

TORU ISHII

Islands and Movement Types*

0. Introduction

Locality conditions on movement have been one of the central issues for generative grammar. A lot of discussion has been around regarding locality conditions on so called A'-movement like overt wh-movement. 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. Ross (1967) proposes island constraints on movement, which state that no elements can be extracted out of domains called "islands." It has been claimed by, among others, Chomsky (1993, 1995b) and Chomsky and Lasnik (1993) that the Wh-island Constraint should be subsumed under the Minimal Link Condition (MLC), the more recent version of Rizzi's (1990) Relativized Minimality (RM). The MLC is sensitive to movement types in the following respects. First, it only regulates the interaction of movements of the same type. For example, in the case of a wh-island violation, two instances of wh-movement are involved. When movements of different types interact, on the other hand, the MLC effects never emerge, as shown below:

- (1) [To whom]_i did John_j seem *t_i* [*t_j* to be referring]?

In (1), NP-movement interacts with wh-movement; there is no violation of the MLC. Second, as observed by, among others, Fukui and Saito (1996), scrambling in Japanese is immune from the MLC (Fukui and Saito (1996:9)):

- (2) [John-ni_j [Bill-ga [[zibun-o_i [Mary-ga *t_j* *t_i* urikonda]] to]
-Dat -Nom self-Acc -Nom sold Comp
omotteiru]] (koto)
think (fact)
Lit. 'To John_j, Bill thinks [that [herself_i, Mary_i sold *t_i* *t_j*]].'

Since the MLC is defined based on the notion of feature-checking, it only regulates feature-driven movement but not non-feature-driven movement. Fukui and Saito argue that the immunity of scrambling from

* I am grateful for Brian Agbayani, Hiroshi Aoyagi, Lisa Cheng, Naoki Fukui, Naomi Harada, Hidehito Hoshi, James Huang, Hajime Ikawa, Utpal Lahiri, and Kazue Takeda for helpful comments and suggestions. Remaining errors and omissions are, of course, the sole responsibility of the author.

ISLANDS AND MOVEMENT TYPES

the MLC straightforwardly follows if we assume that scrambling is not driven by any formal feature.

In contrast to the MLC, the other island constraints (henceforth called the "domain barriers" (DBs)) have been assumed not to be sensitive to movement types. It has been claimed that as far as the overt component is concerned, no element can ever be extracted out of the DBs regardless of what type of movement is involved. Contrary to this wide-spread view, however, I will argue that the DBs are also sensitive to movement types. Specifically, I will argue that exactly like the MLC, the DBs only regulate feature-driven movement like English overt wh-movement but not non-feature-driven movement like Japanese scrambling. I will then propose a new minimalist account of the DBs based on a derivational approach to selectional restrictions (SRs). It is shown that our theory has an advantage over any previous locality theories in that it can account for the hitherto unexplained asymmetry between feature-driven movement and non-feature-driven movement with the DB effects.

The organization of this paper is as follows. Section 1 argues that contrary to what has been claimed, scrambling in Japanese does not exhibit any DB effects. Section 2 presents DB effects in Japanese. I will argue that there is an asymmetry with the DB among movement types. Section 3 proposes a derivational approach to SRs, arguing that it gives a new minimalist account of the DBs. It is also shown that our account of the DBs can accommodate the above-mentioned asymmetry with the DBs. Section 4 considers apparent DB effects with scrambling in Japanese. I will argue that they should be attributed to a condition in the phonological component. Section 5 presents conceptual arguments for our theory. Section 6 makes concluding remarks.

1. No DB Effects with Scrambling in Japanese

As argued by, among others, Harada (1977), Muraki (1979), and Saito (1985), scrambling is a movement operation which is responsible for the relatively free word order in Japanese. At the first glance, scrambling seems to exhibit the DB effects, as the following examples show:¹

¹As argued by, among others, Fukui (1995), Kayne (1984), Lasnik and Saito (1992), Ross (1967), and Saito (1985), the Subject Condition does not hold in Japanese. See these works for detailed discussion of this subject.

TORU ISHII

(3) Complex NP Constraint

a. Relative Clauses

?Mary-ni [John-ga [[t atta] hito]-o sagasite iru rasii
 -Dat -Nom met person-Acc looking-for seem
 'It seems that John is looking for the person who met Mary.'

b. Non-relative Complex NPs

?Bill-ni [John-ga [[Mary-ga t atta to yuu] uwasa]-o
 -Dat -Nom -Nom met Comp say rumor-Acc
 kiita] rasii
 heard seem
 'It seems that John heard a rumor (which says) that Mary met Bill.'

(4) Adjunct Condition

?Mary-ni [John-ga [Bill-ga t somuita node] okotte iru
 -Dat -Nom -Nom acted against because be angry
 rasii
 seem

'It seems that John is angry because Bill acted against Mary.'

All of these examples are mildly deviant. Based on such observations, scrambling has been assumed to be subject to the DBs (see, among others, Saito (1985)).

I argue contra this widely-accepted view that Japanese scrambling is not subject to the DBs. First, although the examples in (3) and (4) are awkward, they are much better than the normal DB violations induced by English overt wh-movement.²

Second, there are cases where scrambling out of an adjunct is legitimate (cf. Ikawa (1996)).³

²See Fukui and Saito (1996) for a similar observation.

³One might argue that in (5a-b), the *ni*-phrase is short-scrambled within the adjunct and the subject of the adjunct is not really empty but is the overt subject in each example. Note, however, that (5a-b) are obligatory control structures and thus the adjunct subjects may not be overt, as shown by (ia-b):

(i) a. *John-ga [Mary-ga sono isu-ni suwari nagara] hon-o
 -Nom -Nom that chair-Dat sit while book-Acc
 yondeita (koto)
 was reading (fact)
 'John was reading a book while Mary was sitting on that chair.'

ISLANDS AND MOVEMENT TYPES

- (5) a. Sono isu-ni; [Johnj-ga [*e*_j *t*_i suwari nagara] hon-o
 that chair-Dat -Nom sit while book-Acc
 yondeita]] (koto)
 was reading (fact)
 'John was reading a book while sitting on that chair.'
- b. Siai-no kekka-ni; [Johnj-ga [*e*_j *t*_i totemo gakkarisite]
 game-Gen result-Dat -Nom very disappointed
 kyuujoo-o atonisita] (koto)
 ball park-Acc left (fact)
 'John left the ball park, disappointed about the result of the
 game.'

In (5a, b), although the clause-initial phrase is extracted out of the adjunct through scrambling, the result is acceptable. The difference between (4) and (5) resides in the fact that while the adjunct clause in the former has an overt subject, the one in the latter has an empty subject. While scrambling out of the adjunct with an empty subject is legitimate, scrambling out of the one with an overt subject is not. It should be noted that in the case of English overt wh-movement, extraction out of an adjunct is illegitimate even when the adjunct has an empty subject:

- (6) *?What did John arrive yesterday, [sad about *t*]?

I take these facts as evidence to suggest that unlike English overt wh-movement, Japanese scrambling is not subject to the DBs. I will argue later that the mildly deviant status of examples like (3) and (4) should be attributed to a condition in the phonological component.

2. The DB Effects in Japanese

I have argued in the last subsection that unlike English overt wh-movement, Japanese scrambling does not exhibit any DB effects. This might lead one to claim that unlike in English, the DBs are inert in Japanese. There is, however, empirical evidence to suggest that the DBs are also operative in Japanese.

-
- b. *John-ga [Mary-ga siai-no kekka-ni totemo gakkarisite]
 -Nom -Nom game-Gen result-Dat very disappointed
 kyuujoo-o atonisita (koto)
 ball-park-Acc left (fact)
 'John left the ball park, with Mary disappointed about the result of
 the game.'

