

TORU ISHII

The "Crossing" Constraint and the Minimal Link Condition*

1 Introduction

This paper investigates the so called "crossing effects" and argues that the MLC as part of the definition of Attract-F straightforwardly accounts for a hitherto unexplained asymmetry with the "crossing effects" between morphologically-driven A'-movement, *i.e.* wh-movement and null operator movement, and the other types of syntactic dependencies. It is shown that contrary to what has been claimed (see, among others, Chomsky and Lasnik (1993) and Chomsky (1995)), wh-island violations are not subsumed under the MLC. I will argue that wh-island violations should be accounted for by the "no extra baggage" condition. The analysis to be advocated provides a minimalist account of the contrast in acceptability between superraising and crossing violations, on the one hand, and wh-island violations, on the other.

The organization of this paper is as follows. Section 2 introduces some basic notions of the Minimalist Program. Section 3 deals with the asymmetry with "crossing effects" and provides an MLC account of them. Section 4 discusses wh-island violations. Section 5 summarizes the discussion.

2 The Minimalist Program

This paper adopts as its theoretical foundation the minimalist program proposed by Chomsky (1993) and further developed by Chomsky (1994, 1995). This section presents the minimalist assumptions which the discussion to follow is crucially based on.

2.1 Full Interpretation and Economy

Under the minimalist conception of language, each linguistic expression, *i.e.* an SD consisting of PF and LF, is "the optimal realization of the interface conditions, where "optimality" is

* I am grateful to Brian Agbayani, Lisa Cheng, Naoki Fukui, James Huang, Hajime Ikawa, and Miyoko Ikawa for helpful comments and discussion. Remaining errors and omissions are, of course, the sole responsibility of the author.

THE CROSSIG CONSTRAINT AND THE MLC

determined by the economy conditions of UG" (Chomsky 1993: 4).

The condition of Full Interpretation (FI) requires that SDs must solely consist of legitimate objects that can receive an interpretation at the relevant interface level. If a derivation (computation) yields an SD consisting of only legitimate objects, it converges; otherwise it crashes. A linguistic expression must also meet the economy conditions. The economy conditions require that the derivation of a linguistic expression must be "optimal."

In determining whether a derivation is "optimal," we have to fix the reference set. A reference set consists of competing derivations for economy. In evaluating derivations for economy, we only consider alternative derivations belonging to the same reference set. Chomsky (1994, 1995) claims that the reference set is determined by an initial numeration N . An initial numeration N is defined as a set of pairs (l, n) , where l is a lexical item in the lexicon and n is its index, interpreted as the number of times that l is selected from the lexicon during derivation. Hence, the economy conditions evaluate convergent derivations with the same initial numeration N , choosing one of them as "optimal" and blocking all the others.

2.2 Formal Features

Each lexical entry contains three collections of features, *i.e.* phonological, formal, and semantic features. Phonological features are stripped away by Spell-Out. The other features are left behind by Spell-Out. Among these three features, only formal features are visible to the computational system.

There are distinctions which cross-cut formal features. First, formal features are either intrinsic or optional. Intrinsic features are explicitly listed in the lexicon or strictly determined by the listed properties. Optional features are added when a lexical item enters into the initial numeration. Formal features are also classified into [+/-interpretable] features. [+Interpretable] features enter into interpretations at LF. [-Interpretable] features, on the other hand, are uninterpretable at LF and must be eliminated for convergence.

2.3 Attract-F and the "No Extra Baggage" Condition

Chomsky (1995) proposes the operation Attract-F, which is a reinterpretation of the operation of movement:

TORU ISHII

(1) Attract-F

K attracts F if F is the closest feature that can enter into a checking relation with a sublabel of K.

(Chomsky 1995: 297)

In (1), K is the target of the operation. A sublabel of K is a feature of the zero-level projection of the head of the target K. Note that since only formal features are visible to the computational system, F is restricted to formal features.

The notion of "closeness" in (1) is defined as follows:

- (2) If β c-commands α and τ is the target of raising, then β is closer to K than α unless β is in the same minimal domain as (a) τ or (b) α . (Chomsky 1995: 356)
- (3) Max (α) is the smallest maximal projection including α . (α is a feature or an X^0 category.) (Chomsky 1995: 299)
- (4) The domain $\delta(\alpha)$ is the set of categories included in Max (α) that are distinct from and do not contain α . (Chomsky 1995: 299)
- (5) The minimal domain Min ($\delta(\alpha)$) of α is the smallest subset K of $\delta(\alpha)$ such that for any $\gamma \in \delta(\alpha)$, some $\beta \in K$ reflexively dominates γ . (Chomsky 1995: 299)

The notion of Attract-F departs from the traditional theory of movement in that what is raised is not a category, but a feature. Chomsky (1995) argues that what is raised should be just the attracted feature 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:

(6) "No Extra Baggage" Condition

F carries along just enough material.

(cf. 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 FI and therefore crash. This is because raising of just F does not carry along any other material.

