A derivational approach to the interpretation of scrambling chains

Mamoru Saito

Department of Anthropology and Philosophy, Nanzan University, 18 Yamazoto-cho, Showa-ku, Nagoya 466-8673, Japan

Abstract

Japanese (and Korean) scrambling is known for its unique properties. It freely applies long-distance, but does not establish an operator-variable relation like the standard cases of A′-movement. It is also distinct from NP-movement in that its trace exhibits strict proper binding effects. Further, as observed by Tada [1990, Scrambling(s). MS., MIT] among others, it shows a pattern similar to Hindi scrambling with respect to binding: A phrase preposed by clause-internal scrambling can serve as an A-binder, but one preposed by long scrambling cannot. The purpose of this paper is to discuss these properties in some detail and to suggest a unified analysis for them. First, I assume, following Tada, that scrambling to sentence-initial positions is subject to total reconstruction, implementing the idea with a slightly modified version of Chomsky's (1993, A minimalist program for linguistic theory. In: Hale, K., Keyser, S.J. (Eds.), The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger. MIT Press, Cambridge, MA) copy and deletion analysis of movement. This accounts for the "semantically vacuous" nature of scrambling. Then, building on the insights of Kitahara (2000, Case and scrambling: a derivational view. MS., Keio University), I propose that the total reconstruction applies derivationally. This explains the A/A′ asymmetry observed with clause-internal scrambling and long-distance scrambling. Finally, I argue that the proper binding effects on scrambling require an independent condition, and suggest its formulation as a constraint on the application of Merge. If the analysis proposed in this paper is on the right track, Japanese (and Korean) scrambling does not have subtypes like A-scrambling and A′-scrambling, but is a uniform operation with well-defined abstract properties. The most fundamental among them is the radical (total) reconstruction property. It is this property that makes scrambling exhibit patterns that are quite different from operator movement and NP-movement.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Scrambling; Radical reconstruction; Binding conditions; Chain interpretation; Feature decomposition; Derivational approach to interpretation; Derivational phase

E-mail address: saito@nanzan-u.ac.jp (M. Saito).

1. Introduction

The purpose of this paper is to present an analysis for the following major properties of Japanese (and Korean) scrambling:

(1)a. Scrambling need not have any effect on the interpretation.
        (the radical reconstruction property)
   b. Clause-internal scrambling and long scrambling are both possible. But only
        a phrase preposed by the former can serve as an "A-binder."
        (the A/A′ problem discussed by Mahajan (1990) and Tada (1990))
   c. A trace of scrambling is constrained strictly by the proper binding condition.
        (the proper binding effect)

I will assume throughout the discussion that Japanese scrambling is uniform, i.e., that it does not have subtypes like A-scrambling and A′-scrambling with different landing sites. This assumption makes the task to explain (1a–c) more challenging. I hope to show that it also has significant consequences.

The analysis in this paper is built on the proposals in two important works, Tada (1990) and Kitahara (2000). Tada proposes that all instances of scrambling to sentence-initial positions are subject to total reconstruction at LF, and attempts to explain the apparent A′-properties of scrambling based on this proposal. Kitahara, on the other hand, argues that the effects of total reconstruction can be explained elegantly within the derivational model, where syntactic objects are interpreted throughout the course of the derivation. The basic idea is that scrambling exhibits the apparent A′-properties when it applies after the scrambled phrase is sent for interpretation.

Extending these ideas, I will propose in this paper that chains are interpreted as they are formed. Chomsky (1993) adopts a copy and deletion analysis of movement in order to secure proper interpretations for examples like (2).

(2)   Whose brother did he see

He suggests that movement leaves behind a copy as in (3a) and deletion creates a proper operator-variable relation as in (3b).

(3)a. [Whose brother] did he see [whose brother]
   b. [Who Ø] did he see [x's brother]

I will entertain the hypothesis that this "deletion for interpretation" applies as soon as a chain is formed in the derivation, and argue that it makes a unified account for the properties of scrambling in (1a–c) possible.

In the following section, I will discuss the basic facts on the radical reconstruction property (1a) and the A/A′ problem (1b), and briefly review Tada's (1990) and Kitahara's (2000) analyses. In Section 3, assuming Kitahara's general approach, I will argue for an alternative based on the hypothesis mentioned above. The proposed

analysis has the same two consequences on binding as Tada (1990). One is that Condition (A) is an "anywhere condition", a hypothesis already proposed on independent grounds by Belletti and Rizzi (1988), Lebeaux (1988) and Epstein et al. (1998), among others. The other is that Condition (C) is an LF condition. Section 4 is concerned with the proper binding effect (1c). I will first argue that this effect cannot be explained by the Minimal Link Condition as suggested in Kitahara (1997), but requires an independent condition. Then, I will propose a formulation of the condition as a constraint on the application of Merge. In Section 5, I will discuss the remaining problems with anaphor scrambling and VP-internal scrambling, and suggest directions toward their solutions. Section 6 concludes this paper.

2. (The absence of) semantic effects with scrambling

2.1. The radical reconstruction property

The most outstanding property of Japanese scrambling is probably its radical reconstruction property. I will start the illustration of this property with examples of *wh* scrambling. Let us first consider the following examples:[1]

(4)a. [$_{TP}$ John-ga   [$_{CP}$[$_{TP}$ dare-ga   sono hon -o   katta] ka] siritagatteiru] (koto)
           -NOM   who -NOM that   book-ACC bought Q 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] siritagatteiru] (koto)
           who -NOM   -NOM that book-NOM bought Q 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 must be contained within a question CP, or more specifically within the CP where it takes scope. This condition applies to *wh*-phrases in English as well, as shown in (5).

(5)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 wonders [$_{CP}$[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 overtly to a CP Spec, it takes scope there. Hence, *who* and *what* in (5a), and *who* and *which* in (5b and c) satisfy the condition trivially. *Whom* is left in-situ in (5a). This *wh* is contained within the matrix CP but not within the embedded CP. Hence, the condition correctly predicts that it can only take matrix scope. On

---

[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.

the other hand, examples like (5b) 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. (5c) contrasts sharply with (5b). 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 (6b) and (7b) in light of the discussion so far.

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

(6b) is derived from (6a) 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 the condition on *wh*-phrases, we expect the example to be as bad as (4b), but it is only slightly marginal. In (7a), the *wh dono hon-o* is located within the most deeply embedded CP. In (7b), 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.

Assuming that the condition on *wh*-phrases illustrated by (4) and (5) applies at LF, I proposed in Saito (1989) that scrambling can be literally undone (i.e., that the scrambled phrase can be totally reconstructed) in the LF component.[2] Given this hypothesis, (6b) and (7b) can satisfy the condition straightforwardly because their LF representations can be identical to those of (6a) and (7a) respectively. This "LF

---

[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.

undoing'' property of Japanese scrambling later came to be called its radical recon-struction property.

2.2. The A/A′ problem

What is often discussed in conjunction with the radical reconstruction property in the recent literature is the A/A′ properties of scrambling. As noted by Tada (1990), among others, a phrase preposed by clause-internal scrambling can serve as an ''A-binder.'' The examples in (8) and (9) illustrate this point with the lexical anaphor *otagai* 'each other'.

(8)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'
(9) ?[$_{TP}$ 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 (8) shows that *otagai* requires a c-commanding antecedent. (9) is derived from (8b) by scrambling the object *karera-o* 'they-ACC' to a position c-commanding *otagai*. The example improves as expected.

The following examples with bound pronouns exhibit the same pattern:

(10)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'
 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 (10a) 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.

Long scrambling, however, exhibits a different pattern with the licensing of ana-phors and bound pronouns. As (11) shows, a phrase preposed by long scrambling cannot serve as the antecedent of a lexical anaphor.

(11)a. *[$_{TP}$[Otagai   -no   sensei] -ga   [$_{CP}$[$_{TP}$ Tanaka-ga   karera-o   hihansita]
   each other-GEN teacher-NOM   -NOM they -ACC criticized
  to] itta] (koto)
  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   -NOM
hihansita] to] itta]] (koto)
criticized that said   fact
'Them$_i$, [each other's teachers] said that Tanaka criticized t_i'

In (11b), *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 (9). Similarly, long scrambling fails to license a bound pronoun as shown in (12).[3]

(12) ?*[$_{TP}$ Dono hon -ni -mo$_i$ [[sono$_i$ tyosya] -ga   [$_{CP}$[$_{TP}$ Hanako-ga t_i keti-o tuketa]
     which book-to-even its   author -NOM   -NOM gave-criticism
   to] itta]]
 that said
 'Every book$_i$, its author said that Hanako criticized 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, (9) and (10b) are grammatical exactly like the English (13a) and (13b).

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

Clause-internal scrambling can be A′ -movement as well, for otherwise, (14) would be incorrectly ruled out by Condition (C).

(14) [$_{TP}$ Zibunzisin-o$_i$ [Taroo-ga t_i semeta]] (koto)
        self   -ACC   -NOM   blamed fact
    'Himself$_i$, Taro blamed t_i'

Long scrambling, on the other hand, can only be A′-movement, and this accounts for the ungrammaticality of (11b) and (12). Mahajan hypothesizes that A-scrambling is movement to AGR Spec while A′ -scrambling involves adjunction.