TORU ISHII

Evidence for the existence of the DB effects in Japanese comes from empty operator movement in Japanese. It has been claimed that empty operator movement is involved in the cleft construction with an NP-Case or PP focus (Hoji (1990)), the *tough* construction with a PP subject (Takazawa (1987)), and the comparative deletion construction (Ishii (1991) and Kikuchi (1987)). Exactly like overt wh-movement in English, empty operator movement involved in these constructions exhibit the DB effects. Let us look at the cleft construction with an NP-Case or PP focus as an example. Empty operator movement in the cleft construction is subject to the CNPC, as shown by (7a) (taken from Hoji (1990:CH5, 31)) and (7b):

- (7) a. Relative Clauses
 ?[Op_i* [*John-ga* [*e_j* *t_i* *atta-koto-ga aru*] *nihonzin_j*]-o oozei
 -Nom have met Japanese-Acc many
sitte iru] no]-wa Russell_i-ni da
 knows Comp-Top Russell-Dat be
 'Lit. It is Russell_i that John knows many Japanese that have
 met *e_j*.'
- b. Non-relative Complex NPs
 ?[Op_i* [*John-ga* [*Mary-ga t_i* *atta to yuu*] *uwasa*]-o
 -Nom -Nom met Comp say rumor-Acc
kiita] no]-wa Bill_i-ni da
 heard Comp-Top -Dat be
 Lit. 'It was Bill_i that John heard a rumor (which says) that
 Mary met *e_j*.'

It should be noted that (7a, b) are as severely deviant as the CNPC violations induced by overt wh-movement in English.

Empty operator movement in the cleft construction is also constrained by the Adjunct Condition, as shown below:

- (8) a. *? [*Op*_i [John_j-ga [_{*e_j*} *t_i* suwari nagara] hon-o
-Nom sit while book-Acc
yondeita] no]-wa [sono isu]_j-ni da
was reading Comp-Top that chair-Dat be
'It was on that chair_j that John was reading the book while
sitting *e_i*.'

ISLANDS AND MOVEMENT TYPES

- b. **[Op_i [John_j-ga [_{e_j} *t_i* totemo gakkarisite] kyuujoo-o
 -Nom very disappointed ball park-Acc
 atonisita] no]-wa [siai-no kekka]-ni da
 left Comp-Top game-Gen result-Dat be
 'It was the result of the game_j that John left the ball park,
 disappointed about *e_j*.'*

It should be noted that in (8), an empty operator is extracted out of the adjunct with an empty subject. The result is as severely deviant as the adjunct condition violation induced by overt wh-movement in English. This is in contrast with the lack of the adjunct condition effects with scrambling in (5).

Another evidence for the existence of the DBs in Japanese can be formulated in relation to focus scrambling. I have shown in the previous section that scrambling does not exhibit any DB effects. When scrambling takes place in what Kuroda (1972, 1979) calls generic sentences, however, the DB effects emerge. Let us first consider the CNPC:⁴

- (9) a. Relative Clauses
 *?Nani-ni otokonoko-ga [[*t* nari tagaru] hito]-o
 what-Dat boy-Nom become want-to person-Acc
 taitei keibetu suru no
 usually look-down-on Q
 Lit. 'What, boys usually look down on the person who wants
 to become *t*'
- b. Non-relative Complex NPs
 *?Nani-ni oya-ga [[kodomo-ga *t* mutyuuni natte iru]
 what-Dat parent-Nom child-Nom mad-about
 to yuu zizitu]-o taitei siranai no
 Comp say fact-Acc usually do-not-know Q
 Lit. 'What, parents do not know the fact that their children
 are mad about *t*'

In (9), the matrix clause, at which level scrambling takes place, counts as generic, since it makes a statement about a habitual state of affairs. It should be noted that (9a, b) are as severely deviant as the CNPC

⁴As will be discussed below, generic non-interrogative sentences are awkward as independent sentences without any contexts due to their obligatory focus readings. To avoid the awkwardness which comes from this factor, generic *wh*-questions are taken to illustrate the locality effects here and in (10).

TORU ISHII

violations induced by overt wh-movement in English. This suggests that unlike scrambling in non-generic sentences, scrambling in generic sentences is subject to the CNPC.

Scrambling in generic sentences is also constrained by the Adjunct Condition:

- (10) a. *?Nani-ni_i onnanoko-ga [*e_j t_i* akogare nagara] taitei
 what-Dat girl-Nom long-for while usually
 ookiku naru no
 grow-up Q
 Lit. 'What, girls usually grow up while longing for *t*?'
 b. *?Nani-ni_i otokonoko-ga [*e_j t_i* katto natte] taitei
 what-Dat boys-Nom angry-at usually
 kenka suru no
 have a quarrel Q
 Lit. 'What, boys usually have a quarrel, angry at *t*?'

In (10), the matrix clause, at which level scrambling takes place, counts as generic. It should be noted that in (10), the element is scrambled out of the adjunct with an empty subject. The result is as severely deviant as the adjunct condition violation induced by overt wh-movement in English. This is in contrast with the lack of the adjunct condition effects in non-generic sentences in (5).

I argue that the contrast between scrambling in generic and non-generic sentences with respect to the DB effects comes from the fact that scrambled phrases receive obligatory focus readings in the former but not in the latter. Before we come to that, it is necessary to consider obligatory focus readings in generic sentences. Kuno (1973) and Kuroda (1972, 1979) observe that Japanese nominative phrases in generic sentences can only be interpreted as having focus readings, presenting examples like (11):⁵

- (11) a. Inu-ga neko-o oikakeru (koto)
 dog-Nom cat-Acc chase (fact)
 'Dogs chase cats.'
 (adapted from Kuroda 1979:7)

⁵Kuno (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.

ISLANDS AND MOVEMENT TYPES

- b. **Kigyoo-ga** seizika-ni kenkin si tagaru (koto)
 company-Nom politician-Dat contribute-money want-to (fact)
 'Companies want to contribute money to politicians.'

(11a, b) are both generic sentences. The nominative phrases in these generic sentences must be interpreted as having focus readings. The more precise translations of (11a, b) are therefore as follows:

- (12) a. It is dogs that chase cats; (Of all the animals we are talking about) dogs and only dogs chase cats.
 b. It is companies that want to contribute money to politicians; (Of all the people or groups of people we are talking about) companies and only companies want to contribute money to politicians.

Kuno and Kuroda observe that examples like (11) 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 numerals or quantifiers. They become natural in contexts which solicit the focus readings of the nominative phrases. For example, (11a, b) are natural when they are given as answers to the following generic questions:

- (13) a. **Nani-ga** neko-o oikakeru no
 what-Nom cat-Acc chase Q
 'What chases cats?'
 b. **Dare-ga** seizika-ni kenkin si tagaru no
 who-Nom politician-Dat contribute-money want-to Q
 'Who wants to contribute money to politicians?'

As observed by Matsuda (1996), focused expressions in generic sentences are not limited to nominative phrases. Accusative and dative phrases also receive obligatory focus readings in generic sentences when they appear in sentence-initial positions:

- (14) a. **Neko-o** inu-ga oikakeru (koto)
 cat-Acc dog-Nom chase (fact)
 'It is cats; that dogs chase *e_i*.'
 b. **Seizika-ni** kigyoo-ga kenkin si tagaru (koto)
 politician-Dat company-Nom contribute-money want-to (fact)
 'It is to politicians; that companies want to contribute money *e_i*.'

TORU ISHII

In (14), the focus readings of the nominative phrases are neutralized. The nominative phrases *inu-ga* 'dog-Nom' and *kigyoo-ga* 'company-Nom' are no longer forced to have focus readings. Instead, the scrambled phrases, *neko-o* 'cat-Acc' in (14a) and *seizika-ni* 'politician-Dat' in (14b), obligatorily have focus readings. This is supported by the fact that (14a, b) are both awkward, if not ungrammatical, as independent sentences without any contexts. They become natural when they are given as answers to the following generic questions:

- (15) a. **Nani-o** inu-ga oikakeru no
 what-Acc dog-Nom chase Q
 'What do dogs chase?'
 b. **Dare-ni** kigyoo-ga kenkin si tagaru no
 politician-Dat company-Nom contribute-money want-to Q
 'Who do companies want to contribute money to?'

