Formal Approaches to Japanese Linguistics 3

Proceedings from FAJL 3

MIT Working Papers in Linguistics 41 (2001)

MIT Working Papers in Linguistics

#1 Papers on Syllable Structure, Metrical Structure and Harmony Processes (1979)
#2 Issues in Japanese Syntax (1980)
#3 Papers on the Grammar of Semitic Languages (1981)
#4 Papers in Formal Syntactic Theory (1982)
#5 Papers in Grammatical Theory (1983)
#6 Papers in Theoretical and Applied Linguistics (1985)
#7 Papers in Morphology (1984)
#8 Papers in Theoretical Linguistics (1986)
#9 The 25th Anniversary of MIT Linguistics (1987)
#10 Functional Heads and Clause Structure (1989)
#11 Papers from the First Student Conference in Linguistics, MIT (1989)
#12 Papers from the Second Student Conference in Linguistics, MIT (1990)
#13 Papers on wh-Movement (1990)
#14 Papers from the Third Student Conference in Linguistics, MIT (1991)
#15 More Papers on wh-Movement (1991)
#16 Papers from the Fourth Student Conference in Linguistics, Ohio State University (1992)
#17 Proceedings of the Kwa Comparative Syntax Workshop, MIT (1992)
#18 Papers on Case and Agreement I (1993)
#19 Papers on Case and Agreement II (1993)
#20 Papers from the 5th Student Conference in Linguistics, University of Washington (1993)
#21 Papers on Phonology and Morphology (1994)
#22 The Morphology Syntax Connection (1994)
#23 Papers from the 6th Student Conference in Linguistics, University of Rochester (1994)
#24 Formal Approaches to Japanese Linguistics: Papers from the May 1994 MIT Conference (1994)
#25 The Interpretive Tract (May 1998)
#26 Papers on Language Processing and Acquisition (1995)
#27 Papers on Minimalist Syntax (1995)
#28 Papers on Language Endangerment and the Maintenance of Linguistic Diversity (1996)
#29 Formal Approaches to Japanese Linguistics: Proceedings of FAJL 2 (1996)
#30 PF: Papers at the Interface (1997)
#31 Proceedings of the Eighth Student Conference in Linguistics (1997)
#32 Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect (May 1998)
#33 Papers on Morphology and Syntax, Cycle One (1999)
#34 Papers on Morphology and Syntax, Cycle Two (1999)
#35 Papers frm the Upenn/MIT Roundtable on the Lexicon (1999)
#36 Proceedings of the 9th Student Conference in Linguistics (2000)
#37 Proceedings of the 10th Student Conference in Linguistics (2001)
#38 Proceedings of the 12th Student Conference in Linguistics (2001)
#39 A Few from Building E-39 (2001)
#40 Proceedings of the 1st HUMIT Student Conference on Language Research (2001)

Toward the Reunification of Japanese Scramblings*

Mamoru Saito
Nanzan University

1. Introduction

Japanese scrambling is known for its mixed properties: there are cases where it exhibits the properties of A-movement, A'-movement, or even PF-movement. Since Mahajan's (1990) influential work on Hindi scrambling, one approach to the mixed properties has been to posit subtypes of Japanese scrambling with different landing sites. This approach, which is tied closely to the feature-checking analysis of scrambling, has been pursued by Miyagawa (1997, 2001) among others. On the other hand, since Tada 1990, attempts have been made to explain the mixed properties on the assumption that Japanese scrambling is uniform. In this paper, I will explore the uniform approach further by developing some proposals made in works such as Tada 1990, Kitahara 2000, and Kuno 2001. The main claim is that the mixed properties provide evidence for the formulation of Condition (A) as an anywhere condition (Belletti and Rizzi 1988, Lebeaux 1988, Epstein, et al. 1998) and for the derivational model of interpretation (Epstein, et al. 1998, Chomsky 1999, Kitahara 2000).

In the following section, I will briefly illustrate the mixed properties of Japanese scrambling. Based on this, I will establish two fundamental assumptions: (i) scrambling is uniformly to a position from where A-binding is possible, and (ii) all scrambling chains are subject to total reconstruction. In Section 3, I will first argue, following Tada 1990, that the ambiguity of clause-internal scrambling with respect to the A- and A'-properties leads to supporting evidence for the hypothesis that Condition (A) is an anywhere condition while Condition (C) applies at LF. Then, adapting Kitahara's (2000) derivational approach to the interpretation of scrambling sentences, I will suggest that scrambling chains are interpreted as they are created (or phase by phase). This explains the fact that long scrambling patterns with A'-movement or PF-movement. Section 4 concerns the proper binding effect on scrambling. There,

---

* This is a shortened version of the paper presented at the 3rd Formal Approaches to Japanese Linguistics Conference. The material in this paper was developed in the syntax seminar at Nanzan University between 1998 and 2000. I would like to thank the participants, especially Mayumi Dejima, Tomoko Kawamura and Mizuki Nishida, for comments and suggestions. I would also like to thank Danny Fox, Hisatsugu Kitahara, David Lebeaux, Shigeru Miyagawa, Anna Szabolcsi and Akira Watanabe for helpful discussions. Special thanks are due to Shin Ishihara for his extensive help with the preparation of this version of the paper. Part of the research reported here was supported by the Nanzan University Pache Research Grant IA.

MIT Working Papers in Linguistics 41, 287–307
Proceedings of the Third Formal Approaches to Japanese Linguistics

I will defend the condition along the lines proposed in Kuno 2001 and suggest a reformulation of the condition as a derivational constraint. Section 5 concludes the paper.

2. The Mixed Properties of Japanese Scrambling

In this section, I will discuss the basic properties of Japanese scrambling: (i) its radical reconstruction property, (ii) its mixed A/A'-properties, and (iii) the fact that its trace shows strict proper binding effects.

2.1 The radical reconstruction property

The most outstanding fact about Japanese scrambling is that it apparently need not have any effect on interpretation. Let us first consider the following examples of *wh*-questions:[1]

(1) a. [$_{TP}$ John-ga [$_{CP}$[$_{TP}$ dare-ga sono hon-o katta] ka]
J.-NOM who-NOM that book-ACC bought Q
siritagatteiru] (koto)
want-to-know fact
'[John wants to know [Q [who bought that book]]]'

b. * [$_{TP}$ Dare-ga [$_{CP}$[$_{TP}$ John-ga sono hon-o katta] ka]
who-NOM J.-NOM that book-ACC bought Q
siritagatteiru] (koto)
want-to-know fact
'[Who wants to know [Q [John bought that book]]]'

Only the embedded clause is a question in these examples. As noted in Harada 1972, contrasts of this kind then indicate that a *wh*-phrase is subject to the condition in (2).

(2) A *wh*-phrase must be contained in the CP where it takes scope.

This condition applies to *wh*-phrases in English as well, as shown in (3).