   However, a research project was initiated by Tada (1990) and others to solve the A/A′ problem under the assumption that scrambling, whether clause-internal or long-distance, is uniform. I will pursue this approach below in Section 3. But before moving on to the main proposal, I will briefly review two important works in this tradition, Tada (1990) and Kitahara (2000), in the following subsection.

---

[3] 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.

2.3. Tada's total reconstruction and Kitahara's derivational approach

Tada (1990) proposes that scrambling is uniformly an adjunction operation, and as noted above, that all cases of scrambling to sentence-initial positions are subject to total reconstruction at LF.[4] The latter hypothesis amounts to saying that radical reconstruction is obligatory. Then, he explains the apparent A/A′ ambiguity of clause-internal scrambling observed in (9) and (14), repeated below as (15a) and (15b), on the assumption that Condition (A) and Condition (C) apply at distinct levels.[5]

(15)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   -NOM   blamed   fact
     'Himself$_i$, Taro blamed t_i'

If scrambling must be "undone" and the scrambled phrase must be reconstructed at its initial position at LF, (15a) cannot satisfy Condition (A) at this level. Then, Condition (A) must be an S-structure condition or an "anywhere condition," i.e., a condition that can be satisfied at any point in the derivation.

On the other hand, the grammaticality of (15b) implies that Condition (C) cannot be an S-structure condition. (15a) shows that the landing site of clause-internal scrambling is a position from where A-binding is possible. Hence, if Condition (C) applies at S-structure, it should exclude (15b). But the example is not problematic if Condition (C) is an LF condition. As the scrambled phrase is reconstructed to its initial position, the example does not violate the condition at this level. Thus, Tada successfully explains (15a and b) without assuming two distinct types of scrambling with different landing sites. I will adapt this analysis with further supporting arguments in Section 3.

But when it comes to the A′-properties of long scrambling, Tada's account seems less principled. The issue is why the scrambled phrase in (11b), repeated in (16) below, cannot serve as the antecedent of the lexical anaphor despite the fact that it is adjoined to TP exactly as the preposed phrase in (15a).

(16) *[$_{TP}$ Karera-o$_i$ [[otagai   -no   sensei] -ga   [$_{CP}$[$_{TP}$ Tanaka-ga t_i
        they -ACC each other-GEN teacher-NOM   -NOM
     hihansita] to] itta]] (koto)
     criticized that said   fact
     'Them$_i$, [each other's teachers] said that Tanaka criticized t_i'

---

[4] According to Tada (1990), the total reconstruction does not apply to VP-internal scrambling. I will put aside this type of scrambling for the moment and return to it in Section 5.

[5] Since anaphor binding and the licensing of bound pronouns more or less show the same pattern, I will mainly use examples with lexical anaphors in the discussion of the A/A′ properties of scrambling from here on.

Tada's solution, roughly, is that a scrambling chain counts as an A-chain when and only when its head and tail are both L-related to a single head. He assumes that V raises to I ($=$T) overtly in Japanese. Then, in some sense, the landing site and the initial site in (15a) are both in the projection of the verb *hihansita* 'criticized'. Hence, the TP-adjoined position, Tada claims, counts as an A-position. This does not apply in the case of (16) because the embedded verb clearly does not raise to the matrix T.

Although Tada (1990) is the most important early work that attempted to solve the A/A′ problem with a uniform analysis of scrambling, it seems fair to say that he left the account for the contrast between (15a) and (16) incomplete.[6] It was, then, still unclear why this contrast obtains given that the landing sites of clause-internal scrambling and long scrambling are identical. An elegant solution to the problem was proposed 10 years later in Katahara (2000), to which I now turn.

Kitahara (2000) extends the derivational model of Epstein et al. (1998) by adding the following:

(17)   Items are interpreted as they become interpretable in the course of the derivation.

More specifically, he proposes that NPs are interpreted and enter into binding relations at the positions where their uninterpretable Case features are checked and deleted. This proposal provides an elegant solution to the A/A′ problem when combined with the hypothesis in (18), which Kitahara adapts from Saito (1992).

(18)   The objective Case in Japanese can be checked at the *v*-projection or at the T-projection.

Let us first reconsider the examples in (15) in light of Kitahara's theory. Given (18), *karera-o* 'they-ACC' in (15a) can move directly to the sentence-initial position and be checked for Case there. In this case, by hypothesis the NP is interpreted at the landing site, which c-commands *otagai* 'each other'. Hence, it can serve as the antecedent of the anaphor. On the other hand, successive-cyclic movement through the *v*-projection can be assumed for (15b). Then, *zibunzisin-o* 'self-ACC' is checked for Case and is interpreted in the *v*-projection. Any further movement is irrelevant for the interpretation. In particular, the movement to the sentence-initial position is "semantically vacuous" and cannot affect the binding relations. Hence, (15b) need not be a Condition (C) violation.

More importantly, Kitahara's theory provides a straightforward analysis for the A′-properties of long scrambling. Let us consider (16) again. If Case checking applies derivationally, then *karera-o* 'they-ACC' must have been checked for Case at the *v*-projection or the T-projection of the embedded clause before it moves into the matrix clause. The NP, then, is interpreted and enters into binding relations within

---

[6] See Nemoto (1993) for a detailed discussion of Tada's (1990) analysis. Tada himself tried to improve on his analysis in Tada (1993), but the discussion of the revised analysis will take us too far afield.

the embedded clause. Hence, further movement of the NP to the sentence-initial position of the matrix clause cannot license *otagai* 'each other'.

According to Kitahara (2000), scrambling exhibits the radical reconstruction property because it can move an NP that is already interpreted. Such movement is necessarily semantically vacuous. If LF is the sole level of interpretation, semantically vacuous movement must be "undone" in the LF component. But there is no need for "undoing" in a model where interpretation takes place as the derivation proceeds. Kitahara's analysis is in accord with the intuition expressed in Chomsky and Lasnik (1977) that scrambling is (in some cases) a stylistic rule applying in the PF side of the grammar.

However, Kitahara inherits a major problem from Saito (1992), which claims that Case plays a crucial role in the explanation of the A/A′ properties of scrambling.[7] As he acknowledges, his Case-based analysis is unable to capture the fact that clause-internal scrambling can have A-properties even when the moved phrase is not the object but a PP. He cites the following examples from Takano (1998).[8]

(19)a. *$[_{TP}[$otagai   -no   hahaoya]-ga   [Taroo-to Hanako]-kara hon -o
      each other-GEN mother -NOM   -and   -from book -ACC
   karita] (koto)
   borrowed fact
   '[Each other's mothers] borrowed books from Taroo and Hanako'
  b. ?$[_{TP}[$Taroo-to Hanako]-kara$_i$ [[otagai   -no   hahaoya]-ga t_i hon -o
          -and   -from   each other-GEN mother -NOM book-ACC
   karita]] (koto)
   borrowed fact
   'From Taroo and Hanako$_i$, [each other's mother borrowed books t_i]'

If the preposed PP in (19b) does not have any uninterpretable feature, it should be interpreted at the position of its trace. The movement, then, should be semantically vacuous exactly as in the case of long scrambling.

In the following section, I will suggest an alternative analysis which is based not on Case but rather on chain interpretation. Although I will not discuss the exact landing site of scrambling, I will assume, as in Kitahara (2000), that it is uniformly a position from where A-binding is possible.[9] The analysis incorporates Tada's analysis of (15a-b) and Kitahara's insight on the contrast between (15a) and (16).

---

[7] The problem is actually already present in Mahajan (1990), where it is proposed that A-scrambling is movement to AGR Spec.

[8] In (19b), the postposition *-kara* 'from' apparently does not prevent *Taroo-to Hanako* 'Taroo and Hanako' from binding *otagai* 'each other'. The "invisibility" of P in binding is a general phenomenon in Japanese. See Hoji (1985) and Murasugi (1991) for discussion.

[9] Among the possibilities raised for the landing site of scrambling are the TP-adjoined position and the higher TP Spec in a multiple Spec structure. See, for example, Saito (1985), Fukui (1986), Kuroda (1988), Miyagawa (1997), Saito and Fukui (1998) for detailed discussions on the relevant issues.

3. Scrambling and the derivational interpretation of chains

In this section, I will first suggest a mechanism for the derivational interpretation of chains. I will then show how it can capture the radical reconstruction property and provide a solution to the A/A′ problem.

3.1. Chain interpretation by copy and deletion

As noted in Section 1, Chomsky (1993) proposes that an operator-variable chain is created by copy and deletion. (20a and b) illustrate this procedure for the simple example in (20).

(20)   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 (20b), *who* is retained in CP Spec as an operator, but its copy in the object position is turned into a variable.

As far as I can see, this procedure can be interpreted as follows. An NP has the D-feature, which makes it possible for the NP to have a "reference" and participate in binding/coreference relations.[10] An overt NP also has phonetic features, say, P-features. A *wh*-phrase, in addition, is equipped with an operator feature O. We may say that the O-feature of *who* is responsible for its interpretation in the CP Spec as *[for which x: x a person]*, and its D-feature for its interpretation in the object position as a variable. Then, what is deleted in the CP Spec in (20b) is the D-feature of *who,* and the other features are retained. In the object position, the O-feature and the P-features delete while the D-feature remains.