In the overt component, however, a category, not just F, raises to the target. Chomsky (1995) argues that this also follows from the "no extra baggage" condition. If only F were raised to the target K 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,

THE CROSSIG CONSTRAINT AND THE MLC

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 FI at PF and therefore crashes. Hence, in the overt component, an "extra baggage" is required for convergence; the whole category, but not just F, is raised. In the case of covert raising, on the other hand, the "no extra baggage condition" requires that only F should be raised, since no PF requirements are imposed in the covert component.¹

It should also be noticed that (1) incorporates the Minimal Link Condition (MLC), which informally states that the link of a chain should be minimal. The MLC, which was one of the economy measures (see, among others, Chomsky (1993, 1994) and Chomsky and Lasnik (1993)), becomes part of the definition of Attract-F. Under this view, the MLC is inviolable, since violation of the MLC is an illegitimate operation.

¹Chomsky (1995) argues that F automatically carries along FF (LI), the set of formal features of LI. Hence, even in the covert component, F never raises alone, but always carries along FF (LI). He presents the following binding and control facts as empirical evidence in favor of this view:

- (i) a. The DA [proved [the defendants to be guilty] during each other's trials]
- b. *The DA [proved [that the defendats are guilty] during each other's trials]
- (ii) 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 (ia), the Case and ϕ -features of *the defendants* are attracted by the matrix V. In (iia), the Case feature and ϕ -features of *there* are attracted by the T. He argues that these attracted features in (i-ii)(a) carry along FF (LI). Since FF (LI) includes A-position properties, which have the ability to serve as a binder or controller, the contrast between (a) and (b) follows.

This empirical argument, however, is inconclusive. First, it is not clear what formal feature counts as having A-position properties. Second, under the minimalist program, 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 LF representations. Rather, LF representations only provide instructions for the interpretive version of binding and control theories. Then, it is conceivable that the attracted Case and ϕ -features in (i-ii)(a) are sufficient to provide appropriate instructions for the 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). Notice that FF (LI) is not carried along for convergence. Hence, it counts as an "extra baggage." Such an "extra baggage" should be banned by the "no extra baggage" condition.

TORU ISHII

Let us finally look at the status of checked formal features. Chomsky (1995) argues that checked features are subject to deletion when possible. Deleted elements are furthermore subject to erasure when possible. Deletion is understood as invisible at LF but accessible to the computational system. Erasure, on the other hand, is understood as not only invisible at LF but also inaccessible to the computational system. According to the general principle of deletion/erasure, [+interpretable] features cannot delete even when they are checked. This is because they are visible at LF and thus deletion of such features leads to a violation of the overriding principle of recoverability of deletion. [-Interpretable] features, on the other hand, always delete when checked, since they are invisible at LF. Deleted [-interpretable] features are furthermore subject to erasure unless it would create an illegitimate syntactic object.

3 The "Crossing Effects"

This section investigates the "crossing effects." It is shown that the "crossing effects" are only observed when two instances of morphologically-driven A'-movement, *i.e.* wh-movement and null operator movement, interact, but not when the other types of syntactic dependencies are involved. I will argue that such an asymmetry with the "crossing effects," which has resisted any principled account, straightforwardly follows from the MLC.

3.1 Asymmetries of the "Crossing Effects"

3.1.1 The "Crossing Effects"

It has been pointed out that there is a clear contrast in acceptability between examples like (7a) and those like (7b) (see, among others, Culicover & Wexler (1977), Fodor (1978), Kuno & Robinson (1972), and Pesetsky (1982)):

- (7) a. ?**which book**_i did you decide [**which boy**_j to persuade **t**_j to buy **t**_i]?
b. ***which boy**_j did you decide [**which book**_i to persuade **t**_j to buy **t**_i]?

While (7b) is completely impossible, (7a) is much less severely deviant. It has been claimed (see, among others, Fodor (1978) and

THE CROSSIG CONSTRAINT AND THE MLC

Kuno & Robinson (1972)) that the difference in acceptability between examples like (7a) and those like (7b) comes from the linear sequences of wh-phrases and their original positions. As shown in (9), (7b), which is completely impossible, has a crossing dependency. (7a), which is less severely deviant, has a nesting dependency, as shown in (8):

(8) **which book_i** did you decide [**which boy_j** to persuade **t_j**
 | _____ |
 to buy **t_i**]

(9) **which boy_j** did you decide [**which book_i** to persuade **t_j** to
 | _____ |
 buy **t_i**]

Such a contrast in acceptability between nesting and crossing dependencies has been called the "crossing effects."