These observations suggest that scrambling in generic sentences obligatorily induce focus readings. One can safely say that the existence of the DB effects with scrambling in generic sentences is due to the obligatory focus readings of the scrambled phrases.⁶

The relation between the DB effects and focus readings is further confirmed by examples like (16, 17):

(16) Complex NP Constraint

- a. Relative Clauses
 *?MARY-NI John-ga [[t atta] hito]-o sagasite iru
 -Dat -Nom met person-Acc looking-for
 (koto)
 (fact)
 Lit. 'It is Mary_i that John is looking for the person who met
 e_i.'

⁶Further support for our view comes from the fact that examples like (i) are awkward:

- (i) ??Neko-o nani-ga oikakeru no
 cat-Acc what-Nom chase Q
 'What chases cats?'

It has been widely accepted that wh-phrases get focus readings. Under our view, since (i) is generic, the scrambled phrase *neko-o* 'cat-Acc' receives an obligatory focus reading. Then, (i) has two focalized constituents, which violates the condition that only one constituent can get focalized within a sentence.

ISLANDS AND MOVEMENT TYPES

- b. Non-relative Complex NPs
 *?BILL-NI John-ga [[Mary-ga *t* atta to yuu]
 -Dat -Nom -Nom met Comp say
 uwasa]-o kiita (koto)
 rumor-Acc heard (fact)
 Lit. 'It was Bill_i that John heard a rumor (which says) that
 Mary met *e_i*.'

(17) Adjunct Condition

- a. *?SONO ISU-NI_i [John-j-ga [*e_j* *t_i* suwari nagara]
 that chair-Dat -Nom sit while
 hon-o yonde ita] (koto)
 book-Acc was reading (fact)
 'Lit. It was on that chair_i that John was reading a book while
 sitting *e_i*.'
- b. *?SLAI-NO KEKKA-NI_i John-j-ga [*e_j* *t_i* totemo
 game-Gen result-Dat -Nom very
 gakkarisite] kyuujoo-o atonisita] (koto)
 disappointed ball park-Acc left (fact)
 Lit. 'It was the result of the game_i that John left the ball park,
 disappointed about *e_i*.'

In (16, 17), the scrambled phrase receives a focus reading through stress assignment. Although (16, 17) are non-generics, they are deviant especially when we pronounce them with a pause after the scrambled phrases. This indicates that the focus readings of the scrambled phrases induce the DB effects.⁷

⁷If 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.'

NP-o 'NP-Acc', however, can function as a kind of topic especially when it gets focalized. In (i), *sono hon-o* 'that book-Acc' is not scrambled but base-generated in its surface position as a topic. The empty category within the relative clause is not a trace but an empty pronominal, as represented below:

- (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.'

This is supported by the fact that clause-initial accusative phrases allow resumptive pronouns (though resumptive pronouns are always marginal in

TORU ISHII

It follows from what has been said that there exists an asymmetry with the DB effects among movement types rather than between English and Japanese. Overt wh-movement, empty operator movement, and focus scrambling obey the DBs. Scrambling, on the other hand, is immune from the DBs. It has been claimed that overt wh-movement and empty operator movement are driven by strong features of C. Furthermore, it is plausible to assume that focus scrambling is driven by a strong feature associated with focus.⁸ Hence, the movement operations which exhibit the DB effects are all feature-driven. Turning to scrambling, Fukui (1993), Fukui and Saito (1996), and Saito (1994) argue that scrambling in Japanese is optional and thus not triggered by any formal feature. If these conjectures are correct, we can say that the asymmetry with the DB effects exists between feature-driven movement and non-feature-driven movement. While the former exhibits the DB effects, the latter does not. No previous theories of the DBs can ever explain this asymmetry. In the next section, I will propose a derivational approach to SRs, arguing that it gives a new minimalist way of explaining the DB effects. It is also shown that our theory of the DBs accommodates the hitherto unexplained asymmetry with the DB effects between the two types of movement.

3. An Account of the DBs

3.1. The Derivational Selectional Restriction

Before turning to an account of the DBs, it is necessary to explicate the notion SR. In the following discussion, I use the notion SR as a cover term for the θ -role assignment properties of thematic items and the categorial selection properties of functional items.⁹ Within the Extended Standard Theory (EST), SRs were assumed to be satisfied at D-structure

Japanese):

(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.'

⁸I claim that focalized phrases also have focus features, which are to be checked with strong focus features by the application of focus scrambling.

⁹It is important to note that the present definition of SR 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].

ISLANDS AND MOVEMENT TYPES

(see, among others, Chomsky (1981)). Within the Minimalist Program (MP) where D-structure is abandoned, SRs should be reformulated either as conditions on interface levels or constraints which apply throughout derivations. Chomsky (1993, 1995b) pursues the former approach, claiming that SRs should be satisfied at LF. I rather pursue the latter, arguing that SRs should be satisfied derivationally. Specifically, I propose the Derivational Selectional Restriction (DSR) (18):

(18) The Derivational Selectional Restriction (DSR)

When a derivation D comes to a stage where it is possible to satisfy an SR, the SR must be satisfied immediately. Otherwise, D is canceled.

In this subsection, I will argue that the DSR enforces cyclic merger of arguments and postcyclic merger of adjuncts. Before we illustrate it, let us present several assumptions which are necessary for the understanding of the following discussion.

First, I use the term "operation" to refer to term-manipulation. Attract and Merge count as operations, since they manipulate terms in phrase structures. Select, on the other hand, does not count as an independent operation. This is because it only introduces a lexical item from a numeration (N) without manipulating any terms in phrase structures. Rather, I assume with Collins (1997) that Select is a part of Merge.

Second, thematic and functional items satisfy their SRs in different ways. Functional items like C, T, a light verb *v*, and D satisfy their SRs by taking specific categories as their complements. For example, a light verb *v* selects V^{\max} as its complement. T selects v^{\max} as its complement. C selects T^{\max} as its complement. Thematic items like N, V, A, and P, on the other hand, satisfy their SRs by the assignment of their θ -roles. I claim that each thematic item has an explicit representation of all of its θ -roles which are to be assigned to its arguments, called a θ -grid in the sense of Stowell (1981). Essentially following Fukui (1986), let us assume that the θ -roles in a θ -grid are structured according to the "closeness" of a θ -role to the computational system CHL .¹⁰ This is represented by the linear order of the θ -role in a θ -grid. The lefthand θ -

¹⁰The structure of θ -roles within a θ -grid corresponds to the traditional argument structure, which is advocated by, among others, Grimshaw (1990). It should be noted that such a structure within a lexical entry is needed in any theory, since it is necessary to make a distinction between "external" and "internal" arguments.

TORU ISHII

role is "closer" than the one to its right in a θ -grid. The assignment of the θ -roles in a θ -grid takes place sequentially from left to right. The rightmost θ -role of transitive and unergative verbs is an "external" θ -role. I assume with Chomsky (1995b) that an "external" θ -role is assigned to an element in the Spec of a light verb v . This may only take place when a verb raises to adjoin to v , forming the amalgamated verbal element [v V- v]. Since this movement takes place overtly during structure-building, we assume that it is triggered by the strong V-feature of v . As an example, let consider the verb *put*, which has the following θ -grid (where L is Locative, Th is Theme, A is Agent):

(19) *put*: <L, Th, A>

According to (19), Locative is "closest" to C_{HL} and thus assigned first. After that, Theme is assigned. Hence, the verb *put* assigns Locative to its complement and Theme to its specifier. Since Agent is an "external" θ -role, it is assigned to the Spec of v after *put* raises to v .

Third, we assume a new derivational definition of the notion "strong feature" (SF). Chomsky (1995b) also proposes a derivational definition of the notion SF, claiming that SFs should be defined as those that derivations "cannot tolerate" in the sense stated in (20):

(20) Suppose that a derivation D has formed a structure containing α with an SF. Then, D is canceled if α is in a category not headed by α .
(adapted from Chomsky (1995b:234))

He claims that (20) can derive two properties of strength. First, it triggers an overt operation. Second, it induces cyclicity.

