S-structure. It was shown that although Chomsky's interface condition approach contributes to the elimination of these two theory-internal linguistic levels and thus counts as a step toward the goal of the MP, it raises a serious conceptual difficulty, i.e., it sneaks in an element of globality into the theory of language. In order to solve this conceptual problem, I have developed Chomsky's (1995) derivational interpretation of strong features, proposing the ICP on UFFs. The ICP states that UFFs must be checked immediately when they become accessible to a computation. It was shown that the ICP approach enables us to capture the D-structure and S-structure properties in the EST in a local fashion without inducing any globality. Finally, I have discussed the consequence of the ICP regarding the composition of phrase structure. It was shown that the ICP coupled with the EP on Select gives rise to the asymmetry with the composition of phrase structure. I have also shown that the ICP and the EP receive strong conceptual support. I have argued that they are "heuristic algorithms" ("computational tricks") which

---

[28]Bures (1992), Jonas and Bobalijk (1993), Branigan and Collins (1993), and Watanabe (1995) present arguments for the view that cyclicity is relevant not only in the overt component but also in the covert component. If they are correct, the ICP coupled with the EP cannot completely subsume strict cyclicity.

give us approximate solutions to the computationally intractable optimization problems induced by the interface conditions on UFFs and Ns. It was shown that our analysis supports the language design that language is fundamentally global and its corresponding optimization problem belongs to Class NP. For its usability in practice, there are language-specific computational devices like the ICP on UFFs and the EP on Select which reduce its fundamental globality to local properties.

In the rest of this thesis, I will argue that apart from this conceptual advantage, our theory of phrase structure also receives strong empirical support from a wide range of facts pertaining to movement constraints, scrambling in Japanese, the distribution of *wh*-elements in-situ, and reconstruction effects.

CHAPTER 3
LOCALITY ON MOVEMENT

3.0 Introduction

The previous chapter has proposed the Immediate Checking Principle (ICP), which requires that uninterpretable formal features (UFFs) should be checked immediately when they become accessible to a computation. It was shown that the ICP subsumes the properties of the two theory-internal linguistic levels in the EST. I have argued that the ICP approach is conceptually more desirable than Chomsky's (1993, 1995) approaches in that the former is local while the latter is global. I have also proposed the Earliness Principle (EP) on Select, which states that lexical items must be selected from a numeration (N) as early as possible. It was shown that the ICP coupled with the EP gives rise to the asymmetry regarding the composition of phrase structure. While arguments are required to be merged cyclically, adjuncts are required to be merged postcyclically. The chapters to follow explore empirical justification for our theory of the composition of phrase structure. I will argue that it receives strong empirical support from a wide range of facts pertaining to movement constraints, scrambling in Japanese, the distribution of *wh*-elements in-situ, and reconstruction effects. These empirical facts also lend support for the view that language is derivational in character rather than in the representational mode. This is because they can only be given a principled minimalist account by making use of information which is available at an intermediate stage of a derivation but later "wiped-out" by an operation before the output representation. Among these empirical facts, I will first consider locality conditions on feature-driven A'-movement. Specifically, this chapter considers the "domain barrier" effects with feature-driven A'-movement, i.e., the Complex NP Constraint, the Adjunct Condition, the Subject Condition, and the non-bridge verb condition, and the ban against feature-driven extraction out of phrases which have undergone feature-driven A'-movement. I will argue that these locality conditions, which are left unexplained under the MP, straightforwardly follow from our theory of phrase structure. It is pointed out that our approach to these locality conditions diverges from all previous approaches in claiming that they should not follow from restrictions on movement but from restrictions on merger.

The organization of this chapter is as follows. Section 3.1 reviews previous generative analyses of locality restrictions on feature-driven A'-movement, taking overt wh-movement in English as an example. Section 3.2 deals with the "domain barriers," which have not been given any principled account under the MP. It is shown that if our theory of phrase structure is adopted, the "domain barrier" effects straightforwardly follow. Section 3.3 considers another locality condition on feature-driven A'-movement, i.e., no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. I will argue that this locality condition also follows from our theory of phrase structure. Section 3.4 discusses locality restrictions on adjunct feature-driven A'-movement. I will argue that they can also be accounted for by our theory of phrase structure. Section 3.5 makes concluding remarks.

3.1 Previous Analyses of Locality on Movement

3.1.1 Standard Theory Approaches

Locality conditions on movement have been one of the central issues for generative grammar. A lot of discussion has especially been around regarding locality conditions on so called feature-driven A'-movement like overt wh-movement and null operator movement. Feature-driven A'-movement is prima facie unbounded in the sense that it can extract an element out of deeply embedded phrases. In fact, however, A'-movement is not unbounded, but is restricted by some constraints on movement. It is Chomsky (1964) which marks the first substantial step toward establishing a general locality condition on movement. Chomsky (1964) proposes the A-over-A principle. The A-over-A principle prohibits movement operations from applying to phrases contained within those of the same category. Although the A-over-A principle is conceptually attractive in that it attempts to reduce all locality conditions to one explanatory principle, it faces serious empirical problems. It has been pointed by, among others, Chomsky (1964, 1968) and Ross (1967) that the A-over-A principle is too strong as well as too weak.[1]

As an alternative to the A-over-A principle, Ross (1967) proposes island constraints on movement. Island constraints informally state that no elements can be extracted out of domains called "islands." Island

---

[1]See Fukui (1997) for a minimalist account of island constraints based on the insight of the traditional A-over-A principle. As Fukui himself admits, however, the classical problems of the traditional version of the A-over-A principle are still left unexplained under his analysis.

constraints which are proposed in Ross (1967) and later works are shown below with their examples:[2]

(1) Complex NP Constraint (CNPC)

a. Relative Clauses

*?**who** do you like [books that criticize ***t***]

b. Non-relative Complex NPs

*?**what** did you study [the evidence that Harry stole ***t***]

(2) Subject Condition

a. Subject NPs (cf. Chomsky (1973))

*?**who** did [pictures of ***t***] please you

b. Sentential Subjects

*?**what** did [that John saw ***t***] surprised Mary

(3) Adjunct Condition

a. Adjunct Clauses

*?**who** do you get jealous [because I spoke to ***t***]

b. Adverbial PPs

*?**which class** did he fall asleep [during ***t***]

c. PP Modifiers of Nouns

*?**which table** did you buy [the pictures [on ***t***]]

d. Secondary Predicates

*?**what** did John arrive yesterday, [sad about ***t***]

(4) Non-bridge Verb Condition (cf. Erteschik-Shir (1973))

?**who** did John whisper [that he saw ***t***]

---

[2]Ross (1967) also proposes the left-branching condition and the Coordinate Structure Constraint. We will put aside these constraints, since the nature of these phenomena is not entirely clear at this point.

(5) Wh-island Constraint (cf. Chomsky (1977))

*?**what$_i$** did you decide [**who$_j$** to persuade $\boldsymbol{t_j}$ to buy $\boldsymbol{t_i}$]

Although these island constraints are empirically adequate, they are conceptually unattractive. This is because the island constraints are construction-specific and thus just observational generalizations rather than abstract explanatory principles.[3]

---

[3]There are some languages which prima facie do not fully obey these island constraints. It has been claimed by, among others, Rizzi (1982), Torrego (1986),and Sportiche (1981) that French does not exhibit any subject condition effects:

(i) le diplimate **dont** [[la secrétaire ***t***] t'a téléphoné]
'the diplomat of whom a secretary called you'

In the relative clause (i), extraction takes place from within the subject, but the result is acceptable. It has also been claimed by, among others, Allwood (1976), Erteschik (1973), Engdahl (1980b, 1982), Hellan and Christensen (1986), and Taraldsen (1978, 1982) that scandinavian languages do not obey the Wh-island Constraint or the CNPC, which is exemplifed by the following Norwegian examples:

(ii) **denne boken** vet vi hvem som har skrevet ***t***
this book know we who that has written
'this book that we know who has written'

(iii) a. **denne boken** kjenner vi den mannen som har skrevet ***t***
this book know we the man that has written
'this book that we know the man who has written'

b. **denne boken** går det rykter om at du har lest ***t***
this book go it rumors about that you have read
'this book that the rumor that you have read is around'
(Hellan and Christensen 1986:4)

Although extraction takes place from within the Wh-island in (ii) and the complex NP in (iii), these examples are all acceptable.

The nature of these phonomena, however, are not entirely clear at this point. Tellier (1990) discusses the Subject Condition in French, pointing out that the immunity from the Subject Condition is only observed with *dont*. The following examples show that French in fact obeys the Subject Condition:

(iii) a. ?*le diplomate **de qui** [[la secrétaire ***t***] t'a téléphoné]
'the diplomat of whom the secretary called you'

b. ?***de qui** [[la secrétaire ***t***]t'a-t-elle téléphoné]
'of whom did the secretary call you'

(Tellier 1990:307)

Turning to the lack of the Wh-island and the CNPC in Scandinavian languages, relativization in Scandinavian languages may not involve movement, as argued for Japanese relative clauses by some literatures (see, among others, Kuno (1973)). We will therefore put aside these phenomena in the discussion to follow.

3.1.2 EST Approaches

Since the advent of the island constraints, a considerable number of studies have been made on reduction of these island constraints to an explanatory principle within the framework of the EST (see, among others, Aoun (1985a, 1985b), Cattell (1976), Chomsky (1973, 1976, 1981, 1986a), Cinque (1990), Culicover and Wilkins (1984), Depréz (1989), Hegarty (1991), Huang (1982), Koster (1978, 1987), Kayne (1983), Lasnik and Saito (1992), Manzini (1992), and Pesetsky (1982)). All of these theories, however, crucially make use of notions which are no longer available in the MP where no structural relations are allowed to be invoked other than those required by BOCs and those induced in a natural way by the derivation itself like local relations to a head in its minimal domain. Hence, these EST analyses are incompatible with the MP. Let us quickly review Chomsky's (1986a) and Kayne's (1983) locality theories as representatives among locality theories in the EST.

Chomsky (1986a) proposes the notion of barrier, which is instrumental for the bounding theory as well as the government theory. The notion of barrier is defined based on the notion of Blocking Category (BC). Maximal projections which are not L-marked constitute BCs for their containing categories. Maximal projections which are L-marked as sisters to θ-role assigners, on the other hand, do not constitute BCs for their containing categories. Any category except IP that is a BC becomes a barrier. Barrierhood is also obtained through the inheritance mechanism. Any category that immediately dominates a BC inherits barrierhood. Movement is constrained by two locality restrictions, i.e., antecedent government and the subjacency condition, which are both defined based on the notion of barrier. According to the antecedent government requirement, a trace left by movement may not be separated by its antecedent by any barriers. If an LF representation contains any trace which violates the antecedent government requirement, then the derivation is excluded by the Empty Category Principle (ECP). The subjacency condition also requires that a trace should not be separated by its antecedent by any barriers. Subjacency violations, however, are less severely deviant than ECP violations, which explains the argument/adjunct asymmetry with the degree of unacceptability of island violations. Although Chomsky's locality theory can subsume the island constraints, it crucially makes use of notions which are not available in the MP like government, antecedent government, L-marking, BC, and barrier. Hence, Chomsky's locality theory is incompatible with the MP.

Kayne (1983) proposes the connectedness condition (CC). The basic idea of the CC is that syntactic dependencies must be mediated by a path of nodes that link the related elements. Syntactic dependencies are only licit if dependents and their antecedents are members of a common well-formed subtree. If we restrict our discussion to a derivation where only one instance of movement is involved, the CC informally states that a trace must have a c-commanding antecedent that is contained within a g-projection of its structural governor. In other words, a trace and its antecedent must be members of a common subtree that is a g-projection of the structural governor of the trace. It then follows that a trace which does not have any structural governor may not be unbounded. Outside the first g-projection, which is determined by X'-principles and structural government, further g-projections are licensed by being in a canonical government configuration with a head. In languages like English where V governs NP to its right, [$_Y$ W Z] is a canonical government configuration, with Z a maximal projection. Hence, in those languages, a trace may not be unbounded within a maximal projection that constitutes a left branch. Kayne's locality theory is conceptually attractive in that it subsumes the subjacency condition and the ECP under the CC. His theory, however, crucially makes use of notions which are no longer available in the MP like g-projection, government, canonical government, and structural government. Hence, it is incompatible with the MP.[4]

3.1.3 Minimalist Approaches

The previous subsection has reviewed the EST approaches to the island constraints. It was pointed out that although they are conceptually more attractive than Ross (1967), they crucially make use of notions which are no longer available in the MP. This subsection reviews previous minimalist approaches to the locality restrictions on feature-driven A'-movement.

Among the island constraints, the Wh-island Constraint has been the main object of study since the advent of the MP. One can cite Chomsky's (1993, 1994, 1995) approach to the Wh-island Constraint. Based on the insight provided by Rizzi's (1990) Relativized Minimality, Chomsky (1993, 1994, 1995) reformulates the Wh-island Constraint as a Minimal Link Condition (MLC), which is one of the economy conditions in Chomsky (1993, 1994) and later incorporated into the definition of

---

[4]Although Kayne's (1983) CC cannot be accommodated under the MP as it is, it can be translated into minimalist terms. One can safely say that the locality theory to be proposed here is a minimalist reinterpretation of Kayne's CC. Leaving the details to the discussion later in this chapter, I argue that structures belonging to different g-projections in the sense of Kayne have not been merged with each other when feature-driven A'-movement takes place. Hence, movement cannot take place across more than one g-projections.

Attract-F in Chomsky (1995).[5] Only few attempts, however, have so far been made at providing a minimalist account of the other island constraints, i.e., the Complex NP Constraint, the Subject Condition, the Adjunct Condition, and the non-bridge verb condition. We will hereafter call these island constraints "domain barriers" as opposed to the relativized minimality effects including the Wh-island Constraint. In the subsections to follow, I will review two previous studies which attempt to give a minimalist account of the "domain barriers," pointing out that they are conceptually problematic.

3.1.3.1 Takahashi (1994)

Takahashi (1994) gives a minimalist account of the island constraints, claiming that the island constraints should be subsumed under the Shortest Movement Condition (SMC) and the Uniformity Corollary on Adjunction (UCA):

(6) The Shortest Movement Condition (SMC)
Make the shortest movement.

(Takahashi 1994:8)

(7) The Uniformity Corollary on Adjunction (UCA)
If $(\alpha_1, ..., \alpha_n)$ is a chain $(1 \leq n)$, then for any i $(1 \leq i \leq n)$, $P(\alpha_i)$.

(Takahashi 1994:20)

---

[5]See Ishii (1997) for the different view that the Wh-island Constraint should not be subsumed under the MLC but under the "no extra baggage" condition. The MLC is instead responsible for the "crossing" constraint effects. It is shown that this analysis gives us a minimalist account of the contrast in acceptability between the "crossing" constraint and superraising effects, on the one hand, and the Wh-island effects, on the other.

In (7), P(α) means that α has property P, which is the property of being adjoined to by X' or being not adjoined to. (7) virtually ensures that adjunction is impermissible to heads of nontrivial chains. This is because if something is adjoined to the head of a chain, then the head becomes distinct from the rest of the chain regarding property P. As for the Wh-island Constraint, Takahashi essentially follows the relativized minimality account advocated by Chomsky (1993, 1994, 1995), Chomsky and Lasnik (1993) and Rizzi (1990), claiming that it is subsumed under the SMC. Hence, not much has to be said about this. The discussion to follow therefore only concerns his analysis of the "domain barriers."

Let us look at how Takahashi's (1994) analysis works, considering the Subject Condition first:

(8) *?**who** did [pictures of ***t***] please you

Since he adopts the VP-internal subject hypothesis, the subject *pictures of who* originates in the Spec of VP and then moves to the Spec of IP, as shown below:

(9) $[_{C'}$ C $[_{IP}$ $\mathbf{[_{D'}\ [_{NP}}$ **pictures of who]]$_j$** $[\boldsymbol{t_j}$ please you]]]

The *wh*-phrase *who* moves from within Spec of IP to the Spec of CP. The SMC requires that *who* should move in a successive cyclic fashion, landing at the nearest possible target at each step of movement, as shown below:

(10) $[_{CP}$ $\mathbf{who_i}$ did $[_{IP}$ $\boldsymbol{t'''_i}$ $[_{IP}$ $[_{DP}$ $\boldsymbol{t''_i}$ $[_{D'}$ $[_{NP}$ $\boldsymbol{t'_i}$ $[_{NP}$ pictures of $\boldsymbol{t_i}$]]]]$_j$ $[t_j$ please you]]]]

In (10), *who* first adjoins to NP and then moves to the Spec of DP. Since Takahashi assumes that specifiers are created by adjunction, the latter step actually involves adjunction to D'. It then adjoins to IP and finally moves to the Spec of CP, which counts as adjunction to C'. This derivation, however, is excluded by the UCA, which applies at LF. This is because *who* is adjoined to D', the head of the non-trivial chain of the subject raising (*pictures of* t_i, t_j). The subject raising chain is therefore non-uniform; this violates the UCA. Hence, the Subject Condition can be accounted for by the SMC and the UCA.

Takahashi furthermore extends the UCA (7), claiming that it applies to coordinate structures as well as chains:

(11) The Uniformity Corollary on Adjunction (Revised)
If (α_1, ..., α_n) is a chain ($1 \leq n$) or α_1, ..., α_n are conjuncts of a coordination, then for any i ($1 \leq i \leq n$), P(α_i).

(adapted from Takahashi 1994:25)

He argues that the basic idea of the UCA is that elements belonging to a group should respect uniformity concerning the adjunction property. Elements of a chain count as belonging to a group and thus chains are subject to the UCA. He argues that it is intuitively correct to say that conjuncts of a coordination also count as belonging to a group. It is then plausible to assume that the UCA applies not only to chains but also to coordinate structures. If this conjecture is correct, then the Coordinate Structure Constraint also follows from the SMC and the UCA:

(12) *?**which book** was John reading ***t*** and driving cars

In the derivation of (12), before wh-movement takes place, the VP *reading which book* and the VP *driving cars* are conjoined. The SMC requires that *which book* adjoins to the former VP on its way to the Spec of CP. Adjunction to a conjunct, however, violates the UCA (11).

Extending this analysis of the Coordinate Structure Constraint, he argues that the Adjunct Condition is also subsumed under the SMC and the UCA. Following Higginbotham (1985), he claims that adjuncts involve coordination. (13a), for example, is assigned semantic representation (13b):

(13) a. John walks slowly

b. $\exists e$ [Walk (John, *e*) & Slow (*e*)]

In (13b), the adjunct *slowly* is coordinated with the main predicate. It then follows that the Adjunct Condition can be accounted for in the same fashion as the Coordinate Structure Constraint:

(14) *?**who** do you get jealous [because I spoke to ***t***]

In the derivation of (14), the SMC requires that *who* adjoins to the adjunct on its way to the Spec of CP. Adjunction to the adjunct, however, is prohibited by the UCA (11), since the adjunct counts as a conjunct of the coordinate structure. Hence, the Adjunct Condition follows.

He furthermore argues that the CNPC can be accounted for in the same way as the Adjunct Condition. Since relative clauses are adjuncts in nature, the relative clause case of the CNPC is subsumed under the Adjunct Condition. Concerning non-relative complex NPs, Takahashi (1994) assumes following Grimshaw (1990) and Stowell (1981) that noun complement clauses are appositive and thus regarded as adjuncts. Then, the non-relative case of the CNPC is also subsumed under the Adjunct Condition.

Although Takahashi's analysis is attractive in reducing all the island constraints including the Coordinate Structure Constraint to the two general principles of grammar, i.e., the SMC and the UCA, his analysis is theoretically problematic. His analysis of the Adjunct Condition (and therefore the CNPC) crucially makes use of the idea that adjuncts involve coordination. As Takahashi himself admits, however, representations like (13b) where adjuncts involve coordination should not count as LF representations but as semantic representations. The information that adjuncts involve coordination can never be read off in the syntactic component. It is only available in the semantic component. I argue that even if it is true that adjuncts semantically involve coordination, his UCA analysis of the Adjunct Condition is untenable. This is because the autonomous syntax thesis requires that the UCA, being syntactic in nature, should not refer to information only available in the semantic component.

3.1.3.2 Toyoshima (1997)

Toyoshima (1997) gives a minimalist account of the locality restrictions which were captured by the Condition of Extraction Domain (CED) within the EST. The CED was originally proposed by Huang (1982) and intended to subsume the Subject Condition and the Adjunct Condition. Toyoshima proposes the Derivation Condition on Extraction Domain (Derivational CED), claiming that it provides a minimalist account for the Adjunct Condition as well as the Subject Condition:

(15) Derivational Condition on Extraction Domain (Derivational CED)
A feature F is accessible for Attract triggered by another feature F' iff F and F' are both introduced as a part of the same process P.
(Toyoshima 1997:510)

The notion of process is defined as below:

(16) Process

A process P is a sequence of operations (OP), Merge or Attract, such that if OP (α, β) = γ is in P, then Merge (δ, γ) is also in P.

(Toyoshima 1997:511)

Let us look at how the Derivational CED works, taking the Subject Condition as an example:

(17) *?**who** did [pictures of ***t***] please you

According to Toyoshima's analysis, the derivation of (17) would proceed in the following way:

(18) ia. Merge (*please*, *you*)iia. Merge (*of*, *who*)

= *please you* = *of who*

iib. Merge (*pictures, of who*)

= *pictures of who*

ib. Merge (*pictures of who*, *please you*)

= *pictures of who please you*

ic. Merge (T, *pictures of who please you*)

= T *pictures of who please you*

id. Attract (T, *pictures of who*)

= ***[pictures of who]*** T ***t*** *please you*

ie. Merge (C, ***[pictures of who]*** *T* ***t*** *please you*)

= C ***[pictures of who]*** T ***t*** *please you*

The next step of this derivation should be to check the strong Q-feature of C through attracting the Q-feature of *who* in the subject phrase. The derivational CED, however, claims that this Attract is illegitimate. Derivation (18) involves two different processes, i.e., (i) and (ii). While C is introduced in process (i), *who* is introduced in process (ii). According to the derivational CED, therefore, the Q-feature of *who* is not accessible for Attract triggered by the Q-feature of C. There is no way of checking the strong Q-feature of C; this derivation crashes. Hence, the Subject Condition follows. The Adjunct Condition can be accounted for in a similar fashion. Since a *wh*-phrase in an adjunct is introduced in a different process from C with a strong Q-feature, the former is inaccessible for Attract triggered by the latter. Hence, the strong Q-feature remains at the interface; the derivation crashes.

Toyoshima's derivational CED, however, has a conceptual problem. The derivational CED raises the problem of globality. Let us consider derivation (18) as an example. When we reach the stage of the derivation where the strong Q-feature of C is to be checked, it is necessary to know whether C and *who* are introduced in the same process or not. This information, however, can never be read off from the phrase structure at this stage of the derivation. In other words, although C and *who* are introduced in different processes at the earlier stages of the derivation, this information is no longer available at the stage where the Q-feature of C is to be checked. Under Toyoshima's analysis, therefore, we need global considerations, looking back to see whether the two elements are introduced in the same process or not. This is theoretically undesirable, since global considerations necessarily induce computational intractability.[6]

---

[6]Toyoshima's analysis also has empirical problems. Let us consider wh-movement out of a passive subject, taking (i) as an example:

(i) *?**who$_i$** were **[pictures of t_i]$_j$** stolen t_j

It has been assumed that (i) is excluded by the Subject Condition. According to Toyoshima, therefore, the deviancy of (i) should be accounted for by the Derivational CED, which is intended to subsume the Subject Condition. The Derivational CED, however, would wrongly predict that examples like (i) are acceptable. In (i), although

To summarize subsection 3.1.3, I have reviewed Takahashi's (1994) and Toyoshima's (1997) minimalist approaches to the island constraints. It was shown that they are confronted with conceptual problems. In the rest of this chapter, I will only consider the "domain barriers," which have not been given any principled account under the MP, putting aside the relativized minimality effects. The organization of the rest of this chapter is as follows. Section 3.2 argues that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure. In section 3.3, I will extend our analysis of the "domain barriers" to another locality phenomenon, i.e., no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. It is shown that this locality condition also follows from our theory of phrase structure. Section 3.4 considers locality restrictions on adjunct A'-movement. It is shown that they also follow from our theory of phrase structure.

---

pictures of who is in the Spec of TP when wh-movement applies, it is originally introduced in the complement position of the verb *steal*. Hence, C and *who* are introduced in the same process. The Derivational CED would claim that the Q-feature of *who* is accessible for Attract triggered by the Q-feature of C, which is undesirable. The same problem arises with wh-movement out of the subject of an unaccusative verb.

In order to solve this problem, Toyoshima claims that overt attraction of a complex category requires what he calls a subparallel process, which "constructs" a complex structure internally and in parallel to a given process. In (i), the subject of the passive verb, being a complex category, undergoes overt attraction. It therefore involves a subparallel process. The Q-feature of the *wh*-phrase within the subject is not accessible to the Q-feature of C due to the subparallel process. The subparallel process, however, is not independently motivated and only needed to solve this problem.

[6]Recall that Chomsky's (1995) definition of Attract/Move requires that what is raised should be c-commanded by its target. Under our analysis where Attract/Move is reinterpreted as consisting of Copy and Merge, Copy is subject to the c-command requirement. The c-command requirement on Copy, however, is not included in the definition of the last resort condition (33), since it is most likely to be derived from some general principles of UG.

3.2 An Account of the "Domain Barriers"

In this section, I will argue that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure. As argued in the previous chapter, the ICP coupled with the EP ensures that adjuncts, whose merger is not triggered by any UFF, are forced to be merged postcyclically. It is shown that the "domain barriers" are exactly those elements which are forced to be merged postcyclically. It then follows that when we come to a stage where the strong Q-feature of C becomes accessible to a computation, the "domain barrier" has not been merged with the main structure which contains the C. The main structure and the "domain barrier" each constitute an independent syntactic object at that stage. No Q-feature inside the "domain barrier" therefore can be attracted by the strong Q-feature of C. This violates the ICP; the derivation is canceled. In the following, we will look at in detail how this analysis works using concrete examples.

3.2.1 The Complex NP Constraint (CNPC)

We will begin by considering the relative clause case of the CNPC, taking (1a) (repeated here as (19)) as an example:

(19) *?**who** do you like [books that criticize ***t***]

This structure can be divided into two parts: the main structure *who do you like books* and the relative clause *that criticize t*. The relative clause can be constructed by checking the UFFs of the selected items in accordance with the ICP and the EP. Given the copy theory of movement and the empty operator movement analysis of English relative clauses, the resultant structure is as follows. Among formal features, we will only pay attention to Q-features here and in the relevant structures to follow, with irrelevant formal features being ignored:

(20) $[_{C^{max}}$ *Op* [C $[_{T^{max}}$ *Op* [T $[_{V^{max}}$ *Op* [criticize $who_{[Q]}$]]]]]]

Turning to the main structure of (19), we first select the verb *like*. The ICP then requires that we should immediately check the internal thematic feature and Accusative Case feature of *like*, since these features, being uninterpretable, are accessible to the computation. In order to check these features, we select *books* and merge *like* with *books*. At this point, the external thematic feature of *like* becomes accessible to the computation. Since it is uninterpretable, it must be checked immediately by combining *you* and *like books*. The resultant structure is as below:

(21) $[_{V^{max}}$ you [like books]]

The relative clause (20) is eventually merged with the D^{max} *books* in order to be properly interpreted at LF. We assume following, among others, Browning (1987) that its interpretation is established through a predication relation between the head noun and the empty operator in the relative clause in the sense of Rothstein (1983) and Williams (1980). There is, however, no legitimate way of combining them before *books* is merged with *like*. This is because when *like* is selected, the ICP requires that the internal thematic feature and Accusative Case feature of *like* should be checked immediately by combining *like* with *books* without being interrupted by any other operation. Hence, merger of *books* and the relative clause (20) may not be applied at this point. Similarly, the ICP prohibits the relative clause (20) from being merged with *books* before merger of *like books* with *you*.

The next step is to select T. Note that the EP prohibits the relative clause (20) from being merged with the main structure before we select T. The ICP then requires that the V-feature of T, being uninterpretable, should be checked immediately by combining T with the V^{max} (21). Then, the Nominative Case feature and D-feature of T become accessible to the computation. Since they are uninterpretable, they must be checked immediately by the raising of *you* to the Spec of T^{max}. When T^{max} is constructed, the next step must be to select C. Note again that due to the EP, we cannot merge the relative clause (20) with the main structure before we select C. The ICP requires that the T-feature of C, being uninterpretable, should be checked immediately by merging C with T^{max}:

(22) $[_{C^{max}}\ C_{[Q]}\ [_{T^{max}}$ you [T $[_{V^{max}}$ you [like books]]]]]

At this point, the strong Q-feature of C becomes accessible to the computation. Since it is uninterpretable, the ICP requires that it should be checked immediately. There is, however, no way of checking the strong Q-feature of C immediately. This is because the relative clause (20), which contains the *wh*-phrase *who*, could not have been merged with the main structure until this point of the derivation due to the ICP and the EP. In other words, the relative clause and the main structure each constitute an independent syntactic object at this stage of the derivation. Since C and *who* belong to different phrase markers at this stage, the former does not c-command the latter. The strong Q-feature of C cannot be checked by the raising of *who* to the Spec of C^{max}. Since the strong Q-feature cannot be checked immediately, it violates the ICP; the derivation is canceled. To put it another way, our analysis claims that the raising of *who* to the Spec of C^{max} would be an instance of "movement across phrase structures," which is prohibited due to the c-command requirement on movement. Therefore, the deviancy of (19), an example of the relative clause case of the CNPC, straightforwardly follows.

Let us next consider the non-relative case of the CNPC, taking (1b) (repeated here as (23)) as an example:

(23) *?**what** did you study [the evidence that Harry stole ***t***]

We follow Grimshaw (1990), Stowell (1981), and Takahashi (1994) in claming that the head nouns of non-relative complex NPs like *evidence* in (23) do not assign any θ-roles to the following clauses.

It first appears that the relation between the noun and the following clause in (25a-c) is parallel to the one between the corresponding verb and its complement in (24a-c):

(24) a. Andrea guesses [that Bill was lying]
b. John claimed [that he would win]
c. Paul explained [that he was temporarily insane]

(25) a. Andrea's guess [that Bill was lying]
b. John's claim [that he would win]
c. Paul's explanation [that he was temporarily insane]
(Stowell 1981:199)

Stowell (1981) argues, however, that these two relations are not parallel to each other. While the verbs *guess*, *claim*, and *explain* refer to the actions of guessing, claiming, and explaining something, the nouns *guess*, *claim*, and *explanation* do not. Rather, the nouns *guess*, *claim*, and *explanation* refer to a thing which is guessed, a thing which is claimed, and a thing which is explained, respectively. In other words, the nouns refer to the same things as the following clauses. Hence, the relation between the nouns and the following clauses in examples like (25a-c) is an appositive one rather than one of a θ-role assignment.

Stowell (1981) presents evidence in support of his appositive analysis. As shown in (26), the identity relation holds between the noun and the following clause:

(26) a. [Adrea's guess] was [that Bill was lying]

b. [John's claim] was [that he would win]

c. [Paul's explanation] was [that he was temporarily insane]

(Stowell 1981:200)

This fact further supports the claim that the head nouns of the non-relative complex NPs do not assign any θ-roles to the following clauses.

Let us return to (23). Since merger of the appositive clause in the non-relative complex NP is not triggered by any UFF, the ICP coupled with the EP ensures that it can only be merged with the main structure postcyclically. It then follows that when we come to the stage of the derivation where the strong Q-feature of C in the main structure is to be checked, the appositive clause which contains *who* has not been merged with the main structure yet. The appositive clause and the main structure each constitute an independent syntactic object at that stage. Since C and *who* belong to different phrase markers, the former does not c-command the latter. The strong Q-feature of C cannot be checked immediately by the raising of *who*; this violates the ICP. This derivation is therefore canceled. The non-relative clause case of the CNPC follows.

3.2.2 The Adjunct Condition

The Adjunct Condition can be analyzed essentially along the same line as the CNPC. Let us consider (3a-d) (repeated here as (27a-d)) as examples:

(27) a. Adjunct Clauses

*?**who** do you get jealous [because I spoke to ***t***]

b. Adverbial PPs

*?**which class** did he fall asleep [during ***t***]

c. PP Modifiers of Nouns

*?**which table** did you buy [the pictures [on ***t***]]

d. Secondary Predicates

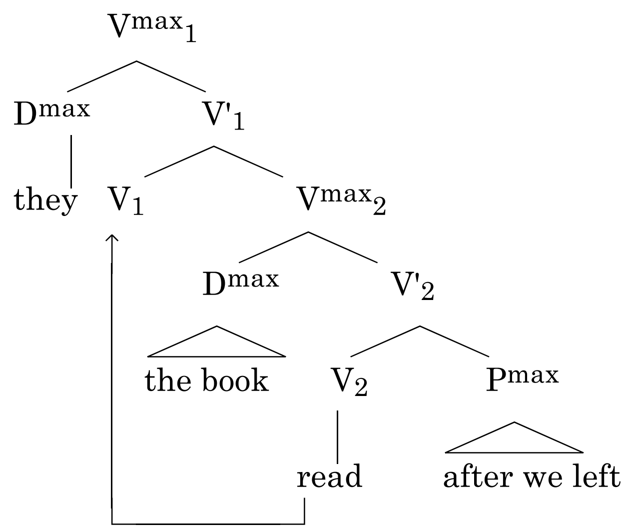
what** did John arrive yesterday, [sad about ***t]

Note that merger of adjuncts like those in (27) and main structures is not triggered by any UFFs. It then follows from the ICP and the EP that adjuncts are required to be merged postcyclically. Hence, when we come to the stage when the strong Q-feature of C in the main structure is to be checked, the adjunct has not been merged with the main structure yet. The adjunct and the main structure each constitute an independent syntactic object at that stage. Hence, C does not c-command the *wh*-phrase in the adjunct clause. The strong Q-feature of C cannot be checked immediately by the raising of the *wh*-phrase in the adjunct clause. The derivation is canceled due to a violation of the ICP. The Adjunct Condition straightforwardly follows.

Recall that according to the MP where the bare phrase structure is assumed, adjuncts may appear in so called "specifier or complement positions," as argued by the Larsonian analysis of adjuncts advocated by, among others, Chomsky (1995), Larson (1988), McConnell-Ginet (1982), and Stroik (1990). For example, under the Larsonian analysis of adjuncts, the structure underlying (28) is (29):

(28) they read the book after we left

(29)



In (29), the adjunct clause *after we left* appears as the sister of V_2 in the underlying structure. The verb *read* raises to V_1. Under this analysis, since adjuncts as well as arguments may appear as the sisters of V, there is no way of making a distinction between the two by means of structural terms. Recall that all the previous analyses of the Adjunct Condition but Takahashi (1994) crucially rely on the assumption that arguments and adjuncts appear in structurally distinct positions. They assume that adjuncts appear in adjoined positions whereas arguments appear in either specifier or complement positions. Hence, those analyses are incompatible with the Larsonian analysis of adjuncts and thus the bare phrase structure.

Our analysis of the Adjunct Condition, on the other hand, is compatible with the Larsonian analysis of adjuncts and thus the bare phrase structure. Recall that our analysis defines the notions of adjunct and complement derivationally. Under our analysis, the difference between adjuncts and complements resides in the fact that the former is required to be merged postcyclically whereas the latter is required to be merged cyclically. Crucially, when the strong Q-feature of C is to be checked, adjuncts have not been merged with main structures; the Adjunct Condition follows. Hence, our analysis of the Adjunct Condition works no matter where adjuncts may appear structurally.

3.2.3 The Non-bridge Verb Condition

Let us next consider the non-bridge verb condition, taking (4) (repeated here as (30)) as an example:

(30) ?**who** did John whisper that he saw ***t***

As first observed by Erteschik-Shir (1973), extraction out of the "complements" of non-bridge verbs like *whisper*, *whine*, and *murmur* is illegitimate. Kayne (1983) and Stowell (1981) claim that non-bridge verbs do not assign any θ-roles to their "complements." In our terms, this observation can be interpreted as follows. Non-bridge verbs do not have any thematic features which are to be checked by their merger with their internal arguments. In other words, the clauses following non-bridge verbs should count as clausal adjuncts rather than as complements.

If this conjecture is correct, the ICP coupled with the EP enables us to account for the non-bridge verb condition. Note that merger of the clausal adjunct with the main structure is not triggered by any UFF. It then follows that the clausal adjunct is forced to be merged postcyclically. Hence, when we come to the stage when the strong Q-feature of C in the main structure is to be checked, the clausal adjunct has not been merged with the main structure yet. Hence, C does not c-command the *wh*-phrases in the clausal adjunct. The strong Q-feature of C cannot be checked immediately. This derivation is canceled due to a violation of the ICP. The non-bridge verb condition straightforwardly follows.

3.2.4 The Subject Condition

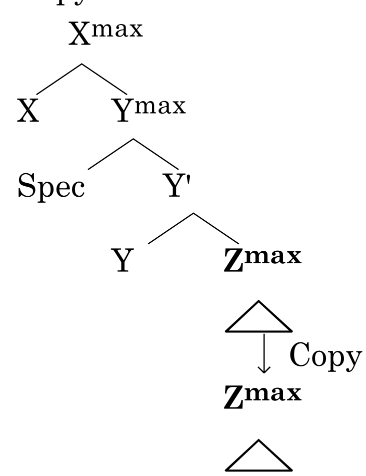
This subsection discusses the Subject Condition, arguing that it also follows from the ICP coupled with the EP. Before turning to the Subject Condition, it is necessary to explicate the notion of Attract/Move. I will argue that Attract/Move is not a primitive operation but a complex operation consisting of two primitive operations, i.e., Copy and Merge. It is shown that unlike the standard Attract/Move approach, our Copy + Merge approach only makes use of the operations of deletion and erasure, entirely dispensing with the operation of checking. The effects of the operation of checking under Chomsky's (1993, 1994, 1995) analysis are brought about by the operation of deletion or erasure which makes UFFs invisible at LF under our analysis. I will argue that this constitutes theoretical support for our Copy + Merge approach, since the operation of checking is untenable within the most recent development of the MP. Since no operation of checking is posited under our Copy + Merge approach, the ICP should be reformulated accordingly. I will argue that the ICP should be reformulated as a condition on LF visibility, which requires that UFFs should be made invisible at LF immediately by deletion or erasure when they become accessible to a computation. Finally, I will propose that erasure should be interpreted as conversion of UFFs to phonological features. It is shown that this conversion ensures that all UFFs must be erased before Spell-Out.

3.2.4.1 Attract/Move = Copy + Merge

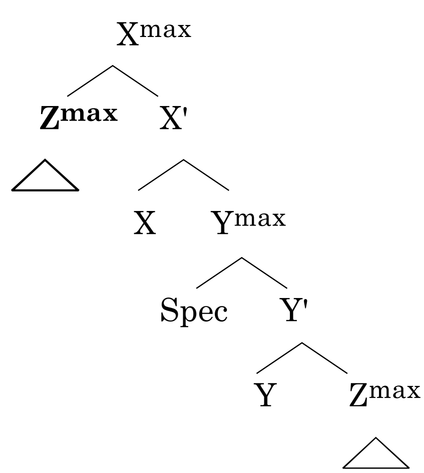
Following Bobaljik (1995) and Collins (1997), I claim that there is no primitive operation Attract/Move. Attract/Move in fact is a complex operation, consisting of two primitive operations, i.e., Copy and Merge. The operation Copy creates a new, independent syntactic object by copying a term within a syntactic object. This copied term is available to be merged with another syntactic object. For example, the raising of Z^{max} from the complement position of Y to the Spec of X^{max} proceeds as depicted below:

(31) Move = Copy + Merge

a. Copy



b. Merge



First, Z^{max} undergoes the copy operation, creating another Z^{max}. This newly created Z^{max} is then merged with the X^{max} in (31a), yielding the larger X^{max} as in (31b).

Let us consider the operation Copy in detail. Suppose that H is a head dominating a feature F_1. I argue that the copy operation is subject to the following last resort condition (though the latter is subsumed under the revised ICP to be proposed below):[7]

(32) α undergoes the copy operation only if F_1 of H enters into a deletion relation with F_2 of α.

In (32), α is the minimal element including F_2 that allows for convergence, as required by the "no extra baggage" condition. Recall that under the standard Attract/Move approach, the raising of α to target K only takes place if F of α enters into a checking relation with a sublabel of K (where a sublabel of K is a feature of the zero-level projection of the head H(K) of K). Under Copy + Merge approach, on the other hand, Copy, the first half of the complex operation Attract/Move, only takes place if a deletion relation, but not a checking relation, is established.

The definition of a deletion relation is shown below:

(33) F_1 enters into a deletion relation with F_2 iff F_1 is deleted. (F_2 may also be deleted.)

Given that deletion only takes place under identity, F_1 deletes only if identical with F_2. Let us assume following Chomsky (1995) that deletion

[7]Recall that Chomsky's (1995) definition of Attract/Move requires that what is raised should be c-commanded by its target. Under our analysis where Attract/Move is reinterpreted as consisting of Copy and Merge, Copy is subject to the c-command requirement. The c-command requirement on Copy, however, is not included in the definition of the last resort condition (33), since it is most likely to be derived from general principles of UG.

makes formal features invisible at LF. Formal features may be made invisible at LF by deletion unless it would contradict the principle of recoverbility, which states that unrecoverable items may not be deleted. Interpretable features receive interpretations at LF. Hence, they may not be made invisible at LF by deletion due to the principle of recoverbility. Uninterpretable features, on the other hand, may be made invisible at LF by deletion. This is because uninterpretable features do not contribute to any content at LF and thus their deletion does not violate the principle of recoverbility. Hence, in the definition of a deletion relation (33), F_1 must be a UFF. When F_2 happens to be a UFF, it is also deleted. It then follows from (32) and (33) that Copy, the first half of the complex operation Attract/Move, only takes place if a UFF enters into a deletion relation.

The UFFs which undergo deletion through Copy are further subject to erasure. Erasure is another operation which makes formal features invisible at LF. Erasure, a stronger form of deletion, not only makes formal features invisible at LF but also inaccessible to a computation (see Chomsky (1995)).[8] I argue that UFFs, whether deleted or undeleted, are required to be erased and thus made inaccessible to a computation before Spell-Out. Formal features can only be erased when they enter into an erasure relation. Suppose that H is a head dominating a feature F_1. The definition of an erasure relation is shown below:

---

[8]Our interpretation of erasure is different from Chomsky's (1995). The latter claims that only deleted formal features are subject to erasure. Our analysis claims, on the other hand, that erasure may apply not only to deleted formal features but also to undeleted formal features.

(34) F_1 enters into an erasure relation with F_2 iff F_2 is in the erasure domain of H and F_1 is erased. (F_2 may also be erased.)

Given that erasure only takes place under identity, F_1 erases only if identical with F_2. Note in passing that exactly like deletion, erasure is subject to the principle of recoverbility, since it makes formal features inaccessible to a computation and thus invisible at LF. Only UFFs therefore may enter into erasure relations. F_1 in the definition of an erasure relation (34) must be a UFF.

The erasure domain of H consists of:

(35) a. X adjoined to H and any features dominated by X

b. Any X^{max} in the Spec of H and any features dominated by X

c. Any X^{max} in the complement of H and any features dominated by X

Note that our notion of erasure domain corresponds to the notion of checking domain in the standard Attract/Move approach. Recall that according to the standard Attract/Move approach, the notion of checking domain is responsible for the fact that a raised term (or a copied term under the Copy theory of movement) with a feature F_2 is merged into the local domain of H with F_1, where F_1 is identical with F_2. Since our Copy + Merge approach does not assume the checking operation, however, the notion of checking domain is no longer available. We therefore have to look for an alternative way of expressing the local relation which holds between F_1 and F_2. The notion of erasure domain plays that role, ensuring that F_1 of H is only erased when F_2 is in the local domain of H through application of Merge. Under our Copy + Merge approach, therefore, the local requirement which is assumed to be imposed on the checking operation in terms of the notion of checking domain in the standard Attract/Move approach is imposed on the erasure operation in terms of the notion of erasure domain. I will later argue that erasure captures the effects which have traditionally been accounted for by agreement and assignment. It has been observed by, among others, Chomsky (1981) that agreement and assignment only take place under a local relation. Hence, it is plausible to claim that erasure is subject to the local requirement.

Let us illustrate how our analysis works, taking (31) as an example. Suppose that X dominates F_1, a UFF. Suppose also that Z^{max} dominates F_2, which is identical with F_1. Then, since F_1 enters into a deletion relation with F_2, Z^{max} undergoes the copy operation given that Z^{max} is the minimal element including F^2 that allows for convergence. After the copy operation, F_1 deletes and thus becomes invisible at LF. If F_2 is uninterpretable, it is also deleted. If F_2 is interpretable, it remains intact. This yields (31a), where there are two independent syntactic objects, X^{max} and Z^{max}. Recall that deleted features, though not visible at LF, are still accessible to a computation. Since UFFs, even if deleted, are required to be erased before Spell-Out, F_1 must enter into an erasure relation with F_2 through merging Z^{max} into the Spec of X, an erasure domain of X. F_1 is erased and made inaccessible to the computation. If F_2 is uninterpretable, it is also erased. If F_2 is interpretable, it remains intact. This yields (31b). Hence, Copy Z^{max} is followed by Merge Z^{max} into the Spec of X, which correctly brings about the effects of the traditional notion of Attract/Move.

This gives validity to our claim that Copy takes place under a deletion relation. As shown above, a UFF serves as a trigger for operations which make an element with the matching feature appear in its local domain. When it discharges its triggering function, it is eliminated (made inaccessible to a computation) through erasure. Suppose that we come across a UFF during a derivation. If there is an element with the matching feature in the N, the UFF requires that element to be merged in its local domain. Since its triggering function is discharged by this merge operation, it is eliminated through erasure. If there is an element with the matching feature not in the N but in a syntactic object already formed, on the other hand, the UFF must first trigger Copy of that element. This is because unless that element undergoes Copy, the UFF can never be eliminated through merger of that element in its local domain. The application of Copy only partially discharges the triggering function of the UFF. It is then natural to assume that the UFF does not get eliminated through erasure by the application of Copy, since it does not completely discharge its triggering function. Rather, it only becomes invisible at LF through deletion, a weaker form of erasure. Hence, it is plausible to claim that Copy takes place under a deletion relation

The salient property of our Copy + Merge approach is that it only makes use of the operations of deletion and erasure, entirely dispensing with the operation of checking.[9] The checking operation was originally proposed by Chomsky (1993), which first set up the MP. Recall that Chomsky (1993) assumes that lexical items are fully inflected in the

[9]See Collins (1997) for a different way of dispensing with the operation of checking.

lexicon. The addition of morphological features to a lexical item in the lexicon involves the simultaneous addition of those features, which play no role at PF or LF. Those features, which are visible at the interface levels, must be eliminated by checking operations prior to PF and LF for convergence. For example, the verb V is taken to be a sequence V = (α, Infl_1, ..., Infl_n) where α is the morphological complex [R-Infl_1-...-Infl_n], R a root and Infl_i an inflectional feature. These inflectional features must be eliminated by the raising of V to the functional categories with the matching features for convergence.

Chomsky (1995) argues, however, that the operation of checking has odd features and thus should be dispensed with. First, the operation of checking is redundant in that the relevant properties can be determined by algorithms at LF. The LF representation which contains an inflected item, say [V-AGR-TENSE], tells us that checking operations have applied to this item, eliminating the inflectional features AGR and TENSE. Such redundant operations should be abandoned because of the notion of "simplicity." Second, the operation of checking is incompatible with the most recent development of the MP where formal features are classified based on [+/- Interpretable]. Recall that among formal features, only uninterpretable features but not interpretable ones may be made invisible at LF due to the principle of recoverbility. According to the operation of checking proposed by Chomsky (1993), however, checked features must be eliminated irrespectively of whether they are interpretable or not. Even interpretable features would be eliminated by checking operations. This would violate the principle of recoverbility. The operation of checking is therefore incompatible with the most recent development of the MP unless we change its original formulation. Hence, our Copy + Merge approach, where no operation of checking is posited, is theoretically desirable.[10]

Since there is no operation of checking under our Copy + Merge approach, the ICP, which presupposes the operation of checking, must be reformulated accordingly. In the next subsection, I will argue that the ICP should be reformulated as a condition on LF visibility, which requires that UFFs should be made invisible at LF immediately when they become accessible to a computation.

3.2.4.2 Elimination of the Operation of Checking

Since no checking operation is available any longer, the ICP, which presupposes the existence of the operation of checking, should be reformulated accordingly. I argue that the ICP should be reformulated as a condition on LF visibility:

---

[10]Chomsky (1995) claims that his system dispenses with the operation of checking. Contrary to his claim, however, the operation of checking is not entirely dispensed with under his system. Let us consider his definition of the last resort condition, which is to be incorporated into the definition of Attract/Move:

(i) Last Resort
Attract/Move F raises F to target K only if F enters into a checking relation with a sublabel of K.
(adapted from Chomsky 1995:280)

Checked features are subject to deletion operations and deleted features are further subject to erasure operations:

(ii) a. A checked feature is deleted when possible.
b. A deleted feature is erased when possible
(Chomsky 1995:280)

Note that his system crucially makes use of the distinction between checked and unchecked features, claiming that only checked features are subject to deletion. Hence, his system still assumes the operational content of checking, though none of his definitions directly make reference to the operation of checking.

(36) The Immediate Checking Principle (Revised)
Uninterpretable formal feature (UFFs) must be made invisible at LF immediately when they become accessible to a computation.

(36) requires that when we come across a UFF during a derivation, we must immediately make it invisible at LF either by deletion or erasure. Through the application of Copy, UFFs are made invisible at LF by deletion. They are made invisible at LF by erasure through being merged into the erasure domain of H, a head, with the matching features. Note that (36) subsumes the last resort condition on Copy (32) and thus makes it possible to maintain the simplest definition of Copy, which only states that Copy creates a new independent syntactic object by copying a term within a syntactic object.

3.2.4.3 Erasure as Conversion of UFFs to Phonological Properties

The above discussion has shown that Copy only takes place when a deletion relation is established. Deleted UFFs are further required to enter into erasure relations before Spell-Out. Since erasure only takes place within an erasure domain, copied terms are required to be merged into that domain. This correctly brings about the effects of Attract/Move. A question now arises what requires deleted UFFs to enter into erasure relations before Spell-Out. I propose that erasure should be interpreted as conversion of UFFs to phonological properties, arguing that erasure of UFFs is required for PF-convergence.[11]

---

[11]See Chomsky (1995) for a similar interpretation of UFFs.