Similar contrasts can be found in every construction that involves "WH-movement" in the sense of Chomsky (1977), as shown in (10)-(14). Following Browning (1987) and Chomsky (1986a, 1986b), I assume that these constructions involve null operator movement to the Spec of CP:

(10) WH-Question and Tough Movement

a. ?I am wondering [**which violin_i** [this sonata is easy
 [**Op_j** to play **t_j** on **t_i**]]]

b. *I am wondering [**what_j** [this violin is easy [**Op_i** to
 play **t_j** on **t_i**]]]

(Cf. Culicover and Wexler 1977: 23, Chomsky 1977: 105, Fodor 1978: 450)

(11) Infinitival Relative and WH-Question

a. ?I finally found a subject [**Op_i** to ask [**who_j** to talk to **t_j**
 about **t_i**]]

b. *I finally found a subject [**Op_i** to ask [**who_j** to talk
 about **t_i** to **t_j**]] (Pesetsky 1982: 269)

(12) Topicalization and WH-Question

a. ?this problem [**Op_i** Mary knows [**who_j** to consult **t_j**
 about **t_i**]]

b. *this specialist [**Op_j** Mary knows [**what problems_i** to
 consult **t_j** about **t_i**]] (Pesetsky 1982: 269)

TORU ISHII

- (13) Cleft and Tough Movement
- a. ?this is the knife [Op_i that the salami is easy [Op_j to cut t_j with t_i]]
 - b. *this is the salami [Op_j that the knife is easy [Op_i to cut t_j with t_i]] (Culicover and Wexler 1977: 23, Kuno and Robinson 1972: 477)
- (14) Too/Enough Movement and Tough Movement
- a. ?the wood is too rough [Op_i for these blunt nails to be easy [Op_j to hammer t_j into t_i]]
 - b. *The nails are too blunt [Op_j for this rough wood to be easy [Op_i for me to hammer t_j into t_i]] (Fodor 1978: 450)

While the (b) examples of (10-14) are completely impossible, the (a) examples of (10-14) are much less severely deviant.

Note that wh-movement and null operator movement can both be characterized as morphologically-driven movements in the sense of Chomsky (1993, 1994, 1995), since they are triggered by some morphological feature of C. Furthermore, they are also characterized as movements into "A'-positions." Then, we can say that the "crossing effects" are observed with the interactions of morphologically-driven A'-movement.

3.1.2 Absence of the Expected Contrast Between Crossing and Nesting

It was observed in the previous section that there is a clear contrast in acceptability between nesting and crossing when dependencies are created by wh-movement or null operator movement. Such a contrast between nesting and crossing, however, is not observed when the other types of syntactic dependencies are involved.

First, when wh-movement interacts with A-binding, control, or A-movement, no "crossing effects" emerge, as exemplified below (see, Chomsky 1977: 192):

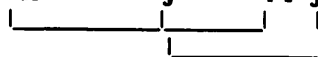
(15) what books_i have those men_j written t_i [about each other_j]?



(16) I told them_j [what books_i PRO_j to read t_i].



(17) [to whom]_i did John_j seem t_i [t_j to be referring]?



TORU ISHII

accusative NP *wairo-o* 'bribe-ACC' to the intermediate clause, the result is (21b):³

- (21) a. **Wairo-o_j** [John-ga [**torihikisaki-ni_i** [katyoo-ga [pro [t_i]
 [t_j watasu]] yoo(ni)] meizita to] syoogen sita
- b. **Torihikisaki-ni_i** [John-ga [**wairo-o_j** [katyoo-ga [pro [t_i]
 [t_j watasu]] yoo(ni)] meizita to] syoogen sita

While (21a) has a nesting dependency, (21b) has a crossing dependency. Both (21)(a) and (b), however, are acceptable, with the stylistic awkwardness being abstracted.

Let us finally look at the interaction of clause-internal and long-distance scramblings:

- (22) John-ga [[Bill-ga [Mary-ni [sono hon-o watasita]]] to]
 -NOM -NOM -DAT that book-ACC gave COMP
 omotteiru
 think
 'John thinks that Bill gave that book to Mary.'

If we scramble *Mary-ni* 'Mary-DAT' to the embedded clause and *sono hon-o* 'that book-ACC' to the matrix clause, we get (23a). If we scramble *sono hon-o* 'that book-ACC' to the embedded clause and *Mary-ni* 'Mary-DAT' to the matrix clause, the result is (23b):

- (23) a. [Sono hon]-o_j [John-ga [[Mary-ni_i [Bill-ga [t_i [t_j
 watasita]]]]] to] omotteiru]
- b. **Mary-ni_i** [John-ga [[[sono hon]-o_j [Bill-ga [t_i [t_j
 watasita]]]]] to] omotteiru]

³As argued in detail in Nemoto (1993), the verb *meizi* cannot take a so called "major object" (cf. Hoji (1991)). Hence, *wairo-o* in (21) cannot be interpreted as a "major object," but can only be interpreted as scrambled out of the most deeply embedded clause.

THE CROSSIG CONSTRAINT AND THE MLC

While (23a) has a nesting dependency, (23b) has a crossing dependency. Both of them, however, are acceptable.