Construed this way, the copy and deletion procedure can be applied straightforwardly in a cyclic fashion as movement takes place.[11] Let us consider the example in (21) to illustrate one way to execute this idea.

(21)   Who do you think John saw

---

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

[11] Chomsky (1999) in fact suggests that information is sent for interpretation phase by phase throughout the derivation. It seems that as far as *wh*-movement and NP-movement are concerned, this idea can be adopted instead of "chain interpretation upon its formation" without any substantial effect. But I postpone the discussion of phase until Section 5 because it does have significant consequences for the analysis of scrambling.

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

(22)   [$_{CP}$ 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 only one position. The P-features must be retained at the head position of the chain. This, if anything, must be part of the definition of overt movement. For the rest, suppose further that deletion is constrained by selection in a broad sense, and that the features can remain only in positions where they are selected.[12] The D-feature is selected in the object position but not in the CP Spec. Hence, it must delete at the CP Spec. Chomsky (1998) proposes that the movement of *who* to the CP Spec takes place because a feature of the C head (call it the EPP-O feature) attracts the O-feature of the *wh*-phrase. Extending the use of the term 'selection', we may say then that the O-feature satisfies a selectional requirement of the C head.[13] As it is clear that the O-feature is not selected in the object position, it can be retained only in the CP Spec. We then obtain (23).

(23)   [$_{CP}$ who [$_{TP}$ John saw who]]
            {P,O}   {D}

   If interpretation takes place as a chain is created, (23) must be interpreted before the derivation proceeds. Here, as Chomsky (1999) notes, the CP Spec cannot be interpreted at this point because it is involved in a further operation: *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. As the O-feature is already deleted at the object position, *who* at the initial site, which only has the D-feature, is interpreted appropriately as a variable. With the next step of the *wh*-movement from the embedded CP Spec to the matrix CP Spec, the following structure is created:

(24)   [$_{CP}$ who [$_{C'}$ do [$_{TP}$ you think [$_{CP}$ who ([$_{TP}$ John saw x])]]]]]
            {P,O}   {P,O}

Note that the moved *wh*-phrase carries only the P-features and the O-feature, since the D-feature is deleted at the embedded CP Spec prior to the movement. 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, can also be retained there. The copy of *who* in the embedded CP Spec disappears as all of its features are deleted.

---

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

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

3.2. Radical reconstruction as a consequence of chain interpretation

In this subsection, I will apply the derivational chain interpretation mechanism to scrambling. I assume as in my previous works that Japanese scrambling is not feature-driven and is truly an optional movement operation.[14] The *wh*-movement in (25a) is triggered by the [+Q] C, while the raising in (25b) is necessiated by the EPP feature on T.

(25)a. What$_i$ did Mary buy t_i
    b.   Bill$_i$ is likely [t_i to be in the house]

The assumption is that there is no feature that triggers scrambling. If we continue to employ the broad usage of the term 'selection', this means that *wh*-movement and NP-movement apply to satisfy selectional requirements of heads, but scrambling has nothing to do with selection.

Another proposal entertained in the literature is that radical (= total) reconstruction is a consequence of the copy and deletion mechanism (see Lee (1994) and Saito (1994)). The idea is that radical reconstruction results when the higher copies in a scrambling chain are deleted. Assuming that scrambling does not satisfy any selectional requirement, I will suggest a specific instantiation of this idea in this subsection. Then, in the following subsection, I will show how it solves the A/A′ problem.

Let us first consider a simple case of scrambling in (26).

(26) [$_{TP}$ Sono hon -o$_i$ [Yamada-ga t_i yonda]] (koto)
        that book -ACC   -NOM read   fact
     'That book$_i$, Yamada read t_i'

Given the copy and deletion analysis, this example has the following initial structure:

(27) [$_{TP}$ Sono hon-o [... sono hon-o ...]]
         {P,D}   {P,D}

The P-features are interpreted at the sentence-initial position. The D-feature cannot be retained in this position because by hypothesis it is not selected there. Then, deletion creates (28).

---

[14] See Saito (1985, 1989), Fukui (1986), Kuroda (1988), Tada (1993), Saito and Fukui (1998) among many others for arguments for this assumption. This view is shared by Bošković and Takahashi (1998), who claim that scrambled phrases are generated directly in their surface positions and undergo covert lowering to receive thematic roles. For the opposing view that scrambling is feature-driven, see, for example, Miyagawa (1997, 2000), Grewendorf and Sabel (1999), and Kawamura (2001). It is actually unclear how crucial the assumption is for the proposals in this paper. Miyagawa's feature-checking analysis of scrambling entails that A-scrambling and A′-scrambling are distinct operations, and hence, contradicts the fundamental hypothesis of this paper that scrambling is uniform. But the situation is less clear with Kawamura's analysis. She develops Grewendorf and Sabel (1999), and argues that Japanese scrambling is triggered uniformly by a "scrambling feature," which deletes completely after it is checked. It may be possible to reconcile her analysis and the proposals in this paper.

(28)   [$_{TP}$ Sono hon-o [... sono hon-o ...]]
         {P}   {D}

This structure already represents radical reconstruction. Only the phonetic features are retained at the scrambled position, and hence, (28) is indistinguishable from cases of PF movement. Scrambling, then, is always semantically vacuous.

   A similar result obtains with long scrambling.

(29)   [$_{TP}$ Sono hon -o$_i$ [Tanaka-ga   [$_{CP}$ Yamada-ga t_i yonda to]   omotteiru]
         that   book-ACC   -NOM   -NOM read   that think
       (koto)
        fact
       'That book$_i$, Tanaka thinks [that Yamada read t_i]'

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

(30) [$_{CP}$ Sono hon-o [$_{TP}$ ... sono hon-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. We then end up with (31).[15]

(31) [$_{CP}$ Sono hon-o [$_{TP}$ ... sono hon-o ...]]
         {P}   {D}

In the matrix clause, *sono hon-o* 'that book-ACC' moves on and the structure in (32) is created.

(32) [$_{TP}$ Sono hon-o [... [$_{CP}$ sono hon-o [$_{TP}$ ... sono hon-o ...]]]]
         {P}   {P}   {D}

The P-features in the embedded C-projection delete to yield (33).

(33) [$_{TP}$ Sono hon-o [... [$_{CP}$ [$_{TP}$ ... sono hon-o ...]]]]
         {P}   {D}

Since only the phonetic features are displaced, we expect the movement to be semantically vacuous in this case also.

---

[15] 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 C-projection. The resulting structure is identical to (31).

So far, we have seen how the copy and deletion mechanism captures the radical reconstruction property, or more specifically, the effects of obligatory total reconstruction in the sense of Tada (1990). One attractive feature of this analysis is that it conforms strictly to the fundamental hypothesis that Japanese scrambling is uniform, whether it is clause-internal or long-distance. Scrambling is always to a position without any selectional relation, and hence, only the phonetic features of the moved phrase are retained at the landing site. But then, a question remains with the A/A′... problem. As discussed above, clause-internal scrambling, and only clause-internal scrambling, exhibits the A-properties with respect to the binding phenomena. The relevant examples in (9) and (11b) are repeated again in (34a) and (34b) below.

(34)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. *[$_{TP}$ Karera-o$_i$ [[otagai -no sensei] -ga [$_{CP}$[$_{TP}$ Tanaka-ga t_i
        they -ACC each other-GEN teacher-NOM   -NOM
     hihansita] to] itta]] (koto)
     criticized that said fact
     'Them$_i$, [each other's teachers] said that Tanaka criticized t_i.'

If scrambling is always semantically vacuous, how is it possible to explain this contrast? I will turn to this question in the following subsection.

3.3. Condition (A) as an 'anywhere condition'

Although the final forms of the chains created by clause-internal scrambling and long scrambling look identical, there is a crucial difference in the ways in which those final forms are obtained. The movement in (34a) initially constructs the chain in (35).

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

Then, the D-feature of the sentence-initial *karera-o* 'they-ACC' and the P-features of the object are deleted to yield (36).

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

This means that at one point of the derivation, the D-feature of *karera-o* appears in a position c-commanding *otagai* 'each other'.

The situation is quite different in the case of (34b). The following chain is created by copy and deletion in the embedded CP:

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

Then, *karera-o* with only phonetic features moves on to the matrix-initial position as in (38).

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

The copy of *karera-o* in the embedded C-projection disappears as its P-features are deleted, and the final form of the scrambling chain, shown in (39), is virtually identical to (36).

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

However, it is clear from (38) and (39) that the D-feature of *karera-o* does not c-command *otagai* 'each other' at any point of the derivation of (34b). Then, the contrast between (34a) and (34b) follows if Condition (A) is an anywhere condition, i.e., a condition that can be satisfied at any point of the derivation. And arguments for this formulation of Condition (A) have already been presented in Belletti and Rizzi (1988), Lebeaux (1988), and Epstein et al. (1998).

Here, I will briefly discuss Belletti and Rizzi's argument, which is based on examples of "backward anaphora" such as (40).

(40) 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 (41).

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

Then, the grammaticality of (40) 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 (42) and (43).

(42)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$_j$]
(43)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 higher clauses. Yet, (42a) and (43a) are clearly much better than them. Belletti and Rizzi's analysis straightforwardly extends to the grammatical cases in (42) and (43). If

anaphors and bound pronouns can be licensed at any point of the derivation, *each other* in (42a) and *his* in (43a) can be licensed before movement applies in these examples.