It is well known that there are cases where UFFs have their phonetic reflexes. The ϕ-features of predicates, which are UFFs, often have their phonetic reflexes. For example, in rich subject-agreement languages like Italian and Spanish, the subject ϕ-features of predicates have their phonetic reflexes. Presented below is the conjugation of the present tense indicative of the Spanish verb *hablar* 'to say':

(37)	a.	habl-o	'I speak'	1s
	b.	habl-as	'you speak'	2s
	c.	habl-a	'he/she speaks'	3s
	d.	habl-amos	'we speak'	1pl
	e.	habl-áis	'you (pl.) speak'	2pl
	f.	habl-an	'they speak'	3pl

The phonetic realizations of the subject ϕ-features of predicates are also observed in languages like English, French, German, and Irish though they have relatively impoverished agreement morphology. Similarly, the object ϕ-features of predicates are phonetically realized in, among others, the participial verbs with wh-movement in Romance languages (Kayne (1989)) and the perfect aspect sentences in Pashto (Huang (1984, 1989)).

The Case features of nominals, being UFFs, also have their phonetic reflexes. The phonetic realizations of the Case features of nominals are clearly observed in languages like Latin. Presented below is the conjugation of the singular form of the Latin masculine noun *puer* 'boy':

(38) a. Nominative: puer

b. Accusative: puerum

c. Genitive: puerî

d. Dative: puero#

e. Ablative: puero#

The phonetic realizations of the Case features of nominals are also observed in languages like English and French though they have relatively impoverished Case morphology.

In order to account for the phonetic realizations of these UFFs, I argue that erasure should be interpreted as conversion of UFFs to their corresponding phonological properties, which are to be stripped away at Spell-Out. I argue that this conversion takes place before Spell-Out and thus the phonetic reflexes of UFFs appear at PF. The uninterpretable ϕ-features of predicates are converted to their corresponding phonological properties through entering into erasure relations with the ϕ-features of nonimals. The Case features of nominals are converted into their corresponding phonological properties through entering into erasure relations with the Case features of verbs. It is well known that languages vary as to whether the ϕ-features of predicates and the Case features of nominals are phonetically realized. Even in languages where they are not phonetically realized, I argue that these UFFs are converted to their corresponding phonological properties, presumably phonetically null, through entering into erasure relations.

Such a conversion is required to take place before Spell-Out by the fact that lexical items with these features have morphological slots to be filled by their phonetic reflexes. Predicates have the morphological slots for ϕ-features which are to be filled by the phonetic reflexes of their uninterpretable ϕ-features. Nominals have the morphological slots for Case features which are to be filled by the phonetic reflexes of their uninterpretable Case features. The information that those lexical items have such morphological slots belongs to their intrinsic property, though what kinds of features are to fill the slots does not. If the slots are not filled, it leads a derivation to crash at PF (or possibly it is ruled out in the morphological component, which stays between Spell-Out and the PF interface).[12]

There are, however, UFFs which never have their phonetic reflexes like the uninterpretable categorial features of functional heads including the Q-feature of C, the thematic features of predicates, and the Case features of verbs. The difference between the UFFs with phonetic reflexes and those without them resides in the fact that while the former is an optional feature, the latter is an intrinsic feature. I argue that UFFs, whether intrinsic or optional, have their corresponding phonological properties, though prima facie intrinsic features are never phonetically realized.

Recall that intrinsic features are part of lexical items within the lexicon. It is then reasonable to claim that the phonological properties of intrinsic UFFs constitute part of those of lexical items themselves. If they are not properly converted to their corresponding phonological properties through conversion, lexical items end up having insufficient phonological information to be properly realized at PF. Such improperly

---

[12]The same can be said about the tense features of verbs, which are also optional UFFs. According to our conversion analysis, verbs are not inflectionally represented but rather bare in the lexicon. They acquire inflectional morphology through syntactic operations. See Lasnik (1994) and Rohrbacher (1993, 1995) for detailed discussion of this subject.

realized lexical items are illegitimate at PF; the derivation would crash. For example, let us consider the verb *see*. It has Accusative Case and Theme features as its intrinsic UFFs. Although the verb *see* is assigned phonological properties in the lexicon, those properties alone are not sufficient to make this verb properly realized at PF. Its intrinsic UFFs must be converted to their corresponding phonological properties through erasure before Spell-Out. The converted phonological properties together with the lexically assigned phonological properties make this verb properly realized at PF. Otherwise, an illegitimate object appears at PF; the derivation crashes at that level. Hence, intrinsic UFFs must also be converted to their corresponding phonological properties through erasure before Spell-Out, though they do not have their own phonetic reflexes. Optional features, on the other hand, are not assigned to lexical items within the lexicon but rather added to them arbitrarily at any point of derivations. It is then reasonable to claim that the phonological properties of optional UFFs do not constitute part of those of lexical items within the lexicon. They are rather added to the lexical items during derivations through conversion. In other words, they constitute inflectional morphemes. Hence, optional UFFs have their phonetic reflexes of their own, which are to be added to the lexical items as their inflectional morphemes, though they may be phonetically null. We can therefore conclude that UFFs, whether intrinsic or optional, must be converted to their corresponding phonological properties through erasure for PF-convergence.[13]

---

[13]This interpretation of erasure differs from Chomsky's (1995) interpretation in that the latter, but not the former, claims that erasure is prohibited if it would create an illegitimate object. Crucial is the case where a term only consists of a strong feature

With the above discussion in mind, let us reconsider how our Copy + Merge approach works. Suppose that we come across a UFF of H, a head, during a derivation. The UFF is required to be immediately made invisible at LF by the ICP. If it is made invisible at LF by erasure through merger of a term with the matching feature into the erasure domain of H, it undergoes conversion to its corresponding phonological property and thus becomes inaccessible to the computation anymore. If it is made invisible at LF by deletion through the application of Copy, on the other hand, it is further subject to erasure. The copied term, which has the matching feature, must be merged into the erasure domain of H. This is because the UFF of H must be converted to its corresponding phonological property for PF-convergence. This ensures that the copied term is merged into the erasure domain of H before Spell-Out.[14, 15]

---

and thus the erasure of the strong feature leads to erasure of the entire term. Chomsky's interpretation of erasure depends on the validity of his analysis of pure expletives. He claims that pure expletives like English *it* only consist of strong D-features. Hence, their strong feature may never be erased. It is not clear at this point, however, whether his analysis of pure expletives is on the right track. The present discussion therefore claims that erasure of a UFF is always mandatory, leaving an analysis of pure expletives for further study,

[14]One might wonder what ensures that subjects are "docked" in the right positions in complex sentences like (i):

(i) John said that Bill read that book

Note that *John* and *Bill* have the same agreement and Case features. Before merger of *John* and *Bill* with the main structure, we have the following:

(ii) a. [$_{T^{max}}$ T [$_{V^{max}}$ John said [$_{C^{max}}$ that [$_{T^{max}}$ T [$_{V^{max}}$ Bill read that book]]]]]
b. John
c. Bill

If we "docked" *John* and *Mary* to the embedded and matrix Spec's of T^{max}, respectively, we would get the following:

(iv) [$_{T^{max}}$ Bill [T [$_{V^{max}}$ John said [$_{C^{max}}$ that [$_{T^{max}}$ John [T [$_{V^{max}}$ Bill read that book]]]]]]]

Chain formation applies in the LF-component, yielding the following two chains:

(v) a. CH = (*Bill*, *Bill*)
b. CH = (*John*, *John*)

These chains, however, are not legitimate at LF. Specifically, these chains do not count as A-chains, since C^{max} intervenes between the head and tail positions. The resultant representation would be ruled out. Note in passing that (vb) also violates the condition on an A-chain which states that the head of an A-chain is in a Case-marked position while its tail is in a θ-position. Hence, we can ensure that the subjects *John* and *Bill* are "docked" into the right positions. The cases where there are two instances of object raising within a sentence can be analyzed in a similar way.

One might also wonder what ensures that *wh*-phrases are "docked" into the right positions in examples like (i):

(i) who wonders what John bought

Before merger of *who* and *what* with the main structure, we have the structures in (ii). Here, we assume just for expository purposes that the subjects have already been merged with the main structure:

(ii) a. [$_{C}$^{max} C [$_{T}$^{max} who [T [$_{V}$^{max} who wonders [$_{C}$^{max} C [$_{T}$^{max} John [T [$_{V}$^{max} John bought what]]]]]]]]

b. who

c. what

If we "docked" *who* and *what* to the embedded and matrix Spec's of C^{max}, respectively, we would get the following:

(iii) [$_{C}$^{max} what [C [$_{T}$^{max} who [T [$_{V}$^{max} who wonders [$_{C}$^{max} who [C [$_{T}$^{max} John [T [$_{V}$^{max} John bought what]]]]]]]]]]

Among the chains created in this derivation, the following chains are relevant here:

(iv) a. CH = (*what*, *what*)

b. CH = (*who*, *who*)

Let us assume following Chomsky (1993) that the non-head positions of chains delete. This yields the following structure:

(v) [$_{C}$^{max} **what$_i$** [C [$_{T}$^{max} ***t'$_j$*** [T [$_{V}$^{max} ***t$_j$*** wonders [$_{C}$^{max} **who $_j$** [C [$_{T}$^{max} John [T [$_{V}$^{max} John bought ***t$_i$***]]]]]]]]]]

In (v), the operator *who* does not have any variable to bind in its c-command domain. This violates the ban against vacuous quantification. Furthermore, the variable t'_j, more precisely the A-chain (t'_j, t_j), does not have any c-commanding operator. Hence, this violates the ban against free variables. Hence, we can ensure that the *wh*-phrases *who* and *what* are "docked" into the right positions.

[15]One could argue that this analysis is conceptually undesirable since it necessarily needs global considerations. Merger of a copied term into an erasure domain of a head H is required for PF-convergence. We have to look at the PF interface level to determine whether to apply Merge to a copied term. The reference to the PF interface level has the "look-ahead" property, which should be avoided on conceptual grounds.

It might be possible to claim, however, that the problem of globality does not arise in this analysis. Recall that the notion of globality is defined as below:

(i) A condition C is global if it cannot determine whether to apply an OP or not only on the basis of information available in Σ.

Note that according to the definition of the notion of globality (i), the problem of globality only arises when we have an option of applying an OP or not during a derivation. If we follow Chomsky (1995) in assuming that Merge counts as costless and thus does not need any motivation for its application, there is no option of not applying Merge to a copied term. Insufficient application of Merge simply generates no derivation. Hence, concerning Merge, the problem of globality never arises if Chomsky's interpretation of

Through the rest of this thesis, unless the distinction between the standard Attract/Move approach and our Copy + Merge approach is relevant, we will pretend for expository purposes that we are assuming the former. Given the established usage, the discussion to follow also sticks to the phrase "checking operation," which is meant to represent a process of making UFFs invisible at LF through either a deletion or erasure operation unless any complication arises. Accordingly, we will stick to the appellation "the Immediate Checking Principle (ICP)," though, to be precise, it does not refer to the operation of checking.

3.2.4.4 An Account of the Subject Condition

Returning to the Subject Condition, let us consider (2) (repeated here as (39)) as an example:

(39) ***who** did [pictures of ***t***] please you

Let us first consider how to construct the subject phrase *pictures of who*. We first select *of*. The selectional restriction feature of *of*, being uninterpretable, must be made invisible at LF immediately by erasure through merger of *of* with *who*. Then, the noun *pictures* is selected. The ICP requires that its selectional restriction feature, being uninterpretable, should be made invisible at LF immediately by erasure through merger of *pictures* with *of who*. Finally, D is selected. Its selectional restriction feature is made invisible immediately at LF by erasure through combining D with *pictures of who*. The resultant structure is as follows:

Merge is correct. If this conjecture is correct, we can say that the inspection of PF only tells us where a copied term should be merged. It does not tell us whether we should apply merger of a copied term or not.

(40) $[_{D^{max}}$ $[_{N^{max}}$ pictures $[_{P^{max}}$ of $[_{D^{max}}$ $who_{[Q]}]]]]$

Let us now turn to consider how to construct the main structure. We first select the verb *please*. The ICP requires that we should immediately make the internal thematic feature and Accusative Case feature of *please* invisible at LF, since these features are uninterpretable. Hence, we merge *please* with *you*, erasing these features. At this point, the external thematic feature of *please* becomes accessible to the computation. Since it is uninterpretable, it must be made invisible at LF immediately by erasure through combining *pictures of who* with *please you*. The resultant structure is as below:

(41) $[_{V^{max}}$ $[_{D^{max}}$ pictures of $who_{[Q]}]$ [please you]]

The next step is to select T. The ICP requires that the selectional restriction feature of T, which states that T takes V^{max} as its complement, should be made invisible at LF immediately by erasure through combining T with the V^{max} *pictures of who please you*, as shown below:

(42) $[_{T^{max}}$ T $[_{V^{max}}$ $[_{D^{max}}$ pictures of $who_{[Q]}]$ [please you]]]

Since the strong D-feature and Nominative Case feature of T become accessible to the computation at this stage, the ICP requires that they should be made invisible at LF immediately. They are made invisible at LF by deletion through application of Copy to the D^{max} *pictures of who* given that the D^{max} is the minimal element including these features that allows for convergence:[16]

---

[16]I will later explicate what counts as a minimal element including a feature F that allows for convergence.

(43) a. $[_{T^{max}}$ T $[_{V^{max}}$ **$[_{D^{max}}$ pictures of who$_{[Q]}$]** [please you]]]

b. **$[_{D^{max}}$ pictures of who$_{[Q]}$]**

The newly created D^{max} (43b) is eventually merged into the Spec of T^{max}, an erasure domain of T, for PF-convergence. It is important to note, however, that merger of the newly created D^{max} (43b) with the main structure (43a) cannot take place at this point of the derivation. Recall that this merger is not required by the ICP, since the UFFs of T have already been made invisible at LF by deletion through application of Copy. We can say that the newly created D^{max} functions as an adjunct in the sense that it is required to be merged postcyclically. The EP then requires that we should select C rather than merge the newly created D^{max} with the main structure at this stage. When C is selected, the ICP requires that its selectional restriction feature, which states that it takes T^{max} as its complement, should be made invisible at LF immediately by erasure through combining C with the T^{max} *pictures of who please you*:

(44) a. $[_{C^{max}}$ $C_{[Q]}$ $[_{T^{max}}$ T $[_{V^{max}}$ $[_{D^{max}}$ pictures of who$_{[Q]}$] [please you]]]]

b. $[_{D^{max}}$ pictures of who$_{[Q]}$]

At this stage of the derivation, the strong Q-feature of C, which is a UFF, becomes accessible to the computation. The ICP requires that it should be made invisible at LF immediately. It should be noted that the strong Q-feature of C cannot be made invisible at LF by deletion through application of Copy to *who* in the D^{max} (44b). The *wh*-phrase *who* is the potential candidate for being subject to the copy operation triggered by the strong Q-feature of C. The *wh*-phrase *who* in the D^{max} (44b), however, is not c-commanded by C. This is because the D^{max} (44b) and the main structure (44a) each constitute an independent syntactic object at this stage of the derivation. Since C does not c-command *who* in the D^{max} (44b), there is no way for the strong Q-feature of C to trigger Copy of that *wh*-phrase. The strong Q-feature of C cannot be made invisible at LF through application of Copy to *who* in the D^{max} (44b), which is eventually merged into the Spec of T^{max}. Hence, we can derive the fact that no extraction is possible out of the subject in the Spec of T^{max}.

There is, however, an alternative candidate for making the strong Q-feature of C invisible at LF. Since we are assuming that Attract/Move consists of Copy and Merge, there exist two instances of the D^{max} *pictures of who* after application of Copy. The one is in the Spec of V^{max} while the other is the D^{max} (44b). Although the strong Q-feature of C cannot trigger Copy of *who* in the D^{max} (44b), it can trigger Copy of *who* within the Spec of V^{max}. This is because C c-commands *who* within the Spec of V^{max}. Then, the resulting structure would be as follows after the strong Q-feature of C is made invisible at LF by deletion through application of Copy:

(45) a. $[_{C^{max}}$ C $[_{T^{max}}$ T $[_{V^{max}}$ $[_{D^{max}}$ pictures of **who**$_{\mathbf{[Q]}}$] [please you]]]]

b. $[_{D^{max}}$ pictures of who$_{[Q]}$]

c. **who**$_{\mathbf{[Q]}}$

(45c) is the newly created *who*. Note that (45c) is the copy of *who* in (45a) but not that of *who* in (45b). The newly created *who* (45c) is not merged with the main structure (45a) at this point. This stage of the derivation, therefore, can be characterized as consisting of the three independent syntactic objects (45a-c).

At the later stage of the derivation, (45b-c) are merged into the right positions in (45a) for PF-convergence. This yields the following structure, with all formal features including Q-features being ignored:

(46) [$_{C}$max who [C [$_{T}$max [$_{D}$max pictures of who] [T [$_{V}$max [$_{D}$max pictures of who] [please you]]]]]]

Although this derivation reaches LF without being canceled, I argue that there is a violation of the ban against vacuous quantification, a condition of FI at LF.

Recall that the operation that copies α introduces α a second time into a syntactic object. The element α appears twice in the syntactic object, in its initial and raised positions. This is the only case in which two terms can be identical, since we distinguish among distinct selections of a single lexical item from the lexicon. This ensures that such pairs that consist of identical terms only arise through Copy. Although the two terms are identical in constitution, they are positionally distinct. Recall also that such terms that are identical in constitution but positionally distinct from each other form a chain in the LF-component. In (46), *pictures of who* is introduced a second time in the syntactic object through Copy . The two occurrences of *pictures of who* are therefore identical in constitution but positionally distinct from each other. Furthermore, *who* in the Spec of C^{max} is introduced by copying *who* within the Spec of V^{max}. These two instances of *who* are also identical in constitution but positionally distinct. Hence, the following two chains are formed during the derivation:

(47) a. CH = (*pictures of who*, *pictures of who*)

b. CH = (*who*, *who*)

Let me stress again that chain (47b) consists of *who* in the Spec of C^{max} and *who* within the Spec of V^{max}, not *who* within the Spec of T^{max}.

Following Chomsky (1993), let us assume that the non-head positions of chains delete within the LF component. Hence, *who* within the Spec of V^{max} and *pictures of who* in the Spec of V^{max} delete. (46) therefore yields the following LF-representation:[17, 18]

(48) [$_{C^{max}}$ **who** [C [$_{T^{max}}$ [$_{D^{max}}$ pictures of who]$_i$ [T [$_{V^{max}}$ t_i [please you]]]]]]

In (48), the operator *who* does not bind any variable and thus violates the ban against vacuous quantification, a condition of the FI. Hence, this derivation crashes at LF.[19]

It is possible to claim that (48) also violates the condition of inclusiveness. Recall that the condition of inclusiveness requires that any structure constructed by a computation from N to LF only consists of elements which are already present in the lexical items selected for the N. This condition holds with the single exception of the terms introduced by Copy. Copy introduces elements which are not present in an N during a derivation. Those introduced by Copy, however, form chains in the LF-component. The non-head positions of the chain are deleted and only the

---

[17]See chapter 6 for detailed discussion of the construction of an operator-variable pair.

[18]Just for expository purposes, we use indices to identify elements of a chain here and in the relevant representations to follow. Recall that the condition of inclusiveness does not allow indices to be assigned to any structure constructed by a computation from N to LF.

[19]See the discussion of the previous chapter for the alternative view that the ban against vacuous quantification does not lead a derivation to crash but only makes an interpretation anomalous. Note that the arguments to follow are valid under either interpretation of the ban against vacuous quantification.

head-position remains. At LF, therefore, the condition of inclusiveness should hold without any exceptions. No new objects are allowed to appear at LF. In LF-representation (48), however, there are two occurrences of *who*, although the N of this derivation only includes one occurrence of *who*. Hence, this derivation adds the new object which is not present in the N; this violates the condition of inclusiveness.

The above discussion showed that the ICP and the EP together with the Copy + Merge approach correctly derives the subject condition effects. Before closing this subsection, let us look at how the present analysis deals with the case where a subject itself is a *wh*-phrase:

(49) **who *t*** bought that book

During the derivation, we come to the following stage:

(50) $[_{T}$max T $[_{V}$max who$_{[Q]}$ [buy $[_{D}$max that book]]]]

At this stage, the ICP requires that the strong D-feature and Nominative Case feature of T should be made invisible at LF immediately by deletion through application of Copy to *who*. The resultant structure is as follows:

(51) a. $[_{T}$max T $[_{V}$max **who$_{[Q]}$** [buy $[_{D}$max that book]]]]

b. **who$_{[Q]}$**

Note that, as argued above, we cannot merge the newly created *who* (51b) with the main structure (51a) at this point of derivation, since the EP requires that C should be selected.

When C is selected, the ICP requires that its selectional restriction feature should be made invisible at LF immediately by erasure through merger of C and the T^{max} *who buy that book*:

(52) a. [$_{C}$max C$_{[Q]}$ [$_{T}$max T [$_{V}$max who$_{[Q]}$ [buy [$_{D}$max that book]]]]]

b. who$_{[Q]}$

At this stage of the derivation, the strong Q-feature of C becomes accessible to the computation. Since it is uninterpretable, it must be made invisible at LF immediately by deletion through copying *who* in the Spec of V^{max}. Note that the strong Q-feature of C cannot trigger Copy of the newly created *who* (52b), since the former does not c-command the latter. The resultant structure is as follows:

(53) a. [$_{C}$max C [$_{T}$max T [$_{V}$max **who$_{[Q]}$** [buy [$_{D}$max that book]]]]]

b. who$_{[Q]}$

c. **who$_{[Q]}$**

Finally, the two instances of *who* are merged with the main structure as the Spec of C^{max}, an erasure domain of C, and the Spec of T^{max}, an erasure domain of T, respectively. The resultant structure is as follows, with all formal features including Q-features being ignored:

(54) [$_{C}$max **who** [C [$_{T}$max **who** [T [$_{V}$max who [buy [$_{D}$max that book]]]]]]]

In (54), since *who* in the Spec of T^{max} is introduced by copying *who* in the Spec of V^{max}, these two occurrences of *who*, though positionally distinct, are identical in constitution. Furthermore, *who* in the Spec of C^{max} is also introduced by copying *who* in the Spec of V^{max}. These two occurrences of *who* are also identical in constitution but positionally distinct. On the assumption that the identity relation is transitive, the three occurrences of *who* in (54) are all identical in constitution, though they are positionally distinct. Hence, *who* in the Spec of C^{max} is identical in constitution with *who* in the Spec of T^{max}, which in turn is identical in constitution with *who* in the Spec of V^{max}. Chain formation applies to (54), yielding the following two chains:[20, 21]

(55) a. CH = (*who*, *who*)

b. CH = (*who*, *who*)

(55) is an operator-variable construction which consists of *who* in the Spec of C^{max} and *who* in the Spec of T^{max}. (55), on the other hand, is an A-chain which consists of *who* in the Spec of T^{max} and *who* in the Spec of V^{max}. Then, the non-head members of the chains, i.e., *who* in the Spec of T^{max} and *who* in the Spec of V^{max}, are deleted. This yields the following LF-representation:

(56) $[_{C^{max}}$ **who** [C $[_{T^{max}}$ ***t′*** [T $[_{V^{max}}$ ***t*** [buy $[_{D^{max}}$ that book]]]]]]]

LF-representation (56) only consists of legitimate objects. Specifically, there is no violation of the ban against vacuous quantification, since *who* properly binds *t′* in the Spec of T^{max} (more precisely, the A-chain (*t*', *t*)) as its variable.

---

[20]Note that this is different from successive cyclic A'-movement from an A-position, which is discussed in Chomsky (1991). The latter yields a chain that is not legitimate, a "heterogeneous chain" which consists of an adjunct chain and an A'-A pair. Chomsky argues that this "heterogeneous chain" must be made a legitimate object, an operator-variable construction, by eliminating the intermediate A'-trace(s) in accordance with the economy condition. In the present case, however, we can construct legitimate objects, an operator-variable chain and an A-chain, without eliminating any intermediate trace. Since the intermediate trace in the Spec of T^{max} need not be deleted, it may not be deleted by the "least effort" principle.

[21]Alternatively, as argued by Chomsky and Lasnik (1993), chain formation applies to (54), yielding the following three-membered chain:

(i) CH = (*who*, *who*, *who*)

Even under this analysis, we get LF-representation (56) after the deletion of the non-head members of the chain.

3.2.4.5 Lack of the Subject Condition Effects in Japanese

It has been observed by, among others, Fukui (1995), Kayne (1984), Kuno (1973), Lasnik and Saito (1992), Ross (1967), Saito (1985, 1992), and Takahashi (1994) that Japanese does not obey the Subject Condition. Let us look at the following examples:

(57) a. The Cleft Construction

[***Op$_i$*** [John-ga [[Mary-ga ***t$_i$*** katta no]-ga
-Nom -Nom bought fact-Nom
mondai da to] omotte iru] no]-wa
problem is Comp think Comp-Top
sono hon$_i$-o da
that book-Acc be
Lit. 'it is that book$_i$ that John thinks that the fact that Mary bought e_i is a problem'

b. The *Tough* Construction

[[zibun-no otooto-kara]$_i$-ga [(John$_j$-nitotte) [***Op$_i$*** [[[*pro$_j$*
self's brother-from-Nom -for
t$_i$ okane-o takusan karite iru no]-ga mondai da]
money-Acc a lot borro fact-Nom problem be
to] mitome] nikui]]
Comp admit hard
Lit. '[from self's brother]$_i$ is hard (for John$_j$) to admit that the fact that he$_j$ has borrowed a lot of money e_i is a problem'

c. The Comparative Deletion Construction

[$\boldsymbol{Op_i}$ [John-ga [Bill-ga $\boldsymbol{t_i}$ katta no]-ga mondai

-Nom -Nom bought fact-Nom problem

da to] itte iru] yorimo] Mary-wa takusan hon-o

be Comp say than -Top many book-Acc

katta

bought

Lit. 'Mary bought more books than John says that the fact that Bill bought is a problem'

As will be discussed extensively in the next chapter, the cleft construction with an NP-Case or PP focus, the *tough* construction with a PP subject, and the comparative deletion construction all involve movement of an empty operator (see, among others, Hoji (1990), Ishii (1991), Kikuchi (1987), and Takezawa (1987)). In (57), although empty operators are extracted out of the subject phrases, the results are acceptable. This indicates that the Subject Condition does not hold in Japanese.

There are some speakers who find the examples in (57) awkward. It is possible to claim, however, that their awkwardness is due to the fact an empty operator is extracted out of a *no* clause. For those speakers, extraction out of an object *no* clause is also awkward, as shown below:

(58) a. The Cleft Construction

[$\boldsymbol{Op_i}$ [John-ga [Mary-ga $\boldsymbol{t_i}$ katta no]-o

-Nom -Nom bought fact-Acc

mondai-ni site iru] no]-wa sono hon$_i$-o da

problem-Dat make Comp-Top that book-Acc be

Lit. 'it is that book$_i$ that John is calling the fact that Mary bought e_i into question'

b. The *Tough* Construction

[[zibun-no otooto-kara]$_i$-ga [(John$_j$-nitotte) [$\boldsymbol{Op_i}$ [pro_j

self's brother-from-Nom -for

[Mary-ga $\boldsymbol{t_i}$ okane-o takusan karite iru no]-o

-Nom money-Acc a lot borrow fact-Acc

mondai-ni si]] nikui]]

problem-Dat make hard

Lit. '[from self's brother]$_i$ is hard (for John$_j$) to call the fact that Mary has borrowed a lot of money e_i into question'

c. The Comparative Deletion Construction

[$\boldsymbol{Op_i}$ [John-ga [Bill-ga $\boldsymbol{t_i}$ katta no]-o

-Nom -Nom bought fact-Acc

mondai-ni site iru] yorimo] Mary-wa takusan

problem-Dat make than -Top many

hon-o katta

book-Acc bought

Lit. 'Mary bought more books than John calls the fact that Bill bought into question'

Hence, even for those speakers who find the examples in (57) awkward, there is no contrast in acceptability between the examples in (57) and those in (58). This suggests that there are no subject condition effects in Japanese.

I argue that the lack of the subject condition effects in Japanese straightforwardly follows from our analysis if we assume following, among others, Fukui (1986), Kuroda (1988), Lasnik and Saito (1992), and Takahashi (1994) that subjects in Japanese stay in-situ at least in the overt component. Recall that subjects in English overtly move from within V^{max} to the Spec of T^{max}. Under our analysis, this overt raising yields the adjuncthood of a subject in the sense that the raised subject is required to be merged with the main structure postcyclically. The subject in English therefore constitutes "domain barriers." Given that the subject in Japanese never undergoes movement, on the other hand, our analysis correctly predicts that the Subject Condition does not hold in Japanese. This is because unmoved subjects count as arguments throughout derivations, being required to be merged cyclically.[22]

Our account of the lack of the subject condition effects in Japanese can be extended to an analysis of extraction from some postverbal subjects in Italian. It has been claimed by, among others, Belletti and Rizzi (1981), Jaeggli (1981), and Rizzi (1982) that some postverbal subjects behave like objects with respect to *ne*-cliticization:

(59) a. ho letto tre libri
'I have read three books'
b. **ne** ho letto tre
'I have read three'

(60) a. sono arrivati tre ragazzi
'three kids have arrived'
b. **ne** sono arrivati tre
'three have arrived'

(Jaeggli 1981:139-140)

---

[22]It should be pointed out that our analysis can also accommodate the lack of the Subject Condition in Japanese under the overt subject raising analysis if subject raising in Japanese is a non-feature-driven movement like scrambling. As will be discussed in detail in chapter 4, our analysis claims that phrases which have undergone non-feature-driven movement do not constitute "barriers" for further extraction.

Ne-cliticization takes place from within the direct object in (59b) and the postverbal subject in (60b). They are both acceptable. This indicates that the postverbal subject in (60) does not constitute a "domain barrier." I assume following, among others, Jaeggli (1981) that such a postverbal subject originates there and thus never undergoes movement. It then follows from our analysis that such a postverbal subject that does not undergo movement never constitutes a "domain barrier." Hence, we can correctly predict that examples like (60b) are acceptable.

3.2.5 Summary

To summarize section 3.2, I have argued that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure, taking English overt wh-movement as an example. It was shown that the "domain barriers" are exactly those elements which are forced to be merged postcyclically. When we come to the stage of a derivation where the strong Q-feature of C is to be made invisible at LF, the "domain barrier" has not been merged with the main structure. The "domain barrier" and the main structure each constitute an independent syntactic object at this stage. No element inside the "domain barrier" may undergo Copy to make the strong Q-feature of C invisible at LF. Since the strong Q-feature cannot be made invisible at LF immediately, this violates the ICP; the derivation is canceled. This analysis gives us a minimalist account of the "domain barrier" effects.

3.3 No Extraction out of Moved Phrases

The analysis of the "domain barriers" proposed in the previous section can be extended to another locality condition on feature-driven A'-

movement, i.e., no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. I will argue that this locality condition can be accounted for along the line of the above mentioned analysis of the Subject Condition.

Let us consider the following examples (cf. Lasnik & Saito (1992)):[23]

(61) Topicalization out of Topic

*?**vowel harmony$_i$**, I think that [[articles about $\boldsymbol{t_i}$]$_j$ [you read t_j]]

(62) Wh-movement out of moved *wh*-phrases

*?**who$_i$** do you wonder [[which pictures of $\boldsymbol{t_i}$]$_j$ [Mary bought t_j]]

(63) Wh-movement out of Topic

*?**who$_i$** do you think that [pictures of $\boldsymbol{t_i}$]$_j$ John wanted t_j

In (61), topicalization applies to *articles about vowel harmony* in the embedded clause, out of which *vowel harmony* is further extracted to the initial position of the matrix clause by topicalization. In (62), wh-movement first applies to *which pictures of who* in the embedded clause. At the matrix clause level, *who* within the embedded Spec of C^{max} is further extracted by wh-movement. In (63), topicalization applies to *pictures of who* in the embedded clause, out of which *who* is further extracted by wh-movement. These examples are all deviant. The observational generalization can be stated as follows. Phrases which have undergone feature-driven A'-movement like topicalization and wh-

---

[23]Lasnik and Saito (1992) claims that although these examples are deviant, they are less severely deviant than the "domain barrier" violation cases. Base on such judgments, they propose that A'-binder do not function as "barriers." I depart from them in claming that there is no substantial contrast in acceptability between these examples and the "domain barrier" violation cases.

movement constitute "barriers" for further feature-driven movement. I will argue that this generalization straightforwardly follows from the ICP coupled with the EP.[24, 25]

Let us consider (61) as an example. Its derivation proceeds as follows. Let us consider the stage where we construct the embedded T^{max}:

---

[24]Based on the observation made by Torrego, Chomsky (1986) claims that extraction out of the *wh*-phrase in the Spec of C^{max} is acceptable:

(i) [de que autora]$_i$ no sabes [[qué traducciones t_i] han ganado premios internacionales]
'by what author don't you know what translations have won international awards'
(Chomsky 1986:26)

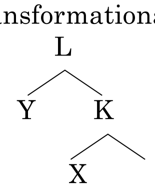
If this observation is correct, then this counts as evidence against this locality condition. Jim Huang (personal communication) suggested to me that the contrast between (i) and (61-63) might be due to the difference in the categorial status of the extracted phrases. While the extracted phrase is a PP in the Spanish example, those in the English examples are NPs. There are cases where PP-movement but not NP-movement is allowed for some unknown reason:

(ii) a. *These students, two of are crazy
b. Of these students, two are crazy

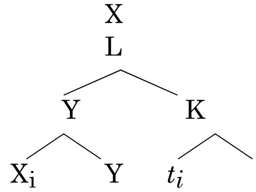
I leave further discussion of this subject for future research.

[25]Note that this locality condition differs from the freezing principle advocated by, among others, Culicover (1976), Culicover and Wexler (1973a, 1973b, 1977), Culicover and Wilkins (1984), and Wexler and Culicover (1980). The freezing principle states that no transformation may apply to any node under a frozen node, where a frozen node is one which does not immediately dominate a base structure. Suppose, for instance, that X is adjoined to Y transformationally:

(i) a.



b.



Suppose further that such an adjunction structure cannot be base-generated. The freezing principle claims that the node L is frozen. No transformation may apply to any node which is dominated by L.

(64) a. [$_{T}$max T [$_{V}$max you [read [$_{D}$max articles about vowel harmony]]]]

b. you

Recall that the ICP together with the EP ensures that *you*, which undergoes Copy for making the Nominative Case feature and D-feature of T invisible at LF, may not be merged with the main structure (64a) at this point of the derivation.

Let us assume that topicalization is triggered by the strong [TOPIC] feature under a functional head F (see, among others, Authier (1992) and Watanabe (1993)). Let us assume that [TOPIC] features are uninterpretable wherever they may appear, since unlike a *wh*-element, no phrase can be interpreted as a topic in-situ. F appears between C and T. In other words, F selects T^{max} and its maximal projection is selected by C. F, which has the following hierarchical structure of features, is selected at this stage:

(65) $F_{[[TOPIC]\ T]}$

The ICP requires that the T-feature of F should be made invisible at LF immediately by erasure through combining F with T^{max}. After the T-feature of F is checked, its strong [TOPIC] feature becomes accessible to the computation. It must be made invisible at LF immediately in accordance with the ICP. This feature is made invisible at LF by deletion through copying *articles about vowel harmony*. The resultant structure is as follows:

(66) a. [$_{F}$max F [$_{T}$max T [$_{V}$max you [read **[$_{D}$max articles about vowel harmony]**]]]]

b. you

c. **[$_{D}$max articles about vowel harmony]**

It should be noted that the newly created D^{max} *articles about vowel harmony* (66c) as well as *you* may not be merged with the main structure (66a) at this stage of derivation due to the ICP and the EP. (66a-c) each constitute an independent syntactic object at this stage.

As the derivation proceeds, we come to the stage where the strong [TOPIC] feature of the matrix F is required to be made invisible at LF immediately by the ICP, as shown below:

(67) a. [$_{F}$max F$_{[TOPIC]}$ [$_{T}$max T [$_{V}$max I [think [$_{C}$max that [$_{F}$max F [$_{T}$max T [$_{V}$max you [read [$_{D}$max articles about vowel harmony]]]]]]]]]]

b. you

c. [$_{D}$max articles about vowel harmony]

d. I

Note that *vowel harmony* within the D^{max} (67c), even if it has a [TOPIC] feature, may not be subject to the copy operation triggered by the strong [TOPIC] feature of the matrix F in (67a). This is because (67a) and (67c) each constitute an independent syntactic object at this point and thus the matrix F in (67a) does not c-command *vowel harmony* within the D^{max} (67c).

The alternative way of making the strong [TOPIC] feature of F invisible at LF is to copy *vowel harmony* within the complement position of the verb *read*, as shown below:

(68) a. [$_{F}$max F [$_{T}$max T [$_{V}$max I [think [$_{C}$max that [$_{F}$max F [$_{T}$max T [$_{V}$max you [read [$_{D}$max articles about **vowel harmony**]]]]]]]]]]

b. you

c. [$_{D}$max articles about vowel harmony]

d. I

e. **vowel harmony**

After combining these five syntactic objects, we get the following structure:

(69) [$_{F}$max vowel harmony [$_{T}$max I [$_{V}$max I [think [$_{C}$max that [$_{F}$max [$_{D}$max articles about vowel harmony] [$_{T}$max you [$_{V}$max you [read [$_{D}$max articles about vowel harmony]]]]]]]]]]

Although this derivation converges, I claim that its LF-representation violates the ban against vacuous quantification as well as the condition of inclusiveness.

Chain formation applies in the LF-component. In (69), *articles about vowel harmony* is introduced a second time into a syntactic object by Copy. The two occurrences of *articles about vowel harmony* are identical in constitution but positionally distinct. Furthermore, *vowel harmony* in the matrix Spec of F^{max} is introduced by copying *vowel harmony* within the complement position of the verb *read*. These two occurrences of *vowel harmony* are also identical in constitution but positionally distinct. The following two chains are therefore created by these two copy operations :

(70) a. CH = (*articles about vowel harmony*, *articles about vowel harmony*)

b. CH = (*vowel harmony*, *vowel harmony*)

It is important to note that (70b) consists of *vowel harmony* in the Spec of F^{max} and *vowel harmony* in the complement position of the verb *read*. This derivation yields the following LF-representation after the deletion of the non-head members of the chains:

(71) [$_{F}{}^{max}$ **vowel harmony** [$_{T}{}^{max}$ I_k [$_{V}{}^{max}$ t_k [think [$_{C}{}^{max}$ that [$_{F}{}^{max}$ [$_{\mathbf{D}}{}^{\mathbf{max}}$ **articles about vowel harmony**]$_{\mathbf{i}}$ [$_{T}{}^{max}$ you$_j$ [$_{V}{}^{max}$ t_j [read $\boldsymbol{t_i}$]]]]]]]]]

This LF-representation violates the ban against vacuous quantification. This is because *vowel harmony*, which counts as an operator in the topic position, does not bind any variable. Hence, this derivation crashes at LF. Furthermore, this LF-represetation can also be excluded by the condition of inclusiveness. This is because there are two occurrences of *vowel harmony* in (71) although the N of this derivation only includes one occurrence of *vowel harmony*. Therefore, the ICP together with the EP correctly predicts that topicalization out of a topic phrase, which is exemplified by (61), is not allowed. Similarly, the ICP coupled with the EP prohibits wh-movement from applying to *wh*-phrases inside moved *wh*-phrases and topicalized phrases as exemplified by (62)-(63). Hence, the ICP together with the EP gives us a minimalist account of the fact that no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement.

3.4 Adjunct A'-movement and the "Domain Barrier" Effects

In the preceding sections, I have argued that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure. It was also shown that our theory of phrase structure accounts for the fact that no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. The discussion in the preceding sections, however, has confined its attention to argument A'-movement. A question now arises how to deal with locality conditions on adjunct A'-movement. In this section, I will argue that the "domain barrier" effects with adjunct A'-movement also follow from our theory of phrase structure.

3.4.1 Adjunct A'-movement

Before turning to the "domain barrier" effects with adjunct A'-movement, let us first consider how to construct simplex *wh*-adjunct questions, taking (72) as an example:

(72) why did he leave

During the derivation of (72), we construct (73), conforming to the ICP and the EP:

(73) a. [$C_{[Q]}$ [T^{max} T [V^{max} he leave]]]

b. he

Since the strong Q-feature of C is uninterpretable, the ICP requires that it should be checked immediately. Note that *why*, whose merger has not been triggered by any UFF, has not been introduced into the derivation but is still in the N. Following Bromberger (1986), Longobardi (1984), Reinhart (1982), and Rizzi (1990), I claim that *why* is the *wh*-version of a sentential adverb. Exactly like other sentential adverbs, *why* is simply required to have the clause it modifies in its immediate c-command domain in order to be properly interpreted at LF. Hence, *why*, being a sentential adverb, is base-generated in the Spec of C^{max} which it modifies. I therefore argue that it is direct merger of *why* into the Spec of C^{max} that checks the strong Q-feature of C:

(74) a. $[_{C^{max}}$ **why$_{[Q]}$** [C $[_{T^{max}}$ T $[_{V^{max}}$ he leave]]]]

b. he

The strong Q-feature of C, being uninterpretable, is erased by this merger. On the other hand, the Q-feature of *why*, being interpretable, remains intact. Finally, *he* is merged into the Spec of T^{max}, yielding (72). Under this analysis, unlike *wh*-arguments like *who* and *what*, there is no trace of *why* within the clause it modifies. In other words, *why* does not have any variable to bind. This is supported by the fact that unlike *wh*-arguments like *who* and *what*, *why* does not range over individuals but over higher order entities, yielding a function as its value (see, among others, Reinhart (1992, 1993) and Szabolsci and Zwarts (1992)).

Let us next consider how our analysis constructs complex *wh*-adjunct questions, taking as an example (75) where *why* modifies the embedded clause:

(75) why do you think that John left

I argue that *why* is first merged into the embedded Spec of C^{max}, where it gets an interpretation at LF, and then raised to the matrix Spec of C^{max}.

Let us look at the derivation of (75) in detail. During the derivation, we come to the following structure:

(76) a. $[C_{[Q]}$ $[_{T^{max}}$ T $[_{V^{max}}$ John left]]]

b. John

Recall that we are assuming that C is divided into two types, [+Q]-C and [-Q]-C. Between these two types of C, while [-Q]-C never has a strong Q-feature, [+Q]-C has a strong Q-feature as its intrinsic feature, though, for simplicity, we have stuck to the expression that C has a strong Q-feature as its intrinsic feature. Since the Q-feature of C is uninterpretable, it never contributes any content to LF. The interrogative/noninterrogative interpretation of a clause depends not on the existence of a Q-feature in C but on the existence of a *wh*-element in the Spec of C^{max} at LF. A clause is interpreted as interrogative when the Spec of C^{max} is occupied by a *wh*-element at LF. Otherwise, it is interpreted as noninterrogative. I claim that the embedded C in (76) is a [+Q]-C, which has a strong Q-feature as its intrinsic property. The strong Q-feature of the embedded C, being uninterpretable, is checked by merger of *why* with (76):

(77) a. $[_{C}{}^{max}$ $\mathbf{why_{[Q]}}$ [C $[_{T}{}^{max}$ T $[_{V}{}^{max}$ John left]]]]

b. John

The strong Q-feature of C, being uninterpretable, is erased by this merger. On the other hand, the Q-feature of *why*, being interpretable, remains intact.

As the derivation proceeds, we come to the stage where the matrix C is selected and merged with the main structure. The matrix C is also a [+Q]-C and thus assigned a strong Q-feature as its intrinsic property. The strong Q-feature must be checked immediately in accordance with the ICP:

(78) a. $[C_{[Q]}$ $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you think $[_{C}{}^{max}$ $why_{[Q]}$ [C $[_{T}{}^{max}$ T $[_{V}{}^{max}$ John left]]]]]]]

b. John

c. you

The strong Q-feature of the matrix C is checked by coping *why* in the embedded Spec of C^{max}, which is eventually merged into the matrix Spec of C^{max} for PF-convergence. We get the following structure after merger of *why*, *you*, and *John* and application of the construction of an operator-variable pair, with all formal features including Q-features being ignored:

(79) $[_{C^{max}}$ **why**$_{\mathbf{i}}$ [C $[_{T^{max}}$ you [T $[_{V^{max}}$ t_k think $[_{C^{max}}$ $\boldsymbol{t_i}$ [C $[_{T^{max}}$ John$_j$ [T $[_{V^{max}}$ t_j left]]]]]]]]]]

Note that under this analysis, although the embedded C has a strong Q-feature, it is erased by merger of *why*. Following Lasnik and Saito (1992), we are assuming that the trace in the Spec of C^{max} left by wh-movement is not interpreted as a *wh*-element. Then, the embedded clause, whose specifier position is not occupied by a *wh*-element at LF, is not interpreted as an interrogative. Since *why* has its trace in the embedded Spec of C^{max}, it has the embedded clause in its immediate c-command domain. Hence, *why* is interpreted as modifying the embedded clause at LF.[26]

---

[26]It is necessary to explicate how our analysis deals with partial wh-movement discussed by, among others, Cheng (1997), Dayal (1994), and McDaniel (1989). In partial wh-movement languages, a *wh*-phrase which is supposed to move to the Spec of an interrogative clause may only move to the Spec of a noninterrogative intermediate clause, as exemplified by the following German examples:

(i) a. **mit wem** glaubst [du [dass [Hans meint [dass [Jakob ***t***
with whom believe you that Hans think that Jakob
gesprochen hat]]]]]
talked has
'with whom do you believe that Hans thinks that Jakob talked'

b. **was** glaubst [du [**mit wem** [Hans meint [dass [Jakob ***t***
WH believe you with whom Hans think that Jakob
gesprochen hat]]]]]
talked has

c. **was** glaubst [du [**was** [Hans meint [**mit wem** [Jakob ***t***
WH believe you WH Hans think with whom Jakob
gesprochen hat]]]]]
talked has

(McDaniel 1989:575)

In (ia), the *wh*-phrase *mit wem* 'with whom' moves to the matrix interrogative clause. In (ib-c), on the other hand, the *wh*-phrase stays in the intermediate clause. The matrix

One might claim that this analysis would allow any intermediate clause of a *wh*-question to be headed by a [+Q]-C with a strong Q-feature. The proliferation of a strong Q-feature, however, is blocked by the following condition proposed by Chomsky (1995):

(80) α enters into a numeration (N) only if it has an effect on output.

(Chomsky 1995:294)

Let us consider (81) as an example:

(81) **what** do you think that John read ***t***

In constructing the N of (81), condition (80) prevents the embedded clause from being headed by a [+Q]-C with a strong Q-feature. This is because the strong Q-feature would not have any effect on PF or LF. If the embedded C had a strong Q-feature, the *wh*-phrase *what* would undergo successive cyclic movement, first moving to the Spec of the embedded C^{max} and then to the Spec of the matrix C^{max}. It would not have any

---

Spec of C^{max} is occupied by the scope marker *was*. The scope marker *was* also appears in every intermediate Spec of C^{max} between the matrix Spec of C^{max} and the *wh*-phrase.

Cheng (1997) argues that in partial wh-movement like (ib-c), the Q-feature of a *wh*-phrase (to be precise, a bundle of feature including the Q-feature) undergoes overt movement, leaving a copy. The Q-feature of a *wh*-phrase (strictly speaking, the bundle of formal features) and its copy are spelled out as the scope marker *was* at PF unless it is followed by category movement. In the latter case, the repair strategy in the sense of Chomsky (1995) takes place at PF between the bundle of formal feature sand the category, yielding overt wh-movement. Hence, in (ib-c), the chain consisting of the *wh*-feature (strictly speaking, the bundle of formal features) is formed in the LF component. Given that the non-head positions of a chain delete, the *wh*-feature only remains in the head position of the chain, i.e., the matrix Spec of C^{max}. Hence, we can correctly predict that the matrix clause, whose Spec is occupied by the *wh*-feature at LF, is interpreted as interrogative while the other clauses, whose Spec is not occupied by any *wh*-feature at LF, are interpreted as noninterrogative. According to our analysis, the intermediate clauses in (ib-c) are all headed by [+Q]-C, which ensures successive cyclic movement of a *wh*-feature. This does not violate (80), since the strong Q-feature of those [+Q]-C's has a PF effect. A question still remains, however, what forces successive cyclic movement of a *wh*-feature. I am indebted to Lisa Cheng (personal communication) for bringing my attention to this subject.

effect on PF, however, since it would not change the phonetic form. It would not have any effect on LF either, since it would not change the LF-representation. Note that the intermediate trace in the Spec of the embedded C^{max} would be eliminated to make the heterogeneous chain created by successive cyclic wh-movement a legitimate object, an operator-variable pair. Hence, the embedded C cannot be a [+Q]-C with a strong Q-feature in (81).

In complex adjunct *wh*-questions like (75), however, the strong Q-feature of the embedded [+Q]-C has an effect on LF. Let us consider (75) again. If the embedded clause is headed by a [+Q]-C with a strong Q-feature, *why* is interpreted as modifying the embedded clause. If the embedded C does not have any strong Q-feature, *why* is not interpreted as modifying the embedded clause but rather as modifying the matrix clause. Hence, the embedded clause in (75) is headed by a [+Q]-C with a strong Q-feature without violating condition (80).[27, 28]

There is empirical evidence to support the base-generation of *why* in the Spec of C^{max}. It has been observed by, among others, Huang (1982) and Lasnik and Saito (1992) that *why* never stays in-situ:

---

[27]One might argue against this analysis, claiming that (80) necessarily calls for global considerations. When constructing an N, we must look ahead to see whether the inclusion of a lexical item into the N has any effect on the interface levels. Recall, however, that the construction of an N is not part of a computation. Hence, even if (80) is global in nature, we can still maintain the view that a computation only needs local information.

[28]As discussed in detail in chapter 1, in languages like Irish where [+Q]-C and [-Q]-C have distinct phonetic forms, the intermediate clause in a *wh*-question may be headed by a [+Q]-C without violating condition (80). This is because the existence of a strong Q-feature has an effect on PF. Morphological realizations of successive cyclic wh-movement are also observed in Chamorro (Chung 1982, 1994), Ewe (Collins 1993), Kikuyu (Clements 1984), Moore (Haïk 1990), and Palauan (Georgopoulos 1985, 1991a, 1991b). Condition (80) also allows a strong Q-feature to be assigned to C in the intermediate clause of a *wh*-question in these languages, since the existence of a strong Q-feature has a PF effect.