(3) a. [$_{CP}$ Who$_i$ [$_{TP}$ t_i asked whom to find out [$_{CP}$ what$_j$ [$_{TP}$ Bill bought t_j]]]]
b. [$_{CP}$ Who$_i$ [$_{TP}$ t_i wonder [$_{CP}$[which picture of whom]$_j$ [$_{TP}$ Bill saw t_j]]]]
c. ??[$_{CP}$[Which picture of whom]$_j$ does [$_{TP}$ Bill wonder [$_{CP}$ who$_i$ [$_{TP}$ t_i saw t_j]]]]

When a *wh* moves to a CP Spec, it takes scope there. Hence, *who* and *what* in (3a), and *who* and *which* in (3b–c) satisfy the condition trivially. *Whom* is left in-situ in (3a). This *wh* is contained within the matrix CP but not within the embedded CP. Hence, the condition predicts correctly that it can only take matrix scope. On the other hand, examples like (3b) are ambiguous as van Riemsdijk and Williams (1981) point out: *whom* can take matrix or embedded scope. This is also predicted by the condition because the *wh* is contained within the embedded CP as well as the matrix CP. (3c) contrasts sharply with (3b). This example is marginal because it is a *wh*-island violation. But its interpretation is clear: it only allows *whom* to take matrix scope, again, as predicted by the condition.

Let us now consider the examples in (4b) and (5b) in light of the discussion above.

(4) a. [$_{TP}$ John-ga [$_{CP}$[$_{TP}$ Mary-ga dono hon-o yonda] ka]
J.-NOM M.-NOM which book-ACC read Q
siritagatteiru] (koto)
want-to-know fact
'[John wants to know [Q [Mary read which book]]]'

b. ?[$_{TP}$ Dono hon-o$_i$ [John-ga [$_{CP}$[$_{TP}$ Mary-ga t_i yonda] ka]
which book-ACC J.-NOM M.-NOM read Q
siritagatteiru]] (koto)
want-to-know fact
'[Which book$_i$, John wants to know [Q [Mary read t_i]]]'

(5) a. [$_{TP}$ John-ga [$_{CP}$[$_{TP}$ minna-ga [$_{CP}$ Mary-ga dono hon-o
J.-NOM all-NOM M.-NOM which book-ACC
yonda to] omotteiru] ka] siritagatteiru] (koto)
read that think Q want-to-know fact
'[John wants to know [Q [everyone thinks [that Mary read which book]]]]'

b. ??[$_{TP}$ [$_{CP}$ Mary-ga dono hon-o yonda to]$_i$ [John-ga
M.-NOM which book-ACC read that J.-NOM
[$_{CP}$ [$_{TP}$ minna-ga t_i omotteiru]ka] siritagatteiru]] (koto)
all-NOM think Q want-to-know fact
'[[That Mary read which book]$_i$, John wants to know [Q [everyone thinks t_i]]]'

(4b) is derived from (4a) by scrambling the *wh dono hon-o* 'which book-ACC' from the embedded object position to the initial position of the matrix clause. The movement clearly places the *wh* out of the CP where it takes scope. Thus, given (2), we expect the example to be as bad as (1b), but it is only slightly degraded. The *wh dono hon-o* in (5a) is located within the most deeply embedded CP. In (5b), this CP is scrambled out of the middle CP, where the *wh* takes scope. Again, we expect the example to be totally out, but it is only marginal.

[1] I will add *koto* 'the fact that' at the end of some example sentences in order to avoid the unnaturalness resulting from the lack of a topic in a matrix clause. But I will ignore *koto* in the "translations" and also in the discussion. The "translations" in single quotes are provided to illustrate the rough structures of the examples and are not meant to be the correct English translations.

Assuming that the condition in (2) applies at LF, I proposed in Saito 1989 that scrambling can be literally undone in the LF component.[2] Given this hypothesis, (4b) and (5b) can satisfy the condition because their LF representations can be identical to those of (4a) and (5a) respectively. This "LF undoing" property of Japanese scrambling later came to be called its radical reconstruction property. In what follows, I will assume that "undoing" is part of the interpretive procedure for scrambling chains, and hence, that it applies obligatorily to all instances of scrambling.[3]

2.2 The binding paradoxes

The discussion in the preceding subsection suggests that scrambling has no effects on interpretation. It is thus tempting to hypothesize that it is a "stylistic rule" applying in the PF component. This proposal is in fact made in Ross 1967, and Chomsky and Lasnik 1977. However, there are also cases in which scrambling does have effects on interpretation. For example, Tada (1990) shows in detail that Japanese scrambling extends the possible binding relations, exactly as in the case of Hindi scrambling discussed by Mahajan (1990).

An important observation in the present context is that a phrase preposed by clause-internal scrambling can serve as an "A-binder." The examples in (6)–(7) illustrate this point with the lexical anaphor *otagai* 'each other'.

(6) a. [$_{TP}$ Karera-ga [otagai-no sensei]-o hihansita] (koto)
they-NOM each other-GEN teacher-ACC criticized fact
'They criticized [each other's teachers]'

b. ?*[$_{TP}$[Otagai-no sensei]-ga karera-o hihansita] (koto)
each other-GEN teacher-NOM they-ACC criticized fact
'[Each other's teachers] criticized them'

(7) ?[$_{TP}$ Karera-o$_i$ [[otagai-no sensei]-ga t_i hihansita]] (koto)
they-ACC each other-GEN teacher-NOM criticized fact
'Them$_i$, [each other's teachers] criticized t_i'

The contrast in (6) shows that *otagai* requires a c-commanding antecedent. (7) is derived from (6b) by scrambling the object *karera-o* 'they-ACC' to a position c-commanding *otagai*, and the example is clearly improved.

The following examples with bound pronouns exhibit the same pattern:

(8) a. ?*[$_{TP}$[Sono$_i$ tyosya]-ga dono hon-ni-mo$_i$ keti-o tuketa]
its author-NOM which book-to-even gave-criticism
'[Its author] criticized every book'

[2] It is actually assumed in Saito 1989 that the relevant condition is the proper binding condition applying to the trace of LF *wh*-movement.

[3] This hypothesis is already proposed in Tada 1990. A possible exception is the very local VP(or vP)-internal scrambling, which is more like object shift. I will not discuss this type of scrambling in this paper.

b. [$_{TP}$ Dono hon-ni-mo$_i$ [[sono$_i$ tyosya]-ga t_i keti-o tuketa]]
which book-to-even its author-NOM gave-criticism
'Every book$_i$, [its author criticized t_i]'

Sono 'its' in (8a) cannot be construed as a bound pronoun because it is not c-commanded by *dono hon-ni-mo* 'to every book' in the object position. But the scrambling of the object to the sentence-initial position makes this binding possible. Examples like (7) and (8b) indicate that we cannot simply assume that scrambling is PF movement.

Interestingly, the improvement in (7) and (8b) is not observed with long scrambling. As (9) shows, a phrase preposed by long scrambling cannot serve as the antecedent of a lexical anaphor.