The argument based on (40), (42) and (43) that Condition (A) is an anywhere condition is not conclusive. Lebeaux (1998) in fact takes (42) and (43) 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):

(44)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 subjects in (44) can reconstruct 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 (45), as Chomsky (1993) points out.

(45)   [The claim that John$_i$ was asleep]$_j$ seemed to him$_i$ [t_j to be correct]

If Condition (C) is an everywhere condition, this example violates it before raising takes place.[16] This example in fact seems to show that Condition (C) must be an LF condition. Given the elimination of D-structure and S-structure with the Minimalist Program, Condition (C), to the extent that it is a syntactic principle, must be an everywhere condition as Lebeaux proposes or an LF condition. And (45) excludes the former option. Once it is established that Condition (C) applies at LF, we can trace Lebeaux's reasoning backwards. (44) should violate Condition (C) at LF. This implies that NP-movement is not subject to reconstruction.[17] Hence, the only way to explain (40), (42) and (43) is to make Condition (A) an anywhere condition.

As noted above, the explanation for the contrast in (34) is straightforward if Condition (A) is an anywhere condition. The D-feature of the preposed NP *karera-o* 'they-ACC' c-commands the lexical anaphor *otagai* 'each other' at one point in the derivation in the case of (34a), but not in the case of (34b). Further, it was

---

[16] This problem is discussed in Lebeaux (1988, 1998), where he suggests, following van Riemsdijk and Williams (1981), that Condition (C) applies (or starts applying) after NP-movement but prior to *wh*-movement.

[17] These conclusions have been proposed in the literature. For example, Chomsky (1993) argues that Binding conditions are in general LF conditions. Arguments for the absence of reconstruction with NP-movement can be found in Mahajan (1990), Chomsky (1995), and Lasnik (1999), among many other places.

concluded above that Condition (C) applies at LF. This automatically solves the initial problem posed by (15b), which is repeated in (46) with another similar example.

(46)a. [$_{TP}$ Zibunzisin-o$_i$ [Taroo-ga t_i semeta]] (koto)
            self   -ACC   -NOM   blamed   fact
        'Himself$_i$, Taro blamed t_i'
   b. [$_{TP}$ Otagai   -o$_i$ [Taroo-to Hanako-ga t_i semeta]] (koto)
            each other-ACC   -and   -NOM   blamed   fact
        'Each other$_i$, Taro and Hanako blamed t_i'

If scrambling is uniformly to a position from where A-binding is possible as I have been assuming, and Condition (C) is an everywhere condition as Lebeaux proposed, then these examples should be Condition (C) violations. But if Condition (C) applies at LF, these examples are correctly predicted to be grammatical. The derivation of (46a) is illustrated in (47).

(47)a. [$_{TP}$ Zibunzisin-o [. . . Taroo-ga . . . zibunzisin-o . . .]]
            {P,D}   {P,D}
   b. [$_{TP}$ Zibunzisin-o [. . . Taroo-ga . . . zibunzisin-o . . .]]
            {P}   {D}

As before, the scrambled NP is initially copied at the sentence-initial position. Then, the D-feature is deleted at the landing site, and the P-features are deleted at the initial site. Thus, after deletion applies, the D-feature of the scrambled NP is not in a position c-commanding the name *Taroo*. Hence, there is no Condition (C) violation.

   I have argued above that the derivational interpretation of chains enables us to maintain the uniform analysis of scrambling and still capture the apparent A/ A′ ambiguity of clause-internal scrambling as well as the A′-properties of long scrambling. The account proposed for (34a) and (46) is virtually identical to Tada's (1990). As noted above, he maintains that radical reconstruction is obligatory, and Condition (A) is an S-structure (or anywhere) condition while Condition (C) applies at LF. I simply proposed that radical reconstruction is a consequence of the derivational chain interpretation mechanism, and provided a supporting argument for his assumptions on the binding conditions. The contrast between (44) and (46) highlights the main proposals here. The landing sites of NP-movement and scrambling are both positions from where A-binding is possible, as indicated by (13) and (34a). Yet, only NP-movement induces Condition (C) violations at LF as observed in (44). This is because scrambling is subject to radical reconstruction, while there is no reconstruction in the case of NP-movement.

   The proposed account for the contrast between clause-internal scrambling and long scrambling in (34) appeals to derivational interpretation like that in Kitahara (2000). It is an improvement on Kitahara's Case-based approach, as it naturally

extends to cases of PP scrambling. The relevant example, (19b), is repeated below as (48).

(48) ?[$_{TP}$[Taroo -to Hanako]-kara$_i$ [[otagai   -no   hahaoya]-ga t_i hon -o
                  -and   -from   each other-GEN mother -NOM   book-ACC
         karita]] (koto)
         borrowed fact
         'From Taroo and Hanako$_i$, [each other's mother borrowed books t_i]'

This example is correctly predicted to be grammatical, since the D-feature of *Taroo-to Hanako* 'Taroo and Hanako' in the scrambled position binds *otagai* 'each other' before it is deleted.

4. The proper binding effect with scrambling

In this section, I will turn to the proper binding effect on scrambling. I will first argue that the effect is real, showing that the relevant facts cannot be accounted for by Chomsky's (1995) Minimal Link Condition, a possibility suggested in Kitahara (1997). Then, I will consider the implications of the effect for the chain interpretation mechanism proposed in the preceding section. Finally, I will suggest a derivational formulation of the proper binding condition.

4.1. Confirmation of the problem

It is noted in Saito (1985) that scrambling is constrained by the proper binding condition. A relevant example is shown in (49).

(49) *[$_{TP}$[Hanako-ga t_i iru to]$_j$ [Sooru-ni$_i$ [Taroo-ga t_j omotteiru]]] (koto)
                  -NOM   be that Seoul -in   -NOM   think   fact
         '[That Hanako lives t_i], in Seoul$_i$, [Taroo thinks t_j]'

This example is derived from (50a) by multiple applications of scrambling.

(50)a. [$_{TP}$ Taroo-ga   [Hanako-ga   Sooru-ni iru to] omotteiru] (koto)
                  -NOM   -NOM Seoul -in be that think   fact
         'Taroo thinks that Hanako lives in Seoul'
     b. [$_{TP}$ Sooru-ni$_i$ [Taroo-ga   [Hanako-ga t_i iru to] omotteiru]] (koto)
            Seoul -in   -NOM   -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 (50b). Then, the embedded CP is scrambled to the position preceding the PP. The resulting sentence

(49) is totally ungrammatical. CP scrambling and multiple scrambling are both allowed in Japanese, as shown in (51a) and (51b).

(51)a. [$_{TP}$[Hanako-ga   Sooru-ni iru to]$_i$ [Taroo -ga t_i omotteiru] (koto)
                  -NOM Seoul -in be that   -NOM think   fact
        '[That Hanako lives in Seoul]$_i$, Taroo thinks t_1'
   b. [$_{TP}$ Sono hon -o$_i$ [Hanako-ni$_j$ [Taroo-ga   [$_{CP}$ Ziroo-ga t_j t_i watasita to]
          that book-ACC   -to   -NOM   -NOM handed that
        omotteiru]]] (koto)
        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 (49) to the unbound trace t_i.

   However, examples like (49) have become problematic because the proper binding account seems untenable under the Minimalist assumptions. Recall here the hypothesis proposed in Saito (1989) that scrambling is subject to radical reconstruction at LF. Given this hypothesis, (49) cannot be ruled out by the proper binding condition at LF because no trace exists after the reconstruction. It was proposed in Saito (1989) then that (49) is ruled out by the S-structure application of the proper binding condition, an analysis that clearly contradicts the Minimalist assumptions.

   An extremely interesting suggestion is made in Kitahara (1997) on this problem. He first provides an explanation in terms of the Minimal Link Condition (MLC) for Müller's (1996) generalization, illustrated in (52) and (53).

(52)a. [$_{VP}$ t_i Gelesen]$_j$ hat das Buch$_i$ keiner t_j
              read   has the book no one
        'No one has read the book'
   b.*daβ [$_{VP}$ t_i gelesen]$_j$ das Buch$_i$ keiner t_j hat
        that   read   the book no one has
        'that no one has read the book'
(53)   ??daβ [$_{VP}$ sich   so richtig t_i dumm vorgekommen] der   Fritz$_i$ noch nicht ist
              that   REFL so really   stupid struck-as   ART   yet   not is
        'that Fritz has not yet struck himself as really stupid'

(52a) shows that in German a phrase containing a trace of scrambling can be topicalized. An unbound trace is apparently allowed in this case. (52b), on the other hand, shows that an unbound trace cannot be created by multiple applications of scrambling. (53) indicates that a phrase containing an NP trace can be scrambled to a position higher than the antecedent of the trace.[18] Given these examples and others, Müller proposes the following generalization:

---

[18] According to Müller, the awkwardness of (53) is due to the VP-scrambling, and has nothing to do with the unbound trace.