(82) a. *who left **why**

b. *who thinks Mary left **why**

In the EST, the unacceptability of examples like (82) has been attributed to the ECP. Since the ECP is not available in the MP anymore, we have to look for an alternative account of the distribution of *why*. Let us consider how our analysis accounts for its distribution, taking (82a) as an example. During the derivation, we construct the following structure:

(83) a. $[C_{[Q]}\ [_{T^{max}}\ T\ [_{V^{max}}\ \text{who}_{[Q]}\ \text{left}]]]$

b. $\text{who}_{[Q]}$

When we come to the stage of the derivation where the strong Q-feature of the matrix C is to be checked, there are two possible continuations. We either copy *who* or select and merge *why*. Among these options, the EP requires that we should choose the latter. The strong Q-feature of C can only be checked by merger of *why* but not by copy of *who*. Hence, there is no legitimate way of deriving (82a). The unacceptability of (82b) can be accounted for in the same way. Our analysis therefore offers a minimalist account of the lack of *why* in-situ.

With the above discussion in mind, the next subsection considers the "domain barrier" effects with adjunct A'-movement.

3.4.2 The "Domain Barrier" Effects with Adjunct A'-movement

Let us consider the "domain barrier" effects with adjunct A'-movement, taking adjunct wh-movement as an example:

(84) Complex NP Constraint (CNPC)

a. Relative Clauses

****why** do you like [books that criticize John]

b. Non-relative Complex NPs

*****why** did you deny [the evidence that Harry stole money]

(85) Subject Condition
*****why** was [that John bought the book] predicted by you

(86) Adjunct Condition

*****why** do you get jealous [because I spoke to Mary]

(87) Non-bridge Verb Condition

*****why** did John whisper [that Bill criticized Mary]

(84)-(87) are all deviant under the readings where *why* modifies the clauses within the islands.

Let us first consider the non-relative clause case of the CNPC, taking (84b) as an example. Considering the appositive clause in (84b), i.e., *that Harry stole money*, we can construct it through checking the UFFs of the selected items in accordance with the ICP and the EP. Note that *why*, whose merger has not been triggered by any UFF, has not been selected or merged yet. The resultant structure is as follows:

(88) a. [$_{C}$^{max} that [$_{T}$^{max} T [$_{V}$^{max} Harry stole money]]]

b. Harry

Note in passing that at this point of the derivation, since merger of *Harry* is not triggered by any UFF, *Harry* has not been merged with (88a) yet. In other words, there are two independent syntactic objects at this stage of the derivation as far as the appositive clause is concerned.

Turning to the main structure, let us consider the stage where we construct the following structure:

(89) a. [$_{C}$^{max} C$_{[Q]}$ [$_{T}$^{max} T [$_{V}$^{max} you [deny the evidence]]]]

b. you

Note that the copied *you* is not merged with (89a) due to the ICP and the EP. The next step of the derivation is to check the strong Q-feature of C. This stage of the derivation is characterized as the following four syntactic objects including the appositive clause:

(90) a. $[_{C}{}^{max}$ $C_{[Q]}$ $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you [deny the evidence]]]]

b. you

c. $[_{C}{}^{max}$ that $[_{T}{}^{max}$ T $[_{V}{}^{max}$ Harry stole money]]]

d. Harry

Note that *why* has not been selected or merged until this stage.

The strong Q-feature of C can be checked if we first select *why* and then merge it into the Spec of C^{max}:

(91) a. $[_{C}{}^{max}$ $\mathbf{why_{[Q]}}$ [C $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you [deny the evidence]]]]]

b. you

c. $[_{C}{}^{max}$ that $[_{T}{}^{max}$ T $[_{V}{}^{max}$ Harry stole money]]]

d. Harry

After merging all these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(92) $[_{C}{}^{max}$ why [C $[_{T}{}^{max}$ you [T $[_{V}{}^{max}$ you [deny the evidence $[_{C}{}^{max}$ that $[_{T}{}^{max}$ Harry [T $[_{V}{}^{max}$ Harry stole money]]]]]]]]]

Although this derivation converges, this is not what we want for (84b), where *why* modifies the appositive clause. Recall that *why* is interpreted at LF as modifying the clause in its immediate c-command domain. In (92), since *why* has the matrix clause in its immediate c-command domain, it is interpreted as modifying the matrix clause but not the appositive clause. Hence, from the N where only the matrix clause is headed by a [+Q]-C with a strong Q-feature, there is no way of generating (84b).

Alternatively, we can think of a derivation based on the N where both the matrix and embedded clauses are headed by [+Q]-C's with strong Q-features. In the derivation, the C of the appositive clause also has a strong Q-feature:

(93) a. $[_{C^{max}}$ $that_{[Q]}$ $[_{T^{max}}$ T $[_{V^{max}}$ Harry stole money]]]

b. Harry

The strong Q-feature is checked if we select *why* and merge it into the Spec of C^{max}:

(94) a. $[_{C^{max}}$ $\mathbf{why_{[Q]}}$ [that $[_{T^{max}}$ T $[_{V^{max}}$ Harry stole money]]]]

b. Harry

After the checking operation takes place, the strong Q-feature of C is erased while the Q-feature of *why* remains. Note that this derivation would yield a representation where *why* modifies the appositive clause, since *why* is merged into the position where it takes the appositive clause in its immediate c-command domain.

Turning to the main structure, let us consider the stage where the strong Q-feature of the matrix C is to be checked:

(95) a. $[_{C^{max}}$ $C_{[Q]}$ $[_{T^{max}}$ T $[_{V^{max}}$ you [deny the evidence]]]]

b. you

Note that the appositive clause (94a) has not been merged with the main structure (95a). The matrix C does not c-command *why*, which stays within a different syntactic object. The former cannot be checked by copying the latter. There is no way of checking the strong Q-feature of the matrix C. The derivation is canceled due to a violation of the ICP.

We cannot generate (84b) either from the N where both the matrix and embedded clauses are headed by [+Q]-C's with strong Q-features or from the one where only the matrix clause is headed by a [+Q]-C with a strong Q-feature. Hence, there is no legitimate way of deriving (84b). The non-relative case of the CNPC with adjunct wh-movement follows. The relative clause case of the CNPC, the Adjunct Condition, and the non-bridge verb condition can be accounted for in similar fashions.[29]

Let us next consider the subject condition effects with adjunct wh-movement, taking (85) (repeated here as (96)) as an example:

(96) ***why** was [that John bought the book] predicted by you

If we construct the N where only the matrix clause is headed by a [+Q]-C with a strong Q-feature, *why* is directly merged into the Spec of the matrix C^{max}. Although this derivation converges, it only derives an LF-representation where *why* modifies the matrix clause but not the embedded clause. Hence, we cannot generate (96) based on that N.

Let us next consider the N where not only the matrix clause but also the sentential subject is headed by a [+Q]-C with a strong Q-feature. Let us first consider the construction of the sentential subject *that John bought the book*. During the derivation, we come to the stage where the strong Q-feature of C of the sentential subject is to be checked:

---

[29]The relative case of the CNPC is worth a mention in passing. One of its logically possible derivations is based on the N where the embedded clause as well as the matrix clause is headed by a [+Q]-C with a strong Q-feature. In that derivation, multiple specifier constructions proposed by Chomsky (1995) come into a play if we conform to the assumption that checking only takes place in specifier positions. Since the relative clause involves empty operator movement to the Spec of the relative clause, *why* must be merged into another Spec. If we assume contra Chomsky that checking may take place in adjoined positions, on the other hand, then *why* is adjoined to the relative clause to check the strong Q-feature. The same is also true of the clausal adjunct case of the Adjunct Condition if Larson's (1987) empty operator movement analysis of a clausal adjunct is correct.

(97) a. [that$_{[Q]}$ [T^{max} T [V^{max} John [bought the book]]]]

b. John

The strong Q-feature of the C is checked immediately by merger of *why* into the Spec of the sentential subject, resulting in the following structure:

(98) a. [C^{max} **why$_{[Q]}$** [that [T^{max} T [V^{max} John [bought the book]]]]]

b. John

The strong Q-feature of C, being uninterpretable, is erased by this merger. The Q-feature of *why*, being interpretable, remains intact.

Turning to the main structure, we come to the stage where the strong feature of the matrix T is to be checked:[30]

(99) a. [T [V^{max} [C^{max} why$_{[Q]}$ [that [T^{max} T [V^{max} John [bought the book]]]]] [was predicted by you]]]

b. John

The strong feature of T is checked by copying the sentential subject:

(100) a. [T [V^{max} **[C^{max} why$_{[Q]}$ that John bought the book]** [was predicted by you]]]

b. John

c. **[C^{max} why$_{[Q]}$ that John bought the book]**

As the derivation proceeds, we come to the stage where the strong Q-feature of the matrix C is to be checked

---

[30]I assume following, among others, Boskovic (1995) and Delahunty (1983), that exactly like D^{max} subjects, sentential subjects originate in the Spec of V^{max} and then move to the Spec of T^{max}. It is not clear at this point, however, what strong feature triggers the raising of a sentential subject to the Spec of T^{max}. See Authier (1991), Koster (1978), and Safir (1985) for a different view.

(101) a. $[C_{[Q]}\ [_{T^{max}}\ T\ [_{V^{max}}\ [_{C^{max}}\ why_{[Q]}$ that John bought the book] [was predicted by you]]]]

b. John

c. $[_{C^{max}}\ why_{[Q]}$ that John bought the book]

The strong Q-feature of the matrix C can only be checked by copying *why* within the Spec of V^{max} but not *why* in (101a). This is because the matrix C c-commands the former but not the latter. This yields the following structure:

(102) a. $[C\ [_{T^{max}}\ T\ [_{V^{max}}\ [_{C^{max}}\ \mathbf{why_{[Q]}}$ that John bought the book] [was predicted by you]]]]

b. John

c. $[_{C^{max}}\ why_{[Q]}$ that John bought the book]

d. $\mathbf{why_{[Q]}}$

After combining these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(103) $[_{C^{max}}\ why_i\ [C\ [[_{C^{max}}$ why that Johnbought the book$]_j$ [T $[[_{C^{max}}\ why_i$ that bought the book$]_j$ [was predicted by you]]]]]]

Chain formation applies in the LF-component. Recall that *why that John bought the book* in the Spec of the matrix T^{max} is introduced by copying *why that John bought the book* in the Spec of V^{max}. These two occurrences are therefore identical in constitution but positionally distinct. *Why* in the Spec of the matrix C^{max} is introduced by copying *why* within the Spec of V^{max}. Hence, these two occurrences of *why* are identical in constitution but positionally distinct. These two copy operations create the following two chains:

(104) a. CH = (*why that John bought that book*, *why that John bought that book*)

b. CH = (*why*, *why*)

If we delete the non-head positions of the chains, we get the following LF-representation:

(105) [$_{C^{max}}$ **why** [C [**[$_{C^{max}}$ why that John bought the book]$_j$** [T [***t$_j$*** [was predicted by you]]]]]]

This apparently violates the condition of inclusiveness. This is because there are two occurrences of *why* in (105) while the N for this derivation only includes one occurrence of *why*. This representation is also excluded by the selectional property of the verb. Recall that we are assuming that the Q-feature of a *wh*-phrase is interpretable while the Q-feature of C is uninterpretable. The clause whose specifier position is occupied by a *wh*-phrase at LF is interpreted as an interrogative. Since the sentential subject has *why* in its specifier position at LF, it is interpreted as an interrogative. It is well known, however, that no verb selects an interrogative in its subject position, not to mention the verb *be*. Note in passing that since *why* does not need any variable to bind, there is no violation of the ban against vacuous quantification in (105).

To summarize this section, I have first argued that an adjunct *wh*-phrase should be directly merged into the Spec of C^{max} to check the strong Q-feature of C. I have then argued that the "domain barrier" effects with adjunct wh-movement straightforwardly follows from our theory of phrase structure.

3.5 Concluding Remarks

This chapter has considered the "domain barrier" effects with feature-driven A'-movement and shown that they straightforwardly follow from our theory of the composition of phrase structure. It was also shown that our analysis of the "domain barriers" can be extended to account for another locality condition which constraints extraction out of moved phrases. I have argued that the pre-minimalist locality theories are incompatible with the MP, since they crucially rely on the notions not available under the MP. I have also argued that our theory has conceptual advantages over the previous minimalist approaches to locality. This is because our locality theory only makes use of the ICP and the EP, which are local "heuristic algorithms" ("computational tricks") for globality induced by the interface conditions on UFFs and Ns. It was shown that our theory of locality diverges from all the previous approaches in claiming that locality facts should not follow from restrictions on movement but from restrictions on merger.

CHAPTER 4

LOCALITY ON SCRAMBLING IN JAPANESE

4.0 Introduction

In chapter 3, I have argued that the "domain barrier" effects with respect to feature-driven A'-movement like English overt wh-movement and topicalization straightforwardly follow from our theory of the composition of phrase structure. It was shown that when we come to the stage where a strong feature is to be checked, the "domain barrier," whose merger is not triggered by any UFF, has not been merged with the main structure. Hence, the strong feature cannot be checked by copying anything within the "domain barrier"; the "domain barrier" effects follow. In this chapter, I will investigate scrambling in Japanese. It is shown that unlike English overt wh-movement and topicalization, Japanese scrambling does not obey the "domain barriers." I will argue that this asymmetry between these two types of movement with the "domain barrier" effects straightforwardly follows from our theory of phrase structure if we assume following Fukui (1993b), Fukui and Saito (1996), and Saito (1994) that Japanese scrambling is not feature-driven. Since this asymmetry only follows from our phrase-structural analysis of the "domain barriers" but not from the previous analyses, it constitutes another empirical support in favor of our theory of phrase structure.

The organization of this chapter is as follows. Section 4.1 considers locality restrictions on scrambling in Japanese. It is pointed out that contrary to Saito's (1985, 1986) observation, scrambling does not exhibit any "domain barrier" effects. Section 4.2 argues that the asymmetry between feature-driven movement like English overt wh-movement and topicalization and non-feature-driven movement like Japanese scrambling concerning the "domain barrier" effects straightforwardly follows from our analysis. It is shown that scrambling, which is not driven by any formal feature, is required to apply postcyclically by the ICP and the EP. It then follows that scrambling may apply after merger of the domain barrier" with the main structure and thus does not exhibit any "domain barrier" effects. Section 4.3 discusses the apparent "domain barrier" effects with scrambling. I will argue that they should be attributed to an A-over-A condition which applies in the PF-component. Section 4.4 considers extraction out of a scrambled phrase. It is shown that extraction out of a scrambled phrase is acceptable regardless of whether it is feature-driven or not. I will argue that the nonbarrierhood of scrambled phrases also follows from our analysis. Section 4.5 deals with scrambling in generic sentences. It is shown that scrambling in generic sentences exhibits the "domain barrier" effects. I will argue that scrambling in generic sentences necessarily induces a focus reading and thus counts as a feature-driven movement. It then follows from our analysis that like English overt wh-movement and topicalization, scrambling in generic sentences exhibits the "domain barrier" effects. Section 4.6 makes concluding remarks.

4.1 Locality Restrictions on Scrambling in Japanese

4.1.1 No "Domain Barrier" Effects with Scrambling

As argued by, among others, Haig (1976), Harada (1977), Miyara (1982), Muraki (1979), and Saito (1985), scrambling is responsible for the relatively free word order in Japanese. For instance, (1b-f) are derived from (1a) through scrambling:[1]

(1) a. John-ga Mary-ni hon-o ageta (koto)
  -Nom -Dat book-Acc gave (fact)

b. John-ga **hon-o** Mary-ni ***t*** ageta (koto)

c. **Mary-ni** John-ga ***t*** hon-o ageta (koto)

d. **Mary-ni**$_j$ **hon-o**$_i$ John-ga $\boldsymbol{t_j}$ $\boldsymbol{t_i}$ ageta (koto)

e. **hon-o**$_i$ John-ga Mary-ni $\boldsymbol{t_i}$ ageta (koto)

f. **hon-o**$_i$ **Mary-ni**$_j$ John-ga $\boldsymbol{t_j}$ $\boldsymbol{t_i}$ ageta (koto)

'John gave a book to Mary'

Apart from short distance (= clause internal) scrambling exemplified by (1b-f), long-distance scrambling is also possible, as shown below:

(2) **sono mura-ni** [John-ga [Mary-ga ***t*** sundeiru to]
that village-Dat -Nom -Nom reside Comp
omotteiru] (koto)
think (fact)
'John thinks that Mary lives in that village'

In (2), *sono mura-ni* 'that village-Dat' is scrambled from within the complement clause to the clause-initial position; the result is acceptable.

Saito (1985, 1986) observes that exactly like overt wh-movement and topicalization in English, scrambling in Japanese is subject to the "domain barriers," presenting the following examples (the judgments are mine):[2]

---

[1]Here and in the relevant examples to follow, *koto* 'the fact that' is added to the end of some examples in order to avoid the unnaturalness resulting from the lack of topic in a matrix clause.

[2]Recall that the Subject Condition does not hold in Japanese due to the lack of overt subject raising.

(3) Complex NP Constraint

a. Relative Clauses

?**ano hon-o** [John-ga [[***t*** katta] hito]-o
that book-Acc -Nom bought person-Acc
sagasite iru rasii]
looking-for seem
'it seems that John is looking for the person who bought that book'

b. Non-relative Complex NPs

?**Bill-o** [John-ga [[Mary-ga ***t*** sakete iru to yuu]
-Acc -Nom -Nom avoiding Comp say
uwasa]-o kiita] (koto)
rumor-Acc heard (fact)
'John heard a rumor (which says) that Mary is avoiding Bill'

(Saito 1985:246)

(4) Adjunct Condition

a. ?**sono hon-o** [John-ga [Mary-ga ***t*** yomioete
that book-Acc -Nom -Nom finish-reading
kara] dekaketa] (koto)
after went-out (fact)
'John went out after Mary finished reading that book'

b. ?**sono hon-o** [John-ga [minna-ga ***t*** kau node]
that book-Acc -Nom all-Nom buy because
tigau hon-o katta] (koto)
different book-Acc bought (fact)
'because everyone buys that book, John bought a different one'

c. ?**Tookyoo-ni** [Mary-ga [John-ga ***t*** ikitagatte iru
Tokyo-Dat -Nom -Nom want-to-go
noni] musisite iru rasii]
although ignoring seem
'it seems that although John wants to go to Tokyo, Mary is ignoring that fact'

(Saito 1985:247)

In (3), the clause-initial phrases are extracted out from the complex NPs. In (4), the clause-initial phrases are extracted out from the adjuncts. All of these examples are mildly deviant. Based on these observations, Saito (1985, 1986) claims that scrambling is subject to the "domain barriers."

I argue contra Saito that there is evidence to suggest that scrambling in Japanese is not subject to the "domain barriers." First, although the examples in (3) and (4) are awkward, they are much better than the normal "domain barrier" violations induced by feature-driven movement like English overt wh-movement and topicalization.[3] Second, there are cases where the adjunct condition effects with scrambling are abrogated. Let us look at the following examples (cf. Ikawa (1996)):

(5) a. **sono hana-ni$_i$** [daremo$_j$-ga [e_j t_i mizu-o yara-
that flower-Dat everyone-Nom water-Acc give
zu] dekakete itta] (koto)
without went-out (fact)
'everyone went out without watering that flower'

b. **sono isu-ni$_i$** [daremo$_j$-ga [e_j t_i suwari nagara]
that chair-Dat everyone-Nom sit while
hon-o yondeita]] (koto)
book-Acc was reading (fact)
'everyone was reading a book while sitting on that chair'

c. **siai-no kekka-ni$_i$** [daremo$_j$-ga [e_j t_i totemo
game-Gen result-Dat everyone-Nom very
gakkarisite] kyuujoo- atonisita] (koto)
disappointed ball park-Acc left (fact)
'everyone left the ball park, disappointed about the result of the game'

In (5a-c), although the clause-initial phrases are extracted out of the adjuncts through scrambling, the results are acceptable. I argue that the difference between (4) and (5) resides in the fact that while the adjunct clauses of the former have overt subjects, those of the latter have empty subjects. While scrambling out of adjuncts with an empty subject is

[3]See Fukui and Saito (1996) for a similar observation.

legitimate, scrambling out of those with an overt subject is not. In other words, when scrambling takes place out of adjuncts with an empty subject, the adjunct condition effects disappear.[4]

I take these facts as evidence to suggest that unlike English overt wh-movement and topicalization, Japanese scrambling is not subject to the "domain barriers." I argue that the apparent "domain barrier" effects with scrambling observed in (3) and (4) should be attributed to a condition applying at the PF interface.

4.1.2 The "Domain Barrier" Effects in Japanese

I have argued in the last subsection that unlike overt wh-movement and topicalization in English, scrambling in Japanese does not exhibit any "domain barrier" effects. This might lead one to claim that unlike English, Japanese does not exhibit any "domain barrier" effects. I will show, however, that a *wh*-adjunct in-situ and empty operator movement in Japanese exhibit the "domain barrier" effects, arguing that the asymmetry with the "domain barrier" effects exists not between English and Japanese but between feature-driven movement and non-feature-driven movement.

---

[4]It should be noted that the immunity of the "domain barriers" as observed in (5) may not be attributed to the existence of the coreference relation between the matrix and adjunct subjects. Scrambling out of the adjunct with an overt subject which corefers with the matrix subject like (i) is deviant:

(i) ?**sono hana-ni**$_i$ [John$_j$-ga [zibun$_j$-ga ***t***$_i$ mizu-o yatte
that flower-Dat -Nom -Nom water-Acc give
kara] dekaketa] (koto)
after went-out (fact)
'John went out after he watered that flower'

As will be extensively argued in the next chapter, *wh*-adjuncts in-situ in Japanese like *naze* 'why' exhibit the "domain barrier" effects, as shown below (cf. Fukui (1988)):

(6) Complex NP Constraint

a. Relative Clauses

*John-wa [Bill-ga **naze** Mary-ni watasita tegami]-o
-Top -Nom why -Dat gave letter-Acc
sagasite iru no
looking for Q
Lit. 'John is looking for the letter which was sent to Mary why'

b. Non-relative Complex NPs

*?John-wa [Bill-ga **naze** sono kuruma-o katta
-Top -Nom why that car-Acc bought
koto]- sonnani okotte iru no
fact-Acc so much be angry Q
Lit. 'John is so angry with the fact that Bill bought that car why'

(7) Adjunct Condition

a. *John$_i$-wa [e_i sono hana-ni **naze** mizu-o yara-

-Top that flower-Dat why water-Acc give

zu] dekakete itta no

without went-o Q

Lit. 'John went out without watering that flower why'

b. *John$_i$-wa [e_i sono isu-ni **naze** suwari nagara]

-Top that chair-Dat why sit while

hon-o yondeiru no

book-Acc is reading Q

Lit. 'John is reading a book while sitting on that chair why'

c. *John$_i$-wa [e_i siken-no kekka-ni **naze**

-Top examination-Gen result-Dat why

gakkarisite] kaette itta no

disappointed went-back Q

Lit. 'John went back, disappointed about the result of the examination why'

(8) The Non-bridge Verb Condition

*?John-wa [Bill-ga **naze** sono hon-o katta tte]

-Top -Nom why that book-Acc bought Comp

tubuyaita no

murmured Q

Lit. 'why did John murmur that Bill bought that book *t*'

In (6-8), the adjunct *wh*-element in-situ *naze* 'why', which is contained within the "domain barrier," may not be associated with the matrix Q-morpheme. The distribution of *wh*-adjuncts in-situ is constrained by the "domain barriers." I will argue in the next chapter that the Q-feature of *naze 'why'*, being strong and thus uninterpretable, is required to raise to the Q-morpheme in order to be checked off. This Q-feature movement is constrained by the "domain barriers." It should be noted that in (7), the *wh*-adjuncts in-situ appear within the adjuncts with an empty subject. The results are still as severely deviant as the adjunct condition violations induced by overt wh-movement in English. This is in contrast with the lack of the adjunct condition effects with scrambling when adjuncts have an empty subject.

The "domain barrier" effects are also observed with empty operator movement, which has been assumed to be triggered by a strong feature of C. There are several constructions which have been argued to involve empty operator movement in Japanese. First, according to Hoji (1990), the cleft construction in Japanese may involve empty operator movement. In the cleft construction in Japanese, an NP, a Case-marked NP, or a PP appears in the focus position, as exemplified below:

(9) The Cleft Construction with an NP Focus

John-ga yonda no wa **sono hon** da

-Nom read Comp Top that book be

'it is that book that John read'

(10) The Cleft Construction with an NP-Case Focus

John-ga yonda no wa **sono hon-o** da

-Nom read Comp Top that book-Acc be

'it is that book that John read'

(11) The Cleft Construction with a PP Focus

John-ga itta no wa **Tookyoo-ni** da

-Nom went Comp Top Tokyo-Dat be

'it is to Tokyo that John went'

Hoji argues that the cleft construction with an NP-Case or PP focus necessarily involves empty operator movement. Under his analysis, therefore, (11), which has a PP focus, is assigned structure (12):[5]

(12) [[$\boldsymbol{Op_i}$ [John-ga $\boldsymbol{t_i}$ itta]] no]-wa **Tookyoo$_i$-ni** da

In (12), the empty operator moves from its original position to the Spec of CP, where it is associated with the PP focus *Tookyoo-ni* 'to Tokyo' through predication.

The cleft construction with an NP focus, on the other hand, involves either empty operator movement or an empty pronoun. Hence, (9), which has an NP focus, is assigned either (13a) or (13b) depending on which strategy is to be employed:

(13) a. [$\boldsymbol{Op_i}$ [John-ga $\boldsymbol{t_i}$ yonda] no]-wa **sono hon$_i$** da

b. [[John-ga $\boldsymbol{pro_i}$ yonda] no]-wa **sono hon$_i$** da

In (13a), the empty operator *Op* moves to the Spec of CP, where it is associated with the NP focus *sono hon* 'that book' through predication. In (13b), the empty pronoun *pro* is base-generated and associated with the NP focus *sono hon* 'that book' through the aboutness relation advocated by Kuno (1973) and Saito (1985). In the following discussion, we will only take the cleft construction with an NP-Case or PP focus but not the one

[5]Following Kikuchi (1987) and Hoji (1990), we assume that an empty operator moves leftward to the clause-initial position, though the present discussion holds regardless of the directionality of empty operator movement.

with an NP focus as an example in order to exclude the possibility of the base-generation of an empty pronoun.

Second, Takezawa (1987) argues that the *tough* construction may involve empty operator movement. Among four types of the *tough* constructions presented by Inoue (1978), he only deals with Type IV, which exhibits different syntactic behaviors than the other types (see Kuroda (1978) and Saito (1982)). The *tough* construction has either an NP or PP as its subject, as exemplified below:

(14) The *Tough* Construction with an NP Subject
sono hon-ga John-nitotte yomi-nikui (koto)
that book-Nom -for hard-to-read (fact)
'it is hard for John to read that book'

(15) The *Tough* Construction with a PP Subject
imooto-kara-ga John-nitotte okane-o kari-yasui
sister-from-Nom -for money-Acc easy-to-borrow
(koto)
(fact)
'it is easy for John to borrow money from his sister'

Following the insight given by the analysis of Japanese topicalization in Saito (1985), Takezawa claims that while the *tough* construction with a PP subject necessarily involves empty operator movement, the one with an NP subject does not. The *tough* construction with an NP subject has two possible derivations. It is derived by either empty operator movement or the base-generation of an empty pronoun.

For example, (15), which has a PP subject, is only assigned structure (16):

(16) [**imooto-kara**$_{\mathbf{i}}$-ga [John$_j$-nitotte [$\boldsymbol{Op_i}$ [e_j $\boldsymbol{t_i}$ okane-o kari]] yasui]]

In (16), the empty operator *Op* moves from its original position to the Spec of CP, where it is associated with the PP subject *imooto-kara* 'from his sister' through predication. On the other hand, (14), which has an NP subject, has the following two derivations:

(17) a. [**sono hon**$_{\mathbf{i}}$-ga [John$_j$-nitotte [$\boldsymbol{Op_i}$ [e_j $\boldsymbol{t_i}$ yomi]] nikui]]

b. [**sono hon**$_{\mathbf{i}}$-ga [John$_j$-nitotte [e_j $\boldsymbol{pro_i}$ yomi] nikui]]

In (17a), the empty operator *Op* moves to the Spec of CP, where it is associated with the NP subject *sono hon* 'that book' through predication. In (17b), the empty pronoun *pro* is base-generated and associated with the NP subject *sono hon* 'that book' through the aboutness relation. In order to exclude the possibility of the base-generation of an empty pronoun, we will only consider the *tough* construction with a PP subject but not the one with an NP subject as an example in the following discussion.

Third, Ishii (1991) and Kikuchi (1987) argue that the comparative deletion construction involves empty operator movement. Although they differ as to the categorial status of the empty operator involved in this construction, we will assume Kikuchi's analysis for expository purposes. It should be noted that the arguments to follow hold under either of the analyses. Under Kikuchi's analysis, (18) is assigned structure (19):

(18) John-ga tabeta yorimo Tom-wa keeki-o takusan tabeta
-Nom ate than -Top cake-Acc many ate
'Tom ate more cakes than John ate'
(Kikuchi 1987:4)

(19) [***Op$_i$*** [John-ga ***t$_i$*** tabeta] yorimo] Tom-wa keeki-o takusan tabeta

In (19), the empty operator *Op* moves from its original position to the Spec of CP.

It has been observed (see the references cited above) that the dependencies involved in these three constructions are unbounded in nature. In other words, so called "long-distance" empty operator movement is possible in these constructions, as shown below:

(20) The Cleft Construction
[***Op$_i$*** [John-ga [Mary-ga kinoo ***t$_i$*** sono syorui-o
-Nom -Nom yesterday that document-Acc
miseta to] omotteiru] no]-wa [ano CIA agent]$_i$-ni da
showed Comp think Comp-Top that CIA agent-Dat be
'it is to that CIA agent that John thinks that Mary showed that document yesterday'
(Hoji 1990: Ch5, 32)

(21) The *Tough* Construction

[[zibun-no otooto-kara]$_i$-ga [(John$_j$-nitotte) [***Op$_i$*** [*e$_j$* [*e$_j$* ***t$_i$***

self's brother-from-Nom -for

okane-o takusan karite iru to] mitome]] nikui]] (koto)

money-Acc a lot borrow Comp admit hard (fact)

Lit. '[from self's brother]$_i$ is hard (for John$_j$) to admit that he$_j$ has borrowed a lot of money *e$_i$*'

(Takezawa 1987:196)

(22) The Comparative Deletion Construction

[***Op$_i$*** [[[John-ga ***t$_i$*** yonda to] iwareteiru to] Tom-ga

-Nom read Comp is said Comp -Nom

uwasa-siteiru] yorimo] Mary-wa takusan hon-o yondeita

rumor make than -Top many book-Acc read

'Mary read more books than Tom made the rumor that it is said that John read'

(Kikuchi 1987:12)

In (20-22), the empty operator is extracted out of the complement clause. The results are acceptable.

Exactly like overt wh-movement in English, the unbounded dependencies involved in these constructions exhibit the "domain barrier" effects. They are subject to the CNPC, as shown in (23-25):

(23) The Cleft Construction

*?[$\boldsymbol{Op_i}$ [John-ga [[e_j $\boldsymbol{t_i}$ atta-koto-ga aru] nihonzin$_j$]-o

-Nom have met Japanese-Acc

oozei sitte iru] no]-wa Russell$_i$-ni da

many know Comp-Top Russell-Dat be

'It is with Russell$_i$ that John knows many Japanese that have met e_i'

(Hoji 1990:Ch5, 31)

(24) The *Tough* Construction

*[[anna taipu-no zyosei-to]$_i$-ga [(John$_j$-nitotte) [$\boldsymbol{Op_i}$ [e_j [[e_k $\boldsymbol{t_i}$

that type of woman-with-Nom -for

kekkon site iru] otoko$_k$]-to hanasi]] nikui]] (koto)

marry man-with talk hard (fact)

Lit. '[with that type of woman]$_i$ is hard (for John) to talk to the man who marry e_i'

(Takezawa 1987:215)

(25) The Comparative Deletion Construction

*[$\boldsymbol{Op_i}$ [[[e_j sono tukue-de $\boldsymbol{t_i}$ yondeita] hito]-o John-ga

that table-on read person-Acc -Nom

nagutta] yorimo] Paul-wa takusan hon-o yondeita

hit than -Top many book-Acc read

Lit. 'Paul has read more books than John hit a person who was reading at that table'

(Kikuchi 1987:13)

It should be noted that (23-25) are as severely deviant as the CNPC violations induced by overt wh-movement in English.

The unbounded dependencies involved in these constructions are also constrained by the Adjunct Condition, as shown below:

(26) The Cleft Construction

a. *[***Op**$_i$* [daremo$_j$-ga [e_j mizu-o ***t**$_i$* yara-zu]
someone-Nom water-Acc give-without
dekakete itta] no]-wa sono hana$_i$-ni da
went-out Comp-Top that flower-Dat be
'It is that flower that everyone went out without watering'

b. *[***Op**$_i$* [dareka$_j$-ga [e_j ***t**$_i$* suwari nagara] hon-o
someone-Nom sit while book-Acc
yondeita] no]-wa [sono isu]$_i$-ni da
was reading Comp-Top that chair-Dat be
'It is on that chair that someone was reading the book while sitting'

c. *[***Op**$_i$* [daremo$_j$-ga [e_j ***t**$_i$* totemo gakkarisite]
everyone-Nom very disappointed
kyuujoo-o atonisita] no]-wa [siai-no
ball park-Acc left Comp-Top game-Gen
kekka]$_i$-ni da
result-Dat be
'It is the result of the game that everyone left the ball park, disappointed about'

(27) The *Tough* Construction

a. *[[kono taipu-no hasigo-kara]$_i$-ga [(syooboosi$_j$-
this type of ladder-from-Nom fireman
nitotte) [$\boldsymbol{Op_i}$ [e_j [e_j $\boldsymbol{t_i}$ asi-o humihazusa-zu]
for miss their footing-without
biru-kara hito-o kyuusyutusi]] yasui]] (koto)
building-from person-Acc rescue easy (fact)
Lit. '[from this type of ladder]$_i$ is easy for firemen to resucue persons from the buidling without missing their footing e_i'

b. *[[kono syu-no kinyuukikan-kara]$_i$-ga [(keieisya$_j$-
this type of financial company-from-Nom manager
nitotte) [$\boldsymbol{Op_i}$ [e_j [e_j $\boldsymbol{t_i}$ syakkin-o si nagara]
for have a loan of money while
kaisya-o tatenaosi]] nikui]] (koto)
company-Acc rebuild hard (fact)
Lit. '[from this type of financial companies]$_i$ is hard (for the manager) to rebuild the company while having a loan of money e_i'

(28) The Comparative Deletion Construction

a. *?[$\boldsymbol{Op_i}$ [Bill$_j$-ga [e_j $\boldsymbol{t_i}$ kiki nagara] benkyoositeita]
-Nom listen-to while was studying
yorimo] John-wa takusan recoodo-o kiita
than -Top many record-Acc listened-to
Lit. John listened to more records than Bill was studying while listening to'

b. *?[$\boldsymbol{Op_i}$ [Bill$_j$-ga [e_j $\mathbf{t_i}$ kangekisite] namida-o nagasita]

-Nom moved tear-Acc dropped

yorimo] John-wa ookuno eiga-ni kangekisita

than -Top many movie-Dat was moved

Lit. 'John was moved by more movies than Bill dropped tears, moved by'

It should be noted that in (26-28), an empty operator is extracted out of the adjunct with an empty subject. The results are still as severely deviant as the adjunct condition violations induced by overt wh-movement in English. This is in contrast with the lack of the adjunct condition effects with scrambling when adjuncts have an empty subject.

Finally, the empty operator movement is subject to the non-bridge verb condition:

(29) The Cleft Construction

*?[$\boldsymbol{Op_i}$ [John-ga [Mary-ga kinoo $\boldsymbol{t_i}$ sono

-Nom -Nom yesterday that

syorui-o miseta to] sasayaita] no]-wa [ano CIA

document-Acc showed Comp whisper Comp-Top that CIA

agent]$_i$-ni da

agent-Dat be

Lit. 'it is to that CIA agent that John whispers that Mary showed that document yesterday'

(30) The *Tough* Construction

*?[[sinyuu-kara]$_i$-ga [(John$_j$-nitotte) [***Op$_i$*** [*e$_j$* [*e$_j$* ***t$_i$***

best friend-from-Nom -for

buzyoku-o uketa to] sakebi]] nikui]]

insult-Acc suffer Comp holler hard

Lit. '[from his best friend]$_i$ is hard (for John$_j$) to holler that he$_j$ suffer an insult *e$_i$*'

(31) The Comparative Deletion Construction

*?[***Op$_i$*** [[Mary-ga ***t$_i$*** katta to] John-ga tubuyaita]

-Nom bought Comp -Nom murmured

yorimo] Susy-wa ookina daiamondo-o katta

than -Top large diamond-Acc bought

Lit. Susy bought a larger diamond than John murmured that Mary bought'

The above observations suggest that unlike scrambling in Japanese, feature-driven movement like empty operator movement and Q-feature movement in Japanese obeys the "domain barriers." The asymmetry with the existence of the "domain barrier" effects therefore exists not between English and Japanese, but between feature-driven movement and scrambling. In the next section, I will argue that this asymmetry straightforwardly follows from our analysis of the "domain barriers."

4.2 Non-feature-driven Movement and the "Domain Barriers"

This section considers the immunity of scrambling from the "domain barriers." I will argue that if we assume following Fukui (1993a), Fukui and Saito (1996), and Saito (1994) that scrambling is not feature-driven, the insensitivity of scrambling to the "domain barriers" straightforwardly follows from our theory of the composition of phrase structure.

4.2.1 Scrambling as Non-feature-driven Movement

Fukui (1993a), Fukui and Saito (1996), and Saito (1994) argue that unlike movement operations like overt wh-movement and topicalization in English, scrambling in Japanese is optional and thus not triggered by any formal feature. They present two arguments in support of this view.[6]

First, as first pointed out by Saito (1986, 1989), scrambling, but not English overt wh-movement and topicalization, is subject to "radical reconstruction":[7]

(32) a. [John-ga [[minna-ga [Mary-ga dono hon-o
-Nom all-Nom -Nom which book-Acc
yonda to] omotteiru] ka] siritagatteiru] (koto)
read Comp think Q want-to-know (fact)
'[John wants to know [Q [everyone thinks [that Mary read which book]]]]'

b. ?[**[Mary-ga dono hon-o yonda to]$_i$** [John-ga [[minna-ga $\boldsymbol{t_i}$ omotteiru] **ka**] siritagatteiru]] (koto)
(Fukui and Saito 1996:4)

In (32b), the complement clause *Mary-ga dono hon-o yonda to* 'that Mary read which book', which contains the *wh*-element *dono hon-o* 'which book',

---

[6]See Poole (1996) for another argument in favor of this view. Miyagawa (1997) argues, on the other hand, that scrambling is not optional. See Takano (1996) for extensive arguments against Miyagawa.

Note that under our analysis where Attract/Move is reinterpreted as Copy + Merge, the optionality of scrambling can be captured by the fact that scrambling is not subject to the last resort condition imposed on Copy proposed in the previous chapter.

[7]See Lee (1995) for further discussion of this subject.

is scrambled to a position outside the domain of the Q-morpheme *ka*. Although (32b) is marginal, it does not have the ungrammatical status of (33), where the *wh*-element *dare-ga* 'who-Nom' is outside the domain of the Q-morpheme *ka*:

(33) *[**dare-ga** [[John-ga sono hon-o katta] **ka**]
who-Nom -Nom that book-Acc bought Q
siritagatteiru] (koto)
want-to-know (fact)
'[who wants to know [Q [John bought that book]]]'
(Fukui and Saito 1996:3)

They argue that the contrast between (32b) and (33) follows from their claim that scrambling is "semantically vacuous" and thus scrambled phrases can be "moved back" to their original position at LF. They first assume an LF-condition which states that the *wh*-element must be within the domain of a Q-morpheme at LF. It then follows that (33) is deviant, since it violates this LF condition. (32b), on the other hand, need not violate this LF condition, since at LF, the scrambled phrase can be "moved back" to its original position and thus the *wh*-element can be within the domain of the Q-morpheme.

Fukui and Saito claim that this "radical reconstruction" property cannot be observed with English wh-movement or topicalization. Let us look at English wh-movement as an example:

(34) a. [who$_i$ [t_i wonders [[which picture of **whom**]$_j$ [Bill bought t_j]]]]

b. ??[[which picture of **whom**]$_j$ does [John wonder [who [t_i bought t_j]]]]

(Fukui and Saito 1996:5)

In (34a), *whom* can be interpreted as taking either matrix or embedded scope. In (34b), on the other hand, *whom* can only take matrix scope but not embedded scope. They assume as a descriptive generalization that a *wh*-element in-situ can be interpreted at a Spec of CP only if it is contained in the CP at LF. Then, the unambiguity of (34b) follows, since the embedded CP does not contain *whom* and thus *whom* cannot take embedded scope. It should be noted that this argument only holds if there is no "radical reconstruction" with overt wh-movement. This is because if the moved *wh*-phrase *which picture of whom* moved back to its original position at LF, then *whom* would be contained by the embedded CP and thus could be interpreted as taking embedded scope. They argue that the lack of "radical reconstruction" with wh-movement is due to the fact that a chain created for the purpose of feature-checking must be retained at LF. If this analysis is correct, we can conclude that scrambling, which allows "radical reconstruction," is not triggered by any formal feature.

Another argument in favor of the non-feature-driven property of scrambling can be formulated based on the lack of the Minimal Link Condition (MLC) effects with scrambling, as argued by Fukui and Saito (1996). The MLC, which is incorporated into the definition of Attract/Move-F in Chomsky (1995), states that when we come across a feature to be checked during a derivation, it attracts and thus enters into a checking relation with the closest appropriate feature. Hence, feature-driven movement like overt wh-movement and topicalization in English is subject to the MLC.

Chomsky (1995) argues that the Wh-island Constraint follows from the fact that wh-movement is subject to the MLC:

(35) ??[**which book**$_j$ [do you wonder [**which boy**$_i$ [to persuade t_i to buy t_j]]]]

During the derivation of (35), we come to a stage where the strong Q-feature of the embedded C is to be checked:

(36) [$C_{[Q]}$ [to persuade which boy to buy which book]]

Since *which boy* c-commands *which book*, the former is closer to C than the latter. Hence, *which boy* raises to the embedded Spec of CP, resulting in the following structure:

(37) [**which boy**$_i$ [to persuade t_i to buy which book]]

As the derivation proceeds, we come to a stage where the strong Q-feature of the matrix C is to be checked:

(38) [$C_{[Q]}$ [you wonder [**which boy**$_i$ [to persuade t_i to buy which book]]]]

Since *which boy* is closer to the matrix C than *which book*, the former raises to the Spec of CP. Note that in Chomsky (1995), the Q-feature of a *wh*-element is assumed to be interpretable and thus still accessible after checked. According to the MLC, therefore, there is no way to raise *which book* to the matrix Spec of CP; the deviant status of (35) follows. After *which boy* raises to the matrix Spec of CP, we get the following:

(39) [**which boy**$_i$ [do you wonder [t'_i [Q [to persuade t_i to buy which book]]]]

Chomsky (1995) claims that although this derivation converges, the result is gibberish.

As pointed out by Fukui and Saito (1996), however, unlike overt wh-movement in English, scrambling does not exhibit any MLC effects. Multiple scrambling like (40b) is acceptable:

(40) a. John-ga [[Bill-ga [Mary-ni [sono hon-o
-Nom -Nom -Dat that book-Acc
watasita]]] to] omotteiru (koto)
gave Comp think (fact)
'John thinks that Bill gave that book to Mary'

b. [**sono hon-o$_j$** [John-ga [[**Mary-ni$_i$** [Bill-ga [t_i [t_j
watasita]]]] to] omotteiru]] (koto)

If scrambling were subject to the MLC, (40b) would be ruled out by the MLC just like the wh-island violation cases like (35). The fact that examples like (40b) are acceptable therefore leads us to claim that scrambling is not subject to the MLC. If feature-driven movement is subject to the MLC, we can conclude that scrambling is not driven by any formal feature.

Given that Japanese scrambling is not feature-driven, it counts as optional movement. Optional movement has been under debate, since it apparently violates the last resort principle, which requires that every movement operation should have a trigger. We therefore need a theory which accommodates optionality under the MP. Fukui (1993a) and Fukui and Saito (1996) argue that certain movement operations count as "costless" for the purpose of economy and thus may take place without having any triggers. Putting the details of their analysis aside, they claim that in head-final languages like Japanese, leftward movement is "costless" while rightward movement is "costly." In head-initial languages like English, on the other hand, leftward movement is "costly" while rightward movement is "costless." Japanese scrambling, being a leftward movement operation in the head-final language, counts as "costless" and thus does not need any trigger. Saito (1994) argues, on the other hand, scrambling is not licensed by feature-checking but by virtue of its contribution to the construction of phrase structure. Under either of these approaches, scrambling is exempt from the last resort principle and thus counts as optional.

The discussion to follow assumes that scrambling in Japanese is optional and thus not driven by any formal feature whatever theory is adopted to accommodate optionality under the MP. I will argue that the lack of "domain barrier" effects with scrambling in Japanese straightforwardly follows form our theory of phrase structure if Fukui and Saito are correct in claiming that scrambling is not triggered by any formal feature.

4.2.2 An Account

It was shown in the last subsection that unlike overt wh-movement and topicalization in English, scrambling in Japanese is not triggered by any formal feature. In this subsection, I will argue that our analysis of the "domain barriers" proposed in the previous chapter can correctly predict that unlike feature-driven movement like English overt wh-movement and topicalization, non-feature-driven movement like Japanese scrambling does not exhibit any "domain barrier" effects.

Let us consider the case where the adjunct condition effects are canceled, taking (5a) (repeated here as (41)) as an example:

(41) **sono hana-ni$_i$** [daremo$_j$-ga [e_j $\boldsymbol{t_i}$ mizu-o yara-zu]
that flower-Dat everyone-Nom water-Acc give-without
dekakete itta] (koto)
went-out (fact)
'everyone went out without watering that flower'

Considering how to construct the the adjunct clause e_j *sono hana-ni mizu-o yara-zu* 'without watering that flower', we can construct it through checking the UFFs of the selected items, conforming to the ICP and the EP, as represented below:

(42) [e_j [[sono hana]-ni mizu-o yara-zu]]
that flower-Dat water-Acc give-without

Turning to consider how to construct the main structure of (41), i.e., *daremo-ga dekakete itta* 'everyone went out', we first select *dekakete itta* 'went out'. Its selectional restriction feature, being uninterpretable, is checked immediately by combining *dekakete itta* 'went out' with *daremo-ga* 'everyone-Nom' in conformity with the ICP:

(43) [daremo-ga dekakete itta]
everyone-Nom went-out

At this stage of derivation, we have two possible continuations. We either apply scrambling to *sono hana-ni* 'that flower-Dat' or combine the main structure *daremo-ga dekakete itta* 'everyone went out' with the adjunct clause *sono hana-ni mizu-o yara-zu* 'without watering that flower'. If scrambling were triggered by a strong feature (say, a [SCRAMBLING] feature), the ICP would require that the next step should be to apply scrambling to *sono hana-ni* 'that flower-Dat' to check the strong [SCRAMBLING] feature, since the latter is uninterpretable. Since *sono hana-ni* 'that flower-Dat' would be within the adjunct clause which would not have been merged with the main structure, the strong [SCRAMBLING] feature could not be checked at this point. This would violate the ICP and thus the derivation would be canceled.

As argued above, however, scrambling is not triggered by any formal feature. Hence, merger of the main structure with the adjunct clause before the application of scrambling does not violate either the ICP or the EP, since neither of these operations involves any UFF. In other words, like merger of adjuncts, scrambling is required to apply postcyclically.[8] We are allowed to first combine the main structure with the adjunct structure and then apply scrambling to *sono hana-ni* 'that flower-Dat'. We can therefore scramble *sono hana-ni* 'that flower-Dat' out of the adjunct clause *sono hana-ni mizu-o yara-zu* 'without watering that flower' without violating the ICP or the EP. The lack of the adjunct condition effect with scrambling in (41) straightforwardly follows. The other examples, where the adjunct condition effects are abrogated, can be accounted for in the same way.

To recapitulate the above discussion, I have argued that the asymmetry between feature-driven and non-feature-driven movement with respect to the "domain barrier" effects straightforwardly follows from the ICP together with the EP. It should be noted that this asymmetry cannot be accounted for by the previous locality theories. Those theories claim that no element can ever be extracted out of a certain domain regardless of whether the movement operation is feature-driven or not. The asymmetry between feature-driven and non-feature-driven movement cannot be accommodated within the previous locality theories without

---

[8]It might be possible to claim that the postcyclic property of scrambling captures the insight given by, among others, Chomsky (1991) that scrambling takes place in the stylistic component. It has been claimed that stylistic rules are those which apply after cyclic rules. It should be noted that the "stylistic" characteristic of scrambling follows from our analysis without assuming any extra component like the stylistic component in grammar. This is because the ICP coupled with the EP requires that scrambling, which is required to apply postcyclically, should apply after cyclic rules.

recourse to any extra devices. Hence, this asymmetry provides strong empirical evidence in support of our analysis of the "domain barriers" and against the previous locality theories.

4.3 The Apparent "Domain Barrier" Effects with Scrambling

I have shown in the last subsection that unlike English overt wh-movement and topicalization, Japanese scrambling is not subject to the "domain barriers," arguing that its immunity from the "domain barriers" immediately follows from our analysis. As presented in section 4.1, however, there are cases where scrambling prima facie exhibits the "domain barrier" effects. I will argue that the apparent "domain barrier" effects with scrambling should be attributed to an A-over-A condition at PF, which refers to [-V] and the notion of closedness. It is also shown that when scrambling takes place from within adjuncts with an empty subject, there is no violation of the A-over-A condition. We can therefore account for the lack of any locality effects with scrambling out of adjuncts with an empty subject.

4.3.1 An A-over-A Condition

This subsection discusses the apparent "domain barrier" effects with scrambling. I will argue that they should not be attributed to the "domain barriers" but to an A-over-A condition which applies at PF.[9]

---

[9]The discussion to follow is based on the judgment that there is a substantial contrast in acceptability between scrambling out of "domain barriers" like (3-4) and scrambling out of complement clauses like (i):

(i)	**Mary-ni**	John-ga	[Bill-ga	***t*** urami-o	motteiru	to]	kanzita
	Mary-Dat	-Nom	-Nom	grudge-Acc	have	Comp	felt
	(koto)						
	(fact)						

The A-over-A condition, which was originally proposed by Chomsky (1964) and further developed by, among others, Bresnan (1976) and Hasegawa (1974), was intended to capture the locality restrictions on overt movement. Although the formulation of the A-over-A condition varies among its advocates, its essential insight is that elements may not be extracted out of those with the same property. In order to account for the apparent "domain barrier" effects with scrambling, I propose the following A-over-A condition:

(44) The A-over-A Condition

A PF representation is ruled out as illegitimate if it contains a structure of the following form:

α ... [$_\beta$ ... t_α ...] ...,

where α and β are both [-V] categories.