Based on the above data, one might argue that unlike English, Japanese lacks the "crossing effects" for some reason. This view is not tenable, since the "crossing effects" are also observed in Japanese when two instances of morphologically-driven A'-movement are involved, as exemplified below:

(24) Taroo-ga [sensei-ga [zibun-no musuko]-ni [pro
 -NOM teacher-NOM self-GEN son -DAT
 gakusyuuzyuku-ni iku yooni] susumeta to] omotteiru
 private school attend recommended COMP think
 'Taro thinks that the teacher recommended his son to attend a
 private school.'

(25) a. ?[Op_j [Taroo-ga [Op_i [seisei-ga t_i [pro t_j iku yooni]

|-----|-----|
 susumeta no wa] [zibun-no musuko]-ni da] to
 be

|-----|
 omotteiru no wa] gakusyuuzyuku-ni da]
 be

b. *[Op_i [Taroo-ga [Op_j [seisei-ga t_i [pro t_j iku yooni]

|-----|-----|
 susumeta no wa] gakusyuuzyuku-ni da] to omotteiru no
 wa] [zibun-no musuko]-ni da]

Following Hoji (1990), let us assume that the cleft construction with an NP-Case focus or a PP focus necessarily involves null operator movement. Then, the contrast in acceptability between (25)(a) and (b) suggests that null operator movement in Japanese also exhibits the "crossing effects."⁴

To recapitulate, there is no contrast in acceptability between nesting and crossing when the dependencies created by NP-movement, A-binding, control, or scrambling are involved. In

⁴In (25), we are assuming that the empty operator moves to the clause-initial position. If the empty operator moves to the clause-final position, the linear crossing constraint would wrongly predict the opposite of what is observed in (25). This is the place where linear approaches to the "crossing effects" like Fodor (1978) and hierarchical approaches like Pesetsky (1982) make different predictions. If it is proved that the empty operator moves to the clause-final position in Japanese, then that would provide evidence in support of the hierarchical approaches to the "crossing effects." Notice that the arguments for the MLC approach, being hierarchical in nature, hold whichever the landing site of the empty operator movement may be.

TORU ISHII

other words, the contrast between nesting and crossing is only observed when two instances of morphologically-driven A'-movement are involved.

3.2 The MLC and Asymmetries of the "Crossing Effects"

This section first proposes a system of clausal typing features. I will then argue that given the proposed clausal typing features, the "crossing effects" are grammatical phenomena and subsumed under the MLC. It is shown that under the MLC analysis of the "crossing effects," we can account for the hitherto unexplained asymmetry with the "crossing effects" between morphologically-driven A'-movement and the other types of syntactic dependencies.

3.2.1 Formal Features of C, Wh-phrases, and Null-operators

Before turning to the discussion of the asymmetry of the "crossing effects," let us consider the formal features of C, wh-phrases, and null operators.

Essentially following Cheng (1991), I assume that the formal features of C determine clause types. I claim that clauses in English can be divided into two types, *i.e.* interrogative and noninterrogative clauses. Each clause type can further be classified into two subtypes, depending on whether or not "WH-movement" in the sense of Chomsky (1977) is involved. "WH-movement" includes both wh-movement and null operator movement. Yes/no questions and wh-questions are both interrogative clauses, but only the latter involves "WH-movement." Declarative and null operator clauses are both noninterrogative clauses, but only the latter involves "WH-movement." Note that null operator clauses are the ones where null operator movement is involved in the sense of Browning (1987) and Chomsky (1977, 1986a, 1986b) such as topicalization, infinitival relative clauses, the tough construction, the cleft construction, and the *too/enough* construction.

I propose two formal features of C, *i.e.* [Q] and [WH], and argue that clause types can be characterized by the combination of these two formal features, as shown below:

(26)	No [Q]	[Q]
No [WH]	Declarative Clauses	Yes/No Questions
[WH]	Null Operator Clauses	Wh-Questions

THE CROSSIG CONSTRAINT AND THE MLC

According to [Q], clauses are classified as interrogative clauses and noninterrogative clauses. Interrogative clauses are headed by the C with a [Q] feature. Noninterrogative clauses, on the other hand, are headed by the C without a [Q] feature. Wh-questions and null operator clauses involve "WH-movement" and thus are headed by the C with a [WH] feature. Yes/No questions and declarative clauses, on the other hand, do not involve "WH-movement" and thus are headed by the C without a [WH] feature.

Based on the above mentioned criterion for [+/- interpretable] features, these formal features of C are [+ interpretable]. This is because the interpretive operations at LF have to know which clause type a clause belongs to. The formal features of C, being [+ interpretable], need not enter into checking relations unless forced by some other factors, since they are legitimate objects at LF. I claim that the [Q] features in fact never enter into any checking relations. The [WH] features in English, on the other hand, always enter into checking relations, since I assume, following Chomsky (1993, 1994, 1995), that they are strong. Let us simply assume that if features are strong, they must enter into checking relations before Spell-Out. How to derive this property of strong features is a question which I want to keep beyond the scope of the present discussion. See Chomsky (1995) for a different view of strong features.