(54)   A phrase containing a trace of movement cannot undergo movement of the same type (operator movement, scrambling, NP-movement).
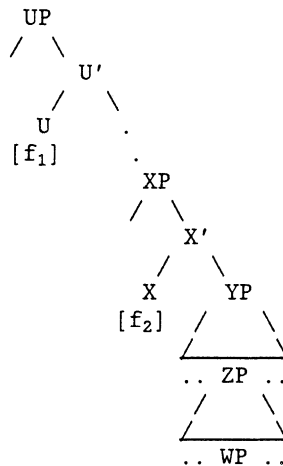
Müller's generalization is confirmed further by the English examples in (55) and (56).

(55)   [How likely [t_i to win the race]]$_j$ is John$_i$ t_j
(56) a.   ??Who$_i$ does John wonder [which picture of t_i]$_j$ Mary liked t_j
     b.   *[Which picture of t_i]$_j$ does John wonder who$_i$ Mary liked t_j

In (55), two distinct kinds of movement, NP-movement and *wh*-movement, apply and create an unbound trace. The result is grammatical as predicted by Müller's generalization. On the other hand, the matrix *wh*-movement in (56b) operates on a phrase that contains a trace of *wh*-movement. The example is hopeless, again, as predicted by the generalization.

Kitahara's (1997) explanation for Müller's generalization is very principled. Let us consider the structure of the relevant cases shown in (57).

(57)

```
                        UP
                       /  \
                          U′
                         /  \
                        U    .
                      [f₁]     .
                                .
                               XP
                              /  \
                                 X′
                                /  \
                               X    YP
                             [f₂]  /    \
                                  /_____\
                                 .. ZP ..
                                  /    \
                                 /_____\
                                .. WP ..
```

First, the phrase WP is attracted by the feature f_2 of X, and it moves to the Spec of XP. Then, the feature f_1 of U attracts ZP, which contains the trace of WP, and pulls it up to the Spec of UP. Here, in the illicit cases, WP and ZP undergo the same type of movement. This means that f_1 is identical to f_2, which in turn implies that either WP or ZP could check the feature f_2. The movement of WP to the Spec of XP, then, violates the MLC, which, roughly, requires a feature to attract the closest phrase with which it can enter into a checking relation. This problem does not arise in the legitimate cases, where WP and ZP undergo different types of movement. In these cases, f_1 and f_2 are distinct features, and hence, f_2 cannot enter into a checking relation with ZP. Thus, it can attract WP without violating the MLC.

Note that grammatical examples like (52a), (53) and (55) are in apparent violation of the proper binding condition. This provides an additional reason to reconsider

the proper binding analysis of (49). Given this situation, Kitahara (1997) suggests the possibility that (49) is also excluded by the MLC. Note that this example falls under Müller's generalization because the CP scrambled to the matrix-initial position contains a trace of scrambling. It seems then that the example can be accounted for in exactly the same way as the ungrammatical (52b) and (56b).

   Although Kitahara's suggestion is quite attractive, it cannot be adopted in this paper. The MLC analysis crucially assumes that Japanese scrambling is feature-driven, an assumption not employed here. And more importantly, there are examples of Japanese scrambling that exhibit the proper binding effect, and yet, cannot be accounted for by the MLC. One of them is shown in (58).

(58) *[$_{TP}$ [PRO t_i iku koto]-ga$_j$ Sooru-made$_i$ Taroo-ni t_j meizirareta]
                  go N   -NOM Seoul-to   -to   ordered-was
        '[To go t_i]$_j$, to Seoul$_i$, was ordered Taroo t_j'

This example involves two instances of movement. First, the PP *Sooru-made* 'to Seoul' is scrambled out of the infinitival complement as illustrated in (59).

(59)a. Hanako-ga Taroo-ni [PRO Sooru-made iku koto]-o   meizita
              -NOM   -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
              -NOM Seoul -to   -to   go N   -ACC ordered
        'Hanako, to Seoul$_i$, ordered Taroo to go t_i'

This type of scrambling is legitimate as (59b) shows.[19] Note that the infinitive complement in (59) is headed by the nominalizer *koto,* and appears in the object position. Thus, it can be passivized as a regular object as shown in (60).

(60)[PRO Sooru-made iku koto]-ga$_j$ Taroo-ni t_j meizirareta
          Seoul -to   go N   -NOM   -to   ordered-was
        '[To go to Seoul]$_i$ was ordered Taroo t_i'

The ungrammatical (58) is derived when this passive applies after the scrambling in (59b).

   (58), like (49), contains an unbound trace of scrambling. But this example does not fall under Müller's generalization, as it is derived by scrambling and passive.

---

[19] The landing site of this scrambling is the matrix *v*P 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 control infinitive like the one in (59) consists of a TP directly embedded under a projection of the nominalizer *koto,* 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 discussions on this point, and Murasugi (1991) for a general discussion on the structure of Japanese complex NPs.

Kitahara's MLC account does not extend to this example either, unless passive and scrambling are triggered by the same feature.[20] It seems then that the proper binding effect on scrambling is real and calls for an explanation. Once it is established that the proper binding effect exists as a real phenomenon, two questions arise. The first is why the effect is absent in examples like (52a), (53) and (55), in particular those cases where a trace of NP-movement is unbound. The second concerns the formulation of the relevant condition. I will discuss these questions in turn in the following two subsections.

4.2. The absence of NP-traces

An answer to the first question is already suggested in Kuno (2001). He adopts Lasnik's (1999) proposal that NP-movement does not produce a trace, and argues that the proper binding condition is satisfied vacuously in examples like (55), repeated below as (61).

(61)   [How likely [t_i to win the race]]$_j$ is John$_i$ t_j

If t_i does not exist, then nothing violates the proper binding condition in this example. Lasnik's proposal, in turn, is based on the absence of scope reconstruction with NP-movement. Since the relevant scope phenomenon requires involved discussion that would take us too far afield, I will instead present an independent argument for his proposal here.[21]

First, recall the conclusion in the preceding section that the examples in (44) are excluded by Condition (C) at LF. (44a and b) are repeated below in (62a and b).

(62)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]

Crucial in this analysis was the hypothesis that there is no reconstruction with NP-movement. But how can this be expressed in precise terms with the copy and deletion analysis of movement? I have hypothesized above that an NP is accompanied by a D-feature and P-features. Then, NP-movement should initially create a chain of the following form:

---

[20] Miyagawa (2000) argues that A-scrambling is triggered by the EPP-feature. It may appear that this makes it possible to analyze (58) as an MLC violation. But as noted in Footnote 19, long scrambling out of a control infinitive can have A′-properties. The following example confirms this point:

(i)   Hanako-ga   otagai   -o$_i$ [Taroo-to Ziroo]-ni [PRO t_i hihansuru koto]-o   meizita
               -NOM each other-ACC   -and   -to   criticize   N -ACC ordered
      'Hanako, each other$_i$, ordered Taroo and Ziroo to criticize t_i'

Hence, even if we assume Miyagawa's theory, the prediction is that (58) should be grammatical with A′-scrambling.

[21] Another argument can be found in Saito and Hoshi (2000), which also considers the possiblity that NP-movement does not produce a trace on yet independent grounds.

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

Here, the P-features are retained at the landing site. And (62a and 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 in only one position, the P-features and the D-feature 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.

  I have argued that the discussion in the preceding section implies that NP-movement leaves no trace. One more thing needs to be said to make the analysis complete. In the discussion of chain interpretation in the preceding section, I hypothesized that a D-feature and an O-feature can be retained only in positions where they are selected. This prevents an O-feature from being retained in an argument position, and a D-feature from being in an operator position. But given this hypothesis, the D-feature in (63) can be retained in either position. If selection is construed in a broad sense to include the feature-checking relation, then the D-feature is selected in both positions. In order to guarantee that the D-feature is retained only at the landing site, I tentatively suggest (64), hoping that it will eventually receive a principled explanation.

(64) Chain interpretation makes the chain minimum.

The P-features in (63) 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, (64) forces the deletion of the D-feature at the initial site, and consequently, prevents NP-movement from leaving a trace.

  It is probably worth noting here that the elimination of NP-traces is quite plausible when the issue is considered from a broader perspective. As Lasnik (1999) explicitly mentions, there is no clear theta-theoretic reason that an NP-trace should exist. An NP can pick up a theta-role, and then, undergo movement without leaving a trace. Then, whether an NP-trace should be assumed or not depends to a large extent on how NP-movement is constrained by the syntactic principles. If the principles explain the relevant phenonena with the aid of NP-traces, the traces should be assumed. If, on the other hand, the traces get in the way in the explanation, they should be eliminated.

  The discussion on NP-movement in Lasnik and Saito (1992) seems quite relevant in this context. There we argued that Chomsky's (1986a) chain condition and our generalized uniformity condition account for most cases of illicit NP-movement. The chain condition requires the head of an A-chain to be its unique Case position. The relevant part of the generalized uniformity condition is shown in (65).

(65) Suppose that β bears a theta-role assigned by α. Then, if γ is a barrier for α, γ dominates β.

This condition states basically that an NP movement cannot cross a barrier in the sense of Chomsky (1986b). One of the conclusions drawn in this discussion was that there is no need to assume that NP-traces are subject to Condition (A) or the classical ECP as formulated in Chomsky (1981), Lasnik and Saito (1984), or any other work around the time. Thus, examples like those in (66), which had been excluded by Condition (A), are explained by the chain condition.