(44) states that no [-V] category may be extracted out from another [-V] category. I argue that the A-over-A condition (44), which is representational in nature, applies at PF. A violation of (44) leads a derivation to crash at PF.[10, 11]

'John felt that Bill have a grudge against Mary

For some speakers including myself, however, it is not entirely clear whether there is a substantial contrast between the two. Compared with short-distance scrambling, long-distance scrambling like (i) is also awkward or at least hard to parse. If the judgment of the latter type is on the right track, we can say that scrambling is immune from any locality conditions. Note that scrambling out of "domain barriers" count as long-distance scrambling in the sense that it crosses a clausal boundary. It follows that scrambling out of "domain barriers," though they are awkward, are in fact grammatical just like long-distance scrambling.

[10]Strictly speaking, the A-over-A condition (44) cannot apply at the PF interface. This is because at the PF interface which only consists of phonetic symbols, there is no relevant structure, not even words. Hence, to be precise, this condition applies at the intermediate level between the point of Spell-Out and the PF interface. Since Chomsky (1995) assumes that the morphological component, which clearly refers to structures, resides on the PF side, it is reasonable to claim that there still remain relevant structures on the PF side. It might be possible to identify this intermediate level as "shallow structure," which dates back to works in the early generative grammar like

Before turning to look at how the A-over-A condition (44) works, let us explicate a categorial feature system which the discussion to follow assumes. Let us assume as in the standard literature that the universal lexicon is divided into two distinct subsets; the set of lexical categories which includes N, V, A, and P and the set of functional categories which includes C, T, and D.[12] Let us assume following Abney (1987) and Fukui (1995) that we state this lexical/functional distinction by postulating a universal feature [+/- F]:[13]

(45) a. Functional Categories: [+F]

b. Lexical Categories: [-F]

It is widely accepted (see, among others, Chomsky (1972, 1981)) that lexical categories are further crossclassified in terms of two primitive features [+/- N] and [+/- V]. The feature specifications of the lexical categories are given below:

(46) Feature Specifications of the Lexical Categories

a. N = [-F, +N, -V]

b. V = [-F, -N, +V]

c. A = [-F, +N, +V]

d. P = [-F, -N, -V]

Turning to functional categories, Abney (1987) introduces the notions of c-projection and s-projection. The c-projection of a category is

---

Lakoff (1971) and Postal (1966). Note in passing that such a constraint may not be formulated as applying at LF, since, as mentioned above, scrambled phrases may be "radically reconstructed" to their original position at LF.

[11] It might be possible to claim that the A-over-A condition (44) does not count as a grammatical constraint but rather as a parsing constraint. Note that the arguments to follow hold regardless of whether this condition is grammatical in nature or not.

[12] Chomsky (1993) assumes another functional category AGR, whose existence is denied in his subsequent works like Chomsky (1995, 1996).

[13] See Jackendoff (1977) and Reuland (1985) for different feature systems of categories.

its syntactic projection in the usual sense. For example, the maximal c-projection of V is VP (= V^{max}). That of T is TP (= T^{max}). The s-projection of a category, which is defined in (47), is the path of nodes along which its descriptive content is "passed along":

(47) β is an s-projection of α iff

a. β = α, or

b. β is a c-projection of an s-projection of α, or

c. β f-selects an s-projection of α

(Abney 1987:57)

F-selection is the syntactic relation between a functional head and its complement. For example, C f-selects a projection of T. D f-selects a projection of N. According to the definition of the notion of s-projection, the maximal s-projection of V is C^{max} via T^{max}. The maximal s-projection of T is also C^{max}. The maximal s-projection of N is D^{max}. Abney argues that this captures the intuition that the verb is the head of a clause while the noun is the head of a nominal without supposing lieterally Clause = V^{max} or Nominal = N^{max}.[14]

Essentially following Abney (1987), I claim that the notion of s-projection can be captured in terms of a feature system where the functional categories are divided based on [+/- N] and [+/- V]. The feature specifications of the functional categories are given below:[15, 16]

---

[14]See, among others, Emonds (1985), Jackendoff (1977), Koster (1978, 1987), and Marantz (1980) for the view that S is the maximal projection of V.

[15]Fukui (1995) proposes a different feature system of functional categories, though his system is also based on [+/- N] and [+/- V].

[16]There are gaps in this feature system of functional categories. There is no functional categoty whose feature specification is [+F, +N, +V] or [+F, -N, -V]. It is possible to claim that AGR, if it really exists, is the category with [+F, +N, +V]. It is clear that AGR is closely related to a verb. AGR has also been assumed to be "nominal" in its nature (see, among others, Chomsky (1981)). Conjunctions like *and* and *or* possibly

(48) Feature Specifications of the Functional Categories

a. D = [+F, +N, -V]

b. C, T = [+F, -N, +V]

These feature specifications of the functional categories capture the fact that while C and T belong to the V system, D belongs to the N system.

In these feature specifications, only N, P, and D have [- V] as its feature. It then follows from the A-over-A condition (44) together with these feature specifications that no category with N, P, or D as its head may be extracted from another category with N, P, or D as its head. As far as movement of a maximal projection is concerend, the A-over-A condition claims that no N^{max}, P^{max}, or D^{max} may be extracted out from another N^{max}, P^{max}, or D^{max}.

Let us consider how the A-over-A condition (44) works. Let us first look at the apparent CNPC effects with scrambling, taking the relative clause case of the CNPC (3a) (repeated here as (49)) as an example:

(49) ?**ano hon-o$_i$** [John-ga [[e_j $\boldsymbol{t_i}$ katta] hito$_j$]-o
that book-Acc -Nom bought person-Acc
sagasite iru rasii]
looking-for seem
'it seems that John is looking for the person who bought that book'

In (49), *ano hon-o* 'that book-Acc' is extracted from the containing phrase *[[e_j ano hon-o katta] hito]-o* 'the person who bought that book'. These

---

count as categories with [+F, -N, -V], since they have no close relation with the N system or the V system. If this conjecture is correct, then we would predict that extraction of a projection of N, P, or D out of the coordinated structure is ruled out by the A-over-A condition (44), given that the coordinate structure is headed by a conjunct (see Zoerner (1995)). The coordinate structure constraint therefore follows from the A-over-A condition (44).

phrases are either N^{max} or D^{max} depending on whether D exists in Japanese or not (see Fukui (1986, 1995) for detailed discussion of this subject). Both of these phrases are [- V] whether they are N^{max} or D^{max}. Hence, extraction of *ano hon-o* 'that book-Acc' out of *[e_j ano hon-o katta] hito]-o* 'the person who bought that book' results in a representation which is ruled out as illegitimate by the A-over-A condition (44). Hence, the apparent CNPC effects with scrambling follows. Note that P^{max} scrambling out of a complex NP can also be ruled out by the A-over-A condition (44):

(50) ?**Bill-ni$_i$** [John-ga [[e_j $\boldsymbol{t_i}$ atta] hito$_j$]-o sagasite iru
-Dat -Nom met person-Acc looking-for
rasii]
seem
'it seems that John is looking for the person who met Bill'

This is because the scrambled P^{max} *Bill-ni* 'Bill-Dat' and the complex NP are both [-V].

Let us turn to the apparent adjunct condition effects with scrambling, taking (4a) (repeated here as (51)) as an example:

(51) ?**sono hon-o** [John-ga [Mary-ga ***t*** yomioete kara]
that book-Acc -Nom -Nom finish-reading after
dekaketa] (koto)
went-out (fact)
'John went out after Mary finished reading that book'

Given that adjuncts like the one in (51) count as P^{max}, *sono hon-o* 'that book-Acc' is extracted out of the P^{max} *Mary-ga sono hon-o yomioete kara* 'after Mary finished reading that book' in (51). Since both of these phrases are [-V], this extraction leads to a representation which violates the A-over-A condition (44). Hence, the apparent adjunct condition effects with scrambling follow. Note that P^{max} scrambling out of an adjunct like (4c) (repeated here as (52)) can also be ruled out as illegitimate by the A-over-A condition (44):

(52) ?**Tookyoo-ni$_i$** [Mary-ga [John-ga t_i ikitagatte iru
Tokyo-Da -Nom -Nom want-to-go
noni] musisite iru rasii]
although ignoring seem
'it seems that although John wants to go to Tokyo, Mary is ignoring that fact'

This is because both the scrambled P^{max} *Tookyoo-ni* 'Tokyo-Dat' and the adjunct P^{max} are both [-V].

The above discussion has shown that the apparent "domain barrier" effects with scrambling in Japanese can be accounted for by the A-over-A condition (44) which applies in the PF component. This analysis gives us a way of explaining the contrast in acceptability with the (apparent) "domain barrier" effects between feature-driven movement, on the one hand, and scrambling, on the other. We have observed that the apparent "domain barrier" effects with scrambling are much weaker than the "domain barrier" effects with feature-driven movement like overt wh-movement and topicalization in English. Recall that under our analysis of the "domain barriers," a derivation violating a "domain barrier" is canceled before the point of Spell-Out due to a violation of the ICP. The derivation therefore never reaches either of the interface levels, specifically LF. Scrambling in Japanese, on the other hand, is not triggered by any formal feature. Scrambling never induces a violation of the ICP and thus never causes a derivation to be canceled before the point of Spell-Out. Even when scrambling exhibits the apparent "domain barrier" effects and thus violates the A-over-A condition (44) within the PF-component, the derivation reaches LF without being canceled and converges at that level. There is no violation of any constraint whatsoever in the course of the computation from N to LF. As argued by Chomsky (1995), computational procedures from N to LF are uniform. The mapping from Spell-Out to PF, on the other hand, has different properties, modifying structures by processes which are different from those permitted in the N -> LF computation. It follows that the N -> LF computation, which is uniform, counts as a core computation in language while the mapping from Spell-Out to PF, which has special properties, counts as peripheral. It is then reasonable to claim that derivations which violate a constraint in the core computation like those violating a "domain barrier" result in severely deviant. On the other hand, derivations which only violate a constraint in the mapping from Spell-Out to PF like those violating the A-over-A condition (44) result in mildly deviant, since they converge as far as the core computation is concerned. Hence, we can correctly predict that the apparent "domain barrier" effects with scrambling are much weaker than the "domain barrier" effects with feature-driven movement like overt wh-movement and topicalization in English.[17]

---

[17]Recall that the derivations which violate the Subject Condition may reach both of the interface levels without being canceled. The subject condition effects, however, are severe. Recall that the Subject Condition is accounted for by the ban against vacuous quantification and the condition of inclusiveness. If the ban against vacuous quantification is relevant to the notion of convergence, its violation leads a derivation to crash at LF and thus in the core computation. Hence, we can correctly predict that the subject condition effects result in severely deviant. As argued in chapter 2, however, a

Apart from the apparent "domain barrier" effects with scrambling, the A-over-A condition (44) receives further empirical support from facts on scrambling out of nominals. If we scramble a phrase out of a nominal, the result is as mildly deviant as the apparent "domain barrier" effects, as shown below:

(53) a. John-ga [[Bill-ga *e*$_j$ kakusi motte ita]
-Nom -Nom had-been-hiding
[Mary-kara-no tegami]$_j$]-o mituketa (koto)
-from-Gen letter-Acc found (fact)
'John found a letter from Mary which Bill had been hiding'

b. ?**Mary-kara-no$_i$** [John-ga [[Bill-ga *e*$_j$ kakusi motte ita]
[***t***$_i$ tegami]$_j$]]-o mituketa (koto)

In (53b), the P^{max} *Mary-kara-no* 'Mary-Dat-Gen' is scrambled out of the nominal phrase *[[Bill-ga e$_j$ kakusi motte ita] [Mary-kara-no tegami]]$_j$-o* 'a letter from Mary which Bill had been hiding'. This violates the A-over-A condition (44), since the scrambled P^{max} and the extraction domain are both [-V]. Hence, we can correctly predict that scrambling out of a nominal like (53b) is as mildly deviant as the apparent "domain barrier" effects due a violation of the A-over-A condition (44).

---

violation of the ban against vacuous quantification might not make a derivation crash but only make an interpretation anomalous. Its derivation, which converges at both PF and LF, converges in the core computation. Under this view, the severely deviant status of the subject condition effects can be accounted for by the fact that its derivation violates not only the ban against vacuous quantification but also the condition of inclusiveness.

4.3.2 Scrambling out of Adjuncts with Empty Subjects

The last subsection has proposed the A-over-A condition (44), arguing that it accounts for the apparent "domain barrier" effects with scrambling. As mentioned above, however, unlike examples like (51) and (52) where extraction takes place from within adjuncts with an overt subject, scrambling out of adjuncts with an empty subject like (41) (repeated here as (54)) is acceptable:

(54) **sono hana-ni$_i$** [daremo$_j$-ga [e_j $\boldsymbol{t_i}$ mizu-o yara-zu]
that flower-Dat everyone-Nom water-Acc give-without
dekakete itta] (koto)
went-out (fact)
'everyone went out without watering that flower'

A question now arises as to why such cases do not violate the A-over-A condition (44). I augue that the A-over-A conditon should be revised to refer not only to [-V] but also to the notion of closeness:

(55) The A-over-A Condition (Revised)
A PF representation is ruled out as illegitimate if it contains a structure of the following form:
α ... [$_\beta$ ... t_α ...] ...,
where α and β are both [-V] and closed categories.

Closed categories are those which count as "saturated" in the sense of Higginbotham (1985) and Speas (1984) and thus function as arguments of a predicate. C^{max} and D^{max} are typical closed categories. Open categories, on the other hand, are those which count as "unsaturated" and thus function as predicates. V^{max} and A^{max} are typical open categories.[18] Following Clark (1985) and Hasegawa (1984/1985), I argue that adjuncts with an empty subject as in (54) involve empty operator movement and thus count as open rather than as closed. Hence, extraction of the closed category like *sono hana-ni* 'that flower-Dat' in (54) out of the adjunct with an empty subject, which is an open category, does not lead to a violation of the A-over-A condition (44). Before turning to an analysis of examples like (54), let us consider the structure of adjuncts with an empty subject.

Kuroda (1965) observes that the empty subject in an adjunct can only refer to the matrix subject but not to someone in the discourse. The relevant example is shown below:

(56) John$_i$-ga [$e_{i/*j}$ hana-ni mizu-o yatte kara]
-Nom flower-Dat water-Acc give after
dekakete itta (koto)
went-out (fact)
'John went out after having watered the flower'

In (56), the empty subject of the adjunct clause can only refer to the matrix subject *John* but not to anyone else. In other words, the empty subject of the adjunct is obligatorily controlled by the matrix subject. Based on this observation, Hasegawa (1984/1985) argues that the empty subject in an adjunct is not an empty prononimal. If it were an empty prononimal, it could refer not only to the matrix subject but also to

[18]It is generally true that categories which are open and thus "unsaturated" in the sense of Higginbotham (1985) and Speas (1984) do not function as "barriers." For instance, in Chomsky's (1986a) system, IP and VP, being "unsaturated," are always devoid of barrierhood due to the exceptional clause in the definition of "barrier" and an adjunction operation, respectively. Then, we might not need to revise the A-over-A condition (44), making the latter only refer to [-V].

someone else in the discourse. She rather argues that the adjunct with an empty subject involves either PRO or empty operator movement, as represented below:

(57) a. **John$_i$**-ga [***PRO$_i$*** hana-ni mizu-o yatte kara] dekakete itta

b. **John$_i$**-ga [***Op$_i$*** [***t$_i$*** hana-ni mizu-o yatte kara]] dekakete itta

What is important to the present discussion is that the adjunct with an empty subject can be analyzed as involving empty operator movement, as in (57b). In (57b), the empty operator moves from the subject position of the adjunct to the peripheral position of the adjunct, where it is associated with the matrix subject *John* through predication.[19]

A similar analysis can be found in Clark (1985). Clark proposes a theory of control, arguing that the control structure always involves empty operator movement. For example, (58), which is a well-known case of obligatory control, is analyzed as in (59):

(58) John tried to enter the race

(Clark 1985:267)

(59) **John$_i$** tried [***Op$_i$*** [***t$_i$*** to enter the race]]

(Clark 1985:268)

In (59), the empty operator moves from its original position to the Spec of the embedded C^{max}, where it is associated with the matrix subject through predication.

Clark also analyzes obligatory control in an adjunct like (60):

---

[19]It is not clear at this point whether adjuncts as in (57) count as C^{max} or P^{max}. If they count as P^{max}, empty operators adjoin to P^{max}. If they count as C^{max}, empty operators move to the Spec of C^{max}.

(60) John$_i$ kissed Mary [after e_i seeing himself$_i$ in the mirror]

(Clark 1985:271)

In (60), the empty subject of the adjunct refers to the matrix subject. In such an environment, the empty subject of an adjunct cannot refer to anything but the matrix subject. In other words, the empty subject is obligatorily controlled by the matrix subject. This is confirmed by the fact that the following examples are deviant:

(61) a. *John kissed Mary$_i$ [after e_i seeing herself$_i$ in the mirror]

(Clark 1985:271)

b. *John kissed Mary [after e_{arb} seeing oneself in the mirror]

In (61a), the empty subject of the adjunct is intended to refer to the matrix object *Mary*. In (61b), the empty subject of the adjunct is intended to refer to an arbitrary person. They are, however, deviant. She argues that obligatory control in adjuncts like (60) also involves empty operator movement, as represented in (62):

(62) **John$_i$** kissed Mary [$\boldsymbol{Op_i}$ [after $\boldsymbol{t_i}$ seeing himself$_i$ in the mirror]]

In (62), the empty operator moves from its original position to the Spec of the adjunct, where it is associated with the matrix subject through predication. Although Clark only deals with English examples, she would say that the control structure in Japanese can be analyzed essentially in the same way. Hence, under her analysis, obligatory control in adjuncts like (56) would be analyzed as involving empty operator movement as in (57b).

Hasegawa (1984/1985) and Clark (1985) agree that the adjunct with an empty subject involves empty operator movement, though they differ as to whether the empty subject can also be identified as PRO. Since the empty operator is interpreted through predication, it is reasonable to claim that the adjunct as a whole, being a predicate, counts as "unsaturated." Hence, if their analysis is correct, the adjunct with an empty subject may count as an open category.

With the above discussion in mind, let us consider scrambling out of adjuncts with an empty subject again, taking (54) (repeated here as (63)) as an example:

(63) **sono hana-ni$_i$** [daremo$_j$-ga [e_j $\boldsymbol{t_i}$ mizu-o yara-zu]
that flower-Dat everyone-Nom water-Acc give-without
dekakete itta] (koto)
went-out (fact)
'everyone went out without watering that flower'

The scrambled phrase *sono hana-ni* 'that flower-Dat' is a closed category. The adjunct, on the other hand, is an open category. This is because it contains an empty subject which is obligatorily controlled by the matrix subject and thus involves empty operator movement. Extraction of the closed category out of the open adjunct does not result in a violation of the A-over-A condition (44), though both of these categories are [-V]. Hence, we can correctly predict that scrambling out of adjuncts with an empty subject like (63) is acceptable.

Note also that unlike scrambling out of adjuncts with an empty subject, scrambling out of adjuncts with an empty object is deviant, as shown below:

(64) a. John-ga [Bill-ga *e* kooen-de nagutta node]

-Nom -Nom park-in hit because

keisatu-ni todokedeta (koto)

police-Dat report (fact)

'John reported to the police because Bill hit him in the park'

b. ?**kooen-de** John-ga [Bill-ga *e* ***t*** nagutta node]

keisatu-ni todokedeta (koto)

In (64b), the P^{max} *kooen-de* 'park-in' is scrambled out of the adjunct with an empty object. The result is mildly deviant. I argue that this also follows from the A-over-A condition (44). Before turning to this issue, let us consider the interpretation of the empty object within an adjunct.

Kuroda (1965) and Hasegawa (1984/1985) observe that exactly like the empty subject in an adjunct, the empty object in an adjunct can only refer to the matrix subject. Let us consider (65) as an example:

(65) John-ga [Mary-ga *e* nagutta node] keisatu-ni

-Nom -Nom hi because police-Dat

uttaeta (koto)

complained (fact)

'John complained to the police because Mary hit him'

They observe that the empty object in the adjunct can only refer to the matrix subject *John* but not to anyone else. Based on this observation, Hasegawa argues that like the adjunct with an empty subject, the one with an empty object may also involve empty operator movement.

Contrary to their observation, however, there is a subject/object asymmetry concerning the interpretation of empty categories within adjuncts, as observed by, among others, Hoji (1985). Within an adjunct, the subject empty category can only refer to the matrix subject while the object empty category can refer to either the matrix subject or someone else in the discourse. In (65), the empty object may either refer to the matrix subject *John* or someone else. I therefore claim following, among others, Hoji (1985) that the object empty category in an adjunct is identified as an empty prononimal, as represented below:

(66) John$_i$-ga [Mary-ga ***pro***$_{i/j}$ nagutta node] keisatu-ni
-Nom -Nom hit because police-Dat
uttaeta (koto)
complained (fact)
'John$_i$ complained to the police before Mary hit him$_{i/j}$'

Returning to (64b), since the adjunct with an empty object does not involve any empty operator movement, it counts as a closed category just like the adjunct without any empty categories. Both the adjunct and the scrambled phrase are closed as well as [-V]. Scrambling of the P^{max} *kooen-de* 'park-in' out of the adjunct therefore results in a violation of the A-over-A condition (44). Hence, we can correctly predict that scrambling out of adjuncts with an empty object like (64b) is mildly deviant.

4.3.3 Against the Base-generation Analysis of Scrambling

Based on the insight given by Hale (1980, 1983), Boskovic and Takahashi (1995) argue that scrambling is not an instance of movement. They rather claim that "scrambled" phrases originate in their surface positions and undergo LF-movement to the positions where they receive their θ-roles.

Let us consider (67) as an example:

(67) **sono hon-o** [John-ga [Mary-ga yonda to] itta (koto)
that book-Acc -Nom -Nom read Comp said (fact)
'John said that Mary read that book'

Under their analysis, (67) would be derived as follows:

(68) a. **sono hon-o** [John-ga [Mary-ga yonda to] itta
that book-Acc -Nom -Nom read Comp said
b. [John-ga [Mary-ga **sono hon-o** yonda to] itta
-Nom -Nom that book-Acc read Comp said

As shown in (68a), the "scrambled" phrase *sono hon-o* 'that book-Acc' is base-generated in its surface position. At LF, *sono hon-o* 'that book-Acc' lowers to the object position of the verb *yonda* 'read' where the former receives its θ-role from the latter. They follow Lasnik and Saito (1992) in claming that movement does not have to leave a trace. This LF-lowering does not leave a trace and thus satisfies the proper binding condition. Note that they are assuming that θ-roles are formal features which count as illegitimate at LF. Thematic features therefore must be eliminated until a derivation reaches LF. Otherwise, the derivation crashes. Hence, θ-roles function as triggers for the LF-movement of "scrambled" phrases.

Boskovic and Takahashi present conceptual and empirical arguments in favor of the base-generation analysis of scrambling. Although we do not go into the details of their arguments, the consistency of their base-generation analysis with the last resort principle is worth mentioning. Under their analysis, "scrambling" counts as a feature-driven movement, since a "scrambled" phrase originates in its surface position and then undergoes LF-movement to check a θ-feature. Unlike the analyses proposed by Fukui and Saito where optional movement is allowed, their analysis is consistent with the last resort principle without recourse to any extra device.

There are, however, conceptual and empirical problems which make us reluctant to adopt the base-generation analysis of scrambling. Conceptually, their base-generation analysis raises a problem of globality as it stands. Recall that they claim that thematic features count as illegitimate objects at LF and thus must undergo elimination before LF for convergence. "Scrambled" phrases undergo LF-movement for eliminating thematic features, which makes the derivation converge. As extensively argued in chapter 2, however, such an interface condition on formal features necessarily needs global considerations. In the present case, when we have an option of applying LF-movement to a "scrambled" phrase, we must look ahead to see whether the application/nonapplication of the LF-movement would yield an LF interface only with legitimate objects. In other words, we cannot decide whether to apply the LF-movement or not only on the basis of information available at that stage. Such global considerations should be avoided, since they necessarily induce much computational complexity.

Turning to its empirical problems, recall that scrambling exhibits some locality effects although they are weaker than those with feature-driven movement like overt wh-movement and empty operator movement. I have argued that they should be attributed to the A-over-A condition at PF. Under the base-generation analysis of scrambling, however, "scrambled" phrases are expected to appear in any base-generated positions. The base-generation analysis would therefore predict that there do not exist any locality effects with scrambling whatsoever. Hence, the base-generation analysis of scrambling is empirically problematic, unless it gives an account of the locality effects with scrambling.

4.4 Extraction out of Scrambled Phrases

It was shown in chapter 3 that when phrases undergo feature-driven movement like overt wh-movement and topicalization, they constitute "barriers" for further feature-driven movement, presenting the following examples:

(69) Topicalization out of Topic
*?**vowel harmony**$_i$, I think that [[articles about $\boldsymbol{t_i}$]$_j$ [you read t_j]]

(70) Wh-movement out of moved *wh*-phrases
*?**who**$_i$ do you wonder [[which pictures of $\boldsymbol{t_i}$]$_j$ [Mary bought t_j]]

(71) Wh-movement out of Topic
*?**who**$_i$ do you think that [pictures of $\boldsymbol{t_i}$]$_j$ John wanted t_j

I have argued that like the "domain barriers," this locality condition also follows from our theory of phrase structure.

Let us quickly go over our analysis of this locality condition, taking (69) as an example. During the derivation, we come to the stage where the strong [TOPIC] feature of the embedded F is to be checked. Since it is uninterpretable, the ICP requires that it should be checked immediately by copying *articles about vowel harmony*, as in (72):

(72) a. [$_F$max F [$_T$max T [$_V$max you read **[$_D$max articles about vowel harmony]**]]]
b. you
c. **[$_D$max articles about vowel harmony]**

The copied D^{max} *articles about vowel harmony* may not be merged with the main structure at this point of the derivation due to the ICP and the EP.

As the derivation proceeds, we come to the stage where the strong [TOPIC] feature of the matrix F is required to be checked immediately by the ICP:

(73) a. [$_{F}$^{max} F$_{[TOPIC]}$ [$_{T}$^{max} T [$_{V}$^{max} I [think [$_{C}$^{max} that [$_{F}$^{max} F [$_{T}$^{max} T [$_{V}$^{max} you read [$_{D}$^{max} articles about vowel harmony]]]]]]]]]

b. you

c. [$_{D}$^{max} articles about vowel harmony]

d. I

It should be noted that *vowel harmony* within the copied D^{max} (73c), which is eventually merged into the Spec of the embedded F, may not be subject to the copy operation triggered by the strong [TOPIC] feature, since the latter does not c-command the former. Hence, we can derive the fact that topicalization out of a topic is illegitimate. Although the strong [TOPIC] feature of the matrix F can attract *vowel harmony* within the D^{max} which stays in its original position, it would violate the ban against vacuous quantification and the condition of inclusiveness.

As observed by Saito (1985), however, scrambling out of a scrambled phrase is acceptable:

(74) [**sono mura-ni**$_i$ [John-ga [[Mary-ga $\boldsymbol{t_i}$ sunde iru to]$_j$

that village-Dat -Nom -Nom live Comp

[Bill-ga t_j itta to]] omotte iru]] (koto)

-Nom said Comp think (fact)

'John thinks that Bill said that Mary lives in that village'

In (74), *Mary-ga sono mura-ni sunde iru to* 'that Mary lives in that village', which is the complement clause of the verb *itta* 'said', is scrambled to the initial position of the embedded clause. The P^{max} *sono mura-ni* 'in that village' is further scrambled out of the scrambled phrase to the initial position of the matrix clause. The result is acceptable.

Furthermore, feature-driven movement like empty operator movement may take place from within a scrambled phrase, as shown below:

(75) The Cleft Construction

[$\boldsymbol{Op_i}$ [[Mary-ga $\boldsymbol{t_i}$ itta to]$_j$ [John-ga t_j omotte iru]]

-Nom went Comp -Nom think

no]-wa Tookyoo$_i$-ni da

Comp-Top Tokyo-Dat be

'it is to Tokyo that John thinks that Mary went'

(76) The *Tough* Construction

[[sooiu taipu-no dansei-kara]$_i$-ga [(Mary$_j$-nitotte)

such a type of man-from-Nom -for

[$\boldsymbol{Op_i}$ [e_j [[e_j $\boldsymbol{t_i}$ purezento-o moratta to]$_k$ [otto-ga

present-Acc was given Comp husband-Nom

t_k gokaisite iru to]] tomodati-ni hanasi]] nikui]] (koto)

misunderstand Comp friend-Dat sa hard (fact)

Lit. '[from such a type of man]$_i$ is hard (for Mary$_j$) to say to her friend that her husband is under the false impression that she was given a present e_i'

(77) The Comparative Deletion Construction

[$\boldsymbol{Op_i}$ [[[John-ga $\boldsymbol{t_i}$ yonda to] iwareteiru to]$_j$

-Nom read Comp is said Comp

Tom-ga t_j uwasa-siteiru] yorimo] Mary-wa takusan

-Nom rumor make than -Top many

hon-o yondeita

book-Acc read

'Mary read more books than Tom made the rumor that it is said that John read'

(Kikuchi 1987:12)

In (75-77), empty operator movement takes place from inside the scrambled phrases. The results are acceptable. I argue that this asymmetry between feature-driven movement, on the one hand, and scrambling, on the other, concerning extraction domains also follows from the ICP coupled with the EP given that scrambling is not feature-driven.

4.4.1 Scrambling out of Scrambled Phrases

Before considering how to account for (74-77), let us look at Saito's (1992) theory of the status of the landing site of scrambling. Modifying Mahajan (1990), Tada (1990), and Webelhuth (1989), Saito argues that scrambling is movement to a non-operator, non-A-position. Since a non-operator, non-A-position is not licensed at LF, however, it is reanalyzed in either of the following ways:

(78) a. The position disappears at LF.

b. The position is reanalyzed as an operator position.

c. The position is reanalyzed as an A-position.

(Saito 1992:99-100)

(78a) is achieved when the scrambled phrase is undone at LF. When the scrambled phrase is not undone, it is either reanalyzed as an operator position or an A-position. It is reanalyzed as an operator position in the case of either clause-internal or long-distance scrambling. It is reanalyzed as an A-position only in the case of clause-internal scrambling but not in the case of long-distance scrambling, given the descriptive generalization that CP (= C^{max}) breaks A-chains (see, among others, Aoun (1981) and Chomsky (1986a)).

With the above discussion in mind, let us first consider scrambling out of a scrambled phrase, taking (74) as an example. Since scrambling is not triggered any formal feature, our theory of phrase structure requires that like merger of adjuncts, scrambling should take place postcyclically. There is therefore no ordering imposed on the application of the two instances of scrambling. We can either scramble the complement clause *Mary-ga sono mura-ni sunde iru to* 'that Mary lives in that village' first and then scramble the P^{max} *sono mura-ni* 'that village-Dat' out of the latter or scramble the P^{max} out of the complement clause first and then scramble the remaining complement clause. Given the copy theory of movement, these two possible derivations are represented below:

(79) a. John-ga [Bill-ga [Mary-ga sono mura-ni
-Nom -Nom -Nom that village-Dat
sunde iru to] itta to] omotte iru
live Comp said Comp think

— Scrambling of the Complement Clause →

b. John-ga [[**[Mary-ga sono mura-ni sunde iru to]$_j$**
[Bill-ga **[Mary-ga sono mura-ni sunde iru to]$_j$** itta]] to] omotte iru

— Scrambling of the P^{max} *sono mura-ni* →

c. **sono mura-ni$_i$** [John-ga [[[Mary-ga **sono mura-ni$_i$** sundeiru to]$_j$ [Bill-ga [Mary-ga sono mura-ni sunde iru to]$_j$ itta]] to] omotte iru]

(80) a. John-ga [Bill-ga [Mary-ga sono mura-ni
-Nom -Nom -Nom that village-Dat
sunde iru to] itta to] omotte iru
live Comp said Comp think

— Scrambling of the P^{max} *sono mura-ni* →

b. **sono mura-ni$_i$** [John-ga [Bill-ga [Mary-ga **sono mura-ni$_i$** sundeiru to] itta to] omotte iru]

— Scrambling of the Complement Clause →

c. sono mura-ni$_i$ [John-ga [[**[Mary-ga sono mura-ni$_i$ sunde iru to]$_j$** [Bill-ga **[Mary-ga sono mura-ni$_i$ sunde iru to]$_j$** itta]] to] omotte iru]

Note that regardless of which derivation is to be adopted, we get the same output except the following point. In (80c), the three occurrences of *sono mura-ni* 'that village-Dat' are all identical in constitution, as we can see from the fact that they are all assigned the same index. Hence, if (80c) reaches LF without any scrambled phrases undone, the following three-membered chain is formed:

(81) CH = (*sono mura-ni*, *sono mura-ni*, *sono mura-ni*)

In (79c), on the other hand, its occurrences in the clause-initial position and within the scrambled complement clause are identical in constitution. Its occurrence in its original position, which is not assigned any index, is not identical in constitution with the other occurrences. Hence, if (79c) reaches LF without any scrambled phrases undone, the following two-membered chain is formed:

(82) CH = (*sono mura-ni*, *sono mura-ni*)

Note that this chain consists of the P^{max} *sono mura-ni* 'that village-Dat' in the clause initial position and the one within the scrambled complement.

It is important to note that as in (79), scrambling of *sono mura-ni* 'that village-Dat' may take place from within the scrambled complement clause. This is in contrast with the fact that movement out of a moved phrase is illegitimate in the case of feature-driven movement like overt wh-movement and topicalization in English. In order to see why there is such a contrast, let us consider how derivation (79) proceeds more precisely under the copy theory of movement. Given that scrambling consists of Copy and Merge, we first copy the complement clause *Mary-ga sono mura-ni sunde iru to* 'that Mary lives in that village':

(83) a. John-ga [Bill-ga **[Mary-ga sono mura-ni sunde iru to]** itta to] omotte iru

b. **[Mary-ga sono mura-ni sunde iru to]**

Here, we have two possible continuations. We either merge the copied complement clause (83b) with the main structure or copy the P^{max} *sono mura-ni* 'that village-Dat' within the complement clause which stays in its original position. Since neither of these operations is triggered any formal feature, we may apply either of them. We therefore apply merger of the complement clause with the main structure at this stage, yielding (79b). We then copy the P^{max} *sono mura-ni* 'that village-Dat' within the scrambled complement clause. We finally apply merger of the copied P^{max} with the main structure, yielding (79c). Hence, scrambling out of the scrambled phrase is possible as in (79). The crucial difference between feature-driven movement and non-feature-driven movement therefore resides in the fact that a copied phrase may be merged immediately with the main structure in the latter, but not in the former.

Let us consider possible continuations of (79c) and (80c) depending on how the scrambled phrases are reanalyzed. According to Saito (1992), since scrambling of *sono mura-ni* 'that village-Dat' is an instance of long-distance scrambling, the scrambled position is either undone at LF or reanalyzed as an operator position. Scrambling of the complement clause *Mary-ga sono mura-ni sunde iru to* 'that Mary lives in that village', on the other hand, is an instance of clause-internal scrambling. The scrambled position can be reanalyzed as any of (78a-c).

Let us first consider the case where both of the scrambled phrases are undone at LF. In such a case, we get a representation where the scrambled phrases both move back to their original positions. There is therefore no violation of any condition. Hence, the derivation converges.

Let us next consider the case where only *sono mura-ni* 'that village-Dat' is undone at LF. In (79c), *sono mura-ni* 'that village-Dat' in the clause-initial position only moves back to its position within the scrambled complement but not to its original position, since its occurrence in the latter position does not count as identical in constitution or form a chain with the other occurrences. In (80c), on the other hand, *sono mura-ni* 'that village-Dat' moves back to its original position through its position within the scrambled complement. From either (79c) or (80c), we get (84):

(84) John-ga [[[Mary-ga sono mura-ni sunde iru to]$_j$ [Bill-ga [Mary-ga sono-mura ni sunde iru to]$_j$ itta]] to] omotte iru

In this derivation, we create the following chain:

(85) CH = (*Mary-ga sono mura-ni sunde iru to*, *Mary-ga sono mura-ni sunde iru to*)

Given that the non-head positions of chains delete within the covert component, *Mary-ga sono mura-ni sunde iru to* 'that Mary lives in that village' in its original position deletes. We get the following representation:

(86) John-ga [[**[Mary-ga sono mura-ni sunde iru to]$_\mathbf{j}$** [Bill-ga $\boldsymbol{t_j}$ itta]] to] omotte iru

Representation (86) is legitimate regardless of whether the chain is interpreted as an A-chain or an operator-variable construction. This derivation converges.

If only the complement clause *Mary-ga sono mura-ni sunde iru to* 'that Mary lives in that village' is undone at LF, on the other hand, we get (87) from either (79c) or (80c):

(87) sono mura-ni$_i$ [John-ga [Bill-ga [Mary-ga sono-mura$_i$ ni sunde iru to] itta to] omotte iru]

In this derivation, we create the following chain:

(88) CH = (*sono mura-ni*, *sono mura-ni*)

The P^{max} *sono mura-ni* 'that village-Dat' in its original position deletes in the covert component. We get the following representation:

(89) **sono mura-ni$_\mathbf{i}$** [John-ga [Bill-ga [Mary-ga $\boldsymbol{t_i}$ ni sunde iru to] itta to] omotte iru]

Representation (89) is legitimate, since the chain can be interpreted as an operator-variable construction. Note that since the C^{max} intervenes between *sono mura-ni* 'that village-Dat' and its trace, the chain may not be interpreted as an A-chain. Hence, this derivation also converges.

Let us finally consider the case where neither of the scrambled phrases is undone at LF. From (79c), we get the following two chains:

(90) a. CH = (*sono mura-ni*, *sono mura-ni*)

b. CH = (*Mary-ga sono mura-ni sunde iru to*, *Mary-ga sono mura-ni sunde iru to*)

Note that chain (90a) consists of the P^{max} *sono mura-ni* 'that village-Dat' in the clause-initial position and the one within the scrambled complement, but not the one in its original position. Given that the non-head positions of chains delete within the covert component, *sono mura-ni* 'that village-Dat' within the scrambled complement clause and the complement clause in its original position delete. We get the following representation:

(91) **sono mura-ni**$_i$ [John-ga [[**[Mary-ga** $\boldsymbol{t_i}$ **sundeiru to]**$_j$ [Bill-ga $\boldsymbol{t_j}$ itta]] to] omotte iru]

In (91), the chain of the P^{max} which undergoes long-distance scrambling may be interpreted as an operator-variable construction. The chain of the complement clause which undergoes clause-internal scrambling may either be interpreted as an A-chain or an operator-variable construction. This derivation converges.

From (80c), on the other hand, we get the following two chains:

(92) a. CH = (*sono mura-ni*, *sono mura-ni*, *sono mura-ni*)

b. CH = (*Mary-ga sono mura-ni sunde iru to*, *Mary-ga sono mura-ni sunde iru to*)

The P^{max} *sono mura-ni* 'that village-Dat' within the scrambled complement clause, the one in its original position, and the complement clause in its original position delete in the covert component. We also get (91). Hence, this derivation also converges.

To summarize the above discussion, no matter how the scrambled phrases may be reanalyzed, there is always a derivation which converges.

Hence, the ICP coupled with the EP correctly predicts that scrambling out of a scrambled phrase like (74) is acceptable.

4.4.2 Empty Operator Movement out of Scrambled Phrases

Let us next consider empty operator movement out of a scrambled phrase, taking (75) (repeated here as (93)) as an example:

(93) The Cleft Construction

[[***Op_i*** [Mary-ga ***t_i*** itta to]_j [John-ga *t_j* omotte iru]]

-Nom went Comp -Nom think

no]-wa Tookyoo_i-ni da

Comp-Top Tokyo-Dat be

'it is to Tokyo that John thinks that Mary went'

We will only consider how to construct the presuppositional part of (93), which is relevant to the present discussion, ignoring the rest.

During its derivation, we construct the following structure:

(94) [John-ga [Mary-ga *Op* itta to] omotte iru]

-Nom -Nom said Comp think

It is important to note that we are not allowed to apply scrambling of the complement clause at this stage of the derivation. This is because scrambling is non-feature-driven and thus required to apply postcyclically by the ICP and the EP. The next step is to select the COMP *no*, as required by the EP. Then, the strong operator feature of C must be checked immediately by copying the operator *Op* within the complement clause, which has not undergone scrambling yet:

(95) a. [[[John-ga [Mary-ga ***Op_i*** itta to] omotte iru] no]

b. ***Op_i***

We then apply merger of the operator (95b) with the main structure and scrambling of the complement clause postcyclically, yielding (96):

(96) [*Op*$_i$ [[Mary-ga *Op*$_i$ itta to]$_j$ [John-ga [Mary-ga *Op*$_i$ itta to]$_j$ omotte iru]] no]

Here we have two possible continuations depending on whether the scrambled phrase is undone. If the scrambled phrase is undone, we get (97):

(97) [*Op*$_i$ [John-ga [Mary-ga *Op*$_i$ itta to] omotte iru] no]

In this derivation, we create the following chain of the operator:

(98) CH = (*Op*, *Op*)

We delete the tail position of the chain, yielding the following representation:

(99) [***Op***$_{\boldsymbol{i}}$ [John-ga [Mary-ga ***t***$_{\boldsymbol{i}}$ itta to] omotte iru] no]

In (99), the chain consisting of the empty operator and its trace is properly interpreted as an operator-variable construction. Hence, this derivation converges.

If the scrambled phrase is not undone, we get (96). The following two chains are created:

(100) a. CH = (*Op*, *Op*)

b. CH = (*Mary-ga Op itta to*, *Mary-ga Op itta to*)

Note that (100a) consists of *Op* in the Spec of C^{max} and *Op* within the scrambled phrase. We delete the non-head positions of the chains, yielding the following representation:

(101) [***Op***$_{\boldsymbol{i}}$ [**[Mary-ga *t*$_{\boldsymbol{i}}$ itta to]$_{\mathbf{j}}$** [John-ga ***t***$_{\boldsymbol{j}}$ omotte iru]] no]

In (101), the chain consisting of the empty operator and its trace within the scrambled complement is properly interpreted as an operator-variable construction. The chain created by scrambling is legitimate regardless of whether it is interpreted as an A-chain or as an operator-variable construction. Hence, the derivation of (93) converges whether the scrambled phrase is undone or not. We can therefore correctly predict that if we extract an element from within a scrambled phrase in terms of feature-driven movement, the result is acceptable.

4.5 Focus Scrambling and the "Domain Barriers"

In section 4.2, I have argued that the lack of the "domain barrier" effects with scrambling straightforwardly follows from our theory of phrase structure. It is not the case, however, that scrambling is always devoid of the "domain barrier" effects. In this section, I will investigate scrambling in what Kuroda (1972, 1979) calls generic sentences. It is first pointed out that scrambling in generic sentences exhibits the "domain barrier" effects. I will then argue that the scrambled phrases in generic sentences are forced to have focus readings. In other words, scrambling in generic sentences counts as an operation triggered by a focus feature. I will argue that the "domain barrier" effects with scrambling in generic sentences straightforwardly follows from our theory of phrase structure, which lends further empirical support for our analysis of the "domain barriers."

4.5.1 Locality on Scrambling in Generic Sentences

In section 4.1, it was shown that scrambling is not subject to the "domain barrier" effects. As evidence in support of this contention, I have shown that scrambling out of the adjunct with an empty subject is acceptable. In what Kuroda (1972, 1979) calls generic sentences, however, scrambling out of adjuncts with an empty subject is deviant, as exemplified below:

(102) a. daremo$_j$-ga [e_j sono hana-ni kizuka-zu]
everyone-Nom that flower-Dat notice-without
itumo toorisugiru (koto)
always go-by (fact)
'everyone goes by without noticing that flower'

b. *?**sono hana-ni$_i$** [daremo$_j$-ga [e_j t_i kizuka-zu] itumo toorisugiru] (koto)

(103) a. dareka$_j$-ga [e_j sono isu-ni suwari nagara]
someone-Nom that chair-Dat sit while
itumo hon-o yomu (koto)
always book-Acc read (fact)
'someone always reads a book while sitting on that chair'

b. *?**sono isu-ni$_i$** [dareka$_j$-ga [e_j t_i suwari nagara] itumo hon-o yomu] (koto)

(104) a. daremo$_j$-ga [e_j John-no ibiki-ni odoroite] itumo
everyone-Nom -Gen snore-Dat surprised always
me-o samasu (koto)
wake-up (fact)
'everyone always wakes up, surprised with John's snore'

b. *?**John-no ibiki-ni$_i$** [daremo$_j$-ga [e_j t_i odoroite] itumo me-o samasu] (koto)

Kuroda (1972, 1979) calls a sentence generic if the sentence makes a statement about a general, habitual, or constant state of affairs. In (102-

104), the matrix clauses count as generic, since they make statements about a habitual state of affairs. As shown in the (b) examples of (102-104), if we extract a phrase from within the adjunct with an empty subject through the application of scrambling at the level of the matrix generic clause, the result is deviant.

It was also shown in the preceding discussion that when a phrase is scrambled from within the adjunct with an overt subject or the complex NP, the result is not as severely deviant as the normal "domain barrier" violation induced by feature-driven movement. I have argued that this should count as evidence to suggest that those cases of scrambling do not exhibit the "domain barrier" effects but the effects induced by the A-over-A condition at PF. In generic sentences, however, scrambling out of the adjunct with an overt subject or the complex NP is as severely deviant as the normal "domain barrier" violation, as exemplified below (cf. Matsuda (1996)):

(105) Scrambling out of the Adjunct with an Overt Subject

a. *?**eki-ni** [Mary-ga [John-ga ***t*** tuite kara]
station-Dat -Nom -Nom arrive after
taitei mukae-ni dekakeru] (koto)
usually leave home to meet him (fact)
'Mary usually leaves home to meet him after John arrives at the station'

b. *?**sono hana-ni** [John-ga [Mary-ga ***t*** mizu-o
that flower-Dat -Nom -Nom water-Acc
yaru node] itumo tigau hana-ni mizu-o
give because always different flower-Dat water-Acc
yaru] (koto)
give (fact)
'because Mary waters that flower, John always waters a different one'

c. *?**oziisan-ni** [Mary-ga [John-ga ***t*** aitagatte iru
grandfather-Dat -Nom -Nom want-to-see
noni] itumo musisuru] (koto)
although always ignore (fact)
'although John wants to see his grandfather, Mary always ignores that fact'

(106) Complex NP Constraint

a. *?**tonari-no ie-ni** [John-ga [[***t*** sundeiru]
next-door-Gen house-Da -Nom reside
hito]-o maiasa mikakeru] (koto)
person-Acc every morning see (fact)
'John sees the person who lives in the next house every morning'

b. *?**sooko-ni** [John-ga [[***t*** hozonsite iru] yasai]-o
storage-Dat -Nom keep vegetable-Acc
mainiti kensa suru] (koto)
everyday inspec (fact)
'everyday John inspects the vegetables which are kept in the storage'

Based on these observations, I argue that unlike scrambling in non-generic sentences, scrambling in generic sentences induces the "domain barrier" effects. I will argue that such a contrast between generic and non-generic sentences regarding the "domain barrier" effects with scrambling straightforwardly follows from the ICP together with the EP. Before we come to that, let us consider focus readings in generic sentences.

4.5.2 The "Domain Barrier" Effects with Focus Scrambling

4.5.2.1 Focus *Ga* in Japanese

As extensively discussed by Kuno (1973) and Kuroda (1972, 1979), Japanese nominative phrases in generic sentences can only be interpreted as having focus readings.[20] Let us consider the following examples:

(107) a. **inu-ga** neko-o oikakeru (koto)
dog-Nom cat-Acc chase (fact)
'dogs chase cats'
(adopted from Kuroda 1979:16)

b. **saru-ga** ningen-no senzo desu
monkey-Nom man-Gen ancestor be
'the monkey is the ancestor of man'
(Kuno 1973: 51)

c. **John-ga** mainiti gakkoo-ni iku (koto)
-Nom everyday school-Dat go (fact)
'John goes to school everyday'

d. **John-ga** mainiti hana-ni mizu-o yaru (koto)
-Nom everyday flower-Dat water-Acc give (fact)
'John waters a flower everyday'

The examples in (107) are all generic in that they make a statement about a general, habitual, or constant state of affairs. The nominative phrases must be interpreted as having focus readings. The more precise translations of the examples in (107) are therefore as follows:

---

[20] Kuno (1973) calls such focus *ga* exhaustive listing *ga*. See Matsuda (1996) for an analysis of focus *ga*, where generic sentences are treated on a par with the cleft construction. See Ogihara (1987) for a semantic analysis of this construction.

(108) a. It is dogs that chase cats; (Of all the animals we are talking about) dogs and only dogs chase cats.

b. It is the monkey that is the ancestor of man; (Of all the animals we are talking about) the monkey and only the monkey is the ancestor of man.

c. It is John who goes to school everyday; (Of all the people we are talking about) John and only John goes to school everyday.

d. It is John who waters a flower everyday; (Of all the people we are talking about) John and only John waters a flower everyday.