(9) a. *[$_{TP}$[Otagai-no sensei]-ga [$_{CP}$[$_{TP}$ Tanaka-ga karera-o
each other-GEN teacher-NOM T.-NOM they-ACC
hihansita] to] itta] (koto)
criticized that said fact
'[Each other's teachers] said that Tanaka criticized them'

b. * [$_{TP}$ Karera-o$_i$ [[otagai-no sensei]-ga [$_{CP}$[$_{TP}$ Tanaka-ga t_i
they-ACC each other-GEN teacher-NOM T.-NOM
hihansita] to] itta]] (koto)
criticized that said fact
'Them$_i$, [each other's teachers] said that Tanaka criticized t_i'

In (9b), *karera-o* 'they-ACC' is scrambled out of the embedded CP to the initial position of the matrix clause, where it c-commands *otagai.* Yet, there is no improvement in sharp contrast with (7). Similarly, long scrambling fails to license a bound pronoun as shown in (10).[4]

(10)?* [$_{TP}$ Dono hon-ni-mo$_i$ [[sono$_i$ tyosya]-ga [$_{CP}$[$_{TP}$ Hanako-ga t_i
which book-to-even its author-NOM H.-NOM
keti-o tuketa] to] itta]]
gave-criticism that said
'Every book$_i$, its author said that Hanako threw cold water on t_i'

Mahajan (1990), who first noted this difference between clause-internal scrambling and long scrambling with Hindi data, proposed that there are two distinct types of scrambling, A and A', with different landing sites. Clause-internal scrambling can be A-movement, and hence, (7) and (8b) are grammatical exactly like the English (11a–b).

[4] Yoshimura (1989) and Saito (1992) argue that when a *wh*-NP is preposed by long scrambling, it can license a bound pronoun. As Daiko Takahashi points out (personal communication, 1994), there seems to be a distinction here between *wh*-phrases and regular quantifiers. I will put aside the *wh* cases in this paper, but see Saito 1995 for a possible analysis of this distinction.

(11) a. They$_i$ seemed to each other$_i$ [t_i to be smart]
b. Everyone$_i$ seems to his$_i$ mother [t_i to be smart]

This type of scrambling can be A'-movement as well, for otherwise, (12) would be incorrectly ruled out by Condition (C).

(12) [$_{TP}$ Zibunzisin-o$_i$ [Taroo-ga t_i semeta]] (koto)
self-ACC T.-NOM blamed fact
'Himself$_i$, Taroo blamed t_i'

Long scrambling, on the other hand, can only be A'-movement, and this accounts for the ungrammaticality of (9b) and (10). Mahajan hypothesizes that A-scrambling is movement to AGR Spec, while A'-scrambling is adjunction.

However, given the assumption that scrambling is uniform, the data discussed above present interesting problems. First, it is unclear why clause-internal scrambling exhibits the A-properties as well as the A'-properties as indicated in (7), (8b) and (12). I will refer to this problem as the A/A' paradox. Second, it is necessary to explain why long scrambling cannot have the A-properties as (9b) and (10) show. This will be called the long scrambling puzzle.

2.3 The proper binding problem

A further problem arises when examples of multiple scrambling are considered. It was noted in Saito 1985 that scrambling is constrained by the proper binding condition. A relevant example is shown in (13).

(13) *[$_{TP}$ [Hanako-ga t_i iru to]$_j$ [Sooru-ni$_i$ [Taroo-ga t_j omotteiru]]] (koto)
H.-NOM be that Seoul-in T.-NOM think fact
'[That Hanako is t_i]$_j$, in Seoul$_i$, [Taroo thinks t_j]'

This example is derived from (14a) by multiple applications of scrambling.

(14) a. [$_{TP}$ Taroo-ga [Hanako-ga Sooru-ni iru to] omotteiru] (koto)
T.-NOM H.-NOM Seoul-in be that think fact
'Taroo thinks that Hanako is in Seoul'

b. [$_{TP}$ Sooru-ni$_i$ [Taroo-ga [Hanako-ga t_i iru to] omotteiru]] (koto)
Seoul-in T.-NOM H.-NOM be that think fact
'In Seoul$_i$, Taroo thinks that Hanako lives t_i'

First, the PP *Sooru-ni* 'in Seoul' is scrambled from the embedded clause to the matrix initial position. This yields a grammatical sentence as shown in (14b). Then, the embedded CP is scrambled to the position preceding the PP. The resulting sentence (13) is totally ungrammatical. CP scrambling and multiple scrambling are both allowed in Japanese, as shown in (15a) and (15b).

(15) a. [$_{TP}$[Hanako-ga Sooru-ni iru to]$_i$ [Taroo-ga t_i omotteiru]] (koto)
H.-NOM Seoul-in be that T.-NOM think fact
'[That Hanako lives in Seoul]$_i$, Taroo thinks t_i'

b. [$_{TP}$ Sono hon-o$_i$ [Hanako-ni$_j$ [Taroo-ga [$_{CP}$ Ziroo-ga t_j t_i
that book-ACC H.-to T.-NOM Z.-NOM
watasita to] omotteiru]]] (koto)
handed that think fact
'That book$_i$, to Hanako$_j$, Taroo thinks that Ziroo handed t_i t_j'

Thus, it seems reasonable to attribute the ungrammaticality of (13) to the unbound trace t_i.

The effect in (13) is quite general and is not limited to the cases of multiple scrambling. In (16), a trace of scrambling is unbound due to a later application of passive.

(16) * [$_{TP}$ [PRO t_i iku koto]-ga$_j$ Sooru-made$_i$ Taroo-ni t_j meizirareta]
go N -NOM Seoul-to T.-to ordered-was
'[To go t_i]$_j$, to Seoul$_i$, was ordered Taroo t_j'

In the derivation of this example, the PP *Sooru-made* 'to Seoul' is first scrambled out of the infinitival complement as illustrated in (17).[5]

(17) a. Hanako-ga Taroo-ni [PRO Sooru-made iku koto]-o meizita
H.-NOM T.-to Seoul-to go N -ACC ordered
'Hanako ordered Taroo to go to Seoul'

b. Hanako-ga Sooru-made$_i$ Taroo-ni [PRO t_i iku koto]-o meizita
H.-NOM Seoul-to T.-to go N -ACC ordered
'Hanako, to Seoul$_i$, ordered Taroo to go t_i'

The infinitival complement in (17) is headed by the nominalizer *koto* and appears in the object position. Hence, it can be passivized as a regular object as shown in (18).

(18) [PRO Sooru-made iku koto]-ga$_j$ Taroo-ni t_j meizirareta
Seoul-to go N -NOM T.-to ordered-was
'[To go to Seoul]$_j$ was ordered Taroo t_i'

---

[5] The landing site of this scrambling is the matrix vP or VP. As noted in Mahajan 1990 for Hindi, scrambling out of an infinitival complement shows both A- and A'-properties, and in this sense, behaves more like clause-internal scrambling than long scrambling out of a tensed CP complement. I assume that a TP is directly embedded under the nominalizer *koto* in a control infinitival complement like the one in (17), and that the absence of C-projection accounts for the observed properties of this type of scrambling. See Nemoto 1993 and Saito 1996 for detailed discussion on this point, and Murasugi 1991 for a general discussion on the structure of Japanese complex NPs.