Chomsky's definition, however, cannot fully capture the properties of strength. Specifically, contrary to Chomsky's claim, there is a pattern of cases where (20) cannot trigger overt operations. (20) cannot trigger root overt operations like overt wh-movement in the matrix clause. Let us consider (21) as an example:

(21) What did you read *t*?

During the derivation of (21), we come to the stage where C with a strong Q-feature appears:¹¹

¹¹Essentially following Chomsky (1995b) and Muysken (1982), we define the notion of maximal projection derivationally. In (22), for instance, the dominating node is assigned the categorial status of C^{max} , since it is the top

ISLANDS AND MOVEMENT TYPES

- (22) [C^{\max} C_[Q] [you read what]]

In order to derive (21), we have to raise the wh-phrase *what* to the Spec of C^{\max} and check the strong Q-feature of C. (20), however, cannot trigger this overt wh-movement. If the wh-phrase *what* did not raise to the Spec of C^{\max} , then the strong Q-feature would remain. According to (20), however, this derivation would not be canceled. This is because the C^{\max} is the root clause and never contained in another category. According to the principle of Procrastinate, which prefers covert operations to overt operations, the relevant feature of *what* should raise in the covert component to check the Q-feature of C. There would be no way to trigger root wh-movement before Spell-Out. Hence, triggering an overt operation, a property of strength, does not follow from (20).

In order to solve this difficulty, I develop Chomsky's derivational definition of SF and propose (23):

- (23) When a derivation D has formed a structure which contains α with an SF, the SF must be checked immediately. Otherwise, D is canceled.

Let us consider structure (22) again. (23) requires that when (22) is formed, the strong Q-feature of C should be checked immediately by the raising of the wh-phrase *what* to the Spec of C^{\max} , correctly triggering root wh-movement before Spell-Out. Hence, (23) can correctly trigger an overt operation, while inducing cyclicity.¹²

Let us now illustrate how the DSR (18) enforces cyclic merger of arguments and postcyclic merger of adjuncts, taking the reading of (24) where the adjunct *yesterday* modifies the embedded clause as an example:

- (24) Bill said that John saw Mary yesterday.

I claim following Collins (1997) that each stage of a derivation can be characterized a set of syntactic objects (phrase structures) already formed

node of the C projection at this stage. If it further projects up, its categorial status will change to an intermediate projection of the C rather than remain as a maximal projection of the C.

¹²When a functional head has more than one SFs, like T in languages with overt subject-raising (attracted by a strong D-feature) and overt V-raising (attracted by a strong V-feature), we assume a structure among SFs which ensures that only one SF is available for checking at one time.

TORU ISHII

and the remaining part of N. At the initial stage of the derivation of (24), there is no formed syntactic object. It only consists of the N of (24), which can be represented as below:

- (25) $N = \{(Bill, 1), (T, 1), (v, 1), (say, 1), (that, 1), (John, 1), (T, 1), (v, 1), (see, 1), (Mary, 1), (yesterday, 1)\}$

Among its members is the verb *see*, which has the θ -grid (26):

- (26) *see*: $\langle Th, A \rangle$

At this initial stage, we have an option of applying the merge operation to *see* and *Mary* for satisfaction of the Theme θ -role of *see*. Recall that although we have to select *see* and *Mary* from the N before merger of these two items, Select does not count as an independent operation but as a part of the merge operation. The DSR requires that we should apply the merge operation immediately. Hence, we apply the operation, resulting in (27):

- (27) $[v^{max} \text{ see Mary}]$

One might claim that when the verb *see* is selected, the DSR would require the verb *see* to be merged with the light verb *v* for satisfaction of the SR of the latter. This is because the verb *see* may count as V^{max} as well as V^0 under the bare phrase structure. I claim, however, that an element in the syntactic object already formed is "closer" to C_{HL} than the one which is still in N.¹³ This prevents the verb *see* from getting merged with the light verb *v* at that stage, since *see* is "closer" to C_{HL} than the light verb *v*.

When we construct (27), the stage of the derivation can be characterized as the syntactic object (27) and the remaining part of the N. At this stage, we have an option of applying merger of *v* and *see Mary* for satisfaction of the SR of *v*. According to the DSR, we should apply that merge operation immediately, yielding (28):

- (28) $[v^{max} v [v^{max} \text{ see Mary}]]$

¹³It is desirable that this assumption should be derived from some deeper principles. I leave this subject for future research.

ISLANDS AND MOVEMENT TYPES

Recall that the Agent θ -role of *see*, being an "external" θ -role, is not available for satisfaction when we construct (27), since the verb *see* has not been adjoined to the light verb yet.

When we construct (28), our definition of SF (23) requires that the verb *see* should raise to adjoin to the light verb v for checking the strong V-feature of v . This yields (29), given the copy theory of movement proposed by Chomsky (1993):

(29) [v^{\max} [v see- v] [v^{\max} see Mary]]

One might claim that when we construct (28), the DSR requires that we should merge T and (28). This is because when we construct (28), we have an option of applying the merge operation for satisfaction of the SR of T, which states that T takes v^{\max} as its complement. Recall, however, that an element in the syntactic object already formed is "closer" to C_{HL} than the one which is still in N. Hence, we should check the strong V-feature of v , preventing the SR of T from being satisfied by merger of T and (28) at that stage.

When we construct (29), we have an option of applying merger to *John* for satisfaction of the Agent θ -role of *see*. According to the DSR, we should apply the merge operation immediately, yielding (30):

(30) [v^{\max} John [[v see- v] [v^{\max} see Mary]]]

Note that we may not apply an operation for satisfaction of the SR of T at this stage, since *see*, which is in the syntactic object already formed, is "closer" to C_{HL} T, which is still in the N.

When we construct (30), the DSR requires that we should apply merger of T and (30) immediately:

(31) [T^{\max} T [v^{\max} John [[v see- v] [v^{\max} see Mary]]]]

Since T has a strong D-feature, our definition of SF (23) requires that it should be checked immediately. Hence, we check the strong D-feature by raising *John* to the Spec of T^{\max} . The resulting structure is (32):

(32) [T^{\max} John [T [v^{\max} John [[v see- v] [v^{\max} see Mary]]]]]

Recall that we cannot satisfy the SR of C by merger of C and (31) when we construct the latter, since T, which is in the syntactic object already formed, is "closer" to C_{HL} than C, which is still in N.

TORU ISHII

Suppose that the adjunct *yesterday* is adjoined to the embedded T^{\max} (32) for its proper licensing.¹⁴ Then, there are two logically possible continuations when we construct (32): (i) Selection of C and merger of C with the T^{\max} (32), and (ii) Merger of the adjunct *yesterday* and the T^{\max} (32). The DSR requires that we should choose the former, since it satisfies the SR of C:

- (33) [C^{\max} C [T^{\max} John [T [v^{\max} John [[v see-v] [v^{\max} see Mary]]]]]]

The derivation proceeds further in accordance with the DSR, resulting in (34):

- (34) [T^{\max} Bill [T [v^{\max} Bill [[v say-v] [v^{\max} say [C^{\max} C [T^{\max} John [T [v^{\max} John [[v see-v] [v^{\max} see Mary]]]]]]]]]]]]]]]]]]]]

It is important to point out that until this final stage of the derivation, the adjunct *yesterday* has not been allowed to be merged with the main structure due to the DSR. At this final stage, we can combine the adjunct *yesterday* with the main structure (34) by adjoining the former to the embedded T^{\max} of the latter. Hence, if we conform to the DSR during a

¹⁴Specifically, I assume with Higginbotham (1985) and Travis (1988) that modification relations like adjective-noun and adverb-verb/clause are established by θ -identification. In the relation between the adverb *yesterday* and the embedded clause in (24), for example, it is conceivable that the event position in the latter is identified with the θ -role of the former. Putting technical details aside, modifiers and modifiees must appear within a certain local configuration for θ -identification. This brings about restrictions on the positioning of adverbs. It ensures that *yesterday* is adjoined to the embedded T^{\max} in (24). It also ensures, for example, that the "VP-adverb" *merely* must appear within the projection of V, as shown in (i):

- (i) a. John is merely being a fool.
b. *Merely John is being a fool.