As observed by Kuno and Kuroda, examples like (107) are awkward, if not ungrammatical, as independent sentences without any contexts due to the obligatory focus readings of the nominative phrases unless the nominative phrases contain numericals or quantifiers. They become natural in contexts which solicit the focus readings of the nominative phrases. For example, (107a-d) are natural when they are given as answers to the following generic questions:

(109) a. **nani-ga** neko-o oikakeru no
what-Nom cat-Acc chase Q
'what chases cats'

b. **nani-ga** ningen-no senzo desu ka
what-Nom man-Gen ancestor be Q
'what is the ancestor of man'

c. **dare-ga** mainiti gakkoo-ni iku no
who-Nom everyday school-Dat go Q
'who goes to school everyday'

d. **dare-ga** mainiti hana-ni mizu-o yaru no
who-Nom everyday flower-Dat water-Acc give Q
'who waters a flower everyday'

There are several points where generic sentences differ from non-generic sentences. First, as Kuroda (1972, 1979) points out, nominative phrases behave differently from non-nominative phrases in generic questions but not in non-generic questions. As shown in (109), the generic questions which require nominative phrases as answers are natural. The generic questions which requires non-nominative phrases as answers, on the other hand, are unnatural, as shown below:

(110) a. ?inu-ga **nani-o** oikakeru no
dog-Nom what-Acc chase Q
'what do dogs chase'

b. ?saru-ga **nani-no** senzo desu ka
monkey-Nom what-Gen ancestor be Q
Lit: 'the monkey is the ancestor of what'

c. ?John-ga mainiti **doko-ni** iku no
-Nom everyday where go Q
'where does John go everyday'

d. ?John-ga mainiti **nani-ni** mizu-o yaru no
-Nom everyday what-Dat water-Acc give Q
'what does John water everyday'

More natural are questions of the form where the nominative phrases are replaced by topic phrases:

(111) a. **inu-wa** **nani-o** oikakeru no
dog-Top what-Acc chas Q
'what do dogs chase'

b. **saru-wa** **nani-no** senzo desu ka
monkey-Top what-Gen ancestor be Q
Lit: 'the monkey is the ancestor of what'

c. **John-wa** mainiti **doko-ni** iku no
-Top everyday where go Q
'where does John go everyday'

d. **John-wa** mainiti **nani-ni** mizu-o yaru no
-Top everyday what-Dat water-Acc give Q
'what does John water everyday'

This is in contrast with non-generic sentences, which make statements about a specific event. In non-generic sentences, no phrases are forced to have focus readings. Not only the non-generic questions soliciting nominative phrases as answers like the (b) examples of (112-113) but also those soliciting non-nominative phrases as answers like the (c) examples of (112-113) are natural:

(112) a. inu-ga neko-o oikaketeiru (koto)
dog-Nom cat-Acc be-chasing (fact)
'the dog is chasing the cat'

b. **nani-ga** neko-o oikaketeiru no
what-Nom cat-Acc be-chasing Q
'what is chasing the cat'

c. inu-ga **nani-o** oikaketeiru no
dog-Nom what-Acc be-chasing Q
'what is the dog chasing'

(113) a. John-ga kinoo gakkoo-ni itta (koto)
-Nom yesterday school-Dat went (fact)
'John went to school yesterday'

b. **dare-ga** kinoo gakkoo-ni itta no
who-Nom yesterday school-Dat went Q
'who went to school yesterday'

c. John-ga kinoo **doko-ni** itta no
-Nom yesterday where went Q
'where did John go yesterday'

Although it is not entirely clear at this point why generic questions like (110), which solicit non-nominative phrases as answers, are unnatural, let us assume following Kuroda (1972, 1979) that the following descriptive generalization is correct:

(114) The generic questions which solicit non-focalized phrases as answers are unnatural.

In generic questions (110), where the nominative phrases are focalized, the non-nominative phrases, which are not focalized, are questioned. Hence, they are unnatural.

Second, Kuno (1973) observes the following generalization:

(115) In generic sentences, no other elements than focalized phrases can get focus readings through stress assignment.

Let us consider the following examples (the judgments are mine):[21]

[21]Stressed elements are represented by bold capitals here and in the relevant examples to follow.

(116) a. John-ga mainiti gakkoo-ni iku (koto)

-Nom everyday school-Dat go (fact)

'John (and only John) goes to school everyday'

b. ?John-ga mainiti **GAKKOO-NI** iku (koto)

'It is to school that John (and only John) goes everyday'

(adapted from Kuno 1973:64)

In (116b), the non-nominative phrase *gakkoo-ni* 'school-Dat' is intended to be focalized through stress assignment; the result is unnatural. It is widely accepted that only one constituent in a given sentence can receive a focus reading. We have observed that the nominative phrases in generic sentences obligatorily receive focus readings. It then follows that (116b) is unnatural. This is because (116b) has more than one focalized constituents within a sentence, i.e., the nominative phrase *John-ga* 'John-Nom' and the stressed element *gakkoo-ni* 'school-Dat'.

This is in contrast with non-generic sentences:

(117) a. **JOHN-GA** Mary-ni okane-o ageta

-Nom -Dat money-Acc gave

'It was John who gave the money to Mary'

b. John-ga **MARY-NI** okane-o ageta

'It was to Mary that John gave the money'

c. John-ga Mary-ni **OKANE-O** ageta

'It was the money that John gave to Mary'

As shown in (117), in a non-generic sentence, any constituent may have a focus reading through stress assignment as far as no other constituent gets stressed within the sentence.

4.5.2.2 Focus Scrambling in Generic Sentences

Keeping the above discussion in mind, let us consider scrambling in generic sentences:

(118) a. **John-no ie-ni** yuubin haitatu-ga mainiti kuru
John's house-Dat mailman-Nom everyday come
(koto)
(fact)
'A mailman comes to John's house everyday'

b. **kono eki-ni** sihatu densya-ga itumo
this station-Dat first-train-Nom always
teisya suru (koto)
stop (fact)
'the first train always stops at this station'

I claim that if we apply scrambling in generic sentences, the focus readings of nominative phrases are neutralized. In (118a-b), the nominative phrases *yuubin haitatu-ga* 'mailman-Nom' and *sihatu densya-ga* 'express train-Nom' are no longer forced to have focus readings. Instead, the scrambled phrases, *John-no ie-ni* 'John's house-Dat' in (118a) and *kono eki-ni* 'this station-Dat' in (118b), obligatorily have focus readings.[22]

This is supported by the contrast in naturalness between (119-120a) and (119-120b):

[22] See Matsuda (1996) for the same observation.

(119) a. **dare-no ie-ni$_i$** yuubin haitatu-ga mainiti t_i kuru

whose house-Dat mailman-Nom everyday come

no

Q

'whose house does a mailman come to everyday'

b. ?John-no ie-ni$_i$ **dare-ga** mainiti t_i kuru no

John's house-Dat who-Nom everyday come Q

'who comes to John's house everyday'

(120) a. **dono eki-ni$_i$** sihatu densya-ga t_i itumo

which station-Dat first-train-Nom always

teisya suru no

stop Q

'which station does the first train always stop at'

b. ?kono eki-ni$_i$ **dono densya-ga** t_i itumo

this station-Dat which train-Nom always

teisya suru no

stop Q

'which train always stops at this station'

The generic scrambled questions soliciting scrambled phrases as answers like (119-120a) are natural. On the other hand, the generic scrambled questions soliciting non-scrambled nominative phrases as answers like (119-120b) are unnatural.

Although the judgments on (119-120b) are subtle, they are clearly less natural than their corresponding non-generic questions:

(121) a. John-no ie-ni **dare-ga** kinoo *t* kita no
John's house-Dat who-Nom yesterday came Q
'who came to John's house yesterday'

b. kono eki-ni **dono densya-ga** *t* kesa
this station-Dat which train-Nom this morning
teisya sita no
stopped Q
'which train stopped at this station this morning'

Furthermore, (119-120b) become natural if we topicalize the dative phrases *John-no ie-ni* 'John's house-Dat' and *kono eki-ni* 'this station-Dat':

(122) a. **John-no ie-ni-wa** dare-ga mainiti *t* kuru no
John's house-Dat-Top who-Nom everyday come Q
'who comes to John's house everyday'

b. **kono eki-ni-wa** dono densya-ga *t* itumo
this station-Dat-Top which train-Nom always
teisya suru no
stop Q
'which train always stops at this station'

On the assumption that the generic questions soliciting non-focused phrases as answers are unnatural, these facts suggest that in generic sentences, scrambled phrases obligatorily have focus readings while the focus readings of nominative phrases are neutralized.

Further support for the obligatory focus readings of scrambled phrases in generic sentences comes from the fact that no elements other than the scrambled phrases can get focus readings through receiving stress:

(123) a. ?John-no ie-ni **YUUBIN HAITATU-GA** mainiti
John's house-Dat mailman-Nom everyday
kuru (koto)
come (fact)
'It is the mailman who comes to John's house (and only to John's house) everyday'

b. ?kono eki-ni **SIHATU DENSYA-GA** itumo
this station-Dat first-train-Nom always
teisya suru (koto)
stop (fact)
'It is the first train that always stops at this station (and only at this station)'

I must admit that the judgments are subtle. It is clear, however, that the examples in (123) are less natural than their corresponding non-generic sentences:

(124) a. John-no ie-ni **YUUBIN HAITATU-GA** kinoo
John's house-Dat mailman-Nom yesterday
kita (koto)
came (fact)
'It is the mailman who came to John's house (and only to John's house) yesterday'

b. kono eki-ni **SIHATU DENSYA-GA** kesa
this station-Dat first-train-Nom this morning
teisya sita (koto)
stop (fact)
'It is the first train that stopped at this station (and only at this station) this morning'

This also suggests that scrambled phrases in generic sentences obligatorily have focus readings while the focus readings of nominative phrases are neutralized.

Let us turn to the relation between long distance scrambling and obligatory focus readings in generic sentences. Let us first consider the following examples:

(125) a. John-ga [sihatu densya-ga kono-eki-ni
-Nom first train-Nom this station-Dat
teisya subekida to] itumo syuchoo suru (koto)
should stop Comp always claim (fact)
'John always claims that the first train should stop at this station'

b. **kono-eki-ni** [John-ga [sihatu densya-ga ***t*** teisya subekida to] itumo syuchoo suru] (koto)

In (125a), while the embedded clause is non-generic, the matrix clause is generic. When we scramble the P^{max} *kono-eki-ni* 'this station-Dat' out of the complement clause to the clause-initial position as in (125b), the scrambled phrase is forced to have a focus reading, with the focus reading of the nominative phrase *John-ga* 'John-Nom' being neutralized.

This is supported by the fact that while a generic question soliciting the scrambled phrase as an answer is natural as in (126a), the one soliciting the nominative phrase as an answer is unnatural as in (126b):

(126) a. **dono eki-ni$_i$** [John-ga [sihatu densya-ga t_i
which station-Dat -Nom first-train-Nom
teisya subekida to] itumo syuchoo suru] no
should stop Comp always claim Q
'which station does John always claim that the first train should stop at'

b. ?kono-eki-ni$_i$ [**dare-ga** [sihatsu densya-ga t_i
this station-Dat who-Nom first-train-Nom
teisya subekida to] itumo syuchoo suru] no
should stop Comp always claim Q
'who always claims that the first train should stop at this station'

Although the judgment is subtle, (126b) is clearly less natural than its non-generic counterpart:

(127) kono-eki-ni$_i$ [**dare-ga** [sihatsu densya-ga t_i
this station-Dat who-Nom first-train-Nom
teisya subekida to] syuchoo site iru] no
should stop Comp claim Q
'who claims that the first train should stop at this station'

Furthermore, like other unnatural generic questions, (126b) becomes natural if we replace *kono eki-ni* 'this station-Dat' by *kono eki-ni-wa* 'this station-Dat-Top':

(128) **kono-eki-ni-wa** [dare-ga [sihatsu densya-ga
this station-Dat who-Nom first-train-Nom
teisya subekida to] itumo syuchoo suru] no
should stop Comp always claim Q
'who always claims that the first train should stop at this station'

Another support for the obligatory focus reading of the scrambled phrase in (125b) comes from the unnaturalness of the following example, where the other element than the scrambled phrase receives a stress:

(129) ?kono-eki-ni$_i$ [**JOHN-GA** [sihatu densya-ga t_i
this house-Dat -Nom first-train-Nom
teisya subekida to] itumo syuchoo suru] (koto)
should stop Comp always cla (fact)
'It is John who always claims that the first train should stop at this this station (only at this station)'

This is in contrast with its non-generic counterpart, which is natural:

(130) kono-eki-ni$_i$ [**JOHN-GA** [sihatu densya-ga t_i
this house-Dat -Nom first-train-Nom
teisya subekida to] kinoo-no kaigi-de syuchoo sita
should stop Comp yesterday's meeting-at claimed
(koto)
(fact)
'It is John who claimed at yesterday's meeting that the first train should stop at this this station (only at this station)'

This fact also indicates that the scrambled phrase in (125b) obligatorily receives a focus reading.

Let us next consider the following examples:

(131) a. John-ga [Mary-ga gakkoo-ni mainiti iku to]
-Nom -Nom school-Dat everyday go Comp
omotteiru (koto)
think (fact)
'John thinks that Mary goes to school everyday'

b. **gakkoo-ni** [John-ga [Mary-ga ***t*** mainiti iku to]
omotteiru] (koto)

(132) a. **doko-ni** [John-ga [Mary-ga *t* mainiti iku to]
where -Nom -Nom everyday go Comp
omotteiru] no
think Q
'where does John think that Mary goes everyday'

b. gakkoo-ni [**dare-ga** [Mary-ga *t* mainiti iku] to
school-Dat who-Nom -Nom everyday go Comp
omotteiru] no
think Q
'who thinks that Mary goes to school everyday'

In (131a), while the embedded clause is generic, the matrix clause is non-generic. When we scramble the P^{max} gakkoo-*ni* 'school-Dat' out of the complement as in (131b), the scrambled phrase is not forced to have a focus reading. This is supported by the fact that the question soliciting the scrambled phrase (132a) and the one soliciting the nominative phrase (132b) are both natural. Hence, we can conclude from the above observations that in the case of long-distance scrambling, scrambled phrases are only forced to have focus readings when the landing-site clauses, i.e., the matrix clauses in the above examples, are generic.

In the next subsection, I will argue that on the assumption that the scrambled phrases at generic clauses obligatorily have focus readings, the "domain barrier" effects with scrambling in generic sentences straightforwardly follow form our theory of phrase structure.

4.5.2.3 An Account

It was shown in section 4.5.1 that scrambling in generic sentences, which obligatorily induces a focus reading of the scrambled phrase, exhibits the "domain barrier" effects. This is in contrast with scrambling in non-generic sentences, which is not subject to the "domain barriers." I will argue that the asymmetry between scrambling in generic and non-generic sentences with respect to the "domain barrier" effects, which cannot be given any principled account under the previous locality theories, can be accounted for by our analysis.

Recall that no "domain barrier" effects with scrambling in non-generic sentences follow from the fact that scrambling in non-generic sentences is not driven by any formal feature. Since scrambling in non-generic sentences does not involve any feature-checking, scrambling is required to apply postcyclically. Crucially, scrambling in a non-generic sentence may apply after merger of the "domain barrier" with the main structure. It then follows that no "domain barrier" effects emerge with scrambling in non-generic sentences.

As shown in the last subsection, however, the scrambled phrases in generic sentences obligatorily have focus readings. It is plausible to assume that unlike scrambling in non-generic sentences, scrambling in generic sentences is driven by a strong [FOCUS] feature. If this is on the right track, the "domain barrier" effects with scrambling in generic sentences follow from our theory of phrase structure exactly like those with overt wh-movement and empty operator movement.

Let us consider the adjunct condition effects with scrambling in generic sentences, taking (102b) (repeated here as (133)) as an example:

(133) *?**sono hana-ni** [daremo$_j$-ga [*e$_j$* ***t*** kizuka-zu]
that flower-Dat everyone-Nom notice-without
itumo toorisugiru] (koto)
always go-by (fact)
'everyone always goes by without noticing that flower'

In (133), *sono hana-ni* 'the flower-Dat' is scrambled out of the adjunct clause; the result is deviant. Concerning the adjunct clause *sono hana-ni kizuka-zu* 'without noticing that flower', it is constructed by checking the UFFs of the selected items in conformity with the ICP and the EP.

Let us next consider how to construct the main structure of (133), i.e., *daremo-ga itumo toorisugiru* 'everyone always goes by'. We first select *toorisugiru* 'go by'. The uninterpretable selectional restriction feature of the verb *toorisugiru* 'go by' is checked by selecting *daremo-ga* 'everyone-Nom' and combining the former with the latter. We yield the following structure:

(134) [daremo-ga toorisugiru]
everyone-Nom go-by

Let us assume that focalization is triggered by a strong [FOCUS] feature under a functional category F, which selects a clause like (134) as its complement. Although the adjunct clause *sono hana-ni kizuka-zu* 'without noticing that flower' should be merged with the main clause (134) for its proper interpretation at LF, they cannot be combined with each other at this stage of the derivation due to the EP. The EP requires that the next step should be to select F rather than to combine the adjunct with the main structure.

When F is selected, the ICP requires that the next step should be to check its selectional restriction through combining F with (134). This yields (135):

(135) [$F_{[FOCUS]}$ [daremo-ga toorisugiru]]
everyone-Nom go-by

The ICP requires that the strong [FOCUS] feature of F, being uninterpretable, should be checked immediately by the application of focus scrambling. The candidate for focus scrambling, i.e., *sono hana-ni* 'the flower-Dat', however, is within the adjunct clause which has not been merged with the main structure. Since F does not c-command *sono hana-ni* 'the flower-Dat,' there is no way of checking the strong [FOCUS] feature of F at this stage. The ICP is violated; this derivation is canceled. We can therefore correctly predict that scrambling of *sono hana-ni* 'the flower-Dat' out of the adjunct clause in (133) is not allowed. The other "domain barrier" effects with scrambling in generic sentences can be accounted for in a similar fashion.

It is crucial in this analysis that unlike scrambling in non-generic sentences, scrambling in generic sentences is driven by a formal feature. There is empirical evidence to suggest that this is on the right track. The first evidence comes from "radical reconstruction" facts. Recall that unlike overt wh-movement and topicalization in English, scrambling in non-generic sentences is subject to "radical reconstruction." Scrambling in generic sentences, on the other hand, is not subject to "radical reconstruction":

(136) a. John-ga [seito-ga asa gohan-ni nani-o
-Nom student-Nom breakfast-Dat what-Acc
tabeta ka] itumo siraberu (koto)
ate Q always investigate (fact)
'John always investigates what the students ate at breakfast'

b. *?**nani-o** [John-ga [seito-ga asagohan-ni ***t*** tabeta ka] itumo siraberu] (koto)

(137) a. John-ga [Bill-ga dare-to atteita ka] okusan-ni
-Nom -Nom who-with met Q wife-Dat
itumo hookokusuru (koto)
always inform (fact)
'John always informs his wife who Bill has met'

b. *?**dare-to** [John-ga [Bill-ga ***t*** atteita ka] okusan-ni itumo hookokusuru] (koto)

In (136-137b), the *wh*-phrases *nani-o* 'what-Acc' and *dare-to* 'who-with' are scrambled out of the embedded interrogative complements; the results are deviant.

(136-137b) may not be as deviant as expected, but they are clearly worse than their non-generic counterparts:

(138) a. ?**nani-o** [John-ga [seito-ga asagohan-ni ***t***
what-Acc -Nom students-Nom breakfast-Dat
tabeta ka] itumo siritagatteiru] (koto)
ate Q always want-to-know (fact)
'John always wants to know what the students ate at breakfast'

b. ?**dare-to** [John-ga [Bill-ga ***t*** atteita ka] itumo
who-Dat -Nom -Nom met Q always
okusan-ni hookoku siteiru] (koto)
wife-Dat inform (fact)
'John has always informed his wife who Bill has met'

Recall that while non-feature-driven movement is subject to "radical reconstruction," feature-driven movement is not. The contrast in acceptability between (136-137b) and (138a-b) suggests that unlike scrambling in non-generic sentences, scrambling in generic sentences is feature-driven.

Second, unlike scrambling in non-generic sentences, scrambling in generic sentences exhibits the relativized minimality effects, though I must admit that the judgments are subtle:

(139) a. John-ga [musume-ga maiasa
-Nom daughter-Nom every morning
sono hana-ni mizu-o wasurezuni yaru to]
that flower-Dat water-Acc not-forget give Comp
itumo zimangeni hanasu (koto)
always boastfully say (fact)
'John always says boastfully that his daughter waters
a flower everyday without forgetting'

b. ??**mizu-o$_j$** [John-ga [[**sono hana-ni$_i$** [musume-ga
maiasa t_j t_i wasurezuni yaru]] to] itumo zimangeni
hanasu] (koto)

(140) a. sono mise-ga [syuzin mizukara-ga maiasa
that shop-Nom boss self-Nom every morning
itiba-kara sakana-o siireru to] itumo
market-from fish-ACC buy Comp always
sendensuru (koto)
advertise (fact)
'that shop always advertises that the boss himself buys fish from the market every morning'

b. ??**sakana-o**$_j$ [sono mise-ga [[**itiba-kara**$_i$ [syuzin mizukara-ga maiasa t_j t_i siireru]] to] itumo sendensuru] (koto)

Recall that feature-driven movement is subject to the MLC. Then, the deviancy of multiple scrambling as in (139-140b) can be accounted for if we assume that like overt wh-movement and empty operator movement, scrambling in generic sentences is feature-driven.

If this analysis is correct, we should expect that the relativized minimality effects disappear when either the matrix or embedded clause is non-generic. This is because according to the MLC, the relativized minimality effects only emerge if more than one instances of feature-driven movement of the same type interact. If either the matrix or embedded clause is non-generic, then feature-driven scrambling interacts with non-feature-driven scrambling. Hence, the relativized minimality effects should not emerge in such cases. This prediction is borne out:

(141) a. **mizu-o$_j$** [John-ga [[**sono hana-ni$_i$** [musume-ga
water-Acc -Nom that flower-Dat daughter-Nom
maiasa t_j t_i wasurezuni yaru]] to] omotteiru]
every morning not-forget give Comp think
(koto)
(fact)
'John thinks that his daughter waters a flower every morning without forgetting'

b. **mizu-o$_j$** [John-ga [[**sono hana-ni$_i$** [musume-ga
water-Acc -Nom that flower-Dat daughter-Nom
t_j t_i wasurezuni yatte ita monoda]] to] itumo
not-forget used-to-give Comp always
zimangeni hanasu] (koto)
boastfully say (fact)
'John says boastfully that his daughter used to water a flower without forgetting'

(142) a. **sakana-o$_j$** [sono mise-ga [[**itiba-kara$_i$** [syuzin
fish-Acc that shop-Nom market-from boss
mizukara-ga maiasa t_j t_i siireru]] to]
self-Nom every morning buy Comp
sendensite iru] (koto)
be advertising (fact)
'that shop is advertising that the boss himself buys fish from the market every morning'

b. **sakana-o$_j$** [sono mise-ga [[**itiba-kara$_i$** [syuzin
fish-Acc that shop-Nom market-from boss
mizukara-ga t_j t_i naganen siireteiru]] to]
self-Nom many years has bought Comp
itumo sendensuru] (koto)
always advertise (fact)
'that shop always advertises that the boss himself has bought fish from the market many years'

In (141-142a), the matrix clause is generic while the embedded clause is non-generic. In (141-142b), on the other hand, the matrix clause is non-generic while the embedded clause is generic. In either case, the relativized minimality effects cannot be observed.

I have shown in the above discussion that scrambling in generic sentences is always triggered by a focus feature, exhibiting the "domain barrier" effects as expected by our theory of phrase structure. This analysis is further supported by the fact that scrambling in a non-generic sentence exhibits the "domain barrier" effects if the scrambled phrase gets stressed and thus has an obligatory focus reading, as shown below:[23]

---

[23]If the focused scrambled phrase is NP-*o* 'NP-Acc', the result is acceptable, as shown below:

(i) **SONO HON-O**, John-ga [[***t*** katta] hito]-o sitte iru (koto)
that book-Acc -Nom bought person-Acc know (fact)
'John knows the person who bought that book'

It seems, however, that NP-*o* 'NP-Acc' can function as a kind of topic especially when it gets focalized:

(ii) ?**AMERIKA(-NO-KOTO)-O$_i$**, John-ga [sono$_i$ rekisi]-ni kuwasii (koto)
America(-Gen-fact)-Acc -Nom its history-Dat familiar (fact)
'John is familiar with the history of America

Hence, it is plausible to claim that in (i), *sono hon-o* 'that book-Acc' is not scrambled but base-generated in its surface position as a topic. Hence, the empty category within the relative clause should not count as a trace but as a empty pronominal, as represented below:

(143) Complex NP Constraint

a. *?**SONO-MURA-NI**, John-ga [[***t*** sunde iru]
that village-Dat -Nom reside
hito]-o sagasiteiru (koto)
person-Acc look for (fact)
'John is looking for a person who lives in that village'

b. *?**HOTERU-KARA**, John-ga [[***t*** dekakeru]
hotel-f -Nom come out
hito]-o matibuse site iru (koto)
person-Acc lie in wait for (fact)
'John is lying in wait for people who come out from the hotel'

(144) Adjunct Condition

a. *?**SONO HANA-NI**, daremo-ga [***t*** mizu-o
that flower-Dat everyone-Nom water-Acc
yara-zu] dekakete itta (koto)
give-without went-out (fact)
'everyone went out without watering that flower'

b. *?**SONO ISU-NI**, dareka-ga [***t*** suwari nagara]
that chair-Dat someone-Nom sit while
hon-o yondeita (koto)
book-Acc was reading (fact)
'someone was reading a book while sitting on that chair'

(iii) **SONO HON-O$_i$**, John-ga [[***pro$_i$*** katta] hito]-o sitte iru (koto)
that book-Acc -Nom bought person-Acc know (fact)
'John knows the person who bought that book'

c. ??**SIAI-NO KEKKA-NI**, daremo-ga [***t*** totemo
game-Gen result-Dat everyone-Nom very
gakkarisite] kyuujoo-o atonisita (koto)
disappointed ball park-Acc left (fact)
'everyone left the ball park, disappointed about the result of the game'

These examples are deviant especially when we pronounce them with a pause after the scrambled phrases. It is plausible to assume that in such cases like (143-144), scrambling is triggered by a focus feature. It then follows from our analysis that just like scrambling in generic sentences, scrambling with stressed scrambled phrases like (143-144) is subject to the "domain barriers."

4.6 Concluding Remarks

To summarize this chapter, it was first pointed out that unlike English overt wh-movement and topicalization, scrambling is not subject to the "domain barriers." I have argued that exactly like merger of adjuncts, scrambling, which is not driven by any formal feature, is required to apply postcyclically by our theory of phrase structure. Scrambling therefore may apply after the "domain barrier" is merged with the main structure. The lack of the "domain barrier" effects with scrambling follows. I have also argued that the apparent "domain barrier" effects with scrambling is due to the A-over-A condition at PF. Scrambling, however, is not always devoid of the "domain barrier" effects. It was shown that scrambling in generic sentences exhibits the "domain barrier" effects, so does scrambling in non-generic sentences where scrambled phrases get stressed. I have argued that scrambling is only subject to the "domain barriers" if it has a focus reading and thus a focus feature as its trigger. This is exactly what is expected under our analysis of the "domain barriers," which claims that feature-driven movement, but not non-feature-driven movement, is subject to the "domain barriers." The presence/absence of the "domain barrier" effects with scrambling therefore lends strong empirical support for our theory of the composition of phrase structure.

CHAPTER 5

THE DISTRIBUTION OF *WH*-ELEMENTS IN-SITU

5.0 Introduction

This chapter considers the distribution of *wh*-elements in-situ. It is shown that there are asymmetries concerning the distribution of *wh*-elements in-situ which have not been given any principled account under the MP. I will propose that the Q-feature of a *wh*-element in-situ should undergo "overt" movement to an interrogative C in Japanese-type languages but not in English-type languages. It is shown that "overt" Q-feature movement coupled with our theory of phrase structure gives us a minimalist account of the asymmetries. If the analysis to be presented is on the right track, the distribution of *wh*-elements in-situ constitutes another empirical evidence in favor of our theory of the composition of phrase structure.

The organization of this chapter is as follows. Section 5.1 presents asymmetries with regard to the distribution of *wh*-elements in-situ which need to be given an account. Section 5.2 reviews previous analyses of *wh*-elements in-situ. I will first review EST analyses of *wh*-elements in-situ, arguing that they are incompatible with the MP and thus cannot be adopted as they stand. I will then review previous minimalist accounts of *wh*-elements in-situ. It is shown that they are confronted with conceptual and empirical difficulties. Section 5.3 considers *wh*-elements in-situ in Japanese-type languages. I will argue that their distribution straightforwardly follows from "overt" Q-feature movement together with our theory of phrase structure. It is also shown that the variation between English-type and Japanese-type languages concerning the existence of overt wh-movement can be attributed to the difference in the make-up processes of *wh*-elements between these two types of languages. Section 5.4 deals with *wh*-elements in-situ in English-type languages. I will argue that unlike in Japanese-type languages, "overt" Q-feature movement does not take place in English-type languages. It is shown that the nonexistence of "overt" Q-feature movement accounts for the immunity of *wh*-elements in-situ in English-type languages from both the "domain barriers" and the relativized minimality. Section 5.5 makes concluding remarks.

5.1 Locality Restrictions on *Wh*-elements In-situ

In languages like English, *wh*-elements must move overtly from their base-generated positions to clause-peripheral positions in singularly interrogatives, as shown in (1):[1]

(1) a. **what** did John buy ***t***

b. *John bought **what**

In (1a), the *wh*-element *what* overtly moves to the clause-initial position. The result is acceptable. In (1b), on the other hand, the *wh*-element *what* stays in-situ. The result is deviant. In multiple interrogatives, on the other hand, *wh*-elements stay in-situ as far as there is another *wh*-element which ends up in a clause-peripheral position, as shown below:

(2) **who *t*** bought **what**

The *wh*-element *what* is not moved to the clause-peripheral position but rather staying in-situ.

---

[1](1b) is acceptable as an echo question, which is irrelevant to the present discussion.

Unlike in languages like English, there is no overt wh-movement in languages like Chinese, Japanese, and Korean. In this type of languages, *wh*-elements never undergo overt wh-movement but rather stay in-situ even in singularly *wh*-interrogatives, as exemplified by the following Japanese examples:

(3) a. John-wa **nani**-o katta no
-Top what-Acc bought Q
'what did John buy'

b. John-wa [Bill-ga **nani**-o katta to] itta no
-Top -Nom what-Acc bought Comp said Q
'what did John say that Bill bought'

In (3a-b), the *wh*-element *nani-o* 'what-Acc' is not dislocated but staying in-situ. (3a-b) can properly be interpreted as matrix *wh*-interrogatives, where the *wh*-phrase *nani-o* 'what-Acc' is associated with the matrix Q-morpheme *no*. This subsection investigates the distribution of *wh*-elements in-situ, presenting asymmetries with locality restrictions on *wh*-elements in-situ which need to be given a principled account.

It has been observed by, among others, Aoun and Li (1993), Kim (1991), Huang (1982), Lasnik and Saito (1984, 1992), May (1985), Nishigauchi (1986, 1990), Pesetsky (1987), Reinhart (1992, 1993), Tiedeman (1990), Tsai (1994), and Watanabe (1992a, 1992b) that the distribution of *wh*-elements in-situ is relatively free in comparison with that of the traces left by overt wh-movement. First, unlike overt wh-movement, *wh*-arguments in-situ are not sensitive to the "domain barriers." *Wh*-arguments in-situ in English-type languages are free of the "domain barrier" effects, as shown by the following English examples:

(4) Complex NP Constraint

a. Relative Clauses

who likes [books that criticize **who**]

b. Non-relative Complex NPs

who studied [the evidence that John stole **what**]

(5) Subject Condition

who thinks that [pictures of **whom**] are on sale

(6) Adjunct Condition

who got jealous [because I talked to **whom**]

(7) Non-bridge Verb Condition

who whispered [that John bought **what**]

In (4-7), although the *wh*-arguments in-situ stay within the "domain barriers," they can properly be interpreted as matrix questions.

Similarly, *wh*-arguments in-situ in Japanese-type languages do not exhibit any "domain barrier" effects either, as shown by the following Japanese examples:[2]

(8) Complex NP Constraint

a. Relative Clauses

John-wa [**nani**-o katta hito]-o sagasite iru no

-Top what-Acc bought person-Acc looking-for Q

Lit. 'John is looking for the person who bought what'

[2]Recall that the Subject Condition does not hold in Japanese.

b. Non-relative Complex NPs

John-wa [Bill-ga **nani**-o katta koto]-o sonnani

-Top -Nom what-Acc bought fact-Acc so much

okotte iru no

be angry Q

Lit. 'John is so angry with the fact that Bill bought what'

(9) Adjunct Condition

John-wa [**nani**-o yonde kara] dekaketa no

-Top what-Acc read after went out Q

Lit. 'John went out after he read what'

(10) The Non-bridge Verb Condition

John-wa [Bill-ga **nani**-o katta tte] tubuyaita no

-Top -Nom what-Acc bought Comp murmured Q

Lit. 'John murmured that Bill bought what'

In (8-10), the *wh*-arguments are within the "domain barriers" while the Q-morphemes are outside of them. The *wh*-arguments are nonetheless properly associated with the Q-morphemes.

Second, *wh*-arguments in-situ in English-type languages are also exempt from the Wh-island Constraint. As first observed by Baker (1970), a *wh*-argument within a wh-island may be associated with an interrogative Comp outside the island, as exemplified by (11-12):

(11) who remembers [why John bought **what**]

(12) who wonders [whether (or not) John bought **what**]

In (11), the *wh*-argument in-situ *what* may take either matrix or embedded scope. In (12), the *wh*-argument in-situ *what* takes matrix scope. This is in contrast with overt wh-movement in English-type languages, which obeys the Wh-island Constraint.

It is not the case, however, that *wh*-elements in-situ never exhibit any locality effects. First, as argued by Nishigauchi (1986, 1990) and Watanabe (1992a, 1992b), *wh*-elements in-situ in Japanese are constrained by the Wh-island Constraint:[3]

(13) Tanaka-kun-wa [**dare**-ga **nani**-o tabeta ka] oboete imasu ka
Tanaka-Top who-Nom what-Acc ate Q remember Q

a. 'does Tanaka know who ate what'

b. NOT 'who is the person *x* such that Tanaka knows what *x* ate'

c. NOT 'what is the thing *x* such that Tanaka knows who ate *x*'

d. NOT 'who is the person *x*, what is the thing *y* such that Tanaka knows whether *x* ate *y*'

(adapted from Nishigauchi 1986:37)

---

[3]Huang (1982) observes that Chinese *wh*-elements in-situ do not exhibit any wh-island effects, pointing out that the following example is acceptable as a matrix question:

(i) ni xiang-ahidao [shei mai-le sheme]
you wonder who buy-Asp what

a. 'who is the person *x* such that you wonder what *x* bought'

b. 'what is the thing *x* such that you wonder who bought *x*'

There is, however, disagreement regarding the wh-island effects with Chinese *wh*-elements in-situ. Lisa Cheng (personal communication) observes that examples like (i) are only acceptable as echo questions not as matrix questions. Tsai (1994), on the other hand, proposes an account of the contrast between Chinese and Japanese based on the observation that the wh-island effects are observed in the latter but not in the former. I will leave this important subject for future research, restricting the following discussion to Japanese *wh*-elements in-situ.

(14) a. ?John-wa [Mary-ga **nani**-o katta ka dooka]
-Top -Nom what-Acc bought whether or not
siritagatte iru no
want to know Q
Lit. 'what does John want to know [whether or not
Mary bought *t*]'
(Watanabe 1992b:257)

b. *John-wa [Mary-ga **naze** sono hon-o katta
-Top -Nom why that book-Acc bought
ka dooka] siritagatte iru no
whether or not want-to-know Q
Lit. 'why does John want to know [whether or not
Mary bought that book *t*]'

(13) has only one interpretation where both of the *wh*-elements in-situ *dare* 'who' and *nani* 'what' are associated with the embedded Q-morpheme *ka*. Neither *dare* 'who' nor *nani* 'what' may be associated with the matrix Q-morpheme *ka*. The entire sentence (13) therefore can only be interpreted as a yes/no question.[4] In (14a), the *wh*-argument in-situ *nani* 'what', which is contained within the embedded *wh*-question, is intended to be associated with the matrix Q-morpheme *no*. The result is deviant, though the degree of its acceptability varies among speakers. In (14b), the *wh*-adjunct in-situ *naze* 'why' is contained within the embedded *wh*-question. It is totally impossible to associate *naze* 'why' with the matrix Q-morpheme *no*. These examples suggest that the association of *wh*-

[4]There are some speakers including myself who find that (13) is more severely deviant on reading (13c) than on readings (13b) and (13d). We will later address ourselves to this contrast in acceptability.

elements in-situ with Q-morphemes is constrained by the Wh-island Constraint in Japanese-type languages. This is in contrast with English-type languages, where *wh*-arguments in-situ are not subject to the Wh-island Constraint.

Second, as pointed out by, among others, Aoun and Li (1993), Fukui (1988), Huang (1982), Lasnik and Saito (1992), Nishigauchi (1986, 1990), Reinhart (1992 1993), Tsai (1994), and Watanabe (1992a, 1992b), unlike *wh*-arguments in-situ, *wh*-adjuncts in-situ in Japanese-type languages are constrained by the "domain barriers":[5]

(15) Complex NP Constraint

a. Relative Clauses

*John-wa [Bill-ga **naze** Mary-ni watasita tegami]-o
-Top -Nom why -Dat gave letter-Acc
sagasite iru no
looking-for Q
Lit. 'John is looking for the letter which Bill gave to Mary why'

b. Non-relative Complex NPs

*?John-wa [Bill-ga **naze** sono kuruma-o katta
-Top -Nom why that car-Acc bought
koto]-o sonnani okotte iru no
fact-Acc so much be angry Q
Lit. 'John is so angry with the fact that Bill bought that car why'

---

[5]Recall that *wh*-adjuncts are never allowed to stay in-situ in English-type language.

(16) Adjunct Condition

*John-wa [Bill-ga **naze** totuzen okoridasita kara]

-Top -Nom why suddenly got angry because

sonnani odoroite iru no

so much be surprised Q

Lit. 'John is so surprised because Bill suddenly got angry why'

(17) The Non-bridge Verb Condition

*?John-wa [Bill-ga **naze** sono hon-o katta tte]

-Top -Nom why that book-Acc bought Comp

tubuyaita no

murmured Q

Lit. 'why did John murmur that Bill bought that book *t*'

In (15-17), the *wh*-adjunct in-situ *naze* 'why', which is contained within the "domain barrier," may not be associated with the matrix Q-morpheme. The association of *wh*-adjuncts in-situ with Q-morphemes is constrained by the "domain barriers." This is in contrast with *wh*-arguments in-situ, which never exhibit any "domain barrier" effects.

Let us summarize the distribution of *wh*-elements in-situ. I have shown that the following four asymmetries exist regarding the distribution of *wh*-elements in-situ:

(18) a. There is an asymmetry between overt argument wh-movement and a *wh*-argument in-situ concerning the "domain barrier" effects. While the former is constrained by the "domain barriers," the latter is not.

b. There is an argument/adjunct asymmetry with *wh*-elements in-situ concerning the "domain barriers." While *wh*-adjuncts in-situ are constrained by the "domain barriers," *wh*-arguments in-situ are not.

c. There is an asymmetry between the "domain barriers" and the Wh-island Constraint with *wh*-arguments in-situ in Japanese-type languages. *Wh*-arguments in-situ in Japanese-type languages are constrained by the Wh-island Constraint, but not by the "domain barriers."

d. There is an asymmetry between English-type and Japanese-type languages regarding the Wh-island Constraint with *wh*-elements in-situ. While *wh*-elements in-situ in Japanese-type languages are constrained by the Wh-island Constraint, those in English-type languages are not.

In the next section, I will review previous approaches to these asymmetries.

5.2 Previous Analyses of *Wh*-elements In-situ

5.2.1 EST Approaches to *Wh*-elements In-situ

Analyses of *wh*-elements in-situ within the framework of the EST can be classified into three groups, i.e., LF-movement analyses, pied-piping analyses, and non-movement analyses. It is shown that apart from their incompatibility with the MP, they are also confronted with empirical difficulties.

5.2.1.1 Huang (1982): An LF-movement Analysis

It is Huang (1982) which marks the first substantial step toward establishing the theory of *wh*-elements in-situ within the EST. Developing the idea advocated by Aoun, Hornstein, and Sportiche (1981), Chomsky (1981), and Jaeggli (1981), Huang argues that *wh*-elements in-situ move in the covert component. According to his view, languages do not differ with respect to whether or not they have wh-movement. They differ in where wh-movement applies. In English-type languages, wh-movement applies both in the overt and covert components. In Japanese-type languages, on the other hand, wh-movement only applies in the covert component. The beauty of his theory resides in the parallelisms that it envisions among languages. In all languages, a *wh*-element is related to an interrogative Comp. *Wh*-interrogatives in all languages therefore look alike at LF.

Huang presents the locality restrictions on *wh*-elements in-situ as evidence in support of LF wh-movement. Following Chomsky (1981), Huang assumes that movement operations are constrained by the subjacency condition and the ECP, whose proper government requirement is satisfied by either lexical government or antecedent government. Huang, however, departs form Chomsky in arguing that while the ECP regulates the mapping both in the overt and covert components, the subjacency condition only regulates the one in the overt movement. He argues that if we assume that *wh*-elements in-situ move covertly, their distribution follows from the subjacency condition and the ECP.

Let us first consider the immunity of *wh*-arguments in-situ from the "domain barriers." Huang assumes that subject as well as object positions in Japanese-type languages are properly governed. Hence, under his analysis, covert argument wh-movement in these languages always satisfies the ECP. Although covert argument wh-movement out of a "domain barrier" crosses more than one cyclic nodes, it vacuously satisfies the subjacency condition. This is because the subjacency condition is not operative in the covert component. Hence, his LF wh-movement analysis can correctly predict that *wh*-arguments in-situ are immune from the "domain barriers."

Huang also argues that the locality restrictions on *wh*-adjuncts in-situ follow from his LF wh-movement analysis. As shown above, *wh*-adjuncts in-situ exhibit the "domain barrier" and wh-island effects. Unlike argument wh-movement, adjunct wh-movement in the covert component violates the ECP. Since the trace left by covert adjunct wh-movement, being in an adjunct position, is not lexically governed, it must be antecedent-governed to satisfy the proper government requirement. The trace, however, is within an island and thus not antecedent-governed; this violates the ECP. Hence, his LF wh-movement analysis can correctly predict that *wh*-adjuncts in-situ are constrained by the "domain barriers" and the Wh-island Constraint.

There is, however, an empirical problem with Huang's LF wh-movement analysis. Let us consider *wh*-arguments in-situ in Japanese. Recall that under Huang's analysis, *wh*-arguments in-situ in Japanese appear in lexically-governed positions and thus always satisfy the ECP. His analysis would then expect that *wh*-arguments in-situ in Japanese never exhibit any locality effects. This is because he assumes that the ECP is the only locality condition that regulates LF movement. As pointed out above, however, *wh*-arguments in-situ in Japanese exhibit the wh-island effects. Hence, his LF wh-movement analysis is empirically inadequate in that it cannot account for the asymmetry between the "domain barrier" effects and the wh-island effects regarding *wh*-arguments in-situ in Japanese.[6]

It should also be noted that Huang's LF wh-movement analysis is incompatible with the MP in the following respects. First, his analysis claims that the subjacency condition regulates the mapping in the overt component but not the one in the covert component. His analysis, however, does not give us any reason why such an overt/covert asymmetry with the effectiveness of the subjacency condition exists. There is no a priori reason why the subjacency condition is operative in the overt component but not in the covert component, not vice versa. His analysis therefore does not count as a true explanation but only as a descriptive generalization.

Second, within the EST, it was possible to claim that overt and covert movement should be subject to different constraints. This is because it would constitute evidence for the existence of the covert component as distinct from the overt component. It should be pointed out, however, that any overt/covert asymmetry is incompatible with the MP where the computation from N to LF should be uniform. There is only one derivation from N to LF, which can be spelled out at any stage. Under the MP, therefore, there is no way to state that up to the branching to PF, we have to obey a certain constraint, but we do not have to from there on.[7]

---

[6]As mentioned in note 3, Huang's (1982) arguments are based on the observation that *wh*-arguments in-situ in Chinese do not exhibit any wh-island effects.

[7]Tsai (1994) points out that making reference to the point of Spell-Out might be allowed in the MP, since it does not necessarily mean that we are associating any property with the point of Spell-Out. As Tsai argues, however, even if we are allowed to make

Third, as extensively argued in chapter 3, the subjacency condition and the ECP themselves are incompatible with the MP, since they crucially make use of notions like cyclic nodes and government which are no longer available under the MP.

5.2.1.2 Pied-piping Analyses

This subsection reviews LF pied-piping analyses advocated by, among others, Choe (1987), Hasegawa (1986), Nishigauchi (1986, 1990), and Watanabe (1992a, 1992b). Contrary to Huang's (1982) LF movement analysis, LF pied-piping analyses claim that like overt wh-movement, *wh*-elements in-situ are also constrained by the subjacency condition. The apparent immunity of *wh*-arguments in-situ from the "domain barriers" in Japanese-type languages are accounted for in terms of LF pied-piping. As their representatives, I will review Nishigauchi (1986, 1990) and Watanabe (1992a, 1992b) in detail. It is shown that apart from their incompatibility with the MP, they are confronted with empirical difficulties.[8]

5.2.1.2.1 Nishigauchi (1986, 1990)

Following the insight given by Kuroda (1965), Nishigauchi (1986, 1990) assumes that *wh*-elements in Japanese-type languages are variables, which must be bound by unselective binders in the sense of Lewis (1975), Kamp (1984), and Heim (1982). The relation between *wh*-

---

reference to the point of Spell-Out, Huang's analysis cannot be accommodated under the MP. This is because Huang's analysis would then assume the arbitrary ordering between the subjacency condition and the point of Spell-Out.

[8]See Fiengo et al. (1988) and Hornstein and Weinberg (1987, 1990) for other syntactic arguments against LF pied-piping. See von Stechow (1993) for syntactic and semantic arguments against LF pied-piping.

elements in-situ and unselective binders is established by the relation of government at LF. Among unselective binders in Japanese is the Q-morpheme *ka/no*. When unselectively bound by the Q-morpheme, *wh*-elements in-situ are interpreted as interrogative operators. In order to satisfy the government requirement on unselective binding, *wh*-elements must move to the Spec of the Q-morpheme *ka/no* at LF, where they are properly interpreted as interrogative operators.

Nishigauchi agrees with Huang (1982) in that *wh*-elements in-situ move in the covert component. The former, however, departs from the latter in arguing that the subjacency condition is operative in the overt and covert components. This account for the fact that the distribution of *wh*-elements in-situ are constrained by the Wh-island Constraint in Japanese-type languages.

Nishigauchi argues that the apparent immunity of *wh*-arguments in-situ from the "domain barriers" can be accounted for by the large-scale pied-piping mechanism at LF. Let us consider the relative clause case of the CNPC, taking (8a) (repeated here as (19)) as an example:

(19) John-wa [**nani**-o katta hito]-o sagasite iru no
-Top what-Acc bought person-Acc looking-for Q
Lit. 'John is looking for the person who bought what'

Since covert wh-movement is restricted by the subjacency condition, the *wh*-phrase *nani* 'what' cannot be extracted from the relative clause. The *wh*-phrase *nani* 'what' instead moves to the embedded Spec of CP, as shown below:

(20) John-wa [$_{\text{NP[WH]}}$ [$_{\text{CP[WH]}}$ **nani** [***t***-o katta]] [$_{\text{N'}}$ hito]]-o
sagasite iru no

Adopting the head-feature percolation convention proposed by Selkirk (1981), Nishigauchi claims that the [WH] feature of *nani* 'what', which is in the Spec of CP, percolates up to the CP. On the assumption that relative clauses stay in the Spec of NP, the [WH] feature assigned to the CP further percolates up to the dominating NP. As a result, the complex NP as a whole gets the [WH] feature. The complex NP as a whole therefore moves to the Spec of the matrix CP without violating the CNPC. The resultant LF-representation is as below:

(21) [**[$_{NP}$ [$_{CP}$ nani [*t*-o katta]] [$_{N'}$ hito]]** [[John-wa ***t***-o sagasite iru]] no]

In (21), since the *wh*-element *nani* 'what' is governed and thus unselectively bound by the Q-morpheme *no*, it is properly interpreted as an interrogative operator. Hence, the apparent immunity of *wh*-arguments in-situ from the subjacency condition follows from the large-scale pied-piping mechanism at LF.

Recall that unlike *wh*-arguments in-situ, *wh*-adjuncts in-situ are constrained by the "domain barriers," as exemplified by (15a) (repeated here as (22)):

(22) *John-wa [Bill-g **naze** Mary-ni watasita tegami]-o
-Top -Nom why -Dat gave letter-Acc
sagasite iru no
looking-for Q
Lit. 'John is looking for the letter which Bill gave to Mary why'

Nishigauchi argues that examples like (22) can be ruled out if we assume that the head-feature percolation convention is subject to the category identity requirement. The category identity requirement claims that a *wh*-element must be identical in syntactic category with the dominating node in order for the [WH] feature of the former to be percolated up to the latter. Specifically, a *wh*-element and its dominating node must share [+/- N] for the [WH] feature to climb up. In (22), *naze* 'why' moves to the Spec of the embedded CP, resulting in the following structure:

(23) John-wa [$_{NP}$ [$_{CP[WH]}$ **naze** [Bill-ga ***t*** Mary-ni watasita]] [$_{N'}$ tegami]]-o sagasite iru no

In (23), the [WH] feature of *naze* 'why', which stays in the Spec of the embedded CP, percolates up to the embedded CP in accordance with the category identity requirement. He claims that C and its projections, not being lexical categories, are neutral with respect to [+/- N]. The [+/- N] status of CP is determined on the basis of the feature associated with an element which occupies C or the Spec of CP. Since *naze* 'why', which is adverbial and thus assigned [-N], is in the Spec of the embedded CP, the embedded CP, which is neutral with respect to [+/- N], inherits the [-N] feature from *naze* 'why'. Hence, the [WH] feature may be percolated up from *naze* 'why' to the embedded CP. The [WH] feature of the embedded CP, however, cannot be percolated up to the dominating NP due to the category identity requirement. This is because the embedded CP has [-N] while the dominating NP has [+N]. The complex NP as a whole therefore cannot be assigned [WH] feature and thus may not be pied-piped to the Spec of the matrix CP. Hence, there is no other way than to move *naze* 'why' itself to the Spec of the matrix CP in order for *naze* 'why' to be governed by the Q-morphme *no*. Movement of *naze* 'why' to the Spec of the matrix CP, however, results in a subjacency violation. Hence, we can correctly predict that examples like (22) are deviant.

Nishigauchi's analysis, however, is confronted with empirical difficulties. His analysis cannot account for the argument/adjunct asymmetry with respect to the adjunct condition and non-bridge verb condition effects. Let us consider the adjunct condition effects, taking (9) and (16) (repeated here as (24) and (25), respectively) as examples:

(24) *Wh*-arguments In-situ

John-wa [**nani**-o yonde kara] dekaketa no

-Top what-Acc read after went out Q

Lit. 'John went out after he read what'

(25) *Wh*-adjuncts In-situ

*John-wa [Bill-ga **naze** totuzen okoridasita kara]

-Top -Nom why suddenly got angry because

sonnani odoroite iru no

so much be surprised Q

Lit. 'John is so surprised because Bill suddenly got angry why'

In order to account for the contrast in acceptability between (24) and (25), we have to ensure that the [WH] feature percolates up to the adjunct in (24) but not in (25). Then, the adjunct in (24) can be pied-piped to the Spec of the matrix CP while the one in (25) cannot be; the contrast between (24) and (25) follows. There is, however, no way for the category identity requirement to account for the contrast between these two examples.