The ungrammatical (16) is derived when this passive applies after the scrambling in (17b).[6]

(13) and (16) are problematic because the proper binding account seems untenable under the Minimalist assumptions. Recall the hypothesis proposed in Saito 1989 that scrambling can be undone in LF. Given this hypothesis, these examples cannot be ruled out by the proper binding condition at LF because no trace exists after the undoing of scrambling. It was proposed in Saito 1989 that (13) is ruled out by the S-structure application of the proper binding condition. But this analysis cannot be maintained within the Minimalist theory, where S-structure is eliminated.

3. A Derivational Approach to the Binding Paradoxes

In this section, I will propose an account for the binding paradoxes illustrated in Section 2.2. For the A/A' paradox, I will basically adopt Tada's (1990) analysis. Among the assumptions in his analysis are (i) scrambling is undone in LF, (ii) Condition (C) applies at LF, and (iii) Condition (A) can be satisfied prior to LF. I will present a version of this analysis with further supporting arguments in Section 3.1. Sections 3.2 and 3.3 concern the long scrambling puzzle. For this, Kitahara (2000) offers an extremely interesting analysis based on the idea that linguistic objects are interpreted throughout the course of the derivation. I will extend this analysis and argue that chains are interpreted as soon as they are created.[7]

3.1 Condition (A) as an anywhere condition

The examples that illustrate the A/A' paradox, (7) and (12), are repeated in (19a) and (19b).

(19) a. ?[$_{TP}$ Karera-o$_i$ [[otagai-no sensei]-ga t_i hihansita]] (koto)
they-ACC each other-GEN teacher-NOM criticized fact
'Them$_i$, [each other's teachers] criticized t_i'

b. [$_{TP}$ Zibunzisin-o$_i$ [Taroo-ga t_i semeta]] (koto)
self-ACC T.-NOM blamed fact
'Himself$_i$, Taroo blamed t_i'

---

[6] Kitahara (1997) proposes an elegant MLC explanation for Müller's (1996) generalization in (i), and suggests further that the account may extend to examples like (13).

(i) A phrase containing a trace of movement cannot undergo movement of the same type (operator movement, scrambling, NP-movement).

However, this account does not cover cases like (16) since they do not fall under Müller's generalization.

[7] I regret that I am unable to discuss Tada and Kitahara's analyses in any detail due to the limitation of space. The reader is refered to the works cited.

(19a) indicates that scrambling is movement to a position from where A-binding is possible. But then, (19b) is problematic because it should be a Condition (C) violation.

Let us first examine (19a) more closely. I hypothesized above that every instance of scrambling is "undone" in LF as part of the interpretive procedure. If this is correct, the example fails to satisfy Condition (A) at LF. Thus, we are led to the conclusion that the condition can be satisfied in the course of the derivation, that is, that it is an anywhere condition. And this formulation of Condition (A) is already proposed in the literature, e.g., in Belletti and Rizzi 1988, Lebeaux 1988 and Epstein, et al. 1998. Belletti and Rizzi's argument is based on examples of "backward anaphora" such as (20).

(20) Pictures of himself$_i$ worry John$_i$

They hypothesize that the NP *pictures of himself*, being the theme argument, originates in a position lower than the experiencer argument *John*, as in (21).

(21) [$_{TP}$[Pictures of himself$_i$]$_j$ [$_{VP}$[$_{V'}$worry t_j] John$_i$]]

Then, the grammaticality of (20) is accounted for if Condition (A) is an anywhere condition. The anaphor is bound by *John* prior to the movement of *pictures of himself* to the subject position.

Lebeaux (1998) presents further examples that support Belletti and Rizzi's analysis. Some of them are shown in (22)–(23).

(22) a. [Each other's$_j$ mothers]$_i$ seem [t_i to please the two boys$_j$]
b. * John$_i$ seemed to each other's$_j$ mothers [t_i to please the two boys$_j$]

(23) a. [His$_j$ first performance]$_i$ seems [t_i' to be expected [t_i to please every composer$_j$]]
b. * The president$_i$ seems to his$_j$ first wife [t_i' to be expected [t_i to please every man$_j$]]

The b-examples show that the object of a psych predicate cannot bind into a higher clause. Yet, (22a) and (23a) are clearly much better than them. Belletti and Rizzi's analysis straightforwardly extends to the grammatical cases in (22) and (23). If anaphors and bound pronouns can be licensed at any point of the derivation, *each other* in (22a) and *his* in (23a) can be licensed before movement applies in these examples.

The argument based on (20) and (22)–(23) that Condition (A) is an anywhere condition is not conclusive. Lebeuax (1998) in fact takes (22)–(23) as evidence instead that an NP can reconstruct in LF to any position of its A-chain. Accordingly, he assumes that Condition (A) is an LF condition. However, the argument can be made complete when Condition (C) effects are taken into consideration. Let us first consider the following examples adapted from Lebeaux 1998:

(24) a. *Himself$_i$ seems to John$_i$ [t_i to be very smart]
b. *Each other$_i$ seem to John and Mary$_i$ [t_i to be very smart]

These examples are plausibly Condition (C) violations. But given Lebeaux's hypothesis that an NP can reconstruct to any position of its A-chain, the matrix subject can reconstruct in LF to the embedded subject position. Consequently, these examples cannot be ruled out by Condition (C) at LF. Based on this, Lebeaux concludes that Condition (C) is an everywhere condition, i.e., a condition that must be satisfied throughout the derivation.

However, this faces a problem with examples like (25), as Chomsky (1993) points out.

(25) The claim that John$_i$ was asleep seemed to him$_i$ [t_i to be correct]

If Condition (C) applies prior to raising, it should exclude this example.[8] This example in fact seems to show that Condition (C) cannot be an everywhere condition but must apply at LF. But then, tracing Lebeaux's reasoning backwards, we arrive based on (24) at the conclusion that NP-movement is not subject to reconstruction. It follows that the only way to properly explain (20) and (22)–(23) is to make Condition (A) an anywhere condition.

The discussion above reconciles the grammaticality of (19a) and the radical reconstruction property of scrambling: Condition (A) is satisfied prior to the undoing of scrambling. Further, it solves the initial problem posed by (19b). (19b) is repeated in (26a) with another relevant example (26b).