Wh-phrases and null operators also have [WH] features. I claim that, contrary to the [WH] features of C, the [WH] features of wh-phrases and null operators are [- interpretable]. This is because we are assuming that [WH] features are the formal features which determine clause types and clausal typing features like [WH] are properties of C, not the properties of wh-phrases and null operators.⁵ It then follows that the [WH] features of wh-phrases and null operators must enter into checking relations with the [WH] features of C and delete and erase when checked, since they are illegitimate objects at LF.

⁵One might argue that the [WH] features of wh-phrases should be [+ interpretable], otherwise wh-phrases may not receive appropriate interpretations at LF. Note, however, that in the present discussion, [WH] features are assumed to be clausal typing features. I only claim that [WH] features, being clausal typing features, are [- interpretable] when they appear in wh-phrases or null operators, since clausal typing is not their property. This does not exclude the possibility that wh-phrases have semantic features which contribute to their interpretations at LF.

3.2.3 The Absence of the Exptected "Crossing Effects" with the Interactions between Wh-movement and NP-movement

Let us consider cases where wh-movement interacts with NP-movement, taking (17) (repeated here as (46)) as an example:

(46) [To whom]_i did John_j seem t_i [t_j to be referring]

Although (46) has a crossing dependency between wh-movement and NP-movement, it evades the "crossing effects." The MLC gives us an account of the absence of the "crossing effects" in (46).

Suppose we come to the following structure during derivation:

(47) [_a T [seem to whom [John to be referring]]]
 [D] [WH] [D]
 [φ] [φ]
 [Case] [Case]

The next step is to fill the Spec of T in order to satisfy its strong [D] feature. There is, however, only one candidate for raising to the Spec of T, *i.e.* *John*. Notice that although *to whom* is closer to the target α than *John*, *to whom* does not have any feature that can enter into checking relations with the sublabels of α . Hence, *to whom* is "invisible" to the attraction by α .

Similar remarks hold when we come to a stage where the Spec of C is to be filled in order to check its strong [WH] feature:

(48) [_b C [John_j [T [seem to whom [t_j to be referring]]]]]
 [Q] [D] [WH]
 [WH] [φ]

The Spec of C can only be filled by *to whom*. Although *John* is closer to the target β than *to whom*, *John* does not have any feature that can enter into checking relations with the sublabels of β .

3.2.4 The Absence of the Exptected "Crossing Effects" with Binding and Control

Let us next consider cases where wh-movement interacts with binding or control. Let us assume that no attraction operations are involved to establish binding and control relations. Then, since the MLC, being part of the definition of Attract-F, only applies to

TORU ISHII

attraction operations, it is irrelevant to binding and control relations. Hence, *wh*-movement is allowed to have crossing dependencies with binding and control relations.

It has been proposed (see, among others, Chomsky (1986b, 1993)), however, that anaphor binding relations are established in terms of attraction operations. Under such an analysis, anaphors raise to their antecedents in the covert component. Suppose this is true. Even under such an attraction analysis, we can still correctly predict the absence of the "crossing effects" with anaphor binding. This is because while *wh*-phrases are attracted by the [WH] features of C, anaphors are attracted by some nominal features of their antecedents. Hence, anaphors are "invisible" to attraction by C and *wh*-phrases, to attraction by nominals. Crossing dependencies with anaphor binding are therefore allowed.

3.2.5 The Absence of the Expected "Crossing Effects" with Scrambling

Let us finally consider the absence of the expected "crossing effects" with scrambling. Following Fukui (1993) and Fukui and Saito (1996), let us assume that unlike English *wh*-movement and NP-movement, scrambling in Japanese is optional and thus not triggered by morphological requirements. In other words, scrambling does not involve any feature-checking operations. Then, scrambling cannot be subsumed under Attract-F, which is contrary to English *wh*-movement and NP-movement whose essential properties follow from Attract-F. Scrambling is by definition exempt from the MLC, which is part of the definition of Attract-F. Hence, provided that the "crossing effects" are subsumed under the MLC, it straightforwardly follows that scrambling in Japanese never exhibits the "crossing effects."

3.2.6 Summary

This section has investigated the hitherto unexplained asymmetry with the "crossing effects" between morphologically-driven A'-movement and the other types of syntactic dependencies. I have argued that given the proposed system of clausal typing features, the asymmetry straightforwardly follows from the MLC.

3.3 Previous Analyses of the "Crossing Effects"

This section reviews two previous analyses of the "crossing effects": Fodor's (1978) parsing analysis and Pesetsky's (1982) path theory. I will first present Fodor's parsing analysis and argue that there is some empirical evidence against her analysis. I will then look at Pesetsky's path theory. It is shown that his theory does not give any principled account of the "crossing effects," since it crucially relies on the stipulation that only morphologically-driven A'-movement creates paths.