Although there is a controversy over the categorial status of adjunct introducing elements like *kara* 'after' and *kara* 'because', they count as either prepositions or complementizers. If they are prepositions, the adjuncts in (24) and (25) count as PP, which is identified as [-N]. The category identity requirement on the head-feature percolation convention would claim that when *naze* 'why', being identified as [-N], moves to the Spec of the embedded CP, its [WH] feature may be percolated up to the adjunct in (25). In (24), on the other hand, when *nani* 'what', being identified as [+N], moves to the Spec of the embedded CP, its [WH] feature may not be percolated up to the adjunct. It would then follow that while the adjunct in (25), being assigned [WH] feature, could be pied-piped to the matrix Spec of CP, the one in (24) could not be. Hence, we would wrongly predict that while (25) is acceptable, (24) is not.

The problem cannot be solved even if we assume that the adjunct introducing elements are complementizers. The adjuncts are then identified as CPs. As mentioned above, since C and CP are neutral with respect to [+/- N], the latter is determined on the basis of an element which occupies C or the Spec of CP. In (24), when *nani* 'what', being identified as [+N], moves to the Spec of the adjunct, the adjunct, being a CP, is also identified as [+N]. In (25), when *naze* 'why', being identified as [-N], moves to the Spec of the adjunct, the adjunct, being a CP, is also identified as [-N]. The [WH] features may percolate from the *wh*-elements to the adjuncts in both cases. We would then expect that the large-scale pied-piping of the adjunct should be possible in both cases. Hence, we would wrongly predict that not only (24) but also (25) is acceptable. The same problem arises with the argument/adjunct asymmetry concerning the non-bridge verb condition effects.

It should also be noted that Nishigauchi's LF pied-piping analysis is incompatible with the MP. First, LF pied-piping is inconsistent with the MP where the operation Move α is reinterpreted as Attract/Move-F. Within the EST, the operation Move selects α and raises it, where α is a category constructed from one or more lexical items. The EST therefore assumes that category movement, which is observed in the overt component, counts as a primitive operation. Given that derivations should be uniform, it is reasonable to claim that category movement, being a primitive operation, takes place not only in the overt component but also in the covert component. We could therefore claim in the EST that since pied-piping is available in the overt component as an instance of category movement, it should also be available in the covert component. Under the MP, however, Move α is reinterpreted as Attract/Move-F. The MP assumes that feature movement should count as a primitive operation. Category movement takes place in the overt component for satisfaction of interface conditions, which counts as a departure from "perfection." It does not count as a primitive operation anymore but as a derivative operation. Hence, although pied-piping, an instance of category movement, is available in the overt component, this cannot be taken as evidence for the existence of LF pied-piping any more under the MP. LF pied-piping is unmotivated under the MP and such an unmotivated mechanism should be abandoned.[9]

Second, recall that Nishigauchi's analysis assumes that a *wh*-element first moves within an island and then the whole island moves to the Spec of an interrogative Comp. It is not clear, however, what motivates wh-movement within an island. This is against the minimalist view that every movement operation is triggered by a formal feature.

---

[9]I will argue later in this chapter that no movement, whether category movement or feature movement, takes place in the covert component.

Third, Nishigauchi's analysis crucially assumes the subjacency condition as a locality principle. As extensively argued in chapter 3, however, the subjacency condition is not compatible with the MP, since it crucially makes use of notions which are not available under the MP.

5.2.1.2.2 Watanabe (1992a, 1992b)

Watanabe (1992a, 1992b) agrees with Nishigauchi (1986, 1990) in claiming that pied-piping in the covert component takes place in Japanese-type languages. The former, however, differs from the latter in arguing that the subjacency condition only regulates the mapping in the overt component but not in the covert component. Watanabe argues that a null wh-operator, which is base-generated in the Spec of DP, is required to move to the Spec of CP in the overt component to be associated with an interrogative Comp. In the covert component, the entire *wh*-phrase moves to the Spec of CP, replacing the chain formed by the S-structure wh-operator movement. Wh-operator movement, which takes place in the overt component, is constrained by the subjacency condition as well as the ECP. Under his analysis, therefore, the wh-island effects with *wh*-elements in-situ in Japanese-type languages are induced by overt wh-operator movement.

Let us look at Watanabe's analysis of the exemption of *wh*-arguments in-situ from the "domain barrier" effects, taking (19) as an example again. Watanabe argues that a null wh-operator may be base-generated in the Spec of the whole complex NP and then moves to the matrix Spec of CP in the overt component, as represented below:

(26) [$_{CP}$ ***Op*** [John-wa [$_{DP}$ ***t*** [nani-o katta hito]]-o sagasite iru] no]

Since the wh-operator moves from outside the complex NP to the matrix Spec of CP, it does not violate the subjacency condition. The whole complex NP from whose Spec position wh-operator movement takes place counts as a *wh*-element and thus undergoes pied-piping to the matrix Spec of CP in the covert component.

Let us next look at Watanabe's analysis of *wh*-adjuncts in-situ. Recall that unlike *wh*-arguments in-situ, *wh*-adjuncts in-situ exhibit the "domain barrier" effects. Under Watanabe's analysis, we should expect that an empty wh-operator originates in the Spec of the whole complex NP and then moves to the matrix Spec of CP. The whole complex NP, which counts as a *wh*-element, should be allowed to undergo pied-piping in the covert component. He argues, however, that unlike *wh*-arguments, *wh*-adjuncts like *naze* 'why' are resistant to unselective binding. On the assumption that pied-piping always involves unselective binding, the "domain barriers" containing *naze* 'why' may not undergo pied-piping. Hence, the "domain barrier" effects with *wh*-adjuncts in-situ follow.

Apart from the incompatibility of LF pied-piping with the MP discussed in the last subsection, Watanabe's analysis suffers from empirical difficulties. Let us first look at the following example:

(27) ?John-wa [[Mary-ga **nani**-o katta ka dooka]
 -Top -Nom what-Acc bought whether or not
 siritagatte iru hito]-o sagasite iru no
 want to know person-Acc looking-for Q
 Lit. 'John is looking for the person who wants to know whether or not Mary bought what'

In (27), the *wh*-element in-situ *nani* 'what' is contained within the indirect *wh*-interrogative, which is further contained within the relative clause. The result is as deviant as an ordinary wh-island violation. As Watanabe (1992b) himself notes, however, his analysis would wrongly predict that examples like (27) are acceptable. Under his analysis, a null wh-operator may originate in the Spec of the whole complex NP and then overtly move to the matrix Spec of CP. Since the wh-operator moves from outside of both the complex NP and the wh-island, there should be no violation of the subjacency condition. In the covert component, the whole complex NP moves to the Spec of CP. Recall that the relation between the interrogative Comp *no* and the *wh*-element *nani* 'what' is established by unselective binding. Since unselective binding is not subject to either the subjacency condition or the ECP, *nani* 'what' could be properly associated with the matrix interrogative Comp *no*, contrary to fact.

Another empirical difficulty with Watanabe's analysis is that it cannot account for the immunity of *wh*-arguments in-situ from the Adjunct Condition and the non-bridge verb condition. Under his analysis, when a *wh*-argument in-situ appears in an adjunct or non-bridge verb "complement," a null wh-operator originates in the Spec of the *wh*-element and then overtly moves to the matrix Spec of CP. Since the null wh-operator is moved out of the adjunct or the non-bridge verb "complement," this movement violates the subjacency condition. Hence, his analysis would wrongly predict that *wh*-arguments may not appear within adjuncts or non-bridge verb "complements." Note that since neither adjunct clauses nor non-bridge verb "complements" are identified as DP, it is highly unlikely that there is a place outside them where we can base-generate a null wh-operator.

5.2.1.3 Aoun and Li (1993): A Non-movement Analysis

Aoun and Li (1993) (hereafter called A&L) argue that *wh*-elements in-situ never undergo movement. Their analysis claims that *wh*-elements in-situ get co-indexed and interpreted with respect to Qu-operators. Qu-operators must appear in the Spec of CP at S-structure in order to take scope. The *wh*-elements in-situ are variables A'-bound by Qu-operators, with the scope of the former being determined by reference to the latter.

A&L argue that the argument/adjunct asymmetry with *wh*-elements in-situ concerning the locality effects follows from their analysis. Given the theory of generalized binding proposed by Aoun (1985a, 1985b), they treat the relation between Qu-operators and *wh*-elements in-situ as a binder-bindee relation. They claim that whatever the exact formulation of the ECP effects may be, the following generalizations should be captured concerning the argument/adjunct asymmetry:

(28) a. A *wh*-in-situ such as *why* in adjunct position must have an antecedent (i.e., must be antecedent governed) in the minimal clause in which it occurs.

b. A *wh*-in-situ such as *who* or *what* in argument position need not have a local antecedent in the minimal clause in which it occurs.

(A&L 1993:219)

Let us first consider the immunity of *wh*-arguments in-situ from the "domain barriers," taking the relative clause case of the CNPC (19) as an example again. Under A&L's analysis, (19) would be assigned the following S-structure:

(29) [$_{CP}$ ***Qu$_i$*** [$_{IP}$ John-wa [**nani$_i$**-o katta hito]-o sagasite iru] no]

In (29), the argument *wh*-element *nani* 'what' stays within the relative clause. The Qu-operator is base-generated in the matrix Spec of CP, which is outside the relative clause. According to (28b) , *nani* 'what', being a *wh*-argument in-situ, need not have a local antecedent within the minimal clause in which it occurs. Hence, *nani* 'what' can be properly interpreted by being A'-bound by the Qu-operator in the matrix Spec of CP.

Let us turn to A&L's analysis of *wh*-adjuncts in-situ, taking the relative clause case of the CNPC (22) as an example again. Under A&L's analysis, there are two possible types of derivations for (22). In the first type, the Qu-operator is base-generated outside the relative clause, as in (30):

(30) $[_{CP}$ ***Qu$_i$*** $[_{IP}$ John-wa [[Bill-ga **naze$_i$** Mary-ni watasita] tegami]-o sagasite iru] no]

In (30), while the *wh*-element is within the relative clause, the Qu-operator is base-generated outside of it. This representation violates (28b), however, since the adjunct *wh*-element in-situ *naze* 'why' does not have any antecedent within the minimal clause in which it occurs. In the second type, on the other hand, the Qu-operator originates within the relative clause and then moves to the matrix Spec of CP. (31a-b) are instances of this type:

(31) a. $[_{CP}$ ***Qu$_i$*** $[_{IP}$ John-wa $[[_{CP}$ ***t$_i$*** [Bill-ga **naze$_i$** Mary-ni watasita]] tegami]-o sagasite iru] no]

b. $[_{CP}$ ***Qu$_i$*** $[_{IP}$ John-wa $[[_{CP}$ ***t'$_i$*** $[_{CP}$ ***t$_i$*** [Bill-ga **naze$_i$** Mary-ni watasita]]] tegami]-o sagasite iru] no]

In (31a), the Qu-operator moves directly to the matrix Spec of CP while in (30b), it first adjoins to the embedded CP and then moves to the matrix Spec of CP. Neither (31a) nor (31b) violates (28b). This is because the adjunct *wh*-element in-situ *naze* 'why' has its antecedent, i.e., the trace of the Qu-operator, within the minimal clause in which it occurs. A&L argue, however, that (31a-b) are both ruled out by the lexical government requirement of the ECP. On the assumption that the head of the relative clause is not a lexical governor, neither t_i in (31a) nor t'_i in (31b) satisfies the lexical government requirement. Hence, there is no legitimate way of deriving (22).

There is, however, empirical evidence which suggests that A&L's analysis is not tenable. Their analysis cannot account for the asymmetry between the "domain barriers" and the Wh-island Constraint with *wh*-arguments in-situ in Japanese-type languages. Let us consider (14a) (repeated here as (32)) as an example:

(32) ?John-wa [Mary-ga **nani**-o katta ka dooka]
-Top -Nom what-Acc bought whether or not
siritagatte iru no
want to know Q
Lit. 'what does John want to know [whether or not Mary bought *t*]'

(32) would be assigned the following representation:

(33) [$_{CP}$ ***Qu$_i$*** [$_{IP}$ John-wa [$_{CP}$ ***Qu$_j$*** [$_{IP}$ Mary-ga **nani$_i$**-o katta] **ka dooka$_j$**] siritagatte iru] no]

Following Huang (1982), A&L claim that the A-not-A question element *ka dooka* 'whether or not' should be treated as an adjunct *wh*-element. It therefore needs a local antecedent, as required by (28a). Since the A-not-A element is locally A'-bound by the Qu-operator in the embedded Spec of CP, it satisfies (28a). The argument *wh*-element in-situ *nani* 'what', on the other hand, is subject to (28b) and thus may be A'-bound by a Qu-operator outside the island. In (33), it is A'-bound by the Qu-operator in the matrix Spec of CP; this satisfies (28b). Hence, A&L would wrongly predict that examples like (32) are acceptable.[10]

It should also be pointed out that A&L's analysis is incompatible with the MP, since it crucially makes use of notions like generalized binding and government which are no longer available under the MP.

5.2.2 Minimalist Approaches to *Wh*-elements In-situ

5.2.2.1 Reinhart (1992, 1993)

Reinhart (1992, 1993) gives a minimalist account of *wh*-elements in-situ, arguing that *wh*-elements in-situ may be interpreted and assigned scope without moving covertly. The mechanism she assumes for the interpretation of *wh*-elements in-situ is what Chomsky (1995) calls absorption, which is originally proposed by Higginbotham and May (1981).

Let us consider how Reinhart's analysis works, taking (34) as an example:

(34) which lady read **which book**

---

[10]It should be noted that there is actually a way of ruling out examples like (32) under A&L's analysis, though they themselves do not mention it. They assume the minimality requirement on the linking of a *wh*-element:

(i) The linking of a *wh*-element with an operator is subject to Minimality. The linking of A with B [... A ... B ...] obeys Minimality iff there is no intervening C [... A ... C ... B ...] such that C is linked to another element D, D ≠ B ≠ A.

(A&L 1993:216)

Given this minimality requirement, the argument *wh*-element *nani* 'what' in (32) may not be linked to the matrix Qu-operator due to the intervening Qu-operator in the embedded Spec of CP. The wh-island effects therefore follow. If we adopt this line of reasoning, we should expect that the wh-island effects always emerge with argument *wh*-elements in-situ. This prediction, however, is invalidated by the immunity of *wh*-arguments in-situ from the Wh-island Constraint in English-type languages.

Following, among others, Karttunen (1977), she assumes that *wh*-elements are translated as existential quantifiers. She claims that existential quantifiers, including *wh*-elements, should be represented in terms of choice functions. The choice function applies to a set, yielding an individual member of the set. The function variable is then bound by an existential operator which may be arbitrarily far away. Under her analysis, therefore, (34) would be assigned the following representation:

(35) for which <x, f> (lady (x) and (x read f (book)))

In (35), the choice function is applied to the *wh*-element in-situ *which book*, yielding f (book). The function variable is then bound by the question operator, which counts as an existential operator. (35) means that (34) denotes the set of true propositions P, each stating for some lady x and for some function f, that x read the book selected by f.

Reinhart's analysis can correctly predict that an argument *wh*-element in-situ may stay within a "domain barrier." This is because the choice function applies to *wh*-arguments in-situ like *dare* 'who' and *nani* 'what', yielding function variables like f (person) and f (thing). The function variable may be bound by the matrix question operator, since there is no locality requirement imposed on the linking between the operator and the function variable. Hence, her analysis can correctly predict that *wh*-arguments in-situ are immune from the "domain barriers."

Concerning the argument/adjunct asymmetry with *wh*-elements in-situ concerning the "domain barrier" effects, Reinhart argues that *wh*-adjuncts are only interpretable in the Spec of CP. This is because unlike *wh*-arguments, *wh*-adjuncts do not have N-sets but rather denote functions ranging over higher-order entities. *Wh*-adjuncts cannot be interpreted in terms of choice functions, which select an individual from a set. Hence, *wh*-adjuncts in-situ must move to the Spec of CP in the covert component, exhibiting the locality effects.

Reinhart's analysis, however, suffers from the following empirical difficulty. Recall that *wh*-arguments in-situ are constrained by the Wh-island Constraint in Japanese-type languages, as exemplified by (32). Under Reinhart's analysis, (32) would be assigned the following representation:

(36) for which *f* (John wants to know whether Mary bought *f* (thing))

In (36), the choice function applies to the *wh*-argument in-situ *nani* 'what', yielding the function variable *f* (thing). Recall that there is no locality requirement imposed on the linking between the operator and the function variable. The function variable is therefore properly bound by the matrix question operator. Hence, her analysis would wrongly predict that examples like (32) are acceptable.[11]

5.2.2.2 Tsai (1994)

Tsai (1994) proposes a minimalist analysis of *wh*-elements in-situ, arguing that their distribution can be accounted for by economy conditions. Following the idea proposed by Kuroda (1965) and further developed by Nishigauchi (1986, 1990) and Watanabe (1992a, 1992b), Tsai claims that

---

[11]It is worth noting that Reinhart's theory of the interpretation of *wh*-elements in-situ, though insufficient to account for the syntactic behavior of *wh*-elements in-situ, is compatible with the analysis to be proposed below. I will propose that only Q-feature but not a *wh*-element as a whole undergoes movement in the case of *wh*-elements in-situ in Japanese-type languages. As Reinhart claims, some interpretation mechanisms are still needed unless *wh*-elements in-situ as a whole move in the covert component. Hence, it is possible to claim that her interpretation rules apply to the output of the syntactic component generated by our Q-feature movement analysis.

wh-elements in Japanese-type languages are free variables. These free variables are properly interpreted through being unselectively bound by Q-operators.

Let us consider how Tsai's analysis works, taking (3a) (repeated here as (37)) as an example:

(37) John-wa **nani**-o katta no
-Top what-Acc bought Q
'what did John buy'

Under Tsai's analysis, (37) would be assigned the following representation:

(38) $[_{CP}$ $\boldsymbol{Op_i}$ $[_{IP}$ John-wa $[_{DP}$ $\boldsymbol{t_i}$ $[_{NP}$ **nani**$_{\mathbf{i}}$]]-o katta] no]

Following Watanabe (1992a, 1992b), Tsai claims that Q-operators originate in the Spec of DP in Japanese. In (38), the Q-operator *Op* originates in the Spec of DP and then moves to the Spec of CP. The *wh*-element *nani* 'what', being a variable, is licensed through being unselectively bound by the Q-operator *Op*. Under his theory, therefore, *wh*-elements in-situ themselves never undergo movement.

Tsai argues that his non-movement analysis of *wh*-elements in-situ is theoretically desirable while giving an account of the language variation concerning the existence of overt wh-movement. He claims that there are two possible strategies of deriving wh-dependencies under the MP: binary generalized transformation (= Merge) as in (39a) and singularly generalized transformation (= Attract/Move) as in (39b):

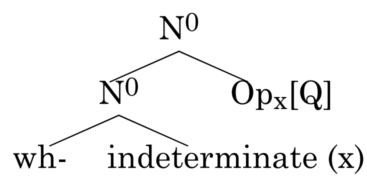
(39) a. $[_{X''}\ \Delta\ [_{X'} \ldots \text{wh} \ldots]] \longrightarrow [_{X''}\ \text{Op}\ [_{X'} \ldots \text{wh} \ldots]]$

b. $[_{X''}\ \Delta\ [_{X'} \ldots \text{wh} \ldots]] \longrightarrow [_{X''}\ \text{wh}_i\ [_{X'} \ldots t_i \ldots]]$

Following Chomsky (1993, 1995), Tsai assumes that Merge is cost-free while Attract/Move is not. The economy condition claims that UG would prefer (39a) to (39b), since the former is less costly than the latter. It then follows that if a language may introduce Q-operators in terms of Merge, that language does not resort to Attract/Move. Since Japanese-type languages may introduce Q-operators, they do not employ Attract/Move in order to form wh-dependencies.

Unlike Japanese-type languages, English-type languages employ overt wh-movement to form wh-dependencies. Following Watanabe (1992a, 1992b), Tsai argues that English-type languages establish an operator-variable pair under X^0-level, as shown below:

(40)



Unlike *wh*-elements in Japanese-type languages, those in English-type languages are not free variables. Although the economy condition tells us that the insertion of a Q-operator should be preferred over movement, a Q-operator may not be inserted in English-type languages. This is because if a Q-operator were inserted, the inserted Q-operator would remain dangling without binding any variable. This would result in a vacuous quantification, which would violate the condition on the operator-variable construction. Hence, overt wh-movement instead of the insertion of a Q-operator takes place in English-type languages. Given that the [+wh] feature of C is strong in English, wh-movement must take place before Spell-Out in order to construct a wh-dependency.

Tsai argues that apart from this theoretical support, his analysis also receives empirical support. First, his analysis can account for the immunity of *wh*-arguments in-situ from the "domain barriers" along the line of the reasonings given by Watanabe (1992a, 1992b).

Second, his analysis can correctly account for the distribution of *wh*-adjuncts in-situ in Japanese-type languages. Tsai claims that unlike *wh*-arguments, *wh*-adjuncts do not count as free variables but as intrinsic operators. In the case of the interrogative with a *wh*-adjunct in-situ, if we inserted a Q-operator, it would result in two instances of vacuous quantification. This is because the *wh*-adjunct in-situ, being an intrinsic operator, as well as the inserted Q-operator would not have any variables to bind. Hence, the *wh*-adjunct in-situ is required to undergo movement. On the assumption that [+wh] features are weak in Japanese, the *wh*-adjunct in-situ undergoes movement in the covert component. This covert movement of the *wh*-adjunct in-situ is constrained by the locality effects.

Third, along the line of reasoning given by Watanabe, Tsai's analysis can account for the fact that *wh*-arguments in-situ are constrained by the Wh-island Constraint in Japanese-type languages.[12]

---

[12]Tsai's analysis can also account for an asymmetry between *wh*-elements in-situ in Chinese and those in Japanese with respect to the Wh-island Constraint if the contrast between the two languages really exists. Following the observation made by Huang (1982), he claims that unlike Japanese *wh*-arguments in-situ, Chinese *wh*-arguments in-situ are not constrained by the Wh-island Constraint:

(i) ni xiang-ahidao [shei mai-le sheme]
you wonder who buy-Asp what
a. 'who is the person *x* such that you wonder what *x* bought'
b. 'what is the thing *x* such that you wonder who bought *x*'
c. 'do you wonder who bought what'

He observes that (i) is acceptable under the wide scope construal of either of the *wh*-elements in-situ. He argues that the difference between Chinese and Japanese resides in the fact that while Q-operators originate in the Spec of DP in the latter, they originate in the Spec of CP in the former. (i) would therefore be assigned the following representations depending on which *wh*-element in-situ is unselectively bound by which Q-operator:

Tsai's analysis, however, is confronted with empirical as well as conceptual difficulties. Let us first consider its empirical problem. First, exactly like Watanabe's analysis, his analysis would wrongly predict that examples like (27) (repeated here as (41)) are acceptable:

(41) ?John-wa [[Mary-ga **nani**-o katta ka dooka]
-Top -Nom what-Acc bought whether or not
siritagatte iru hito]-o sagasite iru no
want to know person-Acc looking-for Q
Lit. 'John is looking for the person who wants to know whether or not Mary bought what'

Under Tsai's analysis, a Q-operator may originate in the Spec of the relative clause as a whole and then move to the matrix Spec of CP. Since this movement does not cross any island, the result should be acceptable. (41), however, is in fact deviant.

Second, exactly like Watanabe's analysis, Tsai's analysis would wrongly predict that *wh*-arguments in-situ exhibit the adjunct condition and non-bridge verb condition effects in Japanese. This is because it is highly unlikely that there is a place outside these islands where a Q-operator may be base-generated.

(ii) a. [$_{CP}$ ***Op$_i$*** [$_{IP}$ ni xiang-ahidao [$_{CP}$ ***Op$_j$*** [$_{IP}$ **shei$_i$** mai-le **sheme$_j$**]]]]
b. [$_{CP}$ ***Op$_j$*** [$_{IP}$ ni xiang-ahidao [$_{CP}$ ***Op$_i$*** [$_{IP}$ **shei$_i$** mai-le **sheme$_j$**]]]]

In (iia-b), the Q-operators in the embedded and matrix Spec's of CP are both base-generated there. Since no movement is involved in this derivation, we can correctly predict that (i) does not violate any locality condition. Hence, the asymmetry between Chinese and Japanese with respect to the existence of the wh-island effects follows. It should be pointed out, however, that there is disagreement among speakers over the judgment concerning the existence of the wh-island effects in Chinese.

Let us turn to its conceptual problems. First, apart from the inherent global property associated with the economy conditions, his analysis needs very global considerations. Recall that under his analysis, the insertion of a Q-operator should always be available unless it would lead a derivation to crash. In order to decide whether to insert a Q-operator at a certain stage of a derivation, we have to look ahead to see whether the insertion of a Q-operator at that stage would result in an LF-interface only with legitimate objects. Such global considerations, however, should be avoided, since it would necessarily induce computational intractability.

Second, his claim that the economy condition prefers insertion of a Q-operator to movement of a *wh*-element is untenable. Recall that the economy conditions only compare the convergent derivations which belong to the same reference set. Chomsky (1993) claims that the derivations with the same LF representation constitute a reference set. Chomsky (1995), on the other hand, claims that the derivations with the same N constitute a reference set. (39a-b), however, never belong to the same reference set whichever definition of the reference set is to be adopted. (39a-b) do not have the same N, since the reference set of the former includes a Q-operator while that of the latter does not. Furthermore, it is clear that (39a-b) do not have the same LF-representation, either. Hence, (39a-b) are not comparable by the economy conditions.

Third, his analysis of overt wh-movement in English-type languages is problematic. He claims that if we did not apply overt wh-movement but rather inserted a Q-operator in English-type languages, it would result in a vacuous quantification. Since this would violate the condition on the operator-variable construction, overt wh-movement must take place. Note that his analysis crucially relies on the assumption that the condition on the operator-variable construction functions as a driving force for movement operations. This is against Chomsky's (1993, 1995) minimalist view that movement operations are only triggered by feature-checking. This is because the construction of an operator-variable pair does not involve any feature-checking. The condition on the operator-variable construction therefore should not function as a driving force for movement operations. Hence, Tsai's analysis of overt wh-movement in English-type languages, which uses the condition on the operator-variable construction as a driving force, is not compatible with the minimalist view.

To summarize this section, I have first reviewed the previous EST analyses of *wh*-elements in-situ. I have argued that they are incompatible with the MP and thus cannot be adopted as they stand. It was also shown that those analyses are empirically problematic. I have then reviewed the previous minimalist approaches to *wh*-elements in-situ. It was shown that they are not tenable on the conceptual and empirical grounds. In the next section, I will consider the distribution of *wh*-elements in-situ in Japanese-type languages and propose "overt" Q-feature movement. It is shown that the "overt" Q-feature movement analysis coupled with our theory of phrase structure enables us to account for the hitherto unexplained asymmetries concerning the distribution of *wh*-elements in-situ in Japanese-type languages.

5.3 *Wh*-elements In-situ in Japanese-type Languages

5.3.1 Indeterminate Elements

Essentially following the idea originally proposed by Kuroda (1965) and further developed by, among others, Nishigauchi (1986, 1990), Tsai (1994), and Watanabe (1992a, 1992b), I claim that unlike *wh*-elements in English-type languages, so called *wh*-elements in Japanese-type languages are not intrinsic interrogative operators. They should rather be identified as free variables which are to be bound by something else. Following Kuroda (1965), let us call so called *wh*-elements in Japanese-type languages indeterminate elements in order to avoid complications that the term "*wh*-element" might induce.

Unlike *wh*-elements in English-type languages which are always construed as interrogative operators, indeterminate elements can be construed as either interrogative or noninterrogative elements depending on the context where they appear. This has been taken as evidence in support of the view that indeterminate elements are free variables rather than intrinsic interrogative operators. Japanese indeterminate elements, for example, can be construed in the following ways:

(42) *Wh*-elements (Interrogative Operators)

a. **dare**-ga sono hon-o kaimasita **ka**
who-Nom that book-Acc bought Q
'who bought that book'

b. John-wa **nani**-o kaimasita **ka**
-Top what
'what did John buy'

c. John-wa sono hon-o **doko**-de kaimasita **ka**
where
'where did John buy that book'

d. John-wa sono hon-o **itu** kaimasita **ka**
when
'when did John buy that book'

e. John-wa sono hon-o **naze** kaimasita **ka**
why
'why did John buy that book'

f. John-wa **dono** hon-o kaimasita **ka**
which
'which book did John buy'

(43) Existential Quantifiers

a. dare-ka 'someone'

b. nani-ka 'something'

c. doko-ka 'somewhere'

d. itu-ka 'sometime'

e. *naze-ka[13] 'for some reason'

---

[13] *Naze-ka* is acceptable, but can only be interpreted as 'do/does not know why' not as a existential quantifier. Hence, examples like (i) are deviant:

(i) *John-wa [Bill-ga naze-ka Mary-to-no konyaku-o kaisyoo sita
-Top -Nom -with-Gen engagement-Acc has broken
to] uwasa siteiru
Comp spread the rumor
'John is spreading the rumor that Bill has broken his engagement with Mary for some reason'

Note that if *naze-ka* could be interpreted as 'for some reason', (i) would be acceptable, as shown by the acceptability of (ii), where *nanraka-no riyuu-de* 'for some reason' is used instead of *naze-ka*:

f. dono N'-ka 'some-N''

(44) Universal Quantifiers (Continuous Cases)[14]

a. DAre-mo 'everyone'

b. ?NAni-mo[15] 'everything'

c. DOko-mo 'everywhere'

d. Itu-mo 'whenever'

e. *NAze-mo 'for whatever reason'

f. DOno-N'-mo 'every N''

(45) Universal Quantifiers (Discontinuous Cases)

a. [**dare**-ga kite **mo**] boku-wa awa-nai

everyone-Nom come I-Top meet-Not

'for all *x*, *x* a person, if *x* comes, I will not meet *x*'

(ii) John-wa [Bill-ga nanraka-no riyuu-de Mary-to-no konyaku-o
-Top -Nom some-Gen reason-for -with-Gen engagement-Acc
kaisyoo sita to] uwasa siteiru
has broken Comp spread the rumor
'John is spreading the rumor that Bill has broken his engagement with Mary for some reason'

[14]The universally-quantified particle *-mo* is isomorphic with the negative polarity particle *-mo*, though they have different pitch patterns. As extensively discussed by McCawley (1968), Japanese is a pitch-accent language and an accent falls on the last syllable of a stretch of high-pitch tones. High-pitch tones are indicated by the upper case and low-pitch tones, by the lower case here and in the relevant examples to follow.

[15]As pointed out by Watanabe (1992a), *NAni-mo* 'everything' can only be used in the fixed expression as in (i), though even examples like (i) are not perfect:

(i) ??NAni-mo-ga iyani natta
-Nom hateful became
'Everything has become hateful/I have become sick with everything'
(Watanabe 1992a:49, the judgment is mine)

In a colloquial speech, *NAni-mo-kamo* instead of *NAni-mo* is used:

(ii) NAni-mo-kamo-ga tetigai darakeda
everything-Nom has gone wrong
'everything has gone wrong with me'

b. [John-ga **nani**-o katte **mo**] boku-wa
-Nom everything-Acc buy I-Top
kamaima-sen
care-Not
'for all *x*, *x* a thing, if John buys *x*, I don't care'

c. [John-ga **doko**-de sono hon-o katte **mo**]
-Nom everywhere-at that book-Acc bought
boku-wa kamaima-sen
I-Top care-Not
'for all *x*, *x* a place, if John buys that book at *x*, I don't care'

d. [John-ga **itu** sono hon-o katte **mo**] boku-wa
-Nom whenever that book-Acc bought I-Top
kamaima-sen
care-Not
'for all *x*, *x* a time, if John buys that book at *x*, I don't care'

e. *[John-ga **naze** sono hon-o katte
-Nom for whatever reason that book-Acc bought
mo] boku-wa kamaima-sen
I-Top care-Not
'for any *x*, *x* a reason, if John buys that book for *x*, I don't care'

f. [John-ga **dono** hon-o katte **mo**] boku-wa

-Nom every book-Acc buy I-Top

kamaima-sen

care-Not

'for all x, x a book, if John buys x, I don't care'

(46) Negative Polarity Items (NPIs)

a. daRE-MO 'anyone'

b. naNI-MO 'anything'

c. doKO-MO 'anywhere'

d. *iTU-MO 'anytime'

e. *naZE-MO 'for any reason'

f. dono N'-MO 'any N''

When indeterminate elements are linked with the Q-morpheme *ka*/*no* in Comp, they are construed as *wh*-elements, i.e., interrogative operators, as exemplified in (42). When combined with the existentially-quantified morpheme *ka*, the universally-quantified morpheme *mo*, and the negative polarity morpheme *mo*, on the other hand, indeterminate elements are construed as existential quantifiers, universal quantifiers, and negative polarity items, respectively, as shown in (43)-(46). Based on these observations, Nishigauchi (1986, 1990) proposes that morphemes like *ka* and *mo* should be analyzed as unselective binders. Under this view, indeterminate elements are construed in various ways through being given a quantificational force by their unselective binder. For example, indeterminate elements are construed as *wh*-elements, i.e., interrogative operators, through being unselectively bound by the Q-morpheme *ka*/*no*.

As the syntactic basis for the above mentioned semantic interpretations of indeterminate elements, I propose that indeterminate elements should be assigned UFFs which are to be checked by the corresponding interpretable features of their unselective binders. In other words, the checking relation provides the syntactic basis for unselective binding:

(47) *Wh*-elements (Interrogative Operators)
dare$_{[Q]}$ ... ka/no$_{[Q]}$ $\longrightarrow$ dare ... ka/no$_{[Q]}$

(48) Existential Quantifiers
dare$_{[E]}$-ka$_{[E]}$ $\longrightarrow$ dare-ka$_{[E]}$

(49) Universal Quantifiers
DAre$_{[U]}$-mo$_{[U]}$ $\longrightarrow$ DAre-mo$_{[U]}$

or

DAre$_{[U]}$ ... mo$_{[U]}$ $\longrightarrow$ DAre ... mo$_{[U]}$

(50) Negative Polarity Items (NPIs)
daRE$_{[N]}$-MO$_{[N]}$ $\longrightarrow$ daRE-MO$_{[N]}$

In (47)-(50), a question feature, an existential quantification feature, a universal quantification feature, and an NPI feature are represented as [Q], [U], [E], and [N], respectively. For a purpose of illustration, let us look at (47) in detail. The indeterminate element which is to be construed as a *wh*-element, i.e., an interrogative operator, is assigned a Q-feature. The Q-feature of the indeterminate element enters into a checking relation with that of the Q-morpheme *ka*/*no*. Note that while the Q-feature of the indeterminate element is uninterpretable, that of the Q-morpheme is interpretable. This view is plausible, since it is the Q-morpheme *ka/no* but not the indeterminate element that has an interrogative quantificational force. Hence, after the checking operation takes place, the Q-feature of the indeterminate element gets erased while that of the Q-morpheme remains intact. This gives us the syntactic basis for the semantic fact that indeterminate elements can only be construed as *wh*-elements (interrogative operators) through being associated with Q-morphemes. The checking processes in (48)-(50) proceed in a similar fashion.[16]

There is an important point which must be kept in mind in the discussion to follow. We observe from (42)-(46) that the indeterminate elements but *naze* are construed in various ways according to their unselective binders. The indeterminate element *dare*, for instance, is construed as a *wh*-element, an existential quantifier, a universal quantifier, and a negative polarity item when it is unselectively bound by the Q-morpheme *ka/no*, the existentially-quantified morpheme *ka*, the universally-quantified morpheme *mo*, and the negative polarity morpheme *mo*, respectively. In other words, *dare* may be assigned a Q-feature, an E-feature, a U-feature, or an N-feature according to the context where it appears. Hence, its UFF, which is to be checked by its unselective binder, should count as an optional feature in the sense of Chomsky (1995). The indeterminate element *naze*, on the other hand, can only be bound by the Q-morpheme *ka/no* and construed as an interrogative operator, as in (42e). As shown in the (e) examples of (43)-(46), it may not be bound by the existentially-quantified morpheme *ka*, the

[16]One might argue that the indeterminate elements should have not only UFFs which are to be checked by their unselective binders but also interpretable features which would represent their semantic import. Considering *wh*-elements as examples, one might say that they also have interpretable [WH]-features. Even if it is correct to assume that *wh*-elements have [WH]-features, the latter, being interpretable, does not have to enter into any checking relation. The arguments to follow therefore hold regardless of whether such interpretable features of the indeterminate elements exist or not.

universally-quantified morpheme *mo*, or the negative polarity morpheme *mo*. In other words, *naze* is always assigned a Q-feature, which is to be checked by the Q-morpheme. Hence, its uniterpretable Q-feature should count as an intrinsic feature in the sense of Chomsky (1995). As we will see in the next section, this difference between *naze* and the other indeterminate elements plays a crucial role in deriving the argument/adjunct asymmetry with respect to the locality restrictions on *wh*-elements in-situ.

5.3.2 "Overt" Q-feature Movement

With the discussion in the previous subsection in mind, this subsection considers how to construct simplex *wh*-interrogatives in Japanese-type languages. I will argue that the Q-feature of an indeterminate element moves overtly to a Q-morpheme in order to be checked off.[17]

Let us first consider how to construct simplex *wh*-interrogatives with *wh*-arguments in-situ, taking (37) (repeated here as (51)) as an example:

(51) John-wa **nani**-o katta no
-Top what-Acc bought Q
'what did John buy'

The indeterminate element *nani* 'what', which is required by the uninterpretable selectional restriction feature of the verb *katta* 'bought', must be inserted in a cyclic manner in accordance with the ICP, yielding

---

[17]Maki (1995) also proposes a feature movement analysis of *wh*-elements in-situ in Japanese, though he assumes that feature movement takes place in the covert component.

nani-o katta 'what-Acc bought'. Recall that the indeterminate element *nani* 'what' in (51), which is to be construed as a *wh*-element (an interrogative operator), has a Q-feature. Although the Q-feature of *nani* 'what' is uninterpretable, it is optional and thus not introduced into the derivation when *nani* 'what' is selected from the N.

As the derivation proceeds, we come to the stage where C is selected:

(52) [[John-wa [nani-o katta]] $C_{[Q]}$]

I argue that exactly like in English-type languages, the Q-feature of C is also strong in Japanese-type languages.[18] The Q-feature of C, being strong and thus uninterpretable, must be checked immediately in conformity with the ICP in Japanese-type as well as English-type languages. Recall that in English-type languages, the strong Q-feature of C is required to be checked immediately by overt wh-movement when the feature becomes accessible to a computation in conformity with the ICP. In Japanese-type languages, on the other hand, the strong Q-feature of C is checked off immediately by merger of the Q-morpheme *no*, which has an interpretable Q-feature:

(53) [[John-wa [nani-o katta]] $\mathbf{no_{[Q]}}$]

After the checking operation takes place, the Q-feature of C, being uninterpretable, is erased while that of the Q-morpheme, being interpretable, remains.

At this stage, we add a Q-feature to *nani* 'what'. Since it is uninterpretable, it must be checked immediately in conformity with the ICP. I argue that the Q-feature of *nani* 'what' enters into a checking

---

[18]See Watanabe (1992a, 1992b) for the claim that the feature of an interrogative Comp is universally strong.

relation with that of the Q-morpheme *no* through feature-movement, as represented in (54a):[19]

(54) [[John-wa [**nani$_{[Q]}$**-o katta]] **no$_{[Q]}$**]

After the checking operation takes place, the Q-feature of *nani* 'what', being uninterpretable, is erased while that of the Q-morpheme *no*, being interpretable, remains intact, resulting in the following structure:

(55) [[John-wa [nani-o katta]] no$_{[Q]}$]

In this way, we can construct interrogatives with *wh*-arguments in-situ like (51).

It should be noted that at stage (52) when the interrogative C is introduced, the EP prevents its strong Q-feature from entering into a checking relation through Q-feature movement from *nani* 'what'. A question now arises what happens if we start with such an N that contains the same elements as the one of (51) except that it does not include any Q-morpheme. In the derivation based on that N, the strong Q-feature of the interrogative C can only be checked by Q-feature movement from *nani* 'what'. Hence, the derivation converges, yielding the following:

(56) *John-wa nani-o katta
-Top what-Acc bought

[19]Recall that we have been assuming that Attract/Move is not an primitive operation but a complex operation consisting of two primitive operations, i.e., Copy and Merge. Checking relations are therefore established by Copy or Merge but not by Attract/Move. Hence, to be precise, the Q-feature of *nani* 'what' first undergoes a copy operation and gets deleted. It is then merged with the Q-morpheme and erased. For expository purposes, however, I pretend to assume that Q-feature undergoes Attract/Move in order to enter into a checking relation.

I argue that although this derivation converges, the resultant LF-representation is gibberish due to the fact that the indeterminate element *nani* 'what' does not have any unselective binder and thus may not be properly interpreted. In other words, the checking relation between the indeterminate element and the Q-morpheme, which provides the syntactic basis for unselective binding, is not properly established.

Note also that Q-feature movement takes place before Spell-Out under our analysis. Recall that UFFs must be converted to their corresponding phonological properties through entering into checking relations before Spell-Out. The Q-feature of an indeterminate element therefore must enter into a checking relation before Spell-Out. More generally, our analysis claims that movement, whether category or feature movement, always takes place in the overt component. No movement ever takes place in the covert component. If this conjecture is correct, the covert component is minimized, remaining only with operations like the construction of an operator-variable pair.[20]

Let us next consider how to construct simplex *wh*-interrogatives with *wh*-adjuncts in-situ like the following example:

(57) John-wa **naze** sono hon-o katta no
-Top why that book-Acc bought Q
'why did John buy that book'

Recall that the indeterminate element *naze* 'why' is assigned an intrinsic uninterpretable Q-feature. Merger of the indeterminate element *naze* 'why', being an adjunct, is not triggered by any UFF. It is therefore forced to be inserted postcyclically. Crucially, when the Q-morpheme *no*

[20]Bobaljik (1995) and Brody (1995) also deny the existence of LF-movement.

is merged to check the strong Q-feature of C, *naze* 'why' has not been selected from the N yet, as represented below:[21]

(58) [[John-wa [sono hon-o katta]] $\text{no}_{[Q]}$]

At this stage, *naze* 'why' is selected from the N. Its uninterpretable Q-feature must be checked immediately in accordance with the ICP. The Q-feature of *naze* 'why' therefore undergoes movement to be checked, as depicted below:

(59) [[John-wa [**naze**$_{\mathbf{[Q]}}$ [sono hon-o katta]]] **no**$_{\mathbf{[Q]}}$]

After the checking operation, the Q-feature of *naze* 'why', being uninterpretable, is erased while that of the Q-morpheme, being interpretable, remains. This yields the following structure:

(60) [[John-wa [naze [sono hon-o katta]]] $\text{no}_{[Q]}$]

In this way, we can construct interrogatives with *wh*-adjuncts in-situ like (57).

5.3.3 Category Movement and Feature Movement

The analysis presented in the previous subsection crucially assumes that feature movement takes place before Spell-Out. This assumption is against Chomsky's (1995) view that a category, not just a feature, is required to move before Spell-Out. Recall that in Chomsky (1995) where the traditional theory of movement is reinterpreted as the theory of feature movement, what is raised should be just a feature unless it would

[21]Recall that the checking relation between an indeterminate element and a Q-morpheme provides the syntactic basis for unselective binding. If the *wh*-adjunct in-situ *naze* 'why' instead of the Q-morpheme *no* were merged into C to check its strong Q-feature, the checking relation between *naze* 'why' and the Q-morpheme *no* would never be established. Hence, *naze* 'why' would end up not having any unselective binder and thus would not be properly interpreted.

result in a crashed derivation. In order to ensure this, he proposes the "no extra baggage" condition, which is one of the economy conditions:

(61) "No Extra Baggage" Condition

F carries along just enough material.

(adapted from Chomsky 1995:262)

According to the "no extra baggage" condition, the derivation that raises just a feature should be chosen as "optimal" unless it would violate the FI and therefore crash. This is because the raising of just a feature does not carry along any "extra baggage." Chomsky argues, however, that unlike in the covert component, a category, not just a feature, should be raised in the overt component. If only a feature moved in the overt component, features of a single lexical item would be scattered. Only the raised feature would be in the checking domain but all the other features would remain in-situ. Chomsky assumes that there is a PF requirement that features of a single lexical item must be within a single X^0. A derivation with such scattered features violates this PF requirement and therefore crashes. Chomsky therefore claims that in the overt component, an "extra baggage" is required for PF-convergence; the whole category, but not just a feature, undergoes a movement operation.

As pointed out by Takano (1996), however, Chomsky's claim that feature movement before Spell-Out leads a derivation to crash at PF is questionable. Recall that Spell-Out is the operation which strips away from an object those elements which are relevant to the phonological component, with the remaining elements being mapped to LF. It is then reasonable to claim that formal features, which are not relevant to PF, do not enter into the phonological component. Then, "overt" formal feature movement has nothing to do with PF. Hence, "overt" formal feature movement does not lead a derivation to crash at PF, contrary to Chomsky's claim.[22]

A question now arises what forces a feature to carry along an "extra baggage" in the case of category movement like English overt wh-movement. I propose that category movement is forced not by the PF-requirement but by the following LF requirement:

(62) The interpretable features of a lexical item should be within the lexical item at LF.

According to (62), if an interpretable feature moves out of the lexical item by feature movement and thus does not stay within the lexical item at LF, the derivation crashes. According to the "no extra baggage" condition, movement of an interpretable feature always carries an "extra baggage" for LF-convergence; the whole category, not just a feature, moves. There is considerable validity to LF-requirement (62) on interpretable features, since it is reasonable to claim that if an interpretable feature of a lexical item does not stay within the lexical item at LF, the lexical item may not be properly interpreted at that level.[23]

---

[22]As discussed in chapter 1, Chomsky (1996) proposes the different view that feature chains cannot be interpreted at PF, which forces category movement in the overt component. This view, however, is untenable. Since formal features do not enter into the PF-component, feature chains never appear at PF.

[23]One might claim that this analysis is conceptually problematic, since it induces a problem of globality. This analysis makes use of the "extra baggage" condition and LF-requirement (62). Note that the "extra baggage" condition, being an economy condition, needs global considerations. Note also that LF-requirement (62) is a global interface condition. It follows that this analysis, which makes use of these global conditions, necessarily induces computational intractability. I argue, however, that this computationally intractable optimization problem can be solved by assuming the following local "heuristic algorithm" ("computational trick"):

(i) Category movement takes place only if an interpretable formal feature is required to undergo Attract/Move.

Hence, a problem of globality does not arise in this analysis.

Let us consider how LF-requirement (62) coupled with the "no extra baggage" condition forces overt wh-movement in English-type languages. Let us first explicate the make-up process of *wh*-elements in English-type languages. Recall that Tsai (1994) and Watanabe (1992a, 1992b) pursue the parallelisms between *wh*-elements in Japanese-type languages and those in English-type languages, arguing that the latter is made up essentially in the same way as the former. We essentially follow Tsai and Watanabe in assuming that *wh*-elements in English-type languages consist of indeterminate elements and their unselective binders, as represented by (40) (repeated here as (63)):[24]

(63) [N^0 [N^0 wh- indeterminate (x)] Op_x [Q]]

Their idea can be reinterpreted under our analysis in the following way. Exactly like indeterminate elements in Japanese-type languages, those in English-type languages are also assigned uninterpretable Q-features which are to be checked by the Q-features of their unselective binders. Unlike in Japanese-type languages, the checking operation between these Q-features take places under N^0-level in English-type languages. The Q-feature of an indeterminate element raises to that of its unselective binder under N^0-level in order to be checked off. After the checking operation takes place, the Q-feature of the indeterminate element, being uninterpretable, is erased while that of its unselective binder, being

---

[24]This idea dates back to earlier generative works like Chomsky (1964) and Katz and Postal (1964).

interpretable, remains intact. English *wh*-elements therefore only have interpretable Q-features, with the uninterpretable Q-features of their indeterminate parts having been "checked off within a lexicon" (or at the level of "microsyntax" in the sense of Hale and Keyser (1993)). Since the Q-features of English *wh*-elements are interpretable, LF-requirement (62) coupled with the "no extra baggage" condition requires that their movement should carry along "extra baggages." This yields category movement.

Turning to Japanese-type languages, recall that the Q-features of indeterminate elements in Japanese-type languages are uninterpretable. They are therefore not subject to LF-requirement (62). Hence, the "no extra baggage" condition requires that feature movement, but not category movement, should take place. Only Q-features undergo movement.

Recall that our analysis claims that the Q-feature of an interrogative C is strong and thus uninterpretable in both English-type and Japanese-type languages (possibly universally as argued by Watanabe (1992a, 1992b)). There is therefore no difference between these types of languages concerning the property of an interrogative C. This is in contrast with Chomsky's (1993, 1995) analysis, where the language variation concerning the existence of overt wh-movement is attributed to the difference in the property of an interrogative C.

According to our analysis, the language variation resides in the way how the uninterpretable Q-feature of an interrogative C is checked off. In English-type languages, the uninterpretable Q-feature of an interrogative C can only be checked off by the Q-feature of a *wh*-element. Since the Q-feature of a *wh*-element in English-type languages is interpretable, LF-requirement (62) together with the "no extra baggage" condition requires that it should carry along an "extra baggage" for LF-convergence; the *wh*-element as a whole, but not just its Q-feature, undergoes movement. In Japanese-type languages where indeterminate elements and Q-morphemes are two independent lexical items, on the other hand, the uninterpretable Q-feature of an interrogative C is checked off by merger of the Q-morpheme *ka/no*, which has an interpretable Q-feature. After this checking operation takes place, the uninterpretable Q-feature of an indeterminate element raises to the Q-morpheme in order to be checked off. LF-requirement (62) together with "no extra baggage" condition requires that Q-feature movement should not carry any "extra baggage." This is because the Q-feature of an indeterminate element in Japanese-type languages is uninterpretable. Hence, only Q-feature rather than a *wh*-element as a whole moves. One may safely say that our analysis is a minimalist reinterpretation of Cheng's (1991) typological theory, where languages without Q-morphemes like English employ overt wh-movement while those with Q-morphemes like Chinese and Japanese do not.[25]

---

[25]There are other areas in which the difference in the property of a functional head has been used to explain variations among languages: the difference between French and English regarding overt verb movement (see, among others, Bobaljik (1995), Chomsky (1991a), Emonds (1976, 1978), Jackendoff (1972), and Pollock (1989)) and the difference among languages regarding overt object shift (see, among others, Chomsky (1993, 1995), Collins (1997), Collins and Thráinsson (1993), Holmberg (1986), Holmberg and Platzack (1995), Johnson (1991), Jonas and Bobaljik (1993), Koizumi (1993), Travis (1992), Ura (1996), Vikner (1995), and Zwart (1993)). I claim that these differences should be explained by the existence of a strong categorial feature under a functional head. Let us condier the existence/nonexistence of overt verb movement. We assume following Chomsky (1993) that while French has a strong V-feature under a functional head, English does not. In French, the strong V-feature should be checked by the V-feature of a verb. Since the latter is interpretable, LF-requirement (62) together with the "no extra baggage" condition requires that the whole category rather than just the feature

5.3.4 The Locality Effects with *Wh*-elements In-situ

With the discussion in the last subsection in mind, this subsection considers the locality restrictions on *wh*-elements in-situ in Japanese-type languages. It is shown that the Q-feature movement analysis coupled with our theory of phrase structure can account for the above-mentioned hitherto unexplained asymmetries concerning *wh*-elements in-situ in Japanese-type languages.