(26) a. [$_{TP}$ Zibunzisin-o$_i$ [Taroo-ga t_i semeta]] (koto)
self-ACC T.-NOM blamed fact
'Himself$_i$, Taroo blamed t_i'

b. [$_{TP}$ Otagai-o$_i$ [Taroo-to Hanako-ga t_i semeta]] (koto)
each other-ACC T.-and H.-NOM blamed fact
'Each other$_i$, Taroo and Hanako blamed t_i'

The contrast between these examples and those in (24) is striking. Since the landing sites of NP-movement and scrambling are both positions from where A-binding is possible, (24) and (26) are both expected to be out if Condition (C) is an everywhere condition. But it was argued above that Condition (C) is an LF condition and that there is no reconstruction with NP-movement. The contrast then automatically follows. (24a–b) are ruled out by Condition (C) at LF, but (26a–b) are not because scrambling is subject to radical reconstruction.

3.2 The derivational application of the copy and deletion analysis

Having settled the A/A' paradox, I will now turn to the second binding problem. The problem is the difference between clause-internal scrambling and long scrambling represented by (7) and (9b), repeated below as (27a–b).

(27) a. ?[$_{TP}$ Karera-o$_i$ [[otagai-no sensei]-ga t_i hihansita]] (koto)
they-ACC each other-GEN teacher-NOM criticized fact

[8] This problem is discussed in Lebeaux 1988 and 1998, where he suggests, following van Riemsdijk and Williams 1981, that Condition (C) applies after NP-movement.

'Them$_i$, [each other's teachers] criticized t_i'

b. * [$_{TP}$ Karera-o$_i$ [[otagai-no sensei]-ga [$_{CP}$[$_{TP}$ Tanaka-ga t_i
they-ACC each other-GEN teacher-NOM T.-NOM
hihansita] to] itta]] (koto)
criticized that said fact
'Them$_i$, [each other's teachers] said that Tanaka criticized t_i'

If A-binding is possible from the landing site of scrambling and Condition (A) is an anywhere condition, it is unclear why (27b) is not as good as (27a). In the remainder of this section, assuming Epstein, et al.'s (1998) derivational model for interpretation, I will present a mechanism for the interpretation of scrambling chains that captures the contrast in (27). But I will first discuss the derivational interpretation of chains in more general terms on the basis of Chomsky's (1993) copy and deletion theory of movement.

Chomsky proposes that an operator-variable chain is created by copy and deletion, as illustrated in (28).

(28) Who$_i$ did John see t_i
a. [$_{CP}$ Who [$_{C'}$ did [$_{TP}$ John see who]]]
| |
b. [$_{CP}$ Who [$_{C'}$ did [$_{TP}$ John see x]]]

In (28b), *who* is retained in CP Spec as an operator, but its copy in the object position is turned into a variable. One way to interpret this proposal would be as follows. An NP has the D-feature, which makes it possible for the NP to have a "reference" and to participate in binding/coreference relations.[9] An overt NP also has phonetic features, say, P-features. A *wh*-phrase, in addition, is equipped with an operator feature O. The O-feature of *who* allows the *wh* to be interpreted as *[for which x: x a person]* in CP Spec and its D-feature is responsible for its interpretation as a variable in the object position. Then, what is deleted in CP Spec in (28b) is the D-feature of *who*, and the other features are retained there. In the object position, the O-feature and the P-features are deleted while the D-feature is retained.

Construed this way, the copy and deletion analysis can be applied straightforwardly in a cyclic fashion as movement takes place.[10] Let us consider the example in (29) to illustrate one way to execute this idea.

(29) Who do you think John saw

[9] It may be that there is an indenpendent feature, say, R-feature, that is closely tied to the categorial feature D and is responsible for the referential properties. For simplicity's sake, I will assume that the D-feature itself enters into binding relations.

[10] Chomsky (1999) in fact suggests that information is sent for interpretation phase by phase throughout the derivation. In the discussion in the text, I will assume that chains are interpreted as they are created, but as far as I can see, the proposals are largely consistent with the phase model as well.

If *who* moves successive-cyclically through CP Spec, we obtain the following structure with the initial movement:

(30) [CP who [TP John saw who]]
{P,O,D} {P,O,D}

Let us suppose, as it seems reasonable, that deletion applies to the features of *who*, P, O and D, so that each of them is retained at one position. The P-features must be retained at the head position of the chain. This, if anything, is part of the definition of overt movement. For the other features, suppose further that deletion is constrained by selectional relation in a broad sense, including the feature-checking relation. More concretely, let us suppose that any feature that enters into a selectional relation can only appear in a position where it is selected.[11] Then, the D-feature in (30) must be deleted at the CP Spec because it is selected only in the object position. Chomsky (1998) proposes that the movement of a *wh*-phrase to an intermediate CP Spec takes place because a feature of the C head (call it the EPP-O feature) attracts the O-feature of the *wh*. With the extended use of the term 'selection', we may say that the O-feature satisfies a selectional requirement of the C head in this case.[12] Then, as the O-feature is not selected in the object position, it is retained in the CP Spec position. This yields the structure in (31).

(31) [CP who [TP John saw who]]
{P,O} {D}

If interpretation takes place as a chain is created, (31) must be interpreted before the derivation proceeds. Here, as Chomsky (1999) notes, the CP Spec cannot be interpreted at this point because it participates in further operations: *who* moves eventually to the matrix CP Spec. Let us assume then that information on the TP, the maximal X^{max} properly contained within the CP, is sent to the interpretive component. Here, *who* in the object position, with only D-feature, is interpreted as a variable as in (32).

(32) [TP John saw x]

Then, after *who* in the embedded CP Spec with the features {P,O} moves to the matrix CP Spec, the following structure obtains:

(33) [CP who [C' do [TP you think [CP who [TP John saw who]]]]]
{P,O} {P,O} {D}

The P-features are retained in the matrix CP Spec, the head of the chain. The O-feature is selected by the matrix [+Q] C, and hence, is also retained there. The

[11] This is a variant of Lee's (1994) idea that only XP positions that participate in feature-checking are retained at LF. See also Kawamura 2001 for a similar proposal.

[12] I assume with Chomsky 1998 that the EPP-O feature on C deletes prior to interpretation after its selectional requirement is satisfied.

copy of *who* in the embedded CP Spec disappears as all of its features are deleted. Thus, the final interpretation will be as in (34).

(34) [CP for which x: x a person [C' do [TP you think [CP [TP John saw x]]]]]

3.3 The long scrambling puzzle resolved

Chomsky's copy and deletion theory in its original form makes a straightforward analysis possible for the radical reconstruction or undoing property of scrambling. It has been suggested in Saito 1994, Lee 1994, and Saito and Fukui 1998, for example, that radical reconstruction results from the total deletion of the higher copies in a scrambling chain in the LF component. Only minor adjustments are necessary to adapt this idea in the derivational model, where there is no independent LF component.

Let us assume, as in the works cited above, that scrambling is not feature-driven.[13] Then, radical reconstruction applies as illustrated in (35).