Note that θ -identification is a symmetric relation between two θ -roles. Both modifiers and modifiees have properties that need to be satisfied. Selection, on the other hand, is an asymmetric relation in the sense that a head, which has a property that needs to be satisfied, asymmetrically selects an element. Crucially, the selected element does not have any property which needs to be satisfied. Hence, it is plausible to claim that SRs include θ -assignment but not θ -identification. Then, since θ -identification is not subject to the DSR, modifiers are merged postcyclically.

ISLANDS AND MOVEMENT TYPES

derivation, arguments are required to be merged cyclically while adjuncts are required to be merged postcyclically, i.e., after argument-of relations are established.¹⁵

In the rest of this section, I will argue that our DSR theory of phrase structure explains the asymmetry with the DB effects between feature-driven movement and non-feature-driven movement, which is most clearly observed in Japanese. Since space is limited, I will only discuss the DBs observed in Japanese, i.e., the CNPC and the Adjunct Condition. See Ishii (1997) for detailed discussion of the DSR analysis of the Subject Condition.

3.2. The DB Effects with Feature-Driven Movement

3.2.1 The Complex NP Constraint (CNPC)

We will begin by considering the relative clause case of the CNPC, taking the English overt wh-movement case (35) as an example:

(35) *?Who do you like [books that criticize *t*]?

(35) can be divided into two parts, i.e., the main structure *you like books* and the relative clause *that criticize who*. Since the relative clause is an adjunct, the DSR requires that it should be merged with the main structure postcyclically. Crucially, the relative clause has not been merged with the main structure when the strong Q-feature of C in the latter is to be checked. In other words, the relative clause and the main structure each constitute an independent syntactic object at that stage of the derivation. 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 by the raising of *who* to the Spec of C^{max} due to the c-command requirement on movement. Since the strong Q-feature cannot be checked immediately, it violates our definition of SF. To put it

¹⁵Our DSR analysis is incompatible with Boskovic and Takahashi (1995), which claims that "scrambled" phrases originate in their surface positions and then undergo LF-movement to the positions where they receive their θ -roles. Their analysis suffers from the following empirical problem. Under their analysis, "scrambled" phrases are expected to appear in any base-generated positions. Hence, there would not be any locality effects with scrambling whatsoever. As pointed out above, however, scrambling exhibits some locality effects, though they are weaker than those induced by feature-driven movement like English overt wh-movement.

TORU ISHII

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 (35), an example of the relative clause case of the CNPC, straightforwardly follows.

Let us next consider the non-relative case of the CNPC, taking the English overt wh-movement case (36) as an example:

(36) *?What did you study [the evidence that Harry stole *t*]?

We assume with Grimshaw (1990) and Stowell (1981) that the head nouns of non-relative complex NPs like *evidence* in (36) do not assign any θ -roles to the following clauses. The relation between the nouns and the following clauses is an appositive one rather than one of a θ -role assignment. Then, (36) can be divided into two parts, i.e., the main structure *you studied the evidence* and the appositive clause *that Harry stole what*. Since the appositive clause is an adjunct, the DSR requires that it should be merged with the main structure postcyclically. Crucially, the appositive clause has not been merged with the main structure when the strong Q-feature of C in the main structure is to be checked. The strong Q-feature of C cannot be checked by the raising of *who*; this violates our definition of SF. The non-relative clause case of the CNPC follows.¹⁶

3.2.2 The Adjunct Condition

The Adjunct Condition can be analyzed essentially along the same line as the CNPC. Let us consider the English overt wh-movement case (37) as an example:

(37) *?Who did you get jealous [because I spoke to *t*]?

(37) can be divided into two parts, i.e., the main structure *you got jealous* and the adjunct clause *because I spoke to who*. According to the DSR, the adjunct clause is required to be merged with the main structure

¹⁶Our analysis leaves unexplained a contrast in acceptability between extraction out of non-relative Complex NPs like (36) and extraction out of real appositives like (i):

(i) *Who did you see John, who loves *t*?
I leave this subject for future research.

ISLANDS AND MOVEMENT TYPES

postcyclically. Hence, the strong Q-feature of C cannot be checked by the raising of the wh-phrase *who* to the Spec of C^{max}. This violates our definition of SF. The Adjunct Condition straightforwardly follows. Although we have only considered the DB effects with English overt wh-movement, those with empty operator movement and focus scrambling can be explained in the same way.¹⁷

3.3. Lack of the DB Effects with Non-Feature-Driven Movement

This subsection explicates how our DSR theory can explain the immunity of non-feature-driven movement from the DBs, taking scrambling in Japanese as an example. Let us consider the case where the adjunct condition effects are canceled, taking (5a) (repeated here as (38)) as an example:

- (38) Sono isu-ni_i [John_j-ga [_{e_j} _{t_i} suwari nagara] hon-o
 that chair-Dat -Nom sit while book-Acc
 yondeita]] (koto)
 was reading (fact)
 'John was reading a book while sitting on that chair.'

(38) can be divided into two parts, i.e., the main structure *John-ga hon-o yondeita* 'John was reading a book' and the adjunct clause *e_j sono isu-ni suwari nagara* 'while sitting on that chair'. According to the DSR, the adjunct clause is required to be merged with the main structure postcyclically. If scrambling were triggered by an SF (say, a [SCRAMBLING] feature), the adjunct would not have been merged with the main structure when the SF is to be checked. The strong [SCRAMBLING] feature could not be checked at that stage of the derivation. This would violate our definition of SF and thus the derivation would be canceled. We are assuming, however, that scrambling is not feature-driven. Scrambling takes place postcyclically exactly like merger of adjuncts.¹⁸ Hence, we may first merge the main

¹⁷Although Kayne's (1984) Connectedness Condition (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 proposed here is a minimalist reinterpretation of Kayne's CC. According to our theory, structures belonging to different g-projections in the sense of Kayne have not been merged with each other when feature-driven movement takes place. Hence, movement cannot take place across more than one g-projections, as Kayne's CC claims.

¹⁸It is possible to claim that the postcyclic application of scrambling captures

TORU ISHII

structure and the adjunct clause and then apply scrambling. We can therefore scramble *sono isu-ni* 'that chair-Dat' out of the adjunct clause *e_j sono isu-ni suwari nagara* 'while sitting on that chair'. The lack of the adjunct condition effects with scrambling straightforwardly follows. The lack of the CNPC effects with scrambling can be explained in a similar way.

To recapitulate the above discussion, I have argued that the asymmetry between feature-driven movement and non-feature driven movement with the DB effects straightforwardly follows from our DSR theory of the DBs. It should be noted that this asymmetry cannot be explained by any previous locality theories such as Chomsky (1986), Kayne (1983), and Lasnik and Saito (1992). 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. They cannot accommodate the asymmetry with the DBs without recourse to any extra devices.

4. The Apparent DB Effects with Scrambling in Japanese

I have shown in the last section that unlike feature-driven movement, non-feature-driven movement like Japanese scrambling is not subject to the DBs, arguing that its immunity from the DBs straightforwardly follows from our DSR analysis. As presented in section 1, however, there are cases where scrambling *prima facie* exhibits the DB effects. In this section, I will argue that the apparent DB effects with scrambling should be attributed to an A-over-A condition in the PF-component.

4. 1. An A-over-A Condition in the PF-component

The A-over-A condition, which was originally proposed by Chomsky (1964), was intended to capture the locality restrictions on overt movement. Although the formulation of the A-over-A condition

the insight given by, among others, Chomsky (1991) that optional operations like scrambling take place in the stylistic component. It has been claimed that stylistic rules are those which apply after cyclic rules like overt *wh*-movement and NP-movement. The "stylistic" characteristic of scrambling therefore follows from our analysis without assuming any extra component like the stylistic component.

ISLANDS AND MOVEMENT TYPES

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 DB effects with scrambling, I propose the following A-over-A condition:

(39) The A-over-A Condition

A PF representation is ruled out as illegitimate if it contains a structure of the following form:

$\alpha \dots [\beta \dots t_{\alpha} \dots] \dots$,

where α and β are both [-V].