5.3.4.1 Argument Wh-movement and a *Wh*-argument In-situ

As extensively argued in chapter 3, the "domain barrier" effects with overt wh-movement straightforwardly follow from our theory of the composition of phrase structure. Let us look at the CNPC effects, taking (64) as an example:

(64) ?***who** do you like [books that criticize ***t***]

It was shown that since the relative clause *that criticize who* is an adjunct, it is required to be merged with the main structure postcyclically. Crucially, when we come to the stage of the derivation where the strong Q-feature of C is to be checked, the relative clause has not been merged with the main structure yet. Hence, there is no way to check the strong Q-feature of C at this stage. This violates the ICP; the derivation is canceled. The deviancy of examples like (64) follows.

Unlike overt wh-movement, however, *wh*-arguments in-situ in Japanese-type languages do not exhibit any "domain barrier" effects. Let

---

should move. Hence, a verb moves overtly in French but not in English. The existence/nonexistence of overt object shift can be explained in a similar fashion.

us consider the relative clause case of the CNPC (19) (repeated here as (65)) as an example:

(65) John-wa [**nani**-o katta hito]-o sagasite iru no
-Top what-Acc bought person-Acc looking-for Q
Lit. 'John is looking for the person who bought what'

Let us first look at how to construct the relative clause. The indeterminate element *nani* 'what' is an argument and thus required to be merged cyclically in order to check the selectional restriction feature of the verb *katta* 'bought':

(66) [nani-o katta]
what-Acc bought

Recall that the Q-feature of *nani* 'what', though uninterpretable, is optional. It is therefore not introduced into the derivation at this stage but rather added later in the derivation. Given that empty operator movement is involved in Japanese relative clauses, we assign the following structure to the relative clause (though the present argument holds regardless of whether the empty operator movement analysis of Japanese relative clauses is correct or not):[26]

(67) [*Op* [*t* nani-o katta]]
what-Acc bought

Let us turn to the main structure. The verb *sagasite iru* 'looking for' has selectional restriction features to be checked off. Those features, being uninterpretable, are checked by first combining the verb *sagasite iru* 'looking for' with *hito* 'person' and then combining *John* with *hito-o sagasite iru* 'looking for the person'. It should be noted that since merger

[26]Murasugi (1991), for example, argues that Japanese relative clauses are identified as IPs, involving base-generated empty pronouns.

of the relative clause (67) is not required by any UFF, it may not be merged with the main structure at this point. The relative clause is required to be merged postcyclically. When we come to the stage where the Q-morpheme *no* is merged with the interrogative C in order to check the Q-feature of the latter, therefore, the relative clause has not been merged with the main structure yet. The main structure and the relative clause each constitute an independent syntactic object at this stage:

(68) a. [[John-wa hito-o sagasite iru] $no_{[Q]}$]
-Top person-Acc looking-for Q

b. [*Op* [*t* nani-o katta]]
what-Acc bought

Note that since the Q-feature of the Q-morpheme *no* is interpretable, it need not enter into a checking relation unless required by some other formal feature.

We then merge the main structure with the relative clause:

(69) [[John-wa [[*Op* [*t* nani-o katta]] hito]-o sagasite iru]
-Top what-Acc bought person-Acc looking-for
$no_{[Q]}$]
Q

After this merger, we add the Q-feature to the indeterminate element *nani* 'what':

(70) [[John-wa [[*Op* [*t* $nani_{[Q]}$-o katta]] hito]-o sagasite iru] $no_{[Q]}$]

Since it is uninterpretable, it must be checked immediately in accordance with the ICP. Since the relative clause has already been merged with the main structure at this stage, the Q-feature of *nani* 'what' raises to the Q-morpheme *no* to be checked off:

(71) [[John-wa [[*Op* [*t* nani$_{[Q]}$-o katta]] hito]-o sagasite iru no$_{[Q]}$]

After the checking operation takes place, the Q-feature of *nani* 'what', being uninterpretable, is erased while that of the Q-morpheme *no*, being interpretable, remains:

(72) [[John-wa [[*Op* [*t* nani-o katta]] hito]-o sagasite iru] no$_{[Q]}$]

Hence, our analysis can correctly predict that (65), a case of the CNPC, is acceptable. The immunity of *wh*-arguments in-situ from the other "domain barriers" can be accounted for in the same way.

It is important to note that this asymmetry between overt wh-movement and a *wh*-argument in-situ counts as empirical evidence in support of our derivational account of the "domain barriers" and against the traditional approaches to the "domain barriers." Under our analysis of the "domain barriers," the "domain barrier" effects are observed with operations which take place before "domain barriers" are merged with main structures but not with those which take place after "domain barriers" are merged with main structures. In the case of overt wh-movement, "domain barriers" have not been merged with main structures when it takes place. Hence, the "domain barrier" effects are observed with overt wh-movement. In the case of *wh*-arguments in-situ, on the other hand, Q-feature movement may take place after "domain barriers" are merged with main structures. Hence, *wh*-arguments in-situ do not exhibit any "domain barrier" effects. The asymmetry between the two therefore follows from the difference in the ordering between the checking of Q-features and merger of "domain barriers" with main structures in the course of a derivation. Such an account is not available under the traditional approaches to the "domain barriers," which claim that extraction out of certain domains is prohibited. This is because those approaches claim that certain domains always count as barriers throughout derivations.

5.3.4.2 Argument and Adjunct *Wh*-elements In-situ

Unlike *wh*-arguments in-situ, *wh*-adjuncts in-situ exhibit the "domain barrier" effects. Let us consider how our analysis can account for this asymmetry, taking the relative clause case of the CNPC (22) (repeated here as (73)) as an example:

(73) *John-wa [Bill-ga **naze** Mary-ni watasita tegami]-o
-Top -Nom why -Dat gave letter-Acc
sagasite iru no
looking-for Q
Lit. 'John is looking for the letter which Bill gave to Mary why'

Let us first consider how to construct the relative clause. We construct structure (74) by checking the UFFs of the selected items in accordance with the ICP and the EP. Note that although we are assuming the empty operator movement analysis of Japanese relative clauses, the validity of the present argument holds regardless whether it is correct or not:

(74) [*Op* [Bill-ga [Mary-ni [*t* watasita]]]]
-Nom -Dat gave

It should be note that *naze* 'why', whose merger has not been required by any UFF, has not been selected from the N at this stage of the derivation.

Turning to the main structure, we construct structure (75) through checking the UFFs of the selected items:

(75) [[John-wa [tegami-o sagasite iru]] $no_{[Q]}$]

-Top letter-Acc looking-for Q

Recall that the Q-feature of the Q-morpheme *no* is interpretable and thus not subject to the ICP.

At this stage, we have two possible continuations. We either merge the relative clause with the main structure or introduce *naze* 'why' into the derivation. Since neither of these is not triggered by any UFF, we can apply either of these only on the ground of the ICP. Recall, however, that the EP requires that lexical items should be selected from an N as early as possible. According to the EP, therefore, the insertion of *naze* 'why' rather than merger of the relative clause with the main structure should be chosen at this stage. The resultant structures are as follows:

(76) a. [[John-wa [tegami-o sagasite iru]] $no_{[Q]}$]

-Top letter-Acc looking-for Q

b. [*Op* [Bill-ga [**naze**$_{\mathbf{[Q]}}$ [Mary-ni [*t* watasita]]]]]

-Nom why -Dat gave

Recall that the Q-feature of *naze* 'why' counts as intrinsic and thus becomes accessible to the computation when *naze* 'why' is selected from the N. Since the Q-feature of *naze* 'why' is uninterpretable, the ICP requires that it should be checked off immediately. Since the relative clause and the main structure each constitute an independent syntactic object at this stage, however, the Q-feature of *naze* 'why' may not raise to the Q-morpheme *no* to be checked off. There is no way to check the Q-feature. This violates the ICP and the derivation is canceled. Hence, we can correctly predict that (73) is deviant. The other "domain barrier" effects with *naze* 'why' can be accounted for in the same fashion.

It should be pointed out that the argument/adjunct asymmetry of *wh*-elements in-situ with the "domain barriers" also constitutes empirical evidence in support of our derivational notion of the "domain barriers." The asymmetry follows from the fact that while the Q-feature of a *wh*-argument in-situ may be introduced after merger of a "domain barrier" with a main structure, that of a *wh*-adjunct in-situ must be introduced before their merger. Such an account is not available under the traditional approaches where movement is never allowed to take place from within the "domain barriers" throughout derivations.[27]

5.3.4.3 The "Domain Barriers" and the Wh-island Constraint

Unlike the "domain barrier" effects, the relativized minimality effects are observed with *wh*-arguments in-situ as well as *wh*-adjuncts in-situ, as exemplified below:

---

[27]Our analysis can correctly predict that neither *itu* 'when' nor *doko* 'where' exhibits any "domain barrier" effects, as exemplified below:

(i) John-wa [Bill-ga **itu** Mary-ni watasita tegami]-o sagasite iru no
-Top -Nom when -Dat gave letter-Acc looking-for Q
Lit. 'John is looking for the letter which Bill gave to Mary when'

(ii) John-wa [Bill-ga **dokode** Mary-ni watasita tegami]-o sagasite iru no
-Top -Nom where -Dat gave letter-Acc looking-for Q
Lit. 'John is looking for the letter Bill gave to Mary where'

Recall that although these *wh*-elements are adverbials, their Q-features are optional. Hence, their Q-features may undergo movement after we merge the "domain barriers" with the main structures. The immunity of these *wh*-elements in-situ from the "domain barriers" follows.

(77) ?John-wa [Mary-ga **nani-o** katta ka dooka]
-Top -Nom what-Acc bought whether or not
siritagatte iru no
want to know Q
Lit. 'what does John want to know [whether or not Mary bought *t*]'

(78) *John-wa [Mary-ga **naze** sono hon-o katta
-Top -Nom why that book-Acc bought
ka dooka] siritagatte iru no
whether or not want-to-know Q
Lit. 'why does John want to know whether or not Mary bought that book *t*'

As argued in the previous section, under the non-movement analysis advocated by, among others, Aoun and Li (1993) and Reinhart (1992, 1993), the wh-island effects with *wh*-arguments in-situ cannot be accounted for. Since our analysis is assuming Q-feature movement, however, the wh-island effects with *wh*-elements in-situ straightforwardly follow.

Let us consider (77) as an example. We construct structure (79) through checking the UFFs of the selected items:

(79) John-wa [Mary-ga nani-o katta ka $dooka_{[Q]}$] siritagatte iru $no_{[Q]}$

Recall that the Q-feature of *nani* 'what' is optional. Hence, its Q-feature is not introduced into the derivation when *nani* 'what' is selected from the N but rather added in the course of the derivation. At this stage, we add a Q-feature to *nani* 'what', yielding the following structure:

(80) John-wa [Mary-ga nani$_{[Q]}$-o katta ka dooka$_{[Q]}$] siritagatte iru no$_{[Q]}$

Since the Q-feature of *nani* 'what' is uninterpretable, the ICP requires that it should be checked immediately by Q-feature movement. There are, however, two possible landing sites for movement of that Q-feature; the embedded Q-morpheme *ka dooka* and the matrix Q-morpheme *no*. Among these two options, we should choose the former, excluding the latter. In order to ensure this, the MLC plays a role.

Following Chomsky (1993, 1995), let us assume that Attract/Move is subject to the MLC:

(81) The Minimal Link Condition (MLC)
Raising of α to β is not allowed if there is γ which can enter into a checking relation with either α or β, where γ is closer to α than β.

Note that our definition of the MLC (81) differs from Chomsky's (1995) in the following respect. Recall that Chomsky (1995) totally eliminates the notion of Move, arguing that the traditional notion of movement should be reinterpreted as Attract-F. Under this view, the locus of the notion is completely shifted from the moved element to the target. Accordingly, Chomsky's definition of the MLC is asymmetric. A closer candidate for raising blocks the raising of α to β, but the raising of α to β over γ that contains a feature that could check α is allowed. We are assuming, on the other hand, that the notion of Move is still needed. The traditional operation of movement is reinterpreted as Attract/Move-F. Accordingly, our definition of the MLC (81) is symmetric. It prevents α from raising to β if there is γ that is a closer candidate for raising. It also blocks α from raising to β if there is γ that contains a feature that could check α.

The notion of closeness is defined as follows. We first define the notion of the smallest maximal projection:

(82) Max (α) is the smallest maximal projection including α (where α is a feature or an X^0 category).

(Chomsky 1995:299)

Based on the notion of smallest maximal projection, we define the notion of domain:

(83) The domain $\delta(\alpha)$ is the set of categories included in Max (α) that are distinct from and do not contain α.

(Chomsky 1995:299)

Based on the notion of the domain, we define the notion of minimal domain:

(84) The minimal domain Min ($\delta(\alpha)$) of α is the smallest subset K of $\delta(\alpha)$ such that for any $\gamma \in \delta(\alpha)$, some $\beta \in$ K reflexively dominates γ.

(Chomsky 1995:299)

Based on the notion of minimal domain, we finally define the notion of closeness:

(85) If β c-commands γ and γ c-commands α, then γ is closer to α than β unless γ is in the same minimal domain as (a) α or (b) β.

(adapted from Chomsky 1995)

Let us consider structure (80) again. In (80), the Q-feature of *nani* 'what' can enter into a checking relation with that of the embedded or matrix Q-morpheme. The matrix Q-morpheme c-commands the embedded Q-morpheme and the embedded Q-morpheme c-commands *nani* 'what'. Furthermore, the embedded Q-morpheme is not in the same minimal domain with either the matrix Q-morpheme or *nani* 'what'. Hence, the embedded Q-morpheme is closer to *nani* 'what' than the matrix Q-morpheme. According to the MLC (81), the Q-feature of *nani* 'what' can only raise to the embedded Q-morpheme but not to the matrix Q-morpheme. We can therefore correctly predict that the indeterminate *nani* 'what' may not be associated with the matrix Q-morpheme in (77). Although the Q-feature of *nani* 'what' may raise to the embedded Q-morpheme to be checked off, the resultant LF-representation gets an anomalous interpretation. This is because the Q-morpheme *ka dooka* 'whether or not' can only introduce a yes/no question but not a *wh*-interrogative. Hence, (77) is deviant under any interpretation. The deviancy of (88) can be accounted for in essentially the same fashion.[28]

---

[28]As pointed out in note 4, there are some speakers who find that (i) is more severely deviant on reading (ic) than on readings (ib) and (id), though (i) is not perfect on any of these readings:

(i) Tanaka-kun-wa [**dare**-ga **nani**-o tabeta ka] oboete imasu ka
Tanaka-Top who-Nom what-Acc ate Q remember Q
a. 'does Tanaka know who ate what'
b. NOT 'who is the person *x* such that Tanaka knows what *x* ate'
c. NOT 'what is the thing *x* such that Tanaka knows who ate *x*'
d. NOT 'who is the person *x*, what is the thing *y* such that Tanaka knows whether *x* ate *y*'

Under our Q-feature movement analysis, this might be accounted for in terms of the linear crossing constraint proposed by Fodor (1978). Given that the Q-feature of an indeterminate element is associated with a Q-morpheme, (i) on reading (ib) or (id) has a nesting dependency, as represented below:

(ii) a. Tanaka-kun-wa [**dare**-ga **nani**-o tabeta **ka**] oboete imasu **ka** (=(ib))
b. Tanaka-kun-wa [**dare**-ga **nani**-o tabeta **ka**] oboete imasu **ka** (=(id))

(i) on reading (ic), on the other hand, has a crossing dependency, as represented below:

(iii) Tanaka-kun-wa [**dare**-ga **nani**-o tabeta **ka**] oboete imasu **ka**

Hence, (ib-d) only violate the Wh-island Constraint. (ic), on the other hand, violates both the Wh-island Constraint and the crossing constraint. Given that cumulative

Note in passing that our analysis can correctly predict that examples like (86) are deviant due to the MLC:

(86) ?John-wa [[Mary-ga **nani**-o katta **ka dooka**]
-Top -Nom what-Acc bought whether or not
siritagatte iru hito]- sagasite iru **no**
want to know person-Acc looking-for Q
Lit. 'John is looking for the person who wants to know whether or not Mary bought what'

Recall that under Tsai (1994) and Watanabe (1992a, 1992b) where a null operator may be base-generated outside the complex NP, examples like (86) could not be ruled out. Under our analysis, on the other hand, the Q-feature of *nani* 'what' moves to be checked off. This feature movement cannot cross over the embedded Q-morpheme due to the MLC. Hence, *nani* 'what' may not be associated with the matrix Q-morpheme. Note that the most deeply embedded clause headed by the Q-morpheme *ka dooka* 'whether or not' may not be properly interpreted as a *wh*-interrogative, though the Q-feature of *nani* 'what' may raise to the Q-morpheme *ka dooka* 'whether or not' to be checked off.

Under our analysis, the asymmetry between the "domain barriers" and the Wh-island Constraint resides in the fact that unlike the former, the latter regulates movement operations whenever they may take place during derivations. For this reason, unlike the "domain barrier" effects, the *wh*-island effects are observed with *wh*-arguments in-situ as well as *wh*-adjuncts in-situ. This presents strong empirical support of our

violations of conditions induce a greater degree of deviance, the difference in acceptability between (ib-d) and (ic) follows. See Saito (1987) for a similar analysis.

locality theory where the "domain barriers" and the Wh-island Constraint are not given a unified account but two different accounts.[29]

To summarize this section, I have proposed the "overt" Q-feature movement analysis of *wh*-elements in-situ in Japanese-type languages. It was shown that "overt" Q-feature movement coupled with the ICP and the EP gives us a minimalist account of the hitherto unexplained asymmetries concerning the distribution of *wh*-elements in-situ in Japanese-type languages. The next section considers *wh*-elements in-

---

[29]Watanabe (1992a, 1992b) observes that the wh-island effects are abrogated in examples like the following:

(i) John-wa [Mary-ga **nani**-o katta ka dooka] **dare**-ni tazuneta
-Top -Nom what-Acc bought whether or not who-Dat asked
no
Q
'who did John ask *t* whether or not Mary bought what'
(Watanabe 1992b:270)

He argues that the wh-island effects are canceled if the following three conditions are met. First, there is another *wh*-element outside the wh-island. Second, the *wh*-element within the wh-island is not c-commanded by the one outside the wh-island. Third, these *wh*-elements take the same scope. He observes that there is a contrast in acceptability between (i) and (ii-iii). Unlike (i), (ii-iii) exhibit the wh-island effects:

(ii) ??John-wa [Mary-ga **nani**-o katta ka dooka] Tom-ni tazuneta no
Lit. 'what did John ask Tom whether or not Mary bought *t*'
(Watanabe 1992b:270)

(iii) ??John-wa **dare**-ni [Mary-ga **nani**-o katta ka dooka] tazuneta no
Lit. 'who did John ask *t* whether or not Mary bought what'
(Watanabe 1992b:271)

In (ii), there is no *wh*-element outside the wh-island. In (iii), the *wh*-element within the wh-island *nani* 'what' is c-commanded by the one outside the wh-island *dare* 'who'.

If Watanabe's observation is correct, this constitutes evidence against our Q-feature movement analysis of *wh*-elements in-situ in Japanese-type languages. This is because under our analysis, every *wh*-element in-situ has an uninterpretable Q-feature which is to be checked off and thus should always exhibit the wh-island effects if it is contained within a wh-island. In other words, our analysis would predict that examples (i-iii) are all deviant due to the Wh-island Constraint.

The status of the above phenomenon, however, is not entirely clear. There might be some extragrammatical factors involved in the judgments of (i-iii). In acceptable cases like (i), the *wh*-element within a wh-island gets focalized and thus may take matrix scope for this reason. In deviant cases like (ii) and (iii), the *wh*-element within a wh-island may not have focal stress. Instead, *Tom-ni* 'Tom-Dat' and *dare-ni* 'who-Dat' have focal stress in (ii) and (iii), respectively. If we destress *nani-o* 'what-Acc' in (i), (i) degrades. See Chomsky (1995) for a similar observation concerning the superiority effects in English.

situ in English-type languages. I will argue that their immunity from the "domain barriers" as well as the Wh-island Constraint straightforwardly follows from our analysis.

5.4 *Wh*-elements In-situ in English-type Languages

Unlike *wh*-elements in-situ in Japanese-type languages, those in English-type languages are not constrained by the "domain barriers" or the Wh-island Constraint. I will argue that this contrast straightforwardly follows from the fact that "overt" Q-feature movement takes place in Japanese-type languages but not in English-type languages.

Let us first consider the immunity of *wh*-elements in-situ from the "domain barrier" effects in English-type languages. Let us consider the relative clause case of the CNPC, taking (4a) (repeated here as (87)) as an example:

(87) who likes [books that criticize **who**]

Recall that the Q-feature of a *wh*-element in English-type languages counts as interpretable and thus need not enter into a checking relation unless required by some other formal feature. Although the matrix interrogative C has a strong Q-feature to be checked off, it enters into a checking relation with the matrix subject *who*. There is nothing else which can trigger movement of the Q-feature of the *wh*-element in-situ *who*. Hence, the Q-feature of the *wh*-element in-situ never undergoes movement, being exempt from the CNPC. Their immunity from the other "domain barrier" effects can be accounted for in the same way.

Let us turn to the immunity of *wh*-elements in-situ from the Wh-island Constraint in English-type languages. As mentioned above, unlike *wh*-elements in-situ in Japanese-type languages, those in English do not exhibit any wh-island effects, as exemplified by (11) (repeated here as (88)):

(88) who remembers [why John bought **what**]

In (88), the *wh*-element in-situ *what* may take either matrix or embedded scope. Recall that unlike in Japanese-type languages, the Q-feature of a *wh*-element is interpretable in English-type languages. The Q-feature of a *wh*-element therefore does not have to move to be checked off in English-type languages. In (88), since the Q-feature of the *wh*-element in-situ *what* does not undergo movement, no wh-island effects emerge.

To summarize this section, I have argued that the immunity of *wh*-elements in-situ from both the "domain barriers" and the Wh-island Constraint in English-type languages straightforwardly follows from our analysis. The asymmetry between English-type and Japanese-type languages concerning the wh-island effects with *wh*-elements in-situ is attributed to the difference in the make-up processes of *wh*-elements in these two types of languages. In both of these two types of languages, a *wh*-element is made up of an indeterminate element and a Q-morpheme. These two types of languages differ as to the level where this make-up process takes place. As mentioned above, in Japanese-type languages, an indeterminate element and a Q-morpheme are two independent lexical items. The make-up process of a *wh*-element therefore takes place in the syntactic component through movement of the uninterpretable Q-feature of a *wh*-element in-situ to a Q-morpheme. This Q-feature movement induces the wh-island effects with *wh*-elements in-situ in this type of languages. In English-type languages, on the other hand, the make-up process of a *wh*-element takes place within the lexicon through movement of the Q-feature of an indeterminate element to its unselective binder. In the syntactic component, no Q-feature movement takes place. Hence, no wh-island effects are observed with *wh*-elements in-situ in this type of languages.

5.5 Concluding Remarks

In this chapter, I have presented the asymmetries regarding the distribution of *wh*-elements in-situ which have not been given any principled account under the MP. It was shown that due to the difference in the make-up processes of *wh*-elements between English-type and Japanese-type languages, "overt" Q-feature movement takes place in the latter but not in the former. I have argued that "overt" Q-feature movement together with our theory of phrase structure accounts for the locality restrictions on *wh*-elements in-situ in Japanese-type languages. I have also argued that the immunity of *wh*-elements in-situ from all the locality conditions in English-type languages follows from the lack of "overt" Q-feature movement in those languages. The distribution of *wh*-elements in-situ therefore constitutes another empirical support in favor of our theory of the composition of phrase structure.

CHAPTER 6
RECONSTRUCTION

6.0 Introduction

The preceding chapters have considered the locality restrictions on movement as empirical evidence in support of our theory of the composition of phrase structure. It was shown that the locality restrictions on movement straightforwardly follow if the composition of phrase structure conforms to the ICP and the EP during derivations. This chapter investigates reconstruction effects with Condition C of the binding theory, variable binding, and the interpretation of *each other*, arguing that they also follow from our theory of phrase structure. If the arguments to follow are on the right track, our theory of phrase structure receives strong empirical support from the fact that its effects are observed in two totally different kinds of relations, i.e., movement and binding relations.

The organization of this chapter is as follows. Section 6.1 discusses reconstruction effects with Condition C of the binding theory. It is shown that there is an argument/adjunct asymmetry concerning reconstruction effects with Condition C of the binding theory. When R-expressions are contained within "fronted" complements, reconstruction effects emerge. When R-expressions are contained within "fronted" adjuncts, anti-reconstruction effects emerge. I will argue that this asymmetry follows from our theory of phrase structure together with the assumption that Condition C applies at LF. Section 6.2 considers reconstruction effects with variable binding. It is shown that reconstruction effects with variable binding occur when pronouns are contained within "fronted" complements but not when they are contained within "fronted" adjuncts. I will argue that this argument/adjunct asymmetry with variable binding also follows from our theory of phrase structure coupled with the assumption that variable binding relations are established at LF. Section 6.3 discusses reconstruction effects with *each other*. It is shown that reconstruction effects with *each other* emerge when *each other* is contained within a "fronted" complement but not when it is contained within a "fronted" adjunct. I will argue that given that a constraint on the interpretation of *each other* applies at LF, this argument/adjunct asymmetry with *each other* follows from our theory of phrase structure. Section 6.4 makes concluding remarks.

6.1 (Anti-)Reconstruction Effects with Condition C

This section investigates reconstruction effects with Condition C of the binding theory. As observed by Lebeaux (1988, 1991) and van Riemsdijk and Williams (1981), there is an argument/adjunct asymmetry concerning reconstruction effects with Condition C of the binding theory. When R-expressions are contained within "fronted" complements, reconstruction effects with Condition C emerge. When R-expressions are contained within "fronted" adjuncts, on the other hand, the reconstruction effects are abrogated. I will argue that this asymmetry concerning the reconstruction effects straightforwardly follows from our theory of phrase structure together with the assumption that Condition C of the binding theory applies at LF.

Before turning to the reconstruction effects, let us consider the status of the binding theory under the MP. It has been assumed in the EST framework (see, among others, Chomsky (1981)) that binding conditions apply at S-structure. Under the MP where S-structure as well as D-structure is eliminated, binding conditions should be reformulated either as interface conditions or as constraints which apply throughout derivations. Following Chomsky (1993) and Chomsky and Lasnik (1993), let us assume that binding conditions are reformulated as interface conditions which hold at LF.[1] Condition C of the binding theory can be formulated as an LF-interface condition, as shown below:[2]

(1) If α is an R-expression, interpret it as disjoint from every c-commanding phrase.

(Chomsky 1993:43)

6.1.1 An Argument/Adjunct Asymmetry with Condition C

Returning to the reconstruction effects with Condition C of the binding theory, let us first look at the following examples:

(2) a. ***he$_i$** likes those pictures of **John$_i$**

b. *Mary says that **he$_i$** feared the examination of **John$_i$**

(3) a. ***he$_i$** likes those pictures near **John$_i$**

---

[1]As discussed in chapter 2, Chomsky (1993) suggests that the binding theory might be outside of the computational system. Under this view, LF-representations only provide instructions for the conceptual-intensional system where linguistic expressions get interpretations. Since the choice between these two views, i.e., whether the binding theory stays inside or outside of the computational system, does not affect the arguments to follow, I assume for expository purposes that binding conditions hold at LF.

[2]For detailed discussion of Condition C of the binding theory, see, among others, Chomsky (1981, 1986b), Huang (1991), Lasnik (1989), and Lasnik and Uriagereka (1988). See Chomsky (1982) and Reinhart (1983) for arguments in favor of the elimination of Condition C. Lasnik (1989) presents extensive arguments against the elimination of Condition C.

b. *Mary says that **he**$_{i}$ peeked at the examination paper near **John**$_{i}$

The examples in (2-3) are all deviant due to violations of Condition C of the binding theory (1), because *John*, being an R-expression, is c-commanded by the coreferent *he*. As observed by Lebeaux (1988, 1991) and van Riemsdijk and Williams (1981), however, the divergence occurs in (4-5):

(4) a. ?*[which pictures [of **John**$_{i}$]] do you think that **he**$_{i}$ likes *t* best

b. *[which examination [of **John**$_{i}$]] did Mary say that **he**$_{i}$ feared *t*

(5) a. [which pictures [near **John**$_{i}$]] do you think that **he**$_{i}$ likes *t* best

b. [which examination paper [near **John**$_{i}$]] did Mary say that **he**$_{i}$ peeked at *t*

While *John* and *he* can be coreferential in (5a-b), they cannot be coreferential in (4a-b). The difference between (4) and (5) resides in the fact that while *John* is the complement of the noun in the former, it is within the adjunct modifying the noun in the latter. In other words, the anti-reconstruction effects are observed in (5) while the reconstruction effects are observed in (4). I will argue that this argument/adjunct asymmetry with respect to the reconstruction effects straightforwardly follows from our theory of phrase structure together with the assumption that Condition C of the binding theory applies at LF.

Let us first consider the case where R-expressions are contained within "fronted" complements, taking (4a) as an example. We first select *of*. The ICP requires that its selectional restriction feature, being uninterpretable, should be checked immediately by selecting *John* and combining *of* with *John*. The next step is to select *pictures*. Its selectional restriction feature, being uninterpretable, must be checked immediately by combining *pictures* with its complement *of John*, conforming to the ICP. We then select the D *which*. The ICP requires that its selectional restriction feature should be checked immediately by combining *which* with *pictures of John*. In this way, we construct the D^{max} *which pictures of John* (6). For expository purposes, we only pay attention to Q-features here and in the relevant structures to follow:

(6) $[_{D}{}^{max}$ which$_{[Q]}$ $[_{N}{}^{max}$ pictures $[_{P}{}^{max}$ of John]]]

As the derivation proceeds, we come to the stage where the strong Q-feature of the matrix C is to be checked:

(7) a. $[_{C}{}^{max}$ C$_{[Q]}$ $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you [think $[_{C}{}^{max}$ that $[_{T}{}^{max}$ T $[_{V}{}^{max}$ he [likes $[_{D}{}^{max}$ which$_{[Q]}$ pictures of John]]]]]]]]]

b. he

c. you

As required by the ICP, the strong Q-feature of C is checked immediately by copying *which pictures of John*, which is the minimal maximal projection containing the Q-feature required for convergence:

(8) a. $[_{C}{}^{max}$ C$_{[Q]}$ $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you [think $[_{C}{}^{max}$ that $[_{T}{}^{max}$ T $[_{V}{}^{max}$ he [likes **$[_{D}{}^{max}$ which$_{[Q]}$ pictures of John]**]]]]]]]]

b. he

c. you

d. **$[_{D}{}^{max}$ which$_{[Q]}$ pictures of John]**

Note that after the checking operation takes place, the Q-feature of C, being uninterpretable, is erased while that of the *wh*-element, being interpretable, remains intact. After merger of these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(9) [**[which pictures of John]$_j$** [C [you$_k$ [T [you$_k$ [think [that [he$_l$ [T [he$_l$ [likes **[which pictures of John]$_j$**]]]]]]]]]]

Recall that we are assuming following Chomsky (1993) that there is an LF-operation for construction of an operator-variable structure. According to that operation, all but the operator phrase must delete in the head position of a chain. In the tail position of a chain, on the other hand, nothing but the operator phrase must delete. Among the chains created in this derivation, the following is relevant to the present discussion:

(10) CH = (*which pictures of John*, *which pictures of John*)

If we apply the operation for construction of an operator-variable structure to this chain, we get the following two LF-representations depending on what counts as an operator:

(11) a. **[which *x*]** [you think that he likes **[*x* pictures of John]**]

b. **[which *x*, *x* pictures of John]** [he liked ***x***]

If only the D *which* counts as an operator, *pictures of John* deletes in the head position of the chain while *which* deletes in the tail position of the chain, resulting in (11a). If the D^{max} *which pictures of John* as a whole counts as an operator, on the other hand, nothing deletes in the head position of the chain while *which pictures of John* deletes in the tail position of the chain, resulting in (11b).

Chomsky (1993) proposes the preference principle for reconstruction, which states that the restriction in the operator position must be minimized unless it would make a derivation crash. The preference principle compares (11a-b) and requires that only the D *which*, but not the whole D^{max} *which pictures of John*, should remain in the operator position. Since both (11a) and (11b) converge, (11a) rather than (11b) is chosen as the LF representation. In (11a), since *John* is c-commanded by *he*, Condition C of the binding theory (1) requires that the former should be disjoint in reference from the latter. Hence, the deviancy of (4a), where *John* and *he* are coreferential, follows.

Let us next consider the case where R-expressions are contained within "fronted" adjuncts, taking (5a) as an example. Unlike in (4a), *John* and *he* may be coreferential in (5a). Let us consider how we construct (5a). We first select the D *which*. The ICP requires that the selectional restriction feature of D, being uninterpretable, should be checked immediately by selecting the noun *pictures* and combining *which* with *pictures*, as shown below:

(12) $[_{D}{}^{max}$ which$_{[Q]}$ $[_{N}{}^{max}$ pictures]]

According to the EP, the next step must be to select the verb *like*. When the verb *like* is selected, the ICP requires that its selectional restriction feature should be checked immediately by combining the verb *like* with the D^{max} *which picture*. The resulting structure is as follows:

(13) $[_{V}{}^{max}$ likes $[_{D}{}^{max}$ which$_{[Q]}$ $[_{N}{}^{max}$ pictures]]]

It should be noted that the adjunct P^{max} *near John*, whose insertion is not triggered by any UFF, has not been merged with the D^{max} *which pictures* at this point of the derivation. In other words, the main structure (13) and the adjunct P^{max} *near John* each constitute an independent syntactic object at this point.

As the derivation proceeds, we come to the stage where the strong Q-feature of the matrix C is to be checked, as shown below:

(14) a. $[_{C^{max}}\ C_{[Q]}\ [_{T^{max}}\ T\ [_{V^{max}}$ you [think $[_{C^{max}}$ that $[_{T^{max}}$ T $[_{V^{max}}$ he [likes $[_{D^{max}}$ which$_{[Q]}$ pictures]]]]]]]]

b. he

c. you

d. $[_{P^{max}}$ near John]

The strong Q-feature of C is checked immediately by copying *which pictures* in accordance with the ICP, as shown below:

(15) a. $[_{C^{max}}\ C_{[Q]}\ [_{T^{max}}\ T\ [_{V^{max}}$ you [think $[_{C^{max}}$ that $[_{T^{max}}$ T $[_{V^{max}}$ he [likes **$[_{D^{max}}$ which$_{[Q]}$ pictures]**]]]]]]]

b. he

c. you

d. $[_{P^{max}}$ near John]

e. **$[_{D^{max}}$ which$_{[Q]}$ pictures]**

These syntactic objects are merged together, resulting in the following structure. Here, we ignore all formal features including Q-features:

(16) [[**[which pictures]$_j$** near John] [C [you$_k$ [T [you$_k$ [think [that [he$_l$ [T [he$_l$ [likes **[which pictures]$_j$**]]]]]]]]]]]

Note that the adjunct *near John* is merged with the D^{max} *which pictures* in the Spec of C^{max} but not the one in its original position.

Among the chains created in this derivation, the following chain is relevant to the discussion:

(17) CH = (*which pictures*, *which pictures*)

Recall that the terms that are identical in constitution but positionally distinct from each other form a chain in the LF-component. In the present derivation, *which pictures* is introduced a second time in the syntactic object through Copy. The two occurrences of *which pictures* are therefore identical in constitution but positionally distinct from each other, forming the chain (17) in the LF-component. It is important to note that the adjunct *near John*, which is merged with *which pictures* after the latter undergoes Copy, is not part of the chain. After application of the LF-operation for construction of an operator-variable structure, we get the following two LF-representations:

(18) a. [**[which *x*]** near John] [you think that he likes **[*x* pictures]**]

b. [**[which *x*, *x* pictures]** near John] [he likes ***x***]

Between these LF-representations, the preference principle chooses (18a), excluding (18b). In (18a), *John* is not c-commanded by *he*. There is no violation of Condition C of the binding theory (1) even if *John* is in the coreference relation with *he*. We can correctly predict that (5a) is acceptable.

To summarize this subsection, our theory of phrase structure requires that the complements of nouns should be merged with main structures cyclically and thus reconstructed to their original positions. It then follows that examples like (4a), where R-expressions are contained within "fronted" complements, show the reconstruction effects with respect to Condition C of the binding theory. The adjuncts modifying nouns, on the other hand, should be merged with main structures postcyclically. There is no way to reconstruct the adjuncts, since they undergo direct insertion to their surface positions and thus do not have any "original" positions. We can therefore correctly predict that examples like (5a), where R-expressions are contained within "fronted" adjuncts, exhibit the anti-reconstruction effects with respect to Condition C of the binding theory.

6.1.2 Complex NPs and the Anti-Reconstruction Effects

As discussed in the previous subsection, the anti-reconstruction effects with Condition C of the binding theory are only observed when R-expressions are contained within "fronted" adjuncts. Recall that adjuncts are those which are forced to be merged with main structures postcyclically. If this analysis is on the right track, we should expect that complex NPs, whose appositive or relative clauses are required to be merged with main structures postcyclically, also exhibit the anti-reconstruction effects. This prediction is borne out, as shown below:[3]

---

[3]Lebeaux (1988, 1991) presents the following examples, claiming that the contrast in acceptability between (i) and (ii) suggests that the non-relative clause cases of the complex NPs should be interpreted as noun-complement structures rather than noun-appositive structures and thus merged with the main structures cyclically:

(i) [which claim that John$_i$ made]$_j$ did he$_i$ later deny t_j

(ii) *[whose claim that John$_i$ likes Mary]$_j$ did he$_i$ deny t_j

Under his analysis, the clause following *claim* in (ii) would be inserted cyclically before application of wh-movement. The LF-representation of (ii) would be as follows:

(iii) [which x, x a person] [he denied [x's claim that John likes Mary]]

In (iii), the R-expression *John* is c-commanded by *he*. Lebeaux's analysis would correctly predict that the former cannot be coreferential with the latter.

As Lebeaux himself points out in the footnote, however, the contrast between (i) and (ii) disappears if we replace *whose* by *which* in (ii). He admits that the deviancy of (ii) may not be due to the fact that the noun *claim* and its following clause constitutes a noun-complement structure but due to the existence of the genitive *wh*-phrase *whose*. It should also be pointed out that there are some speakers who cannot see any contrast between (i) and (ii).

Another possible reason for the deviancy of examples like (ii) for some speakers is that there is no difference in phonetic shape between the noun *claim* and the verb *claim*. The noun *claim* and the verb *claim* are derived one from the other by conversion or zero-derivation. It is reasonable to claim that because of their having the same phonetic

(19) a. [which piece of evidence [that **John$_i$** discovered]] was **he$_i$** willing to discuss *t*

b. [which piece of evidence [that **John$_i$** was asleep]] was **he$_i$** willing to discuss *t*

(Watanabe 1995:290)

(20) a. [which evidence [that **John$_i$** presented to court in order to deceive the attorney]] did **he$_i$** later ignore *t*

b. [which evidence [that **John$_i$** quarreled with his wife at a boathouse]] did **he$_i$** later ignore *t* in court

(21) a. [which explanation [that **John$_i$** offered based on his careful observations]] did **he$_i$** later deny *t*

b. [whose explanation [that **John$_i$** was temporarily mad at his wife]] did **he$_i$** deny *t* in court

(22) a. [which belief [that **John$_i$** discovered]] was **he$_i$** willing to discuss *t*

b. [whose belief [that **John$_i$** was asleep]] was **he$_i$** willing to discuss *t*

In the (a) examples of (19-22), the R-expression *John* is contained within the "fronted" relative clause. In the (b) examples of (19-22), on the other hand, the R-expression *John* is contained within the "fronted" appositive clause of the noun. In all of these examples, *John* can be coreferential with *he*. In other words, the anti-reconstruction effects are observed.

Let us consider (19a-b) as examples. Recall that the relative clause *that John discovered* in (19a) and the appositive clause *that John*

---

form, some speakers interpret the noun *claim* and its following clause as a noun-complement structure by analogy with the verb *claim* and its complement.

was asleep in (19b) are both adjuncts and thus merged with the *wh*-phrases after the latter undergoes Copy. Hence, the LF-representations of (19a-b), which are derived by the operation for construction of an operator-variable structure and chosen by the preference principle, are as follows:

(23) a. [**[which *x*]** that John discovered] [was he willing to discuss **[*x* piece of evidence]**]

b. [**[which *x*]** that John was asleep] [was he willing [to discuss **[*x* piece of evidence]**]

In (23a-b), *John* is not c-commanded by *he*; there is no violation of Condition C of the binding theory even if *John* is in the coreferential relation with *he*. Hence, (19a-b) are acceptable under the coreferential readings between *John* and *he*. The examples in (20-22) can be accounted for in a similar fashion.

To summarize this section, our theory of phrase structure together with the assumption that Condition C of the binding theory applies at LF gives us a minimalist account of the argument/adjunct asymmetry with the (anti-)reconstruction effects with Condition C of the binding theory.

6.2 Reconstruction Effects with Variable Binding

This section considers reconstruction effects with variable binding. It is pointed out that reconstruction effects with variable binding are observed if pronouns are contained within "fronted" complements. If pronouns are contained within "fronted" adjuncts, on the other hand, no reconstruction effects with variable binding are observed. I will argue that such an argument/adjunct asymmetry concerning reconstruction effects with variable binding straightforwardly follows from our theory of phrase structure coupled with the assumption that variable binding relations are established at LF.

6.2.1 An Argument/Adjunct Asymmetry with Variable Binding

Pronouns may take not only referential phrases but also quantificational phrases as their antecedent. In the latter situation, pronouns are used as bound variables, the referential values of which vary with the value-assignment of their quantificational antecedents, as shown below:

(24) **everyone$_i$** loves **his$_i$** mother

In (24), the pronoun *he* may be interpreted as a variable bound by the quantificational phrase *everyone*, as represented in (25):[4]

(25) for every *x*, *x* a person, *x* loves *x*'s mother

Among conditions that regulate bound pronouns, (26) is relevant to the following discussion:

(26) Pronouns can be interpreted as bound variables only if they are c-commanded by quantificational phrases at LF.

This condition is one of the necessary conditions for pronouns to be used as bound variables. In order to see why we need this condition, let us first consider the following examples:

(27) a. [$_{D^{max}}$ a report [$_{P^{max}}$ about [$_{D^{max}}$ **every student**]]] was sent out

[4]See, among others, Chomsky (1981), Koopman and Sportiche (1982/1983), Higginbotham (1980), May (1977, 1985), and Montelbetti (1984) for detailed discussion of bound pronouns.

b. [$_{D}$max this report [$_{P}$max about [$_{D}$max **every student**]]] was sent out

(27a) is ambiguous between what May (1977, 1985) calls an inversely-linked and internal scope readings. In the former reading, the quantificational phrase *every student* has scope over the entire sentence, as paraphrased in (28a). In the latter reading, the quantificational phrase *every student* only has scope over the D^{max} containing it, as paraphrased in (28b):

(28) a. for every student there is a report about him which was sent out

b. a report which contains information about every student was sent out

(27b), on the other hand, only has an internal scope reading due to the specificity condition effect caused by the determiner *this*, meaning that this report which contains information about every student was sent out. In other words, the quantificational phrase *every student* may not take scope over the entire sentence in (27b).

With the above discussion in mind, let us consider the following examples:

(29) a. [$_{D}$max a report card [$_{P}$max about **[$_{D}$max every student]$_i$**]] was sent to **his$_i$** parents

b. *[$_{D}$max this report card [$_{P}$max about **[$_{D}$max every student]$_i$**]] was sent to **his$_i$** parents

In (29a), the pronoun *his* may be interpreted as a variable bound by the quantificational phrase *every student*. In (29b), on the other hand, the pronoun *his* may not be interpreted as a variable bound by the quantificational phrase *every student*. This contrast in the availability of bound variable readings follows from (26). As shown above, the quantificational phrase *every student* may have scope over the entire sentence in (29a) whereas it may not have scope over the entire sentence due to the Specificity Condition in (29b). Given that scope relations are established by c-command relations at LF, the quantificational phrase *every student*, which has scope over the entire sentence, c-commands the pronoun *his* at LF in (29a). Hence, the pronoun *his* may be interpreted as a variable bound by the quantificational phrase *every student*. In (29b), on the other hand, the pronoun *his* is not c-commanded by the quantificational phrase *every student* at LF. This is because *every student* only has scope over the D^{max} but not over the entire sentence due to the Specificity Condition. The pronoun *his* therefore may not be interpreted as a bound variable.

There is, however, a set of examples which apparently does not conform to (26):

(30) a. [which pictures [of **his$_i$** parents]] do you think that **every man$_i$** likes *t*

b. [which attack [on **his$_i$** country]] do you think that **every American$_i$** still remembers *t*

c. [which story [about **him$_i$**]] do you think that **every man$_i$** most often hears *t* from his parents

This type of construction is extensively discussed by Engdahl (1980, 1986). In (30), although the pronouns are moved out of the scope domain of the quantificational phrases, the pronouns can be properly interpreted as bound variables. The phrases which contain the pronouns are interpreted as if they were in their trace sites. In other words, the reconstruction effects with variable binding are observed.

Such reconstruction effects with variable binding, however, are not always available, as shown below:[5]

(31) a. *?[which criticism [because of **$\mathbf{his_i}$** scandal]] do you think that **every $\mathbf{congressman_i}$** remembers *t*

b. *?[which book [around **$\mathbf{him_i}$**]] do you think that **every $\mathbf{man_i}$** most often read *t*

c. ?[which bus stop [near **$\mathbf{his_i}$** house]] do you think that **every $\mathbf{resident_i}$** most often uses *t*

d. *?[[which food [after **$\mathbf{his_i}$** fight]] do you think that **every $\mathbf{boxer_i}$** loves to eat *t*

In (31), the bound variable readings of *he/his* are not available. The difference between the examples in (30) and those in (31) resides in the fact that while the pronouns are contained within the complements in the former, they are contained within the adjuncts in the latter. There exists an argument/adjunct asymmetry concerning the reconstruction effects with variable binding. I will argue that this argument/adjunct asymmetry follows from our theory of phrase structure coupled with the assumption that variable binding relations are established at LF.

Let us first consider the reconstruction effects observed in (30), taking (30a) as an example. In order to construct *which pictures of his*

---

[5]Lebeaux (1991) claims that examples like (31) are not deviant. Based on such observations, he argues that while complements are always inserted cyclically, adjuncts are inserted either cyclically or postcyclically. Contrary to Lebeaux's observations, there is a clear contrast in acceptability between examples like (30) and those in (31), though the degree of deviance of the latter varies among speakers. I therefore claim that the argument/adjunct asymmetry concerning the reconstruction effects with variable binding really exists.

parents, we first select the D *his*. When the D *his* is selected, the ICP requires that its selectional restriction feature, being uninterpretable, should be checked immediately by selecting the noun *parents* and combining *his* with *parents*. Next, the preposition *of* is selected and its selectional restriction feature is checked immediately by combining *of* with *his parents* in accordance with the ICP. Then, the noun *pictures* is selected and its selectional restriction feature is checked by combining *pictures* with *of his parents*. Finally, the determiner *which* is selected and its selectional restriction feature is checked by merger of *which* with *pictures of his parents*. The resultant structure is as below:

(32) $[_{D}{}^{max}$ which$_{[Q]}$ $[_{N}{}^{max}$ pictures $[_{P}{}^{max}$ of $[_{D}{}^{max}$ his $[_{N}{}^{max}$ parents]]]]]

As the derivation proceeds through checking the UFFs of the selected items, we come to the stage where the strong Q-feature of the matrix C is to be checked:

(33) a. $[_{C}{}^{max}$ C$_{[Q]}$ $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you [think $[_{C}{}^{max}$ that $[_{T}{}^{max}$ T $[_{V}{}^{max}$ every man [likes $[_{D}{}^{max}$ which$_{[Q]}$ pictures of his parents]]]]]]]]]

b. every man

c. you

The strong Q-feature is checked by copying *which pictures of his parents*, which is the minimal maximal projection containing the Q-feature required for convergence, as shown below:

(34) a. $[_{C}{}^{max}$ C $[_{T}{}^{max}$ T $[_{V}{}^{max}$ you [think $[_{C}{}^{max}$ that $[_{T}{}^{max}$ T $[_{V}{}^{max}$ every man [likes **$[_{D}{}^{max}$ which$_{[Q]}$ picture of his parents]**]]]]]]]]

b. every man

c. you

d. **[$_{\mathrm{D}}$$^{\mathrm{max}}$ which$_{\mathrm{[Q]}}$ pictures of his parents]**

After merger of these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(35) [**[which pictures of his parents]$_j$** [C [you$_k$ [T [you$_k$ [think [that [every man$_l$ [T [every man$_l$ [likes **[which pictures of his parents]$_j$**]]]]]]]]]]]

Among the chains created by this derivation, the following is relevant to the present discussion:

(36) CH = (*which pictures of his parents*, *which pictures of his parents*)

We apply the operation for construction of an operator-variable structure to this chain. The following two LF-representation are derived:

(37) a. **[which *x*]** [you think that every man likes **[*x* pictures of his parents]**]

b. **[which *x*, *x* pictures of his parents]** [you think that every man likes ***x***]

Between these two representations, the preference principle, which requires the operator position to be minimized, chooses (37a), excluding (37b). In (37a), since the pronoun *his* is c-commanded by the quantificational phrase *every man*, the former can be properly interpreted as a variable bound by the latter. All the other examples in (30) can be accounted for in a similar fashion.