(35) [TP Sono hon-o_i [Yamada-ga t_i yonda]] (koto)
that book-ACC Y.-NOM read fact
'That $book_i$, Yamada read t_i'

a. [TP Sono hon-o [... sono hon-o ...]]
{P,D} {P,D}
b. [TP Sono hon-o [... sono hon-o ...]]
{P} {D}

Scrambling initially copies all features of the moved NP at the landing site as in (35a). When deletion applies, the P-features are retained at the head of the chain as in (35b). The D-feature, on the other hand, cannot be retained there since by hypothesis it does not enter into any selectional relation at that position. It must be in the object position where it is selected by the verb.

Note that only the phonetic features appear at the scrambled position in (33b), and hence, the structure produced by scrambling is indistinguishable from cases of PF movement. This may seem problematic because as noted above, a phrase preposed by clause-initial scrambling may serve as the binder for a lexical anaphor. The relevant example (7) is repeated once again as (36).

(36) ?[TP Karera-o_i [[otagai-no sensei]-ga t_i hihansita]] (koto)
they-ACC each other-GEN teacher-NOM criticized fact
'$Them_i$, [each other's teachers] criticized t_i'

(35b) clearly fails to account for this fact since it should be the D-feature, and not the P-features, of the scrambled NP that licenses the lexical anaphor. But note that although the D-feature of *karera* 'they' in (36) is eventually

[13] See also Saito 1985, 1989, Fukui 1986, Kuroda 1988, Tada 1993, and Bošković and Takahashi 1998 for arguments for this assumption. For the opposing view that scrambling is feature-driven, see, for example, Miyagawa 1997, 2001, Grewendorf and Sabel 1998, and Kawamura 2001.

reconstructed to the object position, it appears in a position c-commanding the lexical anaphor *otagai* 'each other' at one point of the derivation. The derivation of (36) is illustrated in (37).

(37) a. [$_{TP}$ Karera-o [... otagai .. karera-o ...]]
{P,D} {P,D}
b. [$_{TP}$ Karera-o [... otagai .. karera-o ...]]
{P} {D}

Hence, given that Condition (A) is an anywhere condition, *otagai* is successfully licensed at the point the scrambled NP is copied at the sentence-initial position. The analysis of (36) proposed in the preceding subsection, then, can be maintained as such.

Interestingly, the ungrammaticality of (9b), repeated in (38), also follows from this copy and deletion procedure if it is applied cyclically as in the case of *wh*-movement.

(38) * [$_{TP}$ Karera-o$_i$ [[otagai-no sensei]-ga [$_{CP}$[$_{TP}$Tanaka-ga $\underline{t}_i$
they-ACC each other-GEN teacher-NOM T.-NOM
hihansita] to] itta]] (koto)
criticized that said fact
'Them$_i$, [each other's teachers] said that Tanaka criticized $\underline{t}_i$'

If long movement must proceed through the embedded C-projection, this scrambling intially creates the following structure in the embedded CP:

(39) [$_{CP}$ Karera-o [$_{TP}$ ... karera-o ...]]
{P,D} {P,D}

The P-features remain at the CP Spec position, the head of the chain. The D-feature is retained in the object position because this is the only position where the feature is selected. Thus, the "deletion for interpretation" yields (40).[14]

(40) [$_{CP}$ Karera-o [$_{TP}$ ... karera-o ...]]
{P} {D}

In the matrix clause, *karera-o* 'they-ACC' moves on from the embedded CP Spec and the structure in (41) is created.

(41) [$_{TP}$ Karera-o [... otagai ... [$_{CP}$ karera-o [$_{TP}$ ... karera-o ...]]]]
{P} {P} {D}

[14] Nothing prevents the scrambled phrase from moving first to the embedded T-projection. In this case, the deletion of features applies at this point and only the P-features of the scrambled phrase move on to the embedded CP Spec. The resulting structure is identical to (40).

The copy of *karera-o* in the embedded CP Spec disappears as its P-features are deleted. The final form of the scrambling chain, shown in (42), is virtually identical to (37b).

(42) [$_{TP}$ Karera-o [... otagai ... [$_{CP}$[$_{TP}$ ... karera-o ...]]]]
{P} {D}

But there is one important difference between the derivations of (36) and (38). In the case of (38), since the D-feature of *karera-o* is deleted at the embedded CP Spec before the NP moves into the matrix clause, there is no point in the derivation where the feature c-commands the lexical anaphor *otagai*. Thus, the contrast between (36) and (38) is successfully accounted for.

According to the analysis proposed above, clause-internal scrambling and long scrambling apply in exactly the same way: all features are copied at the landing site but only the P-features are retained there after deletion. As a result, they produce identical structures at the end. In fact, there is only one difference between them. That is, long scrambling involves extraction out of a CP (or a phase), and hence, must apply successive-cyclically. Given the derivational interpretation of chains, the contrast between (36) and (38) follows from this single difference.

4. Proper Binding Condition as a Derivational Constraint

In this section, I will examine the proper binding problem illustrated in Section 2.3. A relevant example, (13), is repeated below in (43).

(43) * [$_{TP}$[Hanako-ga $\underline{t}_i$ iru to]$_j$ [Sooru-ni$_i$ [Taroo-ga $\underline{t}_j$ omotteiru]]] (koto)
H.-NOM be that Seoul-in T.-NOM think fact
'[That Hanako is $\underline{t}_i$]$_j$, in Seoul$_i$, [Taroo thinks $\underline{t}_j$]'

The proper binding account for (43) is defended in Kuno 2001. In the following subsection, I will argue that the discussion in the preceding section provides further support for his conclusion. Then, in Section 4.2, I will suggest a derivational formulation of the condition that successfully rules out (43).

4.1 The absence of NP-traces

One of the main concerns of Kuno 2001 is the contrast between (43) and examples of the following kind:

(44) [How likely [$\underline{t}_i$ to win the race]]$_j$ is John$_i$ $\underline{t}_j$

This example apparently contains an unbound trace, and hence, can be taken as evidence against the proper binding condition. If the condition does not exist, then obviously an alternative account must be sought for (43). Kuno, however, argues that the problem posed by (44) is only apparent. More specifically, he

adopts Lasnik's (1999) conclusion that NP-movement does not produce a trace, and argues that (43) is indeed ruled out by the proper binding condition.[15]

The discussion on binding in the preceding section lends further support for Kuno's claim. One of the crucial assumptions in the account for (24) and (25) was that there is no reconsruction with NP-movement. (24) and (25) are repeated below as (45) and (46).

(45) a. *Himself$_i$ seems to John$_i$ [t_i to be very smart]
b. *Each other$_i$ seem to John and Mary$_i$ [t_i to be very smart]

(46) The claim that John$_i$ was asleep seemed to him$_i$ [t_i to be correct]

Following Lebeaux (1998), I assumed that (45a–b) are Condition (C) violations. But I rejected his hypothesis that Condition (C) is an everywhere condition, because it incorrectly rules out (46). The conclusion obtained was that Condition (C) is an LF condition and there is no reconstruction with NP-movement.