(39) states that no [-V] category may be extracted out of another [-V] category. I argue that the A-over-A condition (39), which is representational in nature, applies at PF. A violation of (39) leads a derivation to crash at PF.^{19, 20}

Before we consider how the A-over-A condition (39) works, it is necessary to 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, *v*, and D. Let us assume following Abney (1987) and Fukui (1995) that we state this lexical/functional distinction by postulating a universal feature [+/- F]. It is widely accepted (see, among others, Chomsky (1972, 1981)) that lexical

¹⁹Strictly speaking, the A-over-A condition (39) 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 (1995b) 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 Postal (1966). Note in passing that such a constraint may not be formulated as applying at LF, since, as argued by Fukui and Saito (1996) and Saito (1989), scrambled phrases may be "radically reconstructed" to their original positions at LF.

²⁰It is plausible to claim that the A-over-A condition (39) does not count as a grammatical constraint but rather as a parsing constraint which belongs to the performance system. Note that the arguments to follow hold regardless of whether this condition is grammatical in nature or not. Crucial in the following discussion is the claim that the locality effects with scrambling should be attributed to a constraint residing outside the N → LF core computation.

TORU ISHII

categories are further crossclassified in terms of the two primitive features $[+/- N]$ and $[+/- V]$. The feature specifications of the lexical categories are given below:

(40) 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 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 is the path of nodes along which its descriptive content is "passed along." For example, 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 literally Clause = V^{max} or Nominal = N^{max} .²¹ Essentially following Abney, I claim that the notion of s-projection can be captured in terms of the feature system where the functional categories are also divided based on $[+/- N]$ and $[+/- V]$. The feature specifications of the functional categories are given below:²²

(41) Feature Specifications of the Functional Categories

- a. $D = [+F, +N, -V]$
- b. $C, T, v = [+F, -N, +V]$

These feature specifications of the functional categories capture the fact that while C, T, and v belong to the V system, D belongs to the N system.

²¹See Grimshaw (1991) for a similar insight. See, among others, Emonds (1985) and Jackendoff (1977) for the view that S is the maximal projection of V.

²²There are gaps in this feature system of functional categories. There is no functional category 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 count as categories with $[+F, -N, -V]$, since they have no close relation with the N system or the V system.

ISLANDS AND MOVEMENT TYPES

In these feature specifications, only N, P, and D have [- V] as its feature. It then follows from the A-over-A condition (39) 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 concerned, the A-over-A condition (39) claims that no N^{max} , P^{max} , or D^{max} may be extracted out of another N^{max} , P^{max} , or D^{max} .

Let us consider how the A-over-A condition (39) works. Let us first look at the apparent CNPC effects with scrambling, considering the relative clause case of the CNPC (42) (taken from Saito (1985:246)) as an example (the judgment is mine):

- (42) ?*Ano hon-o*_i [John-ga [*e_j t_i katta*] hito]_j-o sagasite iru
 that book-Acc -Nom bought person-Acc looking-for
 rasii]
 seem
 'It seems that John is looking for the person who bought that book.'

In (42), *ano hon-o* 'that book-Acc' is extracted out of the containing phrase [*e_j ano hon-o katta*] hito]_j-o 'the person who bought that book'. These 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]_j-o 'the person who bought that book' results in a representation which is ruled out as illegitimate by the A-over-A condition (39). Hence, the apparent CNPC effects with scrambling follows. Note that P^{max} scrambling out of the complex NP like (3a) (repeated here as (43)) can also be ruled out by the A-over-A condition (39):

- (43) ?*Mary-ni*_i [John-ga [*e_j t_i atta*] hito]_j-o sagasite iru rasii]
 -Dat -Nom met person-Acc looking-for seem
 'It seems that John is looking for the person who met Mary.'

This is because the scrambled P^{max} *Bill-ni* 'Bill-Dat' and the complex NP are both [-V].²³

²³As correctly pointed out by Kazue Takeda (p.c.), the intermediate adjunction of a scrambled phrase to a [-V] category would void the A-over-A condition effects, which is undesirable. I argue, however, that scrambling moves an element in one swoop. Let us assume that when a scrambled phrase adjoins to a

TORU ISHII

Let us turn to the apparent adjunct condition effects with scrambling, considering (44) (taken from Saito (1985:247)) as an example (the judgment is mine):

- (44) ?*Sono hon-o* [John-ga [minna-ga *t* kau node] tigau
 that book-Acc -Nom all-Nom buy because different
 hon-o katta] (koto)
 book-Acc bought (fact)
 'Because everyone buys that book, John bought a different one.'

Given that adjuncts like the one found in (44) count as P_{max} , *sono hon-o* 'that book-Acc' is extracted out of the P_{max} *minna-ga sono hon-o kau node* 'because everyone buys that book' in (44). Since both of these phrases are [-V], this extraction leads to a representation which violates the A-over-A condition (39). Hence, the apparent adjunct condition effects with scrambling follow. Note that P_{max} scrambling out of an adjunct like (4) (repeated here as (45)) can also be ruled out as illegitimate by the A-over-A condition (39):

- (45) ?*Mary-ni* [John-ga [Bill-ga *t* somuita node] okotte iru
 -Dat -Nom -Nom acted against because be angry
 rasii
 seem
 'It seems that John is angry because Bill acted against Mary.'

This is because the scrambled P_{max} *Mary-ni* 'Mary-Dat' and the adjunct P_{max} are both [-V].

This analysis also gives a way of explaining the contrast in acceptability between the DB effects with feature-driven movement and the apparent DB effects with non-feature-driven movement. Recall that the apparent DB effects with non-feature-driven movement like Japanese scrambling are much weaker than the DB effects with feature-driven movement like English overt wh-movement, empty operator movement, and focus scrambling. Recall that in the case of a DB violation, a derivation is canceled before the point of Spell-Out due to a violation of

category on the way to its final landing site, it leaves a trace there. Let us also assume with Chomsky (1995b) that no element may be adjoined to a maximal projection at LF due to the Full Interpretation. Then, the trace left by scrambling at an intermediate landing site must delete until LF. Given the plausible assumption that deletion is costly, it follows from the economy considerations that scrambling should move an element in one swoop.

ISLANDS AND MOVEMENT TYPES

our definition of SF. The derivation never reaches either of the interface levels. Scrambling in Japanese, on the other hand, is not feature-driven. Scrambling never induces a violation of our definition of SF and thus never causes a derivation to be canceled before the point of Spell-Out. Hence, even when scrambling exhibits the apparent DB effects and thus violates the A-over-A condition (39), the derivation at least 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 (1995b), the 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. 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 when a derivation violates a constraint in the core computation as in the case of a DB violation, the result is severely deviant. On the other hand, when a derivation violates a constraint in the mapping from Spell-Out to PF as in the case of a violation of the A-over-A condition (39), the result is mildly deviant. This is because such a derivation converges in the N → LF core computation. It then follows that the apparent DB effects with non-feature-driven movement like scrambling are much weaker than the DB effects with feature-driven movement like English overt wh-movement, empty operator movement, and focus scrambling.

Apart from the apparent DB effects with scrambling, the A-over-A condition (39) 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 DB effects, as shown below:

- (46) a. John-ga [[Bill-ga *e_j* kakusi motte ita] [Mary-e-no
 -Nom -Nom had-been-hiding -to-Gen
 tegami]_j]-o mituketa (koto)
 letter-Acc found (fact)
 'John found a letter to Mary which Bill had been hiding.'
 b. ?Mary-e-no_i [John-ga [[Bill-ga *e_j* kakusi motte ita]
 [*t_i* tegami]_j]]-o mituketa (koto)

In (46b), the P_{max} *Mary-e-no* 'Mary-to-Gen' is scrambled out of the nominal phrase [[*Bill-ga e_j kakusi motte ita*] [*Mary-e-no tegami*]_j]-o 'a letter to Mary which Bill had been hiding'. This violates the A-over-A

TORU ISHII

condition (39), since the scrambled phrase and the extraction domain are both [-V].²⁴

4.2. Scrambling out of Adjuncts with an Empty Subject

The last subsection has proposed the A-over-A condition (39), arguing that it accounts for the apparent DB effects with Japanese scrambling. As shown above, however, unlike scrambling out of the adjunct with an overt subject, scrambling out of the adjunct with an empty subject is acceptable. A question now arises as to why such cases do not violate the A-over-A condition (39). I argue that the A-over-A condition should be revised to be inert when the extraction domain is a predicate:

(47) The A-over-A Condition (Revised)

A PF representation is ruled out as illegitimate if it contains a structure of the following form:

$\alpha \dots [\beta \dots t_{\alpha} \dots] \dots$,

where α and β are both [-V], and β is not a predicate.