(66)a.   *John$_i$ seems [$_{CP}$ [$_{TP}$ Mary visited t_i]]
   b.   *John$_i$ seems [$_{CP}$ [$_{TP}$ t_i is intelligent]]

And the super-raising example in (67) is accounted for by (65).

(67)   *John$_i$ seems [$_{CP}$ that [$_{TP}$ it is believed [$_{TP}$ t_i to be intelligent]]]

(65)   also accounts for (68), which Lasnik (1985) explained in terms of the locality condition on chains, yielding the condition redundant.

(68)   *John$_i$ seems [$_{CP}$ [$_{TP}$ [$_{NP}$ his$_i$ belief [$_{TP}$ t_i to be intelligent]] is crazy]]

   This discussion on NP-movement rai=0ses serious doubts for the existence of NP-traces. If there were cases of NP-movement that are excluded by Condition (A) or the classical ECP, traces would have been necessary. This is so because these principles apply to traces. The locality condition on chains also requires traces because the relevant chains contain them. But if NP-movement is not constrained by these conditions, but only by the chain condition and (65), then NP-traces seem redundant. (65) clearly does not refer to traces. And Chomsky (1986a) proposes to explain the relevant part of the chain condition by the classical last resort principle, which implies that NP movement applies only when the NP needs to be assigned Case. The classical last resort principle does not refer to traces, either. Note that this conclusion stands even with more advanced analysis. Chomsky (1995), for example, explains many of the examples excluded by (65) in terms of the MLC, and the MLC does not require traces. He also proposes to do away with the last resort principle in favor of a refined theory of formal features. Again, NP-traces are not necessary for this explanation.
   The only empirical facts that made Lasnik and Saito (1992) continue to assume NP-traces are of the kind shown in (69).

(69)a.   *[How likely t_i to be taken of John]$_j$ is advantage$_i$ t_j
   b.   *[How likely t_i to be a riot]$_j$ is there$_i$ t_j

   The configuration of these examples seems identical to that of (61). The only difference, it appears, is that the NP that undergoes raising is a referential NP in (61) while it is an idiom chunk or an expletive in (69). Given the contrast between (61) and (69), there are two possible ways to proceed. One may assume that the configuration

is in general disallowed, hoping that the apparent grammaticality of (61) is to be explained away on independent grounds. This is the possibility pursued in Lasnik and Saito (1992). We argued that (61) and (69) are both ruled out by the proper binding condition, and (61) appears to be allowed because the sentence can have an alternative structure involving control rather than raising. Given this analysis, NP-traces cannot be dispensed with since their presence is assumed crucially in the account for (69). But then, (61) ceases to be a problem for the proper binding condition and it can be assumed that the condition is quite general.

Another way to approach the contrast between (61) and (69) is to assume that *wh*-movement can follow raising generating examples like (61), hoping this time that (69) can be ruled out on independent grounds. This approach implies that (61) and (69) are not proper binding violations. Hence, to the extent that it is viable, examples such as (69a and b) do not constitute decisive evidence for NP-traces.

This is the context in which Lasnik (1999) suggested to eliminate NP-traces. As noted above, the conceptual basis for NP-traces is quite shaky. Further, there is very little empirical reason, if any, to assume NP-traces. Then, if there is evidence that NP-traces do not exist, it provides strong reason to eliminate them. Since the discussion in the preceding section provides such evidence, as shown above, I conclude, like Lasnik (1999) and Kuno (2001), that NP-movement does not leave a trace.

As Kuno (2001) argues, the absence of NP-traces immediately explains the grammaticality of (61) as well as the German (52), repeated below as (70).

(70) ??daβ [$_{VP}$ sich   so richtig t_i dumm vorgekommen] der Fritz$_i$ noch nicht ist
         that   REFL so really   stupid struck-as   ART   not yet   is
         'that Fritz has not yet struck himself as really stupid'

These examples are not proper binding violations because they do not contain unbound traces. The last problematic case is the German (51a), repeated below in (71).

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

In this example, scrambling is followed by topicalization, and hence, it is a trace of scrambling that is apparently unbound. Here, I have little to say about German scrambling, but given the analysis presented so far, the grammaticality of (71) implies that it need not leave a trace. This implies in turn that it should be like NP-movement in the relevant respects. It should be feature-driven, as proposed for example in Grewendorf and Sabel 1999, and further, the feature that attracts the scrambled phrase must be able to enter into checking (and hence, selectional) relation with the categorial feature of the scrambled phrase. If this is

tenable, then (71) also ceases to be a problem for the proper binding condition.[22]

4.3. On the proper formulation of the proper binding condition

I have been assuming Epstein et al. (1998) and Kitahara's (2000) derivational model for interpretation in this paper. 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, which I adopted and argued for, 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 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 straightforwardly as an LF condition. Note that the argument in Saito (1989) that proper binding is an S-structure condition no longer holds. It was proposed there that scrambling is literally undone in LF to yield the radical reconstruction effects. Hence, it was necessary to apply the proper binding condition at S-structure to exclude examples like (49), repeated below as (72).

(72) *$[_{TP}$ [Hanako-ga t_i iru to]$_j$ [Sooru-ni$_i$ [Taroo-ga t_j omotteiru]]] (koto)
                -NOM   be that Seoul -in   -NOM   think   fact
     '[That Hanako is t_i]$_j$, in Seoul$_i$, [Taroo thinks t_j]'

On the other hand, I proposed in this paper that the radical reconstruction effects obtain because of the application of the copy and deletion mechanism in a specific way. Thus, in the scrambling chain in (73a), the D-feature is deleted at the landing site while the P-features are retained there as shown in (73b).

(73)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 remains even at the output of the syntactic computation. Hence, (72) can be ruled out at LF by the proper binding condition, which can be formulated for example as in (74).

---

[22] German scrambling, then, is feature-driven while Japanese (and Korean) scrambling is not. This is not surprising given the many differences between the two movement operations. As is well known, German scrambling is more local and does not allow extraction out of a finite CP. And more importantly, it seems to have ''semantic effects'' unlike Japanese scrambling. It induces the specfic/definite interpretation of the moved phrase as discussed in detail in Diesing (1992), and it does not apply to *wh*-phrases as noted in Grewendorf and Sabel (1999). See Grewendorf and Sabel (1999) and Saito (2000) for discussion on the differences between German scrambling and Japanese scrambling.

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

However, even if we accept the hypothesis that there are LF conditions, we may ask what kinds of conditions qualify as such.[23] 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 a formal mechanism to specify 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 relations 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 the 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 (74), is an LF condition. (74) seems internal to the language system. Then, it seems desirable to reformulate the condition as a derivational constraint. Further, 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 (73b) is constrained at LF. According to the analysis presented above, scrambling chains are headed only by phonetic features after deletion applies. Thus, if (74) applies to them at LF, it should demand the phonetic features to be in a position c-commanding the D-feature they are associated with. However, it is at least strange to say that an LF condition refers to phonetic features. Those features are plausibly invisible at the LF interface.

Based on these considerations, I would like to suggest a reformulation of the proper binding condition as a constraint on the application of Merge.[24] 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 a 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 (75) and (76).[25]

(75) α is subject to Merge only if α is a complete constituent.
(76) α is a *complete constituent* $=_{df}$ (i) α is a term, and

---

[23] The discussion here is quite speculative. A more precise theory of interpretation is needed to make the discussion more concrete. For example, it is unclear at this point what 'LF' exactly means. It may be the output of the syntactic computation, or the collection of the interpretive information obtained throughout the course of the derivation.

[24] Here, I understand Merge as either pure Merge or Merge as part of Move in the sense of Chomsky (1994).

[25] 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 proposes a similar condition to exclude unbound *wh*-traces.

(ii) if a position within α is a member of a chain γ, then every position of γ is contained within α.

This condition prevents the merger of an object that contains a trace but not its antecedent. In particular, it prevents the merger of the CP containing a trace at the matrix TP in (72).

5. Further issues

Before I conclude this paper, I will make brief remarks on two remaining issues. I will first suggest that if an anaphor contains a feature (A or [+ anaphor]) that requires a binder, the feature can be detached from the D-feature by the application of deletion in chains. Secondly, I will discuss VP-internal scrambling in relation to Chomsky's (1998) derivational phase.

5.1. Condition (A) and the radical reconstruction of anaphors

The anywhere nature of Condition (A) was crucial in the solution for the A/A′ problem proposed in Section 3. The following examples discussed in Dejima 1999 confirm this hypothesis.

(77) a. Taroo-ga$_i$ [$_{CP}$ Hanako-ga$_j$ [$_{CP}$ Ziroo-ga$_k$ zibunzisin-o$_{*i,*j,k}$ hihansita to]
      -NOM   -NOM   -NOM self   -ACC criticized that
    itta to] omotteiru (koto)
    said that think   fact
    'Taroo$_i$ thinks [that Hanako$_j$ said [that Ziroo$_k$ criticized self$_{*i,*j,k}$]]'
  b. Taroo-ga$_i$ [$_{CP}$ Hanako-ga$_j$ [$_{CP}$ zibunzisin-o$_{*i,j,k}$ Ziroo-ga$_k$ t hihansita to]
      -NOM   -NOM self   -ACC   -NOM criticized that
    itta to] omotteiru (koto)
    said that think   fact
    'Taroo$_i$ thinks [that Hanako$_j$ said [that self$_{*i,j,k}$ Ziroo$_k$ criticized t]]'
  c. Taroo-ga$_i$ [$_{CP}$ zibunzisin-o$_{i,j,k}$ Hanako-ga$_j$ [$_{CP}$ Ziroo-ga$_k$ t hihansita to]
      -NOM self   -ACC   -NOM   -NOM criticized that
    itta to] omotteiru (koto)
    said that think   fact
    'Taroo$_i$ thinks [that self$_{i,j,k}$ Hanako$_j$ said [that Ziroo$_k$ criticized t]]'

The status of *zibunzisin* 'self' as a local anaphor is controversial since the required locality is not always clear-cut.[26] But the judgments in (77) are reasonably clear on

---

[26] On this point see Nakamura (1996) and the references cited there.

contrastive basis. As indicated, the possible antecedents for *zibunzisin* increase as the anaphor is preposed further.