Then, how can this absence of reconstruction be expressed in precise terms with the copy and deletion analysis of movement? If an NP is accompanied by a D-feature and P-features, as assumed above, then NP-movement should initially create a chain of the following form:

(47) [$_{TP}$ NP ... [... NP ...]]
{P,D} {P,D}

Here, the P-features are retained at the landing site. And (45a–b) show that the D-feature is as well: if this feature can be deleted at the landing site, then these examples cannot be excluded by Condition (C) at LF. Hence, given that features can be retained only in one position, the D-feature and P-features must both delete at the initial site. The initial site is then left with no features, which amounts to saying that there is no trace. Thus, given the copy and deletion analysis of movement, the absence of reconstruction with NP-movement naturally leads to the conclusion that there are no NP-traces. And if NP-traces do not exist, (44) ceases to be a problem for the proper binding condition, as Kuno has argued on independent grounds.

Before I move on to the more precise analysis of (43), one more thing needs to be said to make the copy and deletion analysis of NP-movement complete. In the discussion of chain interpretation in the preceding section, I hypothesized that D-feature and O-feature can be retained only in positions where they are selected. This prevents O-feature from being retained in an

[15] See Saito and Hoshi 2000 for an independent argument against NP-traces. Kuno (2001) generalizes Lasnik's conclusion to the traces of German scrambling to account for the remnant topicalization phenonenon. A relevant example from Müller (1996) is shown in (i).

(i) [$_{VP}$ t_i Gelesen]$_j$ hat das Buch$_i$ keiner t_j
read has the book no one
'No one has read the book'

argument position, and D-feature from being in an operator position. But given this hypothesis, the D-feature in (47) can be retained in either position. If selection is construed in a broad sense to include the feature-checking relation, the D-feature is selected in both positions. Then, it is predicted incorrectly that reconstruction with NP-movement is optional. In order to guarantee that the D-feature is retained only at the landing site, I suggest the following:

(48) Chain interpretation makes the chain minimum.

The P-features in (47) must remain at the landing site. So, if the D-feature is retained at the initial site, the chain ends up having two positions as its members. On the other hand, if it is retained at the landing site, the result will be a singleton chain. Thus, (48) forces the deletion of the D-feature at the initial site, and consequently, prevents NP-movement from leaving a trace. I will leave (48) as a stipulation at this point, hoping that it will eventually be derived from more general considerations.

4.2 On the proper formulation of the proper binding condition

Having defended the proper binding analysis of (43), I will now consider the details of the analysis. Note first that given the discussion in Section 3, the initial problem illustrated in Section 2.3 has a different outlook. It appears at this point that it is indeed possible to rule out (43) by the LF application of the proper binding condition.

I have adapted Epstein, et al. (1998) and Kitahara's (2000) derivational model for interpretation in Section 3. With this model, they propose to eliminate LF as a level of representation where interpretation applies. The hypothesis that Condition (A) is an anywhere condition is in accord with this proposal. At the same time, I argued that Condition (C) is an LF condition. If this is correct, then LF is still needed as a level of representation where some kinds of interpretive procedures apply, contrary to the strong hypothesis of Epstein, et al. and Kitahara. Given this, the proper binding condition can be formulated as an LF condition.

Further, the argument in Saito 1989 that proper binding is an S-structure condition no longer holds in its original form. It was assumed there that scrambling is literally undone in LF to yield the radical reconstruction effect. Hence, it was necessary to apply the proper binding condition prior to LF to exclude examples like (43). However, I proposed in Section 3 that the radical reconstruction effect obtains because "deletion for interpretation" applies in a specific way. Thus, in the scrambling chain in (49a), the D-feature is deleted at the landing site while the P-features are retained there as shown in (49b).

(49) a. [$_{TP}$ NP [... NP ...]]
{P,D} {P,D}
b. [$_{TP}$ NP [... NP ...]]
{P} {D}

Since only P-features remain at the landing site, scrambling is semantically vacuous. But according to this hypothesis, the scrambling chain itself is still

there at the output of the syntactic computation. Hence, (43) can be ruled out at LF by the proper binding condition, which can be formulated, for example, as in (50).

(50) Given a chain $<a_1, ..., a_n>$, a_i c-commands a_{i+1}.

However, there are reasons to doubt that this is the correct approach. First, even if we accept the hypothesis that there are LF conditions, we may ask what kinds of conditions qualify as such.[16] One possibility in this regard is that only those principles that relate to systems external to the language faculty apply at LF. Condition (A) states that an anaphor is interpreted coreferential with a c-commanding NP in its local domain. This condition can be viewed as an interpretive procedure to determine the actual reference of an anaphor or one to determine the anaphoric relation between an anaphor and its linguistic antecedent. The former would require interaction with systems that concern (the knowledge of) the actual world. Given that Condition (A) is derivational, it seems reasonable to assume that the condition has to do with the anaphoric relation between linguistic objects. Two possibilities arise for Condition (C) as well. It can be part of the procedure to determine the actual references of NPs or it can specify a relation between linguistic objects. If Condition (C) is indeed an LF condition, the former interpretation may be plausible. If this speculation is on the right track, it is dubious that the proper binding condition, as formulated in (50), is an LF condition. Since (50) is a constraint on the form of a linguistic object, it should be placed internal to the language system.

Secondly, as Akira Watanabe (personal communication) points out, it probably does not make much sense in the first place to claim that a scrambling chain like (49b) is constrained at LF. According to the analysis presented above, scrambling chains are headed only by P-features after deletion applies. Then, if (50) applies to them at LF, it should demand that P-features c-command the D-feature they are associated with. However, it is at least strange to say that an LF condition refers to P-features. Those features are plausibly invisible at the LF interface.

Then it seems desirable to reformulate the condition as a derivational constraint. I would like to suggest here that it is a condition on the application of Merge.[17] Merge combines two linguistic objects to form a constituent. The two objects to be combined by Merge must also be constituents. Let us then say that an object that contains only part of a chain, e.g. a trace but not its antecedent, does not qualify as a constituent in the relevant sense. This can be stated more formally as in (51)–(52).[18]

[16] The discussion here is speculative. A more precise theory of interpretation is clearly needed to make the discussion more concrete. For example, what we mean by LF is not totally clear at this point. It may be the output of the syntactic computation, or the collection of the interpretive information obtained through the syntactic derivation.

[17] I understand Merge as either pure Merge or Merge as part of Move in the sense of Chomsky 1995.

[18] After the draft of this paper was completed, it was brought to my attention that a similar proposal is made in Ausin 1998. He rejects the classical proper binding condition on independent grounds, and propose a similar condition to exclude unbound *wh*-traces.

(51) α is subject to Merge only if α is a complete constituent.

(52) α is a *complete constituent* $=_{df}$ i) α is a term, and
ii) if a position within α is a member of a chain γ, then every position of γ is contained within α.

This condition, being a derivational constraint, is immune to the problems posed on the LF application of the proper binding condition. In its effects it prevents the merger of an object that contains a trace but not its antecedent. In particular, it prevents the marger of the CP at the matrix TP in (43) because the CP contains a trace but not its antecedent.