Before looking at how (47) works, let us consider the structure of the adjunct 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:

- (48) John_i-ga [_{NP} Mary-o naguru maeni] naiteita (koto)
-Nom -Acc hit before was crying (fact)
'John was crying before he hit Mary.'

In (48), the empty subject of the adjunct clause can only refer to the matrix subject *John* but not to anyone else. In other words, the adjunct with an empty subject involves obligatory control. Based on this observation, Hasegawa (1984/1985) argues that the empty subject in an

²⁴The A-over-A condition (39) would prohibit preposition stranding and extraction out of so called "picture-noun" phrases, which are allowed in languages like English. I claim that these phenomena are peripheral, voiding the effects of the A-over-A condition in terms of "marked" mechanisms like restructuring or readjustment. See, among others, Chomsky (1977) and van Riemsdijk (1978) for further discussion of this subject.

ISLANDS AND MOVEMENT TYPES

adjunct is not an empty pronominal. If it were an empty pronominal, it could refer not only to the matrix subject but also to 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:²⁵

- (49) a. **John_i-ga** [**PRO_i** Mary-o naguru maeni] naiteita
 b. **John_i-ga** [**Op_i** [**t_i** Mary-o naguru maeni]] naiteita

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 (49b). In (49b), the empty operator moves from the subject position of the adjunct to the Spec of the adjunct, where it is associated with the matrix subject *John* through predication. In other words, the adjunct in (49b), which minimally contains the null operator chain, functions as a predicate by virtue of the "open" position created by empty operator movement.

With the above discussion in mind, let us consider scrambling out of the adjunct with an empty subject again, taking like (38) (repeated here as (50)) as an example:

- (50) Sono isu-ni_i [John_j-ga [*e_j* *t_i* suwari nagara] hon-o
 that chair-Dat -Nom sit while book-Acc
 yondeita]] (koto)
 was reading (fact)
 'John was reading a book while sitting on that chair.'

Since the adjunct in (50) may involve empty operator movement, it may function as a predicate. Extraction out of the predicate does not result in a violation of the A-over-A condition (47). Hence, we can correctly predict that scrambling out of the adjunct with an empty subject like (50) is acceptable.²⁶

²⁵A similar analysis can be found in Clark (1985), though the latter claims that the adjunct with an empty subject obligatorily involves empty operator movement.

²⁶One might claim that adjuncts always count as predicates due to their modifying functions. Recall, however, that we are assuming that the relation between an adjunct and its modifiee is established by θ -identification not by predication (see note 14). Hence, adjuncts do not count as predicates unless they involve empty operator movement.

TORU ISHII

Note also that unlike scrambling out of the adjunct with an empty subject, scrambling out of the adjunct with an empty object is mildly deviant, as shown below:

- (51) 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 (51b), the P^{max} *kooen-de* 'park-in' is scrambled out of the adjunct with an empty object. I argue that this also follows from the A-over-A condition (47). Before turning to an analysis of (51b), 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, presenting the following example:

- (52) John-ga [Mary-ga *e* naguru maeni] naiteita
 -Nom -Nom hit before was crying
 'John was crying before Mary hit him.'
 (Hasegawa 1984/1985:290)

They observe that in (52), 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 an empty category within an adjunct, 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 (52), 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 pronominal, as represented below:

ISLANDS AND MOVEMENT TYPES

- (53) John_i-ga [Mary-ga *pro*_{ij} naguru maeni] naiteita
 -Nom -Nom hit before was crying
 'John_i was crying before Mary hit him_{ij}.'

Let us return to (51b). Since the adjunct with an empty object does not involve empty operator movement, it never functions as a predicate exactly like the adjunct without any empty category. Since both the adjunct and the scrambled phrase are [-V], scrambling of the p^{\max} *kooen-de* 'park-in' out of the adjunct results in a violation of the A-over-A condition (47). Hence, we can correctly predict that scrambling out of the adjunct with an empty object like (51b) is mildly deviant.

The reference to the notion of predicate in the A-over-A condition is motivated independently of facts on scrambling out of an adjunct. Nominal predicates, though being [-V], do not function as barriers for scrambling:

- (54) *Dono an-ni* John-ga [N^{\max} *t sansei*] na no
 which proposal-Dat -Nom in favor of be Q
 Lit. 'Which proposal, John is favor of *t*?''

Whatever analysis is to be adopted for nominal predicates, it is clear that *dono an-ni* 'which proposal-Dat', which is the argument of *sansei* 'in favor of', is extracted out of the N^{\max} . Note that examples like (54) would be wrongly ruled out unless we refer to the notion of predicate in the A-over-A condition.

This closes our discussion of locality on scrambling. It was shown that the locality effects with scrambling, which cannot be reduced to the properties of the core computation, should be attributed to the phonological component (or alternatively to the performance system (see note 20)). I must admit, however, that the above discussion is not explanatory but rather descriptive in nature. I leave deeper investigation of this subject for future research.

5. Conceptual Arguments

It was shown in the preceding sections that our derivational approach to SFs and SRs receives empirical support from the locality effects on movement. In this section, I will argue that our derivational approach also receives strong conceptual support.

Let us first consider our definition of SF (23). It only needs local considerations in the sense that it can decide whether to apply an

TORU ISHII

operation OP for checking an SF at a certain stage Σ of a derivation D only on the basis of information available at Σ . This is because (23) simply requires us to check the SF immediately when we form a structure containing the SF.²⁷ It is generally agreed that while global considerations necessarily induce computational complexity, local ones do not (see, among others, Chomsky (1995b)). Then, our definition of SF (23) enables us to avoid the problem of computational complexity, which is conceptually desirable.

The DSR also gains conceptual support if Chomsky (1995a) is correct in claiming that pure Merge for argument-of relations must be forced by SRs. To be specific, suppose that we select the verb *see*. Suppose further that we have an option of applying merger of *see* and *Mary* for satisfaction of the "internal" θ -role of *see*. Under the standard assumption that SRs apply at LF, global considerations are needed to decide whether to apply the merge operation. This is because we have to look ahead to see whether this merger satisfies the SR at LF. The DSR, on the other hand, requires *see* to be merged with *Mary* only on the basis of information available at the present stage. Hence, the DSR only needs local considerations.

I argue that our derivational approach to SFs and SRs supports Chomsky's (1991, 1995a) view that language is fundamentally computationally-intractable and thus unusable due to its inherent global properties, but there are local "computational tricks" which reduce computational burden and facilitate usability of language in practice. In other words, although considerations of computational complexity do not matter for fundamental aspects of language, they do matter for usable parts of language. The MP assumes that language is subject to "bare output conditions" (BOCs), which ensure that linguistic expressions, i.e. PF and LF, are legible to the external systems at the interface. Hence, under the BOC-driven view of language, both SFs and SRs should be subject to interface conditions as their fundamental properties. SFs are illegitimate at PF (see Chomsky (1993)). SRs must be satisfied at LF. These interface conditions, however, necessarily need global considerations and their corresponding optimization problems are computationally intractable. I argue that the DSR and our definition of SF serve as local "computational tricks" to solve these intractable problems induced by their fundamental properties, making that part of language usable in practice.

²⁷Note that Chomsky's definition of SF (20) is not local. It must look ahead to see whether the application/non-application of OP would result in a structure which violates (20) at the next stage of D.