Another contrast from Dejima (1999) is shown in (78).

(78)a. *Karera-ga [CP Hanako-ga [CP Ziroo-ga otagai   -o
     they -NOM   -NOM   -NOM each other -ACC
     sonkeisiteiru to] itta to] omotteiru (koto)
     respect   that said that think   fact
     'They think [that Hanako said [that Ziroo respect each other]]'
  b. Karera-ga [CP otagai   -o Hanako-ga [CP Ziroo-ga *t* sonkeisiteiru
     they -NOM each other-ACC   -NOM   -NOM respect
     to] itta to] omotteiru (koto)
     that said that think   fact
     'They think [that each other, Hanako said [that Ziroo respect *t*]]'

Again, it is controversial whether *otagai* is a local anaphor, but the contrast is clear. (78b) shows that when the anaphor is scrambled to the sentence-initial position of the second clause, the matrix subject qualifies as the antecedent.

What is particular striking here is the grammaticality of (78b) and the fact that *zibunzisin* 'self' in (77c) can take the matrix subject as its antecedent. If the classical radical reconstruction analysis is assumed, these facts constitute straightforward evidence that Condition (A) is an anywhere condition. The scrambled anaphors must be reconstructed at LF, and hence, cannot satisfy Condition (A) at this level. On the other hand, if Condition (A) is an anywhere condition, it can be satisfied prior to the reconstruction.

However, refinements are necessary for a more precise account of the examples in (77) and (78). Recall first that a phrase preposed by long scrambling cannot serve as the antecedent of a lexical anaphor. The relevant example (34b) is repeated below in (79).

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

According to the analysis proposed above, the movement of *karera-o* 'they-ACC' from the embedded C-projection to the matrix initial position is movement of the phonetic features alone as shown in (38), repeated as (80).

(80) [TP Karera-o [… otagai … [CP karera-o [TP … karera-o …]]]]
      {P}   {P}   {D}

Then, the final step of the scrambling of *otagai-o* 'each other-ACC' in (78b) should also be movement of phonetic features. But it cannot be that the phonetic features establish the desired binding relation. Then, how is the anaphor in (78b) licensed and interpreted? What (78b) and (79) show after all is a kind of a binder/bindee

asymmetry: long scrambling of an anaphor can create a new binding possibility but long scrambling of the antecedent cannot.

What seems to be playing a role here is an anaphoric feature (A-feature or simply [+anaphor]) that needs to be bound by a D-feature. If an anaphor has this feature and this feature can move along with the P-features of the anaphor, then (78b) can successfully be explained as in (81).

(81)   [$_{TP}$ Karera-ga . . . [$_{CP}$ [$_{TP}$ otagai-o . . . [$_{CP}$ otagai-o [$_{TP}$ . . . otagai-o . . .]]]]]
           {P,D}   {P,A}   Ø   {D}

Otagai-o 'each other-ACC' originates in the object position of the most deeply embedded clause with the features {P, D, A}. The initial movement copies it in the lower C-projection. After deletion, {P, A} are in the C-projection and {D} is in the object position. The second movement copies *otagai-o* with {P, A} in the initial position of the middle clause. Then, {P, A} are deleted in the lower C-projection. After *karera-ga* 'they-NOM' is introduced into the matrix subject position, its D-feature can serve as the local antecedent of the A-feature of *otagai*. Thus, *otagai* is successfully licensed and interpreted in (78b).

The proposal that an A-feature, unlike a D-feature, can move along with the P-features may sound like a mere stipulation. But this is in a sense what we expect given the discussion so far. It has been assumed that P-features are always retained at the head of the chain. As noted above, this may be considered part of the definition of overt movement. For D-features and O-features, it was hypothesized that they can be retained only in positions where they are selected (in a broad sense). Then, what constraint could apply to A-features? Note that they do not participate in any selectional relation. For example, it is not part of the general selectional properties of verbs to have or to not have an anaphor in the object position. It seems then that where an A-feature is retained in a chain cannot be constrained by selection. Consequently, it is plausible, to say the least, to hypothesize that an A-feature can be retained at any position of a chain.[27]

The analysis of (78b) outlined above extends to well-known examples like (82).

(82)   John wonders [$_{CP}$[which picture of himself]$_i$[$_{TP}$ Mary thinks [$_{CP}$[$_{TP}$ Susan liked t_i]]]]

If the analysis in Section 3 is assumed, only the O-feature and the P-features of the *wh*-phrase move from the most deeply embedded CP Spec to the final landing site, as illustrated in (83).

(83)   John wonders [$_{CP}$[which picture of himself] [$_{TP}$ Mary thinks
                               {O,P}
         [$_{CP}$[which picture of himself] [$_{TP}$ Susan liked [which picture of himself]]]]]
             {O,P}   {D}

---

[27] This remark may also apply to the *wh*-feature of the Japanese in situ *wh*-phrases if it requires a [+Q] C as an "antecedent."

Eventually, the intermediate copy disappears as all of its features are deleted. Here, if *himself* has the A-feature and the feature moves along with the P-features of the noun, (83) is more precisely as in (84).

(84)   John wonders [$_{CP}$[which picture of himself] [$_{TP}$ Mary thinks
                    {O,P}   {A}
       [$_{CP}$[which picture of himself] [$_{TP}$ Susan liked [which picture of himself]]]
           {O,P}   {A}   {D}

In this structure, the A-feature of *himself* is ''bound'' by the D-feature of *John* in its local domain. The grammaticality of (82) is, then, correctly predicted exactly as in the cases of (77c) and (78b).

5.2. VP-internal scrambling and Chomsky's (1998) phase

Up till now, I have concentrated on scrambling across the subject and have postponed the discussion of VP-internal scrambling. In the remainder of this section, I will present a tentative analysis for this type of very local scrambling. I will then show that the suggested analysis makes it possible to embed the proposals made so far in Chomsky's (1998) theory of phase.

It has standardly been assumed that (85a) represents the ''basic word-order'' in Japanese, and that (85b) is derived by scrambling the direct object across the indirect object.[28]

(85)a. Hanako-ga   Taroo-ni sono hon -o   watasita (koto)
            -NOM   -to that book-ACC handed   fact
   b. Hanako-ga   sono hon -o$_i$ Taroo-ni t_i watasita (koto)
            -NOM that book-ACC   -to   handed   fact
     'Hanako handed that book to Taroo'

Tada (1990), applying Mahajan's (1990) work on Hindi to Japanese, shows that this kind of scrambling patterns strictly like A-movement with respect to binding. The examples in (86) indicate that a phrase preposed by VP-internal scrambling can serve as an antecedent for *otagai* 'each other', as in the case of clause-internal scrambling across the subject.

(86) a.   ?*Hanako-ga   [otagai -no   ryoosin]-ni [Taroo-to Ziroo]-o
                 -NOM each other-GEN parents-to   -and   -ACC
         syookaisita (koto)
         introduced   fact
         'Hanako introduced Taroo and Ziroo to each other's parents'

---

[28] Actually, it has been somewhat controversial whether VP-internal scrambling indeed exists as an operation. Miyagawa (1997), for example, argues that (85a and b) are both base-generated, while Yatsushiro (1999) defends the movement analysis. I will assume the operation here in part because the issues it raises arise on independent grounds within the theory of phase, as will be discussed directly.

   b.   Hanako-ga   [Taroo-to Ziroo]-o$_i$ [otagai   -no   ryoosin]-ni t_i
             -NOM   -and   -ACC each other-GEN parents-to
        syookaisita (koto)
        introduced   fact

The ungrammaticality of (87b), on the other hand, suggests that this kind of scrambling must always have the A-properties, unlike scrambling across the subject. The examples in (87) are from Nemoto (1993).

(87) a. Maikeru-ga   [Keeto-to Zyoo]-ni otagai   -o   syookaisita (koto)
             -NOM   -and   -to each other-ACC introduced   fact
        'Michael introduced each other to Kate and Joe'
     b. *Maikeru-ga   otagai   -o$_i$ [Keeto-to Zyoo]-ni t_i syookaisita (koto)
             -NOM each other-ACC   -and   -to   introduced fact

If VP-internal scrambling always has the A-properties, then (87b) can be ruled out by Condition (C), exactly as (62b), repeated below as (88).

(88)   *Each other$_i$ seem to John and Mary$_i$ [t_i to be smart]

(87b) contrasts sharply with (46b), repeated below as (89), which shows that clause-internal scrambling across the subject can pattern with A′ -movement.