3.3.1 Fodor's (1978) Parsing Analysis

Fodor (1978) attributes the contrast in acceptability between nesting and crossing to a parsing strategy. She proposes the Nested Dependency Constraint (NDC), which is a parsing constraint on linear dependencies which applies to sentences that are fully ambiguous with respect to gap filling:

(49) The Nested Dependency Constraint (NDC)

If there are two or more filler-gap dependencies in the same sentence, their scope may not intersect if either disjoint or nested dependencies are compatible with the well-formedness conditions of the language.

(Fodor 1978: 448)

There is, however, evidence to show that the NDC, a linear parsing constraint, is on the wrong track. As mentioned above, the "crossing effects" are not observed with scrambling in Japanese. Let us consider (24) again. If we scramble *gakushuuzyuku-ni* 'private school-DAT', which is the locative PP in the most deeply embedded clause, to the matrix clause and *zibun-no musuko-ni* 'self-GEN son-DAT', which is the PP argument of the verb *susume* 'recommend', to the intermediate clause, the result is (50). On the other hand, if we scramble *gakusyuuzyuku-ni* 'private school-DAT' to the intermediate clause and *zibun-no musuko-ni* 'self-GEN son-DAT' to the matrix clause, the result is (51):

(50) *gakusyuuzyuku-ni* [Taroo-ga [[*zibun-no musuko*]-*ni*

| _____ |
| _____ |
[sensei-ga t_i [pro t_j iku yooni] *susumeta to*] *omotteiru*
| _____ |

TORU ISHII

(51) [zibun-no musuko]-ni_i [Taroo-ga [gakusyuuzyuku-ni_j
|_____| |_____|
[sensei-ga t_i [pro t_j iku yooni] susumeta to] omotteiru
|_____| |_____|
|_____|

Both (50) and (51) are fully ambiguous in the sense of Fodor (1978). This is because they have two gaps of the same syntactic category, *i.e.* PP, created by two instances of scrambling, and those two gaps are in positions accessible to either instance of scrambling. Since (50) has a nesting dependency while (51) has a crossing dependency, the NDC would predict that the latter is unacceptable. With stylistic awkwardness being abstracted, both (50) and (51) are in fact acceptable.

3.3.2 Pesetsky's (1982) Path Theory

Pesetsky (1982) proposes the Path Containment Condition (PCC) and argues that it can account for the "crossing effects":

(52) Path Containment Condition (PCC)

If two paths overlap, one must contain the other.

(Pesetsky 1982: 309)

Although Pesetsky (1982) is correct in claiming that the "crossing effects" should be accounted for by grammatical principles rather than parsing strategies, his PCC account has two drawbacks. First, as pointed out above, the "crossing effects" can be observed only when two instances of *wh*-movement or null operator movement interact. When *wh*-movement or null operator movement interacts with NP-movement, binding, or control, however, the "crossing effects" cannot be observed. In order to account for the absence of the "crossing effects," Pesetsky's PCC account has to stipulate that only A'-movement creates paths and thus is subject to the PCC. Hence, his path theory cannot provide any principled account of the absence of the "crossing effects" with NP-movement, binding, and control.

Second, as mentioned above, long-distance scrambling in Japanese does not exhibit any "crossing effects." It has been claimed (see, among others, Mahajan (1990) and Saito (1992)) that long-distance scrambling is a case of A'-movement. According to Pesetsky's definition of the notion of "path," A'-movement always creates paths. Hence, the PCC is descriptively inadequate in that it

THE CROSSIG CONSTRAINT AND THE MLC

would wrongly predict that crossing dependencies created by long distance scrambling are unacceptable. In order to avoid this problem, Pesetsky has to stipulate that optional movement, *i.e.* non-morphologically-driven movement, does not create any paths even if it is an instance of A'-movement.

4 Wh-island Violations, the "No Extra Baggage" Condition, and the Functional Determination of Feature Weight

As pointed out repeatedly, when two instances of morphologically-driven A'-movement interact, there is a clear contrast in acceptability between nesting and crossing dependencies. While crossing dependencies are completely impossible, nesting dependencies are much less severely deviant. It has been claimed that the less deviant status of nesting dependencies is due to wh-island violations. However, no studies have ever tried to provide a minimalist account of the less deviant status of wh-island violations.

In this section, I propose a minimalist account of wh-island violations and argue that wh-island violations should be attributed to the "no extra baggage" condition. It is also shown that the less severely deviant status of wh-island violations straightforwardly follows from the "no extra baggage" condition. I will also argue that this analysis sheds a new light on not only the contrast in acceptability between nesting and crossing dependencies but also the one between wh-island violations and superraising violations.

4.1 C-Features and the Reference Set for Economy

Since the "no extra baggage" condition is one of the economy conditions, it evaluates convergent derivations belonging to the same reference set. Chomsky (1994, 1995) claims that the reference set is determined by an initial numeration *N*. As far as C-features are concerned, however, let us assume that the reference set does not refer to all the C-features included in an initial numeration, though it is still determined by the initial numeration. I argue that the reference set only refers to the primary clausal typing, *i.e.* interrogative vs. noninterrogative clauses, but not to the further clausal typing, *i.e.* whether or not "WH-movement" is involved. In other words, the [Q] features of C always enter into the reference set if they are included in initial numerations. The [WH] features of C, on the other hand, never enter into the reference set even if they are included in initial numerations.