6. Conclusion

I have argued that contrary to what has been claimed, there is an asymmetry with the DB effects among movement types. Specifically, it was shown that while feature-driven movement like English overt *wh*-movement, empty operator movement, and focus scrambling obeys the DBs, non-feature-driven movement like Japanese scrambling does not. I have proposed the DSR theory of the DBs, arguing that it gives a minimalist account of this asymmetry. As is well known, there are another two asymmetries with the DBs among movement types, which I have not discussed for lack of space. First, while overt argument *wh*-movement exhibits the DB effects, "covert argument *wh*-movement" (a *wh*-argument in-situ) does not. Second, unlike *wh*-arguments in-situ, *wh*-adjuncts in-situ exhibit the DB effects. See Ishii (1997) for detailed discussion of these two asymmetries with the DBs.

Our analysis also raises a lot of interesting issues. I will briefly point out two of them here. First, under our analysis, the DB and RM effects are given different treatments. While the former follows from the DSR, the latter like the *wh*-island constraint follows from the MLC. This dichotomy would be further supported if there is a type of movement which only obeys one of these but not the other. *Wh*-arguments in-situ in Japanese might come under this type. As mentioned above, *wh*-arguments in-situ do not exhibit any DB effects. Nishigauchi (1990) and Watanabe (1992) observe that *wh*-arguments in-situ in Japanese exhibit the *wh*-island effects. If their observations are correct, *wh*-arguments in-situ in Japanese obey the RM but not the DBs, providing further evidence for our analysis.²⁸

Second, our DSR analysis enables us to capture the argument/adjunct distinction in a minimalist way. It has been assumed in the pre-minimalist period (see, among others, Chomsky (1972) and Jackendoff (1977)) that the argument/adjunct distinction should be 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 examples:

(55) a. John hit Mary.

²⁸See Fukui (1997) for a different view on the dichotomy of islands. Under his view, the *wh*-island constraint and the CNPC are explained in terms of the feature version of the A-over-A principle. The CED (the Subject Condition and the Adjunct Condition), on the other hand, are explained in terms of the Principle of Unique Licensing and the MLC as proposed by Fukui and Saito (1996).

TORU ISHII

- b. John speaks eloquently.

While *Mary* in (55a) is the argument of the verb *hit*, *eloquently* in (55b) is an adjunct. Under Chomsky's (1972) X-bar theory where the uniform two-level hypothesis is adopted, for instance, (55a, b) are represented as in (56a, b), respectively, with the irrelevant parts being ignored:

- (56) a. [_V hit Mary]
 b. [_V [_V speaks] eloquently]

While *Mary* in (56a), being an argument, is attached under V', *eloquently* in (56b), being an adjunct, is attached under V".

Such a representational argument/adjunct distinction, however, is no longer available in the MP, where phrase structures should be "bare." Crucially, neither non-branching nodes 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. Our DSR 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 in conformity with bare phrase structure. Questions still remain how we explain the ECP effects with *wh*-variants of the adverbials based on our derivational notion of adjunct. I leave this question for future research.

References

- Abney, Steven 1987. *The English noun phrase in its sentential aspect*. Doctoral dissertation, MIT.
Boskovic, Zeljko and Daiko Takahashi 1995. *Scrambling and last resort*. Unpublished manuscript, University of Connecticut and Tohoku University.
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 1972. *Studies on semantics in generative grammar*. The Hague: Mouton.
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 1981. *Lectures on government and binding*. Dordrecht: Foris.

ISLANDS AND MOVEMENT TYPES

- Chomsky, Noam 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam 1991. 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 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 1995a. "Fall Lectures," ms., MIT. [Transcribed and edited by Robert Freidin.]
- Chomsky, Noam 1995b. *The minimalist program*. Cambridge, MA: MIT Press.
- 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.
- Clark, Robin 1985. *Boundaries and the treatment of control*. Doctoral dissertation, University of California, Los Angeles.
- Collins, Chris 1997. *Local economy*. Cambridge, MA: MIT Press.
- Emonds, Joseph 1985. *A unified theory of syntactic categories*. Dordrecht: Foris.
- Fukui, Naoki 1986. *A theory of category projection and its applications*. Doctoral dissertation, MIT.
- Fukui, Naoki 1993. Parameters and optionality. *Linguistic Inquiry*, 24, 399-420.
- 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 1997. Attract and the A-over-A principle. *UCI working papers in linguistics* 3, 51-67.
- Fukui, Naoki and Mamoru Saito 1996. *Order in the theory of phrase structure and movement*. University of California, Irvine and Nanzan University.
- Grimshaw, Jane 1990. *Argument structure*. Cambridge, MA: MIT Press.
- Grimshaw, Jane 1991. *Extended projection*. Unpublished manuscript, Brandeis University.
- Harada, Shin-Ichi 1977. "Nihongo-ni henkei-wa hituyoo-da" (There are transformations in Japanese). *Gengo*, 6, 88-103.
- Hasegawa, Nobuko 1984/1985. On the so-called 'zero pronouns' in Japanese. *The Linguistic Review*, 4, 289-341.
- Higginbotham, James 1985. On semantics. *Linguistic Inquiry*, 16, 547-593.
- 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.
- Ikawa, Hajime 1996. LF as a level without functional categories and its parallelisms with the overt syntax of a language without functional categories. *UCI working papers in linguistics* 1, 39-62.

TORU ISHII

- Ishii, Toru 1997. *An asymmetry in the composition of phrase structure and its consequences*. Doctoral dissertation, University of California, Irvine.
- Ishii, Yasuo 1991. *Operators and empty categories in Japanese*. Doctoral dissertation, University of Connecticut.
- Jackendoff, Ray 1977. *X-bar syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard 1984. *Connectedness and binary branching*. Dordrecht: Foris.
- Kikuchi, Akira 1987. *Comparative deletion in Japanese*. Unpublished manuscript, Yamagata University.
- 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. 1972. The categorial and the thetic judgments. *Foundations of language*, 9, 153-185.
- Kuroda, S.-Y. 1979. *The 'whole of the doughnut*. Belgium: E. Story-scientia P.V.B.A.
- Lasnik, Howard and Mamoru Saito 1992. *Move alpha: Conditions and its application and output*. Cambridge, MA: MIT Press.
- Matsuda, Yuki 1996. *A syntactic analysis of focus sentences in Japanese*. Unpublished manuscript, University of Southern California.
- Muraki, Masatake 1979. "On the rule scrambling in Japanese" In George Bedell, Eichi Kobayashi, and Masatake Muraki (eds.), *Explorations in linguistics: Papers in honor of Kazuko Inoue*. Tokyo: Kenkyusha
- Muysken, Pieter 1982. Parametrizing the notion "Head". *Journal of Linguistic Research*, 2, 57-75.
- Nishigauchi, Taisuke 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer.
- Riemsdijk, Henk van 1978. *A case study in syntactic markedness*. Foris: Dordrecht.
- Rizzi, Luigi 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Postal, Paul 1966. *The method of universal grammar*. Unpublished paper.
- Ross, John 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT.
- Saito, Mamoru 1985. *Some asymmetries in Japanese and their theoretical implications*. Doctoral dissertation, MIT.
- 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 1994. Improper adjunction. *MIT working papers in linguistics* 24, 263-293.
- Stowell, Timothy 1981. *Origins of phrase structure*. Doctoral dissertation, MIT.
- Takezawa, Koichi 1987. *A configurational approach to case-marking in Japanese*. Doctoral dissertation, University of Washington.
- Travis, Lisa 1988. "The Syntax of Adverbs," McGill Working Papers in Linguistics, 280-310, McGill University.

ISLANDS AND MOVEMENT TYPES

Watanabe, Akira 1992. *Wh-in-situ, subjacency, and chain formation*. MIT occasional papers in linguistics, Number 2.

Common Courses of Human Sciences
Kitami Institute of Technology
165 Koen-cho, Kitami, Hokkaido 090 JAPAN
ISHII-Toru/comm@king.cc.kitami-it.ac.jp