Turning to the unavailability of bound variable readings in (31), let us consider (31a) (repeated here as (38)) as an example:

(38) *?[which criticism [because of **his$_i$** scandal]] do you think that **every congressman$_i$** remembers *t*

Since the P^{max} *because of his scandal* is an adjunct, our theory of phrase structure requires that it should be merged postcyclically. Especially, when we come to the stage where the strong Q-feature of the matrix C is to be checked, the P^{max} *because of his scandal* has not been merged with the D^{max} *which criticism*, as shown below:

(39) a. [$_{C}$max $C_{[Q]}$ [$_{T}$max T [$_{V}$max you [think [$_{C}$max that [$_{T}$max T [$_{V}$max every congressman [remembers [$_{D}$max which$_{[Q]}$ criticism]]]]]]]]]

b. every congressman

c. you

d. [$_{P}$max because of his scandal]

Hence, the strong Q-feature is checked by copying *which criticism*, which is the minimal maximal projection containing the Q-feature required for convergence, as shown below:

(40) a. [$_{C}$max C [$_{T}$max T [$_{V}$max you [think [$_{C}$max that [$_{T}$max T [$_{V}$max every congressman [remembers **[$_{D}$max which$_{[Q]}$ criticism]**]]]]]]]]

b. every congressman

c. you

d. [$_{P}$max because of his scandal]

e. **[$_{D}$max which$_{[Q]}$ criticism]**

After merger of these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(41) [[**[which criticism]$_j$** because of his scandal] [C [you$_k$ [T [you$_k$ [think [that [every congressman$_l$ [T [every congressman$_l$ [remembers **[which criticism]$_j$**]]]]]]]]]]

Among the chains created by this derivation, the following is relevant to the present discussion:

(42) CH = (*which criticism*, *which criticism*)

It should be noted that the adjunct *because of his scandal*, which is merged with *which criticism* after the latter undergoes Copy, is not part of the chain. We then apply the operation for construction of an operator-variable structure and get the following two LF-representations:

(43) a. [**[which *x*]** because of his scandal] [you think that every congressman remembers **[*x* criticism]**]

b. [**[which *x*, *x* criticism]** because of his scandal] [you think that every congressman remembers ***x***]

Between these two representations, the preference principle chooses (43a), excluding (43b). In LF-representation (43a), the pronoun *his* is not c-commanded by the quantificational phrase *every congressman*. Hence, we can correctly predict that *his* cannot be interpreted as a variable bound by *every congressman*. The other examples in (31) can also be accounted for in a similar way. Hence, our theory of phrase structure together with the assumption that variable binding relations are established at LF can account for the argument/adjunct asymmetry concerning the reconstruction effects with variable binding.

It is important to note that Lebeaux's (1988, 1991) theory of phrase structure cannot account for this argument/adjunct asymmetry concerning the reconstruction effects with variable binding. Specifically, his theory would wrongly predict that bound variable readings are available in examples like (31), where the pronouns are contained within the "fronted" adjuncts. Let us consider (38) again as an example. Recall that under Lebeaux's theory, adjuncts are introduced either cyclically or postcyclically. If the adjunct *because of his scandal* is introduced cyclically, we get the LF-representations in (44):

(44) a. **[which *x*]** [you think that every congressman remembers **[*x* criticism because of his scandal]**]

b. **[which *x*, *x* criticism because of his scandal]** [you think that every congressman remembers ***x***]

If the adjunct *because of his scandal* is introduced postcyclically, the LF-representations in (45) are derived:

(45) a. [**[which *x*]** because of his scandal] [you think that every congressman remembers **[*x* criticism]**]

b. [**[which *x*, *x* criticism]** because of his scandal] [you think that every congressman remembers ***x***]

Among these four LF-representations, the preference principle chooses (44a) and (45a), where the operator positions are minimized, and excludes (44b) and (45b). Although the pronoun *his* is not c-commanded by the quantificational phrase *every congressman* in (45a), the former is c-commanded by the latter in (44a). Hence, Lebeaux's theory would predict that the pronoun *his* may be interpreted as a variable bound by the quantificational phrase *every congressman*, contrary to fact.

6.2.2 Complex NPs and the Reconstruction Effects

If the analysis presented above is correct, we should expect that bound variable readings are not available when pronouns are contained within "fronted" complex NPs. This is because the appositive and relative clauses within complex NPs are adjuncts and thus forced to be merged with main structures postcyclically. This prediction is borne out:[6]

(46) Relative Clauses

a. *?[which evidence [that **he$_i$** presented in court]] do you think that **every congressman$_i$** lied about *t*

b. ??[which evidence [that **his$_i$** attorney presented in court]] do you think that **every congressman$_i$** still remembers *t*

c. *?[which claim [that **he$_i$** made in court]] do you think that **every congressman$_i$** lied about *t*

d. ??[which claim [that **his$_i$** attorney made in court]] do you think that **every congressman$_i$** still remembers *t*

(47) Non-Relative Complex NPs

a. *?[which evidence [that **he$_i$** received a bribe from the company]] do you think that **every congressman$_i$** will ignore *t*

---

[6]Lebeaux (1991) claims that the complex NP cases like (46-47) are not deviant. Contrary to Lebeaux's observations, there is a clear contrast in acceptability between the complement cases like examples in (30) and the complex NP cases like those in (46-47), though the degree of deviance of the latter varies among speakers.

b. ??[which evidence [that **his$_i$** secretary received a bribe from the company]] do you think that **every congressman$_i$** will ignore *t*

c. *?[which story [that **he$_i$** kills the monster]] do you think that **every student$_i$** likes *t* best

d. ?*[which story [that **his$_i$** friend kills the monster]] do you think that **every student$_i$** likes *t* best

Let us consider (46a) and (47a) as examples. The relative clause *that he presented in court* in (46a) and the appositive clause *that he received a bribe from the company* in (47a) are both adjuncts and thus required to be merged with the main structures postcyclically. Especially, when we come to the stage of the derivations where the strong Q-feature of C is to be checked, those clauses have not been merged with the main structures. They are merged with the *wh*-phrases after the latter undergoes Copy to check the Q-feature of C. LF-representations (48) and (49) are assigned to (46a) and (47a), respectively, after application of the operation for construction of an operator-variable structure and the preference principle:

(48) [**[which *x*]** that he presented in court] [you think that every congressman lied about **[*x* evidence]**]

(49) [**[which *x*]** that he received a bribe from the company] [you think that every congressman will ignore **[*x* evidence]**]

In neither (48) nor (49), the pronoun *he* is c-commanded by the quantificational phrase *every congressman*. Hence, *he* cannot be interpreted as a variable bound by *every congressman*. The other examples in (46) and (47) can be accounted for in the same fashion.

To recapitulate this section, our theory of phrase structure together with the assumption that variable binding relations are established at LF can account for the argument/adjunct asymmetry concerning the reconstruction effects with variable binding. It was also shown that this asymmetry constitutes evidence in favor of our theory, where adjuncts are forced to be merged postcyclically, and against Lebeaux's (1988, 1991) theory, where adjuncts are merged either cyclically or postcyclically.

6.3 Reconstruction Effects with *Each Other*

This section considers reconstruction effects with *each other*. It is shown that there is an argument/adjunct asymmetry concerning reconstruction effects with *each other*. When *each other* is contained within "fronted" complements, the reconstruction effects are observed. When *each other* is contained within "fronted" adjuncts, on the other hand, the reconstruction effects are abrogated. I will argue that this argument/adjunct asymmetry straightforwardly follows from our theory of phrase structure together with the assumption that the c-command requirement at LF constitutes a necessary condition for the licensing of *each other*.

Before we come to the reconstruction effects, let us consider the interpretation of *each other*. There have been three fundamental approaches to *each other* in generative grammar, i.e., the binding theory approach, the linking theory approach, and the LF-movement approach. The binding theory approach is explored by, among others, Chomsky (1981) and Huang (1983). The linking theory approach was advocated by, among others, Higginbotham (1983) and Montalbetti (1984). The LF-movement approach was explored by, among others, Chomsky (1986b), Lebeaux (1983), and Pica (1987). It is beyond the scope of this study to consider which approach to *each other* should be preferred over the others. For the purpose of the present discussion, it is sufficient to claim that *each other* is subject to the following condition:

(50) *Each other* must have a c-commanding antecedent in a certain domain for its proper interpretation.

Every approach assumes this condition as a necessary condition for the proper interpretation of *each other*, though its exact formulations vary among them. This condition is responsible for ruling out examples like (51), where *each other* is not c-commanded by its antecedent *them* at any point of the derivation:

(51) ***each other$_i$** thinks that John admire **them$_i$**

The discussion to follow assumes that condition (50) applies at LF.

Returning to the reconstruction effects with *each other*, let us first consider the following examples:

(52) **they$_i$** saw [$_D$max pictures [$_P$max of **each other$_i$**]]

(53) a. **they$_i$** were watching [$_D$max the bags [$_P$max around **each other$_i$**]] at the airport

b. **they$_i$** were looking at [$_D$max the girls [$_P$max near **each other$_i$**]] on the beach

In (52), *each other* appears within the complement P^{max} and can take *they* as its antecedent. In (53a-b), on the other hand, *each other* is in the adjunct P^{max}. Although the judgments vary, there are some speakers who accept examples like (53).

If we "front" the D^{max} containing *each other* by wh-movement, however, an argument/adjunct asymmetry emerges:

(54) [which pictures [of **each other$_i$**]] did **they$_i$** see *t*

(55) a. ?[which of the bags [around **each other$_i$**]] were **they$_i$** watching *t* at the airport

b. ?[which of the girls [near **each other$_i$**]] were **they$_i$** looking at *t* on the beach

(54) prima facie violates condition (50), since *each other* is moved out of the c-command domain of its antecedent *they*. Nonetheless, (54) is acceptable. In (54), the phrase which contains *each other* is interpreted as if it were in its trace site. In other words, the reconstruction effects concerning the interpretation of *each other* are observed. (55a-b), on the other hand, are deviant even for the speakers who accept (53a-b). In (55), the reconstruction effects are abrogated. The difference between (54) and (55) resides in the fact that *each other* is contained within the complement of the noun in the former while it is contained within the adjunct in the latter. I will argue that this argument/adjunct asymmetry concerning the reconstruction effects with *each other* straightforwardly follows from our theory of phrase structure.

Let us first consider (54). According to our theory of phrase structure, *of each other* in (54), being an argument, is required to be merged with the main structure in a cyclic manner. Especially, it has already been merged with *which pictures* when the strong Q-feature of C is to be checked, as shown below:

(56) a. [$_{C^{max}}$ $C_{[Q]}$ [$_{T^{max}}$ T [$_{V^{max}}$ they [see [$_{D^{max}}$ $which_{[Q]}$ pictures of each other]]]]]

b. they

Since *which pictures of each other* is the minimal maximal projection containing the Q-feature required for convergence, the strong Q-feature of C is checked by copying *which pictures of each other*:

(57) a. [$_{C}$max C [$_{T}$max T [$_{V}$max they [see **[$_{D}$max which$_{[Q]}$ pictures of each other]**]]]]

b. they

c. **[$_{D}$max which$_{[Q]}$ pictures of each other]**

After merger of these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(58) [**[which pictures of each other]$_{j}$** [C [they$_{k}$ [T [they$_{k}$ [see **[which pictures of each other]$_{j}$**]]]]]]

Among the chains created by this derivation, (59) is relevant to the present discussion:

(59) CH = (*which pictures of each other*, *which pictures of each other*)

We then apply the operation for construction of an operator-variable structure. Among the LF-representations generated by this operation, the preference principle chooses (60), where the operator position is minimized:

(60) **[which *x*]** [they see **[*x* pictures of each other]**]

In (60), *each other* is c-commanded by its antecedent *they* and thus properly interpreted. Hence, we can correctly predict that the reconstruction effects with *each other* occur in (54).

Let us next consider the unavailability of the reconstruction effects, taking (55a) (repeated here as (61)) as an example:

(61) ?[which of the bags [around **each other$_i$**]] were **they$_i$** watching *t* at the airport

According to our theory of phrase structure, *around each other* in (61), being an adjunct, must be merged with the main structure postcyclically. Especially, at the stage when the strong Q-feature of C is to be checked, the adjunct *around each other* has not been merged with *which of the bags*:

(62) a. [$_C$max C$_{[Q]}$ [$_T$max T [they were watching [$_D$max which$_{[Q]}$ of the bags]]]]

b. they

c. [$_P$max at the airport]

d. [$_P$max around each other]

The strong Q-feature of C is therefore checked by copying *which of the bags*:

(63) a. [$_C$max C [$_T$max T [they were watching **[$_D$max which$_{[Q]}$ of the bags]**]]]

b. they

c. [$_P$max at the airport]

d. [$_P$max around each other]

e. **[$_D$max which$_{[Q]}$ of the bags]**

After merger of these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

(64) [[**[which of the bags]$_j$** [around each other]] [C [they$_k$ [T [they$_k$ [were watching **[which of the bags]$_j$** at the airport]]]]]]

Among the chains created in this derivation, (65) is relevant to the present discussion:

(65) CH = (*which of the bags*, *which of the bags*)

Note that the adjunct *around each other* is not part of the chain, since it is merged with *which of the bags* after the latter undergoes Copy to check the strong Q-feature of C. The operation for construction of an operator-variable structure applies to (65). Among the representations derived by this operation, the preference principle chooses (66):

(66) [**[which *x*]** around each other] [they were watching **[*x* of the bags]** at the airport]

In (66), *each other* is not c-commanded by its antecedent *they*; this violates condition (50). Hence, we can correctly predict that examples like (61) are deviant.

6.4 Concluding Remarks

This chapter has considered the argument/adjunct asymmetries concerning the reconstruction effects with Condition C of the binding theory, variable binding, and the interpretation of *each other*. It was shown that the asymmetries straightforwardly follow from our theory of the composition of phrase structure together with the assumption that binding relations are established at LF. Recall that the preceding chapters have extensively argued that the locality restrictions on movement follow from our theory of phrase structure. Our theory of phrase structure therefore receives strong empirical support from the fact that its effects are observed in the two totally different kinds of relations, i.e., movement and binding relations.

REFERENCES

Abney, Steven (1987). *The English noun phrase in its sentential aspect*. Doctoral dissertation, MIT.

Allwood, Jens (1976). The complex NP constraint as a non-universal rule and some semantic factors influencing the acceptability of Swedishsentences which violate the CNPC. In J. Stillings (ed.), *University of Massachusetts occasional papers in linguistics II.* Amherst: University of Massachusetts, Amherst.

Aoun, Joseph (1981). *The formal nature of anaphoric relations*. Doctoral dissertation, MIT.

Aoun, Joseph (1985a). *A grammar of anaphora.* Cambridge, MA: MIT Press.

Aoun, Joseph (1985b). *Generalized binding*. Dordrecht: Foris.

Aoun, Joseph, Norbert Hornstein, and Dominique Sportiche (1981). Some aspects of wide scope quantification. *Journal of Linguistic Research*, *1*, 69-95.

Aoun, Joseph and Yen-hu Audrey Li (1993). Wh-elements in situ: Syntax or LF? *Linguistic Inquiry*, *24*, 199-238.

Authier, J.-Marc (1991). V-governed expletives, case theory, and the projection principle. *Linguistic Inquiry*, *22*, 721-740.

Authier, J.-Marc (1992). Iterated CPs and embedded topicalization. *Linguistic Inquiry*, *23*, 329-336.

Baker, Chris (1970). Notes on the description of English questions: The role of an abstract question morpheme. *Foundations of language*, *6*, 197-219.

Baker, Mark (1988). *Incorporation*. Chicago: University of Chicago Press.

Baker, Mark (1996). *The polysynthesis parameter*. Oxford: Oxford University Press.

Battistella, Edwin (1989). Chinese reflexivization: A movement to INFL approach. *Linguistics*, *27*, 987-1012.

Belletti, Adriana and Luigi Rizzi (1981). The syntax of *ne*: Some theoretical implications. *The Linguistic Review*, *1*, 117-154.

Bobaljik, Jonathan David (1995a). In terms of merge: Copy and head movement. In Rob Pensalfini and Hiroyuki Ura (eds.), *Papers on minimalist syntax: MIT working papers in linguistics 27* (pp. 41- 64). Cambridge, MA: MITWPL.

Bobaljik, Jonathan David (1995b). *Morphosyntax: The syntax of verbal inflection*. Doctoral dissertation, MIT.

Boskovic, Zeljko (1995). Case properties of clauses and the greed principle. *Studia Linguistica*, *49*, 32-53.

Boskovic, Zeljko and Daiko Takahashi (1995). *Scrambling and last resort*. Unpublished manuscript, University of Connecticut and Tohoku University.

Branigan, Phil and Chris Collins (1993). Verb movement and the quotative construction in English. In Jonathan David Bobaljik and Colin Phillips (eds.), *Papers on case and agreement I: MIT working papers in linguistics 18* (pp. 1-13). Cambridge, MA: MITWPL.

Bresnan, Joan (1976). On the form and functioning of transformations. *Linguistic Inquiry*, *7*, 3-40.

Brody, Michael (1995). *Lexico-logical form: A radically minimalist program*. Cambridge, MA: MIT Press.

Bromberger, Sylvain (1986). *What we don't know when we don't know why*. Unpublished manuscript, MIT.

Browning, Marguerite (1987). *Null operator constructions*. Doctoral dissertation, MIT.

Bures, Anton (1992). There is an argument for an LF cycle here. In Jeannette Marshall Denton, Grace P. Chan, and Costas P. Canakis (eds.), *CLS 28, Volume 2* (pp. 14-35). Chicago: Chicago Linguistic Society.

Cattell, Ray (1976). Constraints on movement rules. *Language*, *52*, 18-50.

Cheng, Lisa Lai-Shen (1991). *On the typology of wh-questions*. Doctoral dissertation, MIT.

Cheng, Lisa Lai-Shen (1997). "Partial" wh-movement. In Luther Chen-Sheng Liu and Kazue Takeda (eds.), *UCI working papers in linguistics, volume 3* (pp. 27-50). Irvine: Irvine Linguistics Students Association, University of California, Irvine.

Choe, Jae W. (1987). LF movement and pied-piping. *Linguistic Inquiry*, *18*, 348-353.

Chomsky, Noam (1964). *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, Noam (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam (1968). *Language and mind*. New York: Harcourt Brace Jovanovich.

Chomsky, Noam (1972). *Studies on semantics in generative grammar*. The Hague: Mouton.

Chomsky, Noam (1973). Conditions on transformations. In Stephan R. Anderson and Paul Kiparsky (eds.), *A festschrift for Morris Halle* (pp. 232-286). New York: Holt, Rinehart, and Winston.

Chomsky, Noam (1975). *Reflections on language*. New York: Pantheon.

Chomsky, Noam (1976). Conditions on rules of grammar. *Linguistic Analysis*, *2*, 303-351.

Chomsky, Noam (1977). On WH-movement. In Peter Culicover, Thomas Wasow, and Adrian Akmajian (eds.), *Formal syntax* (pp. 71-132). New York: Academic Press.

Chomsky, Noam (1980). *Rules and representations*. New York: Columbia University Press.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

Chomsky, Noam (1986a). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1986b). *Knowledge of language: Its nature, origin, and use*. New York: Prager.

Chomsky, Noam (1991a). Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in comparative grammar* (pp. 417-454). Cambridge, MA: MIT Press.

Chomsky, Noam (1991b). Linguistics and cognitive sciences: Problems and mysteries. In Asa Kasher (ed.), *The Chomskyan turn* (pp. 3-25). Cambridge, MA: Blackwell.

Chomsky, Noam (1993). A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser (eds.), *The view from building 20* (pp. 1-52). Cambridge, MA: MIT Press.

Chomsky, Noam (1994). *Bare phrase structure*. MIT occasional papers in linguistics, Number 5. Cambridge, MA: MITWPL.

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

Chomsky, Noam (1996). *Some observations on economy in generative grammar*. Unpublished manuscript, MIT.

Chomsky, Noam and Howard Lasnik (1993). The theory of principles and parameters. In Joahim Jacobs, Armin von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research* (pp. 506-569). Berlin: Walter de Gruyter.

Chung, Sandra (1982). Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry*, *13*, 39-77.

Chung, Sandra (1994). *Wh*-agreement and "referentiality" in Chamorro. *Linguistic Inquiry*, *25*, 1-44.

Chung, Sandra and Carol Georgopoulos (1988). Agreement with gaps in Chamorro and Palauan. In Michael Barlow and Charles A. Ferguson (eds.), *Agreement in natural language: Approaches, theories, and descriptions* (pp. 251-267). Stanford: Center for the study of language and information.

Cinque, Guglielmo (1990). *Types of A'-dependencies*. Cambridge, MA: MIT Press.

Clark, Robin (1985). *Boundaries and the treatment of control*. Doctoral dissertation, University of California, Los Angeles.

Clements, George N. (1984). Binding domains in Kikuyu. *Studies in the linguistic sciences*, *14*, 37-56.

Cole, Peter, Gabriella Hermon, and Li-May Sung (1990). Principles and parameters of long-distance reflexives. *Linguistic Inquiry*, *21*, 1-22.

Collins, Chris (1993). *Topics in Ewe syntax*. Doctoral dissertation, MIT.

Collins, Chris (1997). *Local economy*. Cambridge, MA: MIT Press.

Collins, Chris and Höskuldur Thráinsson (1993). Object shift in double object constructions and the theory of case. In Colin Phillips (ed.), *Papers on case and agreement II: MIT working papers in linguistics 19* (pp. 131-174). Cambridge, MA: MITWPL.

Culicover, Peter and Kenneth Wexler (1977). Some syntactic implications of a theory of language learnability. In Peter Culicover, Thomas Wasow, and Adrian Akmajian (eds.), *Formal syntax* (pp. 7-60). New York: Academic Press.

Culicover, Peter W. (1976). *Syntax*. New York: Academic Press.

Culicover, Peter W. and Ken Wexler (1973a). *An application of the freezing principle to the dative in English*. Social sciences working paper 48, University of California, Irvine.

Culicover, Peter W. and Ken Wexler (1973b). *Three further applications of the freezing principle in English*. Social sciences working paper 48, University of California, Irvine.

Culicover, Peter W. and Wendy Wilkins (1984). *Locality in linguistic theory*. New York: Academic Press.

Dayal, Veneeta Strivastav (1994). Scope marking as indirect WH dependency. *Natural Language Semantics*, *2*, 137-170.

Delahunty, Gerald (1983). But sentential subjects do exist. *Linguistic Analysis*, *12*, 379-398.

Depréz, Viviane (1989). *On the typology of syntactic positions and the nature of chains*. Doctoral dissertation, MIT.

Di Sciullo, Anna Maria and Edwin Williams (1987). *On the definition of word*. Cambridge, MA: MIT Press.

Emonds, Joseph (1976). *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.

Emonds, Joseph (1978). The verbal complex of V'-V in French. *Linguistic Inquiry*, *9*, 151-175.

Emonds, Joseph (1985). *A unified theory of syntactic categories*. Dordrecht: Foris.

Engdahl, Elisabet (1980a). *The syntax and semantics of questions in Swedish*. Doctoral dissertation, University of Massachusetts, Amherst.

Engdahl, Elisabet (1980b). WH constructions in Swedish and the relevance of subjacency. In J. Jensen (ed.), *Proceedings of NELS 10: Cahiers linguistiques d'Ottawa* (pp. 98-108). Ottawa: Department of Linguistics, University of Ottawa.

Engdahl, Elisabet (1982). Restrictions on unbounded dependencies in Swedish. In Elisabet Engdahl and Eva Ejerhed (eds.), *Readings on unbounded dependencies in Scandinavian languages* (pp. 151-174). Stockholm: Almquist and Wiksell International.

Engdahl, Elisabet (1986). *Constituent questions: The syntax and semantics of questions with special reference to Swedish*. Dordrecht: Reidel.

Epstein, Samuel David (1992). Derivational constraints on A'-chain formation. *Linguistic Inquiry*, *23*, 235-259.

Erteschik-Shir, Nomi (1973). *On the nature of island constraints*. Doctoral dissertation, MIT.

Fiengo, Robert, C.-T. James Huang, Howard Lasnik, and Tanya Reinhart (1988). The syntax of wh-in-situ. In Hagit Borer (ed.), *Proceedings of the seventh west coast conference on formal linguistics* (pp. 81-98). Stanford: Stanford Linguistics Association by the Center for the Study of Language and Information.

Fodor, Janet Dean (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, *9*, 427-473.

Foley, William, A. and Robert D. Jr. Van Valin (1988). *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

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

Fukui, Naoki (1988). Extraction of *naze*: Some theoretical implications. *Natural Language and Linguistic Theory*, *6*, 503-526.

Fukui, Naoki (1993a). Parameters and optionality. *Linguistic Inquiry*, *24*, 399-420.

Fukui, Naoki (1993b). A note on improper movement. *The Linguistic Review*, *10*, 111-126.

Fukui, Naoki (1995). Comparative syntax of English and Japanese: A principles-and-parameters approach to typology. In Theodora Bynon and Masayoshi Shibatani (eds.), *Approaches to language typology* (pp. 327-372). Oxford: Oxford University Press.

Fukui, Naoki (1996). On the nature of economy in language. *Cognitive studies*, *3*, 51-71.

Fukui, Naoki (1997). Attract and the A-over-A principle. In Luther Chen-Sheng Liu and Kazue Takeda (eds.), *UCI working papers in linguistics, volume 3* (pp. 51-67). Irvine: Irvine Linguistics Students Association, University of California, Irvine.

Fukui, Naoki and Mamoru Saito (1996). *Order in the theory of phrase structure and movement*. University of California, Irvine and Nanzan University.

Georgopoulos, Carol (1985). Variables in Palauan syntax. *Natural Language and Linguistic Theory*, *3*, 59-94.

Georgopoulos, Carol (1991a). On A- and A'-agreement. *Lingua*, *85*, 135-169.

Georgopoulos, Carol (1991b). *Syntactic variables*. Dordrecht: Kluwer.

Grimshaw, Jane (1979). Complement selection and lexicon. *Linguistic Inquiry*, *10*, 279-326.

Grimshaw, Jane (1990). *Argument structure*. Cambridge, MA: MIT Press.

Haig, John (1976). Shadow pronoun deletion in Japanese. *Linguistic Inquiry*, *7*, 363-371.

Hale, Kenneth (1980). Remarks on Japanese phrase structure: Comments on the papers on Japanese syntax. In Yukio Otsu and Ann Farmer (eds.), *MIT working papers in Linguistics 2* (pp. 185- 201). Cambridge, MA: MITWPL.

Hale, Kenneth (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory*, *1*, 5-47.

Hale, Kenneth and Samuel Jay Keyser (1993). On argument structure and the lexical expression of syntactic relations. In Kenneth Hale and Samuel Jay Keyser (eds.), *View from building 20* (pp. 53-109). Cambridge, MA: MIT Press.

Harada, Shin-Ichi (1977). "Nihongo-ni henkei-wa hituyoo-da" (There are transformations in Japanese). *Gengo*, *6*, 88-103.

Hasegawa, Kinsuke (1974). Generalized A-over-A principle. *Eigo Seinen (The Rising Generation)*, *119*, 736-737, 808-810.

Hasegawa, Nobuko (1984/1985). On the so-called 'zero pronouns' in Japanese. *The Linguistic Review*, *4*, 289-341.

Hasegawa, Nobuko (1986). More arguments for the pied-piping analysis of wh-questions in Japanese. In Nobuko Hasegawa and Yoshihisa Kitagawa (eds.), *University of Massachusetts occasional papers in linguistics 11* (pp. 79-107). Amherst: University of Massachusetts, Amherst.

Haïk, Isabelle (1990). Anaphoric, pronominal and referential INFL. *Natural Language and Linguistic Theory*, *8*, 347-374.

Hegarty, Michael V. (1991). *Adjunct extraction and chain configurations*. Doctoral dissertation, MIT.

Heim, Irene (1982). *The semantics of definite and indefinite noun phrases*. Doctoral dissertation, University of Massachusetts, Amherst.

Hellan, Lars and Kirsti Koch Christensen (1986). Introduction. In Lars Hellan and Kirsti Koch Christensen (eds.), *Topics in Scandinavian syntax* (pp. 1-29). Dordrecht: Reidel.

Higginbotham, James (1980). Pronouns and bound variables. *Linguistic Inquiry*, *11*, 679-708.

Higginbotham, James (1983). Logical form, binding, and nominals. *Linguistic Inquiry*, *14*, 395-420.

Higginbotham, James (1985). On semantics. *Linguistic Inquiry*, *16*, 547-593.

Higginbotham, James and Robert May (1981). Questions, quantifiers and crossing. *The Linguistic Review*, *1*, 41-80.

Hoji, Hajime (1985). *Logical form constraints and configurational structures*. Doctoral dissertation, University of Washington.

Hoji, Hajime (1990). *Theories of anaphora and aspects of Japanese syntax*. Unpublished manuscript, University of Southern California.

Holmberg, Anders (1986). *Word order and syntactic features in the Scandinavian languages and English*. Doctoral dissertation, University of Stockholm.

Holmberg, Anders and Christer Platzack (1995). *The role of inflection in Scandinavian syntax*. Oxford: Oxford University Press.

Hopcroft, John and Jeffrey Ullman (1979). *Introduction to automata theory, languages, and computation*. Menlo Park: Addison-Wesley.

Hornstein, Nobert and Amy Weinberg (1987). Superiority and generalized binding. In J. McDonough and B. Plunkett (eds.), *Proceedings of NELS 17* (pp. 311-324). Amherst: GLSA, University of Massachusetts, Amherst.

Hornstein, Nobert and Amy Weinberg (1990). The necessity of LF. *The Linguistic Review*, *7*, 129-167.

Huang, C.-T. James (1982). *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT.

Huang, C.-T. James (1983). A note on the binding theory. *Linguistic Inquiry*, *14*, 554-561.

Huang, C.-T. James (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, *15*, 53-74.

Huang, C.-T. James (1989). Pro-drop in Chinese: A generalized control theory. In Osvaldo Jaeggli and Kenneth J. Safir (eds.), *The null subject parameter* (pp. 185-214). Dordrecht: Kluwer.

Huang, C.-T. James (1991). Remarks on the status of the null object. In Robert Freidin (ed.), *Principles and parameters in comparative grammar* (pp. 56-76). Cambridge, MA: MIT Press.

Huang, C.-T. James (1994). *On lexical structure and syntactic projections*. Unpublished manuscript, University of California, Irvine.

Huang, C.-T. James and Jane Tang (1991). The local nature of the long-distance reflexive in Chinese. In Jan Koster and Eric Reuland (eds.), *Long-distance anaphora* (pp. 263-282). Cambridge: Cambridge University Press.

Ikawa, Hajime (1996). LF as a level without functional categories and its parallelisms with the overt syntax of a language without functional categories. In Brian Agbayani, Kazue Takeda, and Sze-Wing Tang (eds.), *UCI working papers in linguistics 1* (pp. 39-62). Irvine: Irvine Linguistics Students Association.

Inoue, Kazuko (1978). "Tough sentences" in Japanese. In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics* (pp. 122-154). Tokyo: Kaitakusha.

Ishii, Toru (1997). The "crossing" constraint and the minimal link condition. In Luther Chen-Sheng Liu and Kazue Takeda (eds.), *UCI working papers in linguistics, volume 3* (pp. 69-98). Irvine: Irvine Linguistics Students Association, University of California, Irvine.

Ishii, Yasuo (1991). *Operators and empty categories in Japanese.* Doctoral dissertation, University of Connecticut.

Jackendoff, Ray (1972). *Semantic interpretation of generative grammar.* Cambridge, MA: MIT Press.

Jackendoff, Ray (1977). *X-bar syntax: A study of phrase structure.* Cambridge, MA: MIT Press.

Jaeggli, Osvaldo (1981). *Topics in Romance syntax.* Dordrecht: Foris.

Johnson, David (1990). A catalogue of complexity classes. In Jan van Leeuwen (ed.), *Handbook of theoretical computer science, Volume A: algorithms and complexity* (pp. 67-161). Cambridge: MIT Press/Elsevier.

Johnson, Kyle (1991). Object positions. *Natural Language and Linguistic Theory*, *9*, 577-636.

Jonas, Dianne and Jonathan David Bobaljik (1993b). Specs for subjects: The role of TP in Icelandic. In Jonathan David Bobaljik and Colin Phillips (eds.), *Papers on case and agreement I: MIT working papers in linguistics 18* (pp. 59-98). Cambridge, MA: MITWPL.

Kamp, Hans (1981). A theory of truth and discourse representation. In J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Formal methods in the study of language* (pp. 1-41). Amsterdam: Matematisch Centrum.

Karttunen, Lauri (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, *1*, 3-44.

Katada, Fusa (1991). The LF representation of anaphors. *Linguistic Inquiry*, *22*, 287-313.

Katz, Jerold and Paul Postal (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.

Kayne, Richard (1983). *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard (1989). Facets of romance past participle agreement. In Paola Benincà (ed.), *Dialect variation and the theory of grammar* (pp. 85-103). Dordrecht: Foris.

Kayne, Richard (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kikuchi, Akira (1987). *Comparative deletion in Japanese*. Unpublished manuscript. Yamagata University.

Kim, Soo Won (1991). *Chain scope and quantification structure*. Doctoral dissertation, Brandies University.

Kitagawa, Yoshihisa (1985). Small but clausal. In William H. Eilfort, Paul D. Kroeber, and Karen L. Peterson (eds.), *CLS 21* (pp. 210- 220). Chicago: Chicago Linguistic Society.

Kitagawa, Yoshihisa (1986). *Subjects in Japanese and English*. Doctoral dissertation, University of Massachusetts.

Kitahara, Hisatsugu (1995). Target alpha: Deducing strict cyclicity from derivational economy. *Linguistic Inquiry*, *26*, 47-77.

Kitahara, Hisatsugu (1997). *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.

Koizumi, Masatoshi (1995). *Phrase structure in minimalist syntax*. Doctoral dissertation, MIT.

Koopman, Hilda and Dominique Sportiche (1982/1983). Variables and the bijection principle. *The Linguistic Review*, *2*, 139-160.

Koopman, Hilda and Dominique Sportiche (1986). A note on long extraction in Vata and the ECP. *Natural Language and Linguistic Theory*, *4*, 357-374.

Koster, Jan (1978a). *Locality principles in syntax.* Dordrecht: Foris.

Koster, Jan (1978b). Why subject sentences don't exist. In Samuel Jay Keyser (ed.), *Recent transformational studies in European languages* (pp. 53-64). Cambridge, MA: MIT Press.

Koster, Jan (1986). *Domains and dynasties: The radical autonomy of syntax.* Dordrecht: Foris.

Kuno, Susumu (1973). *The structure of the Japanese language.* Cambridge, MA: MIT Press.

Kuroda, S.-Y. (1965). *Generative grammatical studies in Japanese language.* Doctoral dissertation, MIT.

Kuroda, S.-Y. (1988). Whether you agree or not: Rough ideas about the comparative grammar of English and Japanese. In William Poser (ed.), *Japanese syntax* (pp. 103-143). Stanford: Center for the Study of Language and Information.

Kuroda, S.-Y. (1972). The categorial and the thetic judgments. *Foundations of language*, *9*, 153-185.

Kuroda, S-Y (1978). Case marking, canonical sentence patterns, and counter-equi in Japanese. In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics* (pp. 30-51). Tokyo: Kaitakusha.

Kuroda, S.-Y. (1979). *The 'w'hole of the doughnut.* Belgium: E. Story-scientia P.V.B.A.

Larson, Richard (1987). Extraction and multiple selection in PP. In Marguerite Browning, Ewa Czaykowska-Higgins, and Elizabeth Ritter (eds.), *MIT working papers in linguistics 9* (pp. 119-135). Cambridge, MA: MITWPL.

Larson, Richard (1988). On the double object construction. *Linguistic Inquiry*, *19*, 335-391.

Lasnik, Howard (1989). *Essays on anaphora*. Dordrecht: Kluwer.

Lasnik, Howard (1994). *Verbal morphology*. Unpublished manuscript, University of Connecticut.

Lasnik, Howard and Joseph J. Kupin (1977). A restrictive theory of transformational grammar. *Theoretical linguistics*, *4*, 173-196.

Lasnik, Howard and Mamoru Saito (1984). On the nature of proper government. *Linguistic Inquiry*, *15*, 235-289.

Lasnik, Howard and Mamoru Saito (1992). *Move alpha: Conditions and its application and output*. Cambridge, MA: MIT Press.

Lasnik, Howard and Juan Uriagereka (1988). *A course in GB syntax: Lectures on binding and empty categories*. Cambridge, MA: MIT Press.

Lebeaux, David (1983). A distributional difference between reciprocals and reflexives. *Linguistic Inquiry*, *14*, 723-730.

Lebeaux, David (1988). *Language acquisition and the form of the grammar*. Doctoral dissertation, University of Massachusetts, Amherst.

Lebeaux, David (1991). Relative clauses, licensing, and the nature of the derivation. In Susan Rothstein (ed.), *Syntax and semantics 25, Perspectives on phrase structure: Heads and licensing* (pp. 209-239). San Diego: Academic Press.

Lee, Rhanghyeyun K. (1995). *Economy of representation*. Doctoral dissertation, University of Connecticut.

Levin, Beth and Malka Rappaport (1986). The formation of adjectival passives. *Linguistic Inquiry*, *17*, 623-662.

Lewis, David (1975). Adverbs of quantification. In Edward Keenan (ed.), *Formal semantics of natural language* (pp. 3-15). Cambridge: Cambridge University Press.

Longobardi, Giuseppe (1984). *Movement, scope and island constraints*. Unpublished manuscript, University of Venice.

Mahajan, Anoop (1990). *The A/A-bar distinction and movement theory*. Doctoral dissertation, MIT.

Maki, Hideto (1995). *The syntax of particles*. Doctoral dissertation, University of Connecticut.

Manzini, Rita (1992). *Locality: A theory and some of its empirical consequences*. Cambridge, MA: MIT Press.

Marantz, Alec (1980). English S is the maximal projection of V. In J. Jensen (ed.), *Proceedings of NELS 10: Cahiers linguistiques d'Ottawa*. Ottawa: Department of Linguistics, University of Ottawa.

Marantz, Alec (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Matsuda, Yuki (1996). *A syntactic analysis of focus sentences in Japanese*. Unpublished manuscript, University of Southern California.

May, Robert (1977). *The grammar of quantification*. Doctoral dissertation, MIT.

May, Robert (1985). *Logical form: Its structure and derivation*. Cambridge, MA: MIT Press.

McCawley, James (1968). *The phonological component of a grammar of Japanese*. The Hague: Mouton.

McCloskey, James (1979). *Transformational syntax and model theoretic semantics*. Dordrecht: Reidel.

McCloskey, James (1989). Resumptive pronouns, A'-binding, and levels of representation in Irish. In Randall Hendrick (ed.), *Syntax and semantics 23: The syntax of the modern Celtic languages* (pp. 199-248). New York: Academic Press.

McConnell-Ginet, Sally (1982). Adverbs and logical form. *Language*, *58*, 144-184.

McDaniel, Dana (1989). Partial and multiple wh-movement. *Natural Language and Linguistic Theory*, *7*, 565-604.

Miyagawa, Shigeru (1997). Against optional scrambling. *Linguistic Inquiry*, *28*, 1-25.

Miyara, Shinsho (1982). Reordering in Japanese. *Linguistic Analysis*, *9*, 307-348.

Montalbetti, Mario (1984). *After binding: On the interpretation of pronouns*. Doctoral dissertation, MIT.

Murasugi, Keiko (1991). *Noun phrases in Japanese and English*. Doctoral dissertation, University of Connecticut.

Muysken, Pieter (1982). Parametrizing the notion "Head". *Journal of Linguistic Research*, *2*, 57-75.

Nishigauchi, Taisuke (1984). Control and the thematic domain. *Language*, *60*, 215-250.

Nishigauchi, Taisuke (1986). *Quantification in syntax*. Doctoral dissertation, University of Massachusetts.

Nishigauchi, Taisuke (1990). *Quantification in the theory of grammar*. Dordrecht: Kluwer.

Ogihara, Toshiyuki (1987). 'Obligatory focus' in Japanese and type-shifting principles. In Megan Crowhurst (ed.), *Proceedings of the west coast conference on formal linguistics, Volume 6* (pp. 213-227). Stanford: Stanford Linguistics Association.

Pesetsky, David (1982). *Paths and categories*. Doctoral dissertation, MIT.

Pesetsky, David (1987). Wh-in-situ: Movement and unselective binding. In Eric Reuland and Alice Meulen (eds.), *The representation of (in)definiteness* (pp. 98-129). Cambridge, MA: MIT Press.

Pesetsky, David (1989). *Language-particular processes and the earliness principle*. Unpublished manuscript, MIT.

Pica, Pierre (1987). On the nature of the reflexivization cycle. In J. McDonough and B. Plunkett (eds.), *Proceedings of NELS 17* (pp. 483-499). Amherst: GLSA, University of Massachusetts, Amherst.

Pollock, Jean-Yves (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, *20*, 365-424.

Poole, Geoffrey (1996). Optional movement in the minimalist program. In Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart (eds.), *Minimal ideas: Syntactic studies in the minimalist program* (pp. 199-216). Amsterdam: John Benjamins Publishing Company.

Postal, Paul (1966). *The method of universal grammar*. Unpublished paper.

Randall, Janet H. (1988). Inheritance. In Wendy Wilkins (ed.), *Syntax and semantics 21: Thematic relations* (pp. 129-146). San Diego: Academic Press.

Rappaport, Malka and Beth Levin (1988). What to do with theta-role. In Wendy Wilkins (ed.), *Syntax and semantics 21: Thematic relations* (pp. 7-36). San Diego: Academic Press.

Reinhart, Tanya (1982). A second Comp position. In Adriana Belletti, Luciana Brandi, and Luigi Rizzi (eds.), *Theory of markedness in generative grammar*. Pisa: Scuola Normale Superiore.

Reinhart, Tanya (1983). Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy*, *6*, 47-88.

Reinhart, Tanya (1992). *Interpreting wh-in-situ*. Unpublished manuscript, Tel Aviv University.

Reinhart, Tanya (1993). *Wh-in-situ in the framework of the minimalist program*. Unpublished manuscript, Tel Aviv University.

Reuland, Eric (1985). A feature system for the set of categorial heads. In Pieter Muysken and Henk van Riemsdijk (eds.), *Features and projections* (pp. 41-88). Dordrecht: Foris.

Rizzi, Luigi (1982). *Issues in Italian syntax*. Dordrecht: Foris.

Rizzi, Luigi (1986). On chain formation. In Hagit Borer (ed.), *Syntax and semantics 19: The grammar of pronominal clitics* (pp. 65-95). New York: Academic Press.

Rizzi, Luigi (1990). *Relativized minimality*. Cambridge, MA: MIT Press.

Robert, Ian (1994). Two types of head movement in Romance. In David Lightfoot and Norbert Hornstein (eds.), *Verb movement* (pp. 207-42). Cambridge: Cambridge University Press.

Rohrbacher, Bernhard (1993). *The Germanic VO languages and the full paradigm: A theory of V to I raising*. Doctoral dissertation, University of Massachusetts, Amherst.

Rohrbacher, Bernhard (1995). Explaining the syntactic consequences of 'rich' agreement morphology: On the licensing of V-to-AgrS raising and pro. In Raul Aranovich, William Byrne, Susanne Preuss, and Martha Senturia (eds.), *The proceedings of the thirteenth west coast conference on formal linguistics* (pp. 350-364). Stanford: Stanford Linguistics Association by the Center for the Study of Language and Information.

Ross, John (1967). *Constraints on variables in syntax*. Doctoral dissertation, MIT.

Rothstein, Susan (1983). *The syntactic form of predication*. Doctoral dissertation, MIT.

Safir, Kenneth (1985). *Syntactic chains*. Cambridge: Cambridge University Press.

Saito, Mamoru (1982). *Case marking in Japanese*. Unpublished manuscript, MIT.

Saito, Mamoru (1985). *Some asymmetries in Japanese and their theoretical implications*. Doctoral dissertation, MIT.

Saito, Mamoru (1986). Three notes on syntactic movement in Japanese. In Takashi Imai and Mamoru Saito (eds.), *Issues in Japanese linguistics* (pp. 301-350). Dordrecht: Foris.

Saito, Mamoru (1987). *An extension of K. I. Harada's WH-Q binding analysis*. Paper presented at the 5th National Conference of the English Linguistic Society of Japan.

Saito, Mamoru (1989). Scrambling as semantically vacuous A'-movement. In Mark Baltin and Anthony Kroch (eds.), *Alternative conceptions of phrase structure* (pp. 182-200). Chicago: University of Chicago Press.

Saito, Mamoru (1992). Long distance scrambling in Japanese. *Journal of East Asian Linguistics*, *1*, 69-118.

Saito, Mamoru (1994). Improper adjunction. In Masatoshi Koizumi and Hiroyuki Ura (eds.), *Formal approaches to Japanese linguistics 1: MIT working papers in linguistics 24* (pp. 263-293). Cambridge, MA: MITWPL.

Schwartz, Linda (1988). Thematic relations and case linking in Russian. In Wendy Wilkins (ed.), *Syntax and semantics 21: Thematic relations* (pp. 167-189). San Diego: Academic Press.

Selkirk, Elisabeth (1981). *The syntax of words*. Cambridge, MA: MIT Press.

Speas, Margaret (1984). Saturation and phrase structure. In Diana Archangeli, Andrew Barss, and Richard Sproat (eds.), *MIT working papers in linguistics 6* (pp. 174-198). Cambridge, MA: MITPWL.

Sportiche, Dominique (1981). Bounding nodes in French. *The Linguistic Review*, *1*, 219-246.

Sportiche, Dominique (1983). *Structural invariance and symmetry in syntax*. Doctoral dissertation, MIT.

Stowell, Timothy (1981). *Origins of phrase structure*. Doctoral dissertation, MIT.

Stroik, Thomas (1990). Adverbs as V-sisters. *Linguistic Inquiry*, *21*, 654-661.

Szabolcsi, Anna and Frans Zwarts (1992). *Weak islands and an algebraic semantics for scope-taking*. Unpublished manuscript, University of California, Los Angeles and Rijksuniversiteit Groningen.

Tada, Hiroaki (1990). *Scrambling(s).* Paper presented at workshop on Japanese syntax, Ohio State University, MIT.

Takahashi, Daiko (1994). *Minimality of movement*. Doctoral dissertation, University of Connecticut.

Takano, Yuji (1996). *Movement and parametric variation in syntax*. Doctoral dissertation, University of California, Irvine.

Takeda, Kazue (1997). A note on locality of feature movement and category movement. In Luther Chen-Sheng Liu and Kazue Takeda (eds.), *UCI working papers in linguistics, volume 3* (pp. 183-202). Irvine: Irvine Linguistics Students Association, University of California, Irvine.

Takezawa, Koichi (1987). *A configurational approach to case-marking in Japanese*. Doctoral dissertation, University of Washington.

Taraldsen, Knut Tarald (1978). The scope of *wh* movement in Norwegian. *Linguistic Inquiry*, *9*, 623-640.

Taraldsen, Knut Tarald (1982). On the theoretical interpretation of a class of 'marked' extractions. In Adriana Belletti, Luciana Brandi, and Luigi Rizzi (eds.), *The theory of markedness in generative grammar* (pp. 475-516). Pisa: Scuola Normale Superiore.

Tellier, Christine (1990). Condition violations in French. *Linguistic Inquiry*, *21*, 306-311.

Tiedeman, Robyne (1990). An S-structure/LF asymmetry in subject extraction. *Linguistic Inquiry*, *21*, 661-667.

Torrego, Esther (1986). *On empty categories in nominals*. Unpublished manuscript, University of Massachusetts, Boston.

Toyoshima, Takashi (1996). Derivational CED: A consequence of the bottom-up parallel-process of merge and attract. In Brian Agbayani and Sze-Wing Tang (eds.), *Proceedings of the fifteenth west coast conference on formal linguistics* (pp. 505-519). Stanford: Stanford Linguistics Association by the Center for the Study of Language and Information.

Travis, Lisa deMena (1984). *Parameters and effects of word order variation*. Doctoral dissertation, MIT.

Travis, Lisa deMena (1992). Inner aspect and the structure of VP. In K. Broderick (ed.), *Proceedings of NELS 22*. Amherst: GLSA, University of Massachusetts, Amherst.

Tsai, Wei-Tien (1994). *On economizing the theory of A-bar dependencies*. Doctoral dissertation, MIT.

Ura, Hiroyuki (1995). Towards a theory of "strictly derivational" economy condition. In Rob Pensalfini and Hiroyuki Ura (eds.), *Papers on minimalist syntax: MIT working papers in linguistics 27* (pp. 243-267). Cambridge, MA: MITWPL.

Ura, Hiroyuki (1996). *Multiple feature-checking: A theory of grammatical function splitting*. Doctoral dissertation, MIT.

van Riemsdijk, Henk and Edwin Williams (1981). NP structure. *The Linguistic Review*, *1*(2), 171-217.

Vikner, Sten (1995). *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

von Stechow, Armin (1993). *Against LF pied-piping*. SfS-report-12-93, Universität Tübingen.

Watanabe, Akira (1992a). Subjacency and S-structure movement of WH-in-situ. *Journal of East Asian Linguistics*, *1*, 255-291.

Watanabe, Akira (1992b). *Wh-in-situ, subjacency, and chain formation*. MIT occasional papers in linguistics, Number 2. Cambridge, MA: MITWPL.

Watanabe, Akira (1993). Larsonian CP recursion, factive complements, and selection. In Amy J. Schafer (ed.), *Proceedings of NELS 23* (pp. 523-537). Amherst: GLSA, University of Massachusetts, Amherst.

Watanabe, Akira (1995). Conceptual basis of cyclicity. In Rob Pensalfini and Hiroyuki Ura (eds.), *Papers on minimalist syntax: MIT Working Papers in Linguistics 27* (pp. 269-291). Cambridge, MA: MITWPL.

Webelhuth, Gert (1989). *Syntactic saturation phenomena and the modern Germanic languages*. Doctoral dissertation, University of Massachusetts, Amherst.

Wexler, Ken and Peter W. Culicover (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Wilkins, Wendy (1988). Thematic structure and reflexivization. In Wendy Wilkins (ed.), *Syntax and semantics 21: Thematic relations* (pp. 191-213). San Diego: Academic Press.

Williams, Edwin (1980). Predication. *Linguistic Inquiry*, *11*, 203-238.

Williams, Edwin (1981). Argument structure and morphology. *The Linguistic Review*, *1*, 81-114.

Williams, Edwin (1994). *Thematic structure in syntax*. Cambridge, MA: MIT Press.

Zoerner, Ed (1995). *Coordination: The syntax of &P*. Doctoral dissertation, University of California, Irvine.

Zubizarreta, Maria Luisa (1987). *Levels of representation in the lexicon and in the syntax*. Dordrecht: Foris.

Zwart, C. Jan Wouter (1993). *Dutch syntax: A minimalist approach*. Doctoral dissertation, University of Groningen.