TORU ISHII

Hence, as far as C-features are concerned, the initial numerations with the same value of the [Q] feature constitute the same reference set for the economy conditions irrespectively of whether they include [WH] features.

Let us furthermore assume that as far as C is concerned, the reference set does not include any phonological features of lexical items. The interrogative complementizers, *i.e.* *whether* and *if*, have [Q] features as well as phonological features. The noninterrogative complementizer *that* has phonological features, but not any [Q] feature. These complementizers enter into initial numerations with their phonological and [Q] features, if they have ones, when they are selected from the lexicon. Their phonological features, however, never enter into the reference set. Hence, whether complementizers are overt or null is irrelevant to the reference set. In other words, the initial numerations with overt complementizers and those with null complementizers belong to the same reference set provided that all the other lexical choices are the same.

4.2 Wh-island Violations

Given these assumptions, let us consider wh-island violations involving two instances of wh-movement, considering (53) as an example:

(53) ?*which book*_i did you decide [*which boy*_j [John should persuade *t*_j to buy *t*_i]]?

In the initial numeration of (53), both the matrix and embedded Cs have [Q] and [WH] features. Since [WH] features never enter into the reference set, the derivations based on this initial numeration belong to the same reference set as those based on the following three initial numerations: (i) Both matrix and embedded Cs have only [Q], (ii) The matrix C has both [Q] and [WH] while the embedded C has only [Q], and (iii) The embedded C has both [Q] and [WH] while the matrix C has only [Q]. Hence, the derivations based on these four initial numerations are comparable by the economy conditions.

Notice that only initial numerations (i) and (ii) include overt interrogative complementizers (*whether* or *if*) as their members, since the embedded Cs only with [Q] must always be overt. Remember, however, that the phonological features of complementizers never enter into the reference set. Hence, overt complementizers in the initial numerations are ignored for the

TORU ISHII

- (56) **which boy_i** [C [you decide [C [John should persuade t_i to
 [Q] [Q]
 [WH]
 buy **which book**]]]]
 [WH]

In (56), *which boy* rather than *which book* raises to the matrix Spec of C in order to check the strong [WH] feature of the matrix C in accordance with the MLC. The [WH] feature of *which boy* deletes and erases through this checking operation. The [WH] feature of *which book*, on the other hand, remains unchecked at this stage of derivation. But, it can be attracted by the [WH] feature of the matrix C in the covert component, as shown below:

- (57) **which boy_i** [C [you decide [C [John should persuade t_i
 [Q] [Q]
 [WH]
 to buy **which book**]]]]

Remember that the [WH] feature of C is [+ interpretable] and thus survives throughout a derivation even after having entered into checking relations. Since this attraction by the matrix C occurs in the covert component, just the [WH] feature, instead of the whole category *which book*, raises to the matrix C in order to be checked off. The [WH] feature of *which book* deletes and erases through this checking operation. Hence, this derivation converges and yields (58), which is acceptable:⁷

- (58) Which boy_i did you decide whether/if John should persuade t_i
 to buy which book?

Let us finally consider a derivation based on the initial numeration where the embedded C has both [Q] and [WH] while the matrix C has only [Q]. Based on this numeration, we form the following structure at the point of Spell-Out:

⁷For some speakers, examples like (58) are slightly deviant. It should be noticed, however, that even for those speakers, there is still a clear contrast between examples like (58) and those violating the wh-island constraint.

TORU ISHII

those induced by two instances of operator movement can be accounted for in a similar fashion.⁸

Remember that unlike nesting dependencies, crossing dependencies are excluded as illegitimate by the MLC. Let us assume, following Chomsky (1995), that superraising violations as in (62) should also be attributed to the MLC:

(62) *John_i seems [that it was told t_i [that Mary is a genius]]]

I argue that the present analysis gains further support from the fact that it gives us a minimalist account of the contrast in acceptability between superraising and "crossing" violations, on the one hand, and wh-island violations, on the other. It has been observed that while superraising and "crossing" violations are totally impossible, wh-island violations are less severely deviant. No minimalist account, however, has even been given to this contrast. Under the no "extra baggage" condition analysis of the wh-island constraint, derivations violating the wh-island constraint converge, but are only excluded by the "no extra baggage" condition, an economy condition. This is in contrast with superraising and "crossing" violations, where derivations never converge but are canceled by the MLC. It is then plausible to claim that derivations blocked by the economy conditions, being convergent, would result in less severe violations than canceled derivations, which never reach the interface levels. If this conjecture is correct, we can correctly predict that wh-island violations like (53) are less severely deviant than superraising and "crossing" violations.⁹

⁸Since the economy conditions compare convergent derivations, they necessarily involve global considerations in comparing more than one phrase-markers. One might argue that such global conditions should be totally eliminated to reduce problems of computational complexity. As Chomsky (1991, 1993) argues, however, there is no a priori reason to suppose that language is "usable" or "conducive to efficient use." Language could possibly be computationally intractable due to its global properties. Language can be used in practice thanks to heuristics, that is, what Chomsky (1991) calls "computational tricks."