5. Summary and Conclusion

In this paper, I explored the hypothesis that Japanese scrambling is a uniform operation, whether it is clause-internal or long-distance. Its landing site is always a position from where A-binding is possible, and every instance of scrambling is subject to radical reconstruction. I argued on independent grounds that Condition (A) is an anywhere condition, as proposed by Belletti and Rizzi (1988), Lebeaux (1988) and Epstein, et al. (1998), and that Condition (C) applies at LF. This automatically explains the fact that clause-internal scrambling exhibits both A- and A'-properties, basically along the lines suggested in Tada 1990. Next, extending Epstein, et al. (1998) and Kitahara's (2000) derivational theory, I proposed that chains are interpreted as they are created. The fact that long scrambling exhibits only A'-properties was derived as a consequence of this proposal. Finally, I pointed out that the discussion on binding presented in this paper provides support for Lasnik's suggestion that there are no A-traces, and hence, for Kuno's (2001) hypothesis that Japanese scrambling is subject to the proper binding condition. I suggested a formulation of the condition as a constraint on the application of Merge.

If the analysis proposed in this paper is on the right track, Japanese scrambling is uniform in the most strict sense. There is no difference between the landing sites of clause-internal scrambling and long scrambling. Japanese scrambling exhibits mixed properties because its landing site is similar in properties to that of NP-movement but it is subject to reconstruction like *wh*-movement. As a result, this movement operation offers unique evidence for the distinct ways in which Condition (A) and Condition (C) apply, for the derivational model of interpretation, and for the proper binding condition as a derivational constraint.

References

Ausin, Adolfo (1998). A minimalist approach to the remnant movement constraint. Ms., University of Connecticut.

Belletti, Adriana and Luigi Rizzi (1988). Psych-verbs and θ-theory. *Natural Language & Linguistic Theory* 6: 291–352.

Bošković, Zeljko and Daiko Takahashi (1998). Scrambling and last resort. *Linguistic Inquiry* 29: 347–366.
Chomsky, Noam (1993). A minimalist program for linguistic theory. In Ken Hale and S. Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, Mass.: MIT Press.
Chomsky, Noam (1995). Bare phrase structure. In Gert Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program*, Oxford: Blackwell.
Chomsky, Noam (1998). *Minimalist Inquiries: The Framework.* MIT Occasional Papers in Linguistics #15.
Chomsky, Noam (1999). *Derivation by Phase.* MIT Occasional Papers in Linguistics #18.
Chomsky, Noam and Howard Lasnik (1977). Filters and control. *Linguistic Inquiry* 8: 425–504.
Epstein, Samuel D., Erich Groat, Ruriko Kawashima, and Hisatsugu Kitahara (1998). *A Derivational Approach to Syntactic Relations.* Oxford: University Press.
Fukui, Naoki (1986). *A Theory of Category Projection and its Applications.* Ph.D. dissertation, MIT.
Grewendorf, Günter and Joachim Sabel (1999). Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language and Linguistic Theory* 17: 1–65.
Harada, Kazuko (1972). Constraints on WH-Q binding. *Studies in Descriptive and Applied Linguistics* 5: 180–206.
Kawamura, Tomoko (2001). *A Feature-Checking Analysis of Japanese Scrambling.* M.A. thesis, Nanzan University.
Kitahara, Hisatsugu (1997). *Elementary Operations and Optiomal Derivations.* Cambridge, Mass.: MIT Press.
Kitahara, Hisatsugu (2000). Case and scrambling: A derivational view. Ms., Keio University.
Kuno, Masakazu (2001). Why does movement sometimes leave a trace and sometimes not? Ms., University of Tokyo.
Kuroda, S.-Y. (1988). Whether we agree or not: A comparative syntax of English and Japanese. In William Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, Stanford: CSLI Publications.
Lasnik, Howard (1999). Chains of arguments. In Samuel D. Epstein and Norbert Hornstein (eds.), *Working Minimalism*, Cambridge, Mass.: MIT Press.
Lebeaux, David (1988). *Language Acquisition and the Form of the Grammar*, Ph.D. dissertation, University of Massachusetts, Amherst.
Lebeaux, David (1998). Where does the binding theory apply? Ms., NEC Research Center, Princeton.
Lee, Rhanghyeyun K. (1994). *Economy of Representation.* Ph.D. dissertation, University of Connecticut.
Mahajan, Anoop (1990) *The A/A-Bar Distinction and Movement Theory.* Ph.D. dissertation, MIT.
Miyagawa, Shigeru (1997). Against optional scrambling. *Linguistic Inquiry* 28: 1–25.
Miyagawa, Shigeru (2001). The EPP, scrambling, and *wh*-in-situ. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, Mass.: MIT Press.
Müller, Gereon (1996). A constraint on remnant movement. *Natural Language and Linguistic Theory* 14: 355–407.
Murasugi, Keiko (1991). *Noun Phrases in Japanese and English.* Ph.D. dissertation, University of Connecticut.
Nemoto, Naoko (1993) *Chains and Case Positions: A Study from Scrambling in Japanese.* Ph.D. dissertation, University of Connecticut.
Riemsdijk, Henk van and Edwin Williams (1981). NP-Structure. *The Linguistic Review* 1: 171–217.

Ross, John R. (1967). *Constraints on Variables in Syntax.* Ph.D. dissertation, MIT.
Saito, Mamoru (1985). *Some Asymmetries in Japanese and their Theoretical Implications.* Ph.D. dissertation, MIT.
Saito, Mamoru (1989). Scrambling as semantically vacuous A'-movement. In Mark Baltin and Anthony Kroch (eds.), *Alternative Conceptions of Phrase Structure*, Chicago: University of Chicago Press.
Saito, Mamoru (1992). Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1: 69–118.
Saito, Mamoru (1994). Improper adjunction. *MIT Working Papers in Linguistics* 24: 263–293.
Saito, Mamoru (1995). Radical reconstruction. Ms., University of Connecticut.
Saito, Mamoru (1996). Barriers for A-adjunction. *Japanese/Korean Linguistics* 5: 1–36.
Saito, Mamoru and Naoki Fukui (1998). Order in phrase structure and movement. *Linguistic Inquiry* 29: 439–474.
Saito, Mamoru and Hiroto Hoshi (2000). The Japanese light verb construction and the minimalist program. In Roger Martin, et al. (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, Mass.: MIT Press.
Tada, Hiroaki (1990). Scrambling(s). Ms., MIT.
Tada, Hiroaki (1993). *A/A-bar Partition in Derivation.* Ph.D. dissertation, MIT.
Yoshimura, Noriko (1989). Parasitic pronouns. Ms., University of Southern California..

Mamoru Saito
Department of Anthropology and Philosophy
Nanzan University
12 Yamazato-cho
Showa-ku, Nagoya 466-8673
Japan

saito@nanzan-u.ac.jp