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 (σ_1, σ_2, ..., σ_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

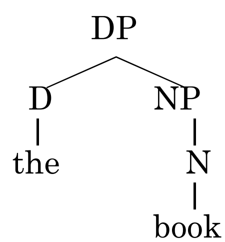
b. K = {γ, {α, β}}, where α, β are objects and γ is the label of K.

(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]

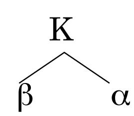
(Chomsky 1995:272-274)

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:

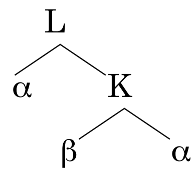
(11) a.



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 $<\alpha, \beta>$, where β is the co-constituent of the original α. α in the raised position, on the other hand, is identified as the pair $<\alpha, K>$, 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 = ($<\alpha, K>$, $<\alpha, \beta>$), 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.