⁹It has been observed, however, that extraction of adjunct wh-phrases out of wh-islands results in severe violations just as superraising, which I would like to leave for further study.

4.3 Against Chomsky's (1995) Account of Wh-island Violations

Chomsky (1995) argues that wh-island violations as well as superraising violations should follow from the MLC. Chomsky's MLC analysis, however, has a serious drawback. As mentioned above, there is a clear contrast in acceptability between superraising and wh-island violations. While superraising violations are totally impossible, wh-island violations are less severely deviant. Under Chomsky's analysis, however, both superraising and wh-island violations are excluded as illegitimate by the MLC. Hence, it does not provide any account of the difference in acceptability between these two violations.

5 Conclusion

This paper has argued that the hitherto unexplained asymmetry with the "crossing effects" between morphologically-driven A'-movement and the other types of syntactic dependencies straightforwardly follows from the MLC given the proposed system of clausal typing features. I have then argued that wh-island violations should be attributed to the "no extra baggage" condition rather than the MLC. It was shown that the analysis advocated in this paper provides a minimalist account of the difference in acceptability between superraising and "crossing" violations, on the one hand, and wh-island violations, on the other.

References

- Browning, Marguerite 1987. *Null operator constructions*. Doctoral dissertation, MIT.
- Cheng, Lisa Lai-Shen 1991. *On the typology of wh-questions*. Doctoral dissertation, MIT.
- 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 1986a. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam 1986b. *Knowledge of language: Its nature, origin, and use*. New York: Prager.
- Chomsky, Noam 1991. Some notes on economy of derivation and representation. In Robert Freidin (Ed.), *Principles and parameters in comparative grammar*. Cambridge, MA: MIT Press.

TORU ISHII

- Chomsky, Noam 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser (Eds.), *The view from building 20* (pp. 1-52). Cambridge, MA: MIT Press.
- Chomsky, Noam 1994. *Bare phrase structure*. MIT working papers in linguistics.
- Chomsky, Noam 1995. *Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam and Howard Lasnik 1993. The theory of principles and parameters. In Joachim Jacobs, Armin von Stechow, Wolfgang Sternefeld and Theo Vennemann (Eds.), *Syntax: An international handbook of contemporary research* (pp. 506-569). Berlin: Walter de Gruyter.
- Culicover, Peter and Kenneth Wexler 1977. Some syntactic implications of a theory of language learnability. In Peter Culicover, Thomas Wasow and Adrian Akmajian (Eds.), *Formal syntax* (pp. 7-60). New York: Academic Press.
- Fodor, Janet Dean 1978. Parsing strategies and constraints on transformations. *LI*, 9, 427-473.
- Fukui, Naoki 1993. Parameters and optionality. *LI*, 24, 399-420.
- Fukui, Naoki and Mamoru Saito 1996. *Order in the theory of phrase structure and movement*. Unpublished manuscript, University of California, Irvine and Nanzan University.
- Hoji, Hajime 1985. *Logical form constraints and configurational structures*. Doctoral dissertation, University of Washington.
- Hoji, Hajime 1986. Weak crossover and Japanese phrase structure. In Takashi Imai and Mamoru Saito (Eds.), *Issues in Japanese linguistics* (pp. 163-202). Dordrecht: Foris.
- Hoji, Hajime 1990. *Theories of anaphora and aspects of Japanese syntax*. Unpublished manuscript, University of Southern California.
- Hoji, Hajime 1991. *Raising-to-object, ECM and the major object in Japanese*. Paper read at Japanese syntax workshop at University of Rochester.
- Kuno, Susumu and Jane Robinson 1972. Multiple WH-questions. *LI*, 3, 463-487.
- Larson, Richard 1988. On the double object construction. *LI*, 19, 335-391.
- Mahajan, Anoop 1990. *The A/A-bar distinction and movement theory*. Doctoral dissertation, MIT.
- Nemoto, Naoko 1993. *Chains and case positions: A study from scrambling in Japanese*. Doctoral dissertation, University of Connecticut.
- Pesetsky, David 1982. *Paths and categories*. 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 1992. Long distance scrambling in Japanese. *JEAL*, 1, 69-118.

THE CROSSIG CONSTRAINT AND THE MLC

Webelhuth, Gert 1989. *Syntactic saturation phenomena and the modern germanic languages*. Doctoral dissertation, University of Massachusetts, Amherst.

Department of Linguistics
3151 Social Science Plaza
University of California, Irvine
Irvine, CA 92697-5100
and
Common Courses of Human Sciences
Kitami Institute of Technology
165 Koen-cho Kitami 090 JAPAN
n65000@cc.kitami-it.ac.jp