(89)   [$_{TP}$ Otagai   -o$_i$ [Taroo-to Hanako-ga t_i semeta]] (koto)
             each other-ACC   -and   -NOM   blamed   fact
        'Each other$_i$, Taro and Hanako blamed t_i'

The data discussed above indicate that VP-internal scrambling should be analyzed exactly like NP-movement. But how is this possible? I argued in Section 4 that in the configuration of NP-movement shown in (90), all features are deleted from the lower copy and hence, the lower copy itself disappears.

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

The P-features are retained at the landing site. The D-feature by hypothesis can be retained only at a position where it is selected (in the broad sense). It is selected in both positions in (90), but the principle in (64), repeated in (91), forces it also to be retained at the landing site.

(91)   Chain interpretation makes the chain minimum.

Consequently, in (88) the D-feature of *each other* c-commands that of *John and Mary* at LF, and the example is ruled out by Condition (C). The situation was quite different in the case of scrambling across the subject. The D-feature of the scrambled

NP is not selected at the landing site, and hence, cannot be retained there. As a result, it is deleted at the landing site while the P-features are retained there. Then, the scrambling chain remains, but with the D-feature of the scrambled NP "reconstructed." Since the D-feature of *otagai* in (89) is no longer at the landing site at LF, the example is not a Condition (C) violation.

Deletion clearly cannot apply to VP-internal scrambling chains in the same way it does to chains created by scrambling across the subject. If it does, (87b) should not be a Condition (C) violation. Given the mechanism proposed so far, the only way to make the example violate Condition (C) is by somehow forcing the D-feature of the VP-internally scrambled NP to remain at the landing site. This is, in turn, possible if the landing site of VP-internal scrambling counts as a selected position. Then, (91) will force the D-feature to be retained at the landing site exactly as in the case of NP-movement.

The target of VP-internal scrambling is either VP or *v*P. Let us first consider VP. In this case, the scrambled phrase remains within the projection of its theta-role assigner. Thus, the following will yield the desired result:

(92) If an NP is combined with a projection of its theta-role assigner by Merge, the position of the NP counts as a selected position.

Given (92), scrambling to VP should fail to leave a trace exactly as in the case of NP-movement. This account can be extended to scrambling to *v*P if V raises to *v* in Japanese. We can say that when V raises to *v,* the phrases headed by *v* become projections of the V–*v* complex. Then, the scrambling of, say, an object NP to *v*P counts as movement to a selected position, again, exactly as in the case of NP-movement.[29]

The account presented above for VP-internal scrambling is still stipulative in flavor. But it has a desirable consequence when we try to reevaluate the proposals in the preceding sections against Chomsky's (1998) theory of phase. Chomsky proposes that a derivation proceeds phase by phase, where CP and *v*P constitute phases. One central condition in this theory is the Phase Impenetrability Condition (PIC), given in (93).

(93) In phase α with head H, the domain of H is not accessible to operations outside α, but only H and its edge.

---

[29] The similarity between this account of VP-internal scrambling and Tada's (1990) analysis of the A-properties of clause-internal scrambling should be obvious. If scrambling to *v*P is indeed possible, it should be guaranteed that the scrambling applies before the external argument is merged. Otherwise, since Condition (A) is an anywhere condition, the object preposed by *v*P-internal scrambling should be able to serve as the antecedent of an anaphor contained within the subject in examples like (i).

(i) ?*Otaigai -no   sensei -ga   karera-o$_i$ minna-ni t_i syookaisita (koto)
     each other-GEN teacher-NOM they -ACC all -to   introduced fact
     'Each others' teachers, them$_i$, introduced t_i to everyone'

I tentatively assume that this 'ordering' is necessary so that the subject can be properly attracted to TP Spec. Note that a similar problem arises with object shift if it is movement to the *v*-projection as proposed in Chomsky (1999). See Ura (2000) and the references cited there for much relevant discussion.

This condition states that only the head and the Spec of a phase is accessible to an operation taking place in a higher phase. It forces successive-cyclic *wh*-movement, as illustrated with the CP phase in (94).

(94)   [$_{CP}$ wh [$_{TP}$ ... [$_{CP}$ wh [$_{TP}$ ... wh ...]]]]

The *wh* cannot move directly from within the embedded TP to the matrix CP Spec, because the initial position is not accessible to operations in the matrix CP. Then, the *wh* must first move to the embedded CP Spec, the edge of the embedded CP phase. The position is accessible to operations in the matrix clause, and the *wh* can move on to the matrix CP Spec. To make the initial step of this movement possible, Chomsky suggests that a head of a phase may be assigned a feature that attracts a *wh*-phrase. In the discussion above, I have called this feature the EPP-O feature.

   The analysis of *wh*-movement proposed in Section 3 above is consistent with this theory of phase. It was proposed there that deletion applies to chains as they are created. Then, for (94), deletion yields the chain in (95) at the point the *wh* moves into the embedded CP Spec.

(95)   [$_{CP}$ [$_{TP}$ ... [$_{CP}$ wh   [$_{TP}$ ... wh ...]]]]
                    {P,O}   {D}

Only the P-features and the O-feature of the *wh* move on to the matrix CP Spec, and after deletion, the chain in (96) obtains.

(96) [$_{CP}$ wh   [$_{TP}$ ... [$_{CP}$ wh [$_{TP}$ ... wh ...]]]]
        {P,O}   Ø   {D}

As noted above, Chomsky (1998) proposes that not only CPs but also *v*Ps constitute phases. The evidence for the phasehood of *v*Ps seems less direct, but its adoptation does not affect the analysis in any case. The initial step of the movement will, then, be as in (97) instead of (95).

(97)   [$_{CP}$ [$_{TP}$ ... [$_{vP}$ ... [$_{CP}$ [$_{TP}$ ... [$_{vP}$ wh [$_{VP}$ ... wh ...]]]]]]]]
                                    {P,O}   {D}

The P-features and the O-feature of the *wh* move on successive-cyclically to the matrix CP Spec.

   Chomsky (1999) suggests further that interpretation also applies phase by phase. The suggestion, roughly, is that information is sent to interpretation at each phase for that phase excluding its head and its edge. It is possible to abandon the idea that chains are interpreted as they are created and adopt "interpretation by phase" instead, without much effect on the discussion presented above. As far as I can see, they yield the same results for the cases of *wh*-movement. Or put more precisely, "interpretation by phase" implies derivational interpretation of chains. Proper interpretation for the VP in (97) can be

obtained at the most deeply embedded *v*P only if the O-feature is already deleted from the initial position of the *wh* at this point.

However, an interesting problem arises when the theory of phase is applied to scrambling. Let us consider again (34a), repeated in (98a), and the initial form of its scrambling chain in (98b).

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

The analysis proposed above was that Condition (A) is satisfied at the point the structure in (98b) is created. With this structure, the D-feature of the scrambled phrase *karera-o* 'they-ACC' locally c-commands the anaphor *otagai* 'each other'. Later, the chain is interpreted by the deletion of the D-feature at the landing site and of the P-features at the initial site. The resulting chain is shown in (99).

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

But this does not affect the grammaticality of (98a), given that Condition (A) is an anywhere condition.

This analysis, however, cannot be maintained straightforwardly if *v*P is a phase as suggested in Chomsky (1998). The scrambled phrase in (98a) will then first move to the *v*-projection as in (100a).

(100)a.   [$_{vP}$ NP [... NP ...]]
        {P,D}   {P,D}
    b.   [$_{vP}$ NP [... NP ...]]
         {P}   {D}

If chain interpretation applies at this point and yield (100b), then what moves on to the T-projection across the subject is just the P-features of the scrambled phrase. Hence, it will be predicted incorrectly that (98a) is ungrammatical. But recall here the suggestion above that scrambling to *v*P is exactly like NP-movement as its landing site counts as a selected position. Given this, the scrambling chain in (100a) is interpreted as (101) instead of (100b).

(101)[$_{vP}$ NP   [... NP ...]]
      {P,D}   Ø

Thus, the analysis of (98a) proposed in Section 3 can be maintained as such even with the *v*P phase.

6. Conclusion

In this paper, I assumed following Tada (1990) that Japanese scrambling is a uniform operation, whether it is clause-internal or long-distance, and provided an analysis for its radical reconstruction property, the A/A′ distinction it exhibits, and the proper binding effect it induces. I first suggested that radical reconstruction results from the interpretation of scrambling chains by the deletion of D-feature from the landing site. Then I proposed that this chain interpretation takes place derivationally, modifying the idea of Kitahara (2000). This made it possible to solve the A/A′ problem while maintaining that scrambling is a uniform operation. I also presented an analysis for the proper binding effect based on the proposed chain interpretation mechanism. Finally, I briefly discussed VP-internal scrambling, and suggested a way to embed the proposals in this paper under Chomsky's (1998) theory of phase.

Japanese (and Korean) scrambling is potentially theoretically significant because of its unique properties. The radical reconstruction property is quite distinctive. If scrambling is in fact a uniform operation, it is unique also in that it exhibits both A and A′ properties. I argued that these unique properties provide evidence for a specific mechanism of chain interpretation and for specific proposals on the binding conditions; Condition (A) is an anywhere condition whereas Condition (C) is an LF condition. The proper binding effect distinguishes Japanese (and Korean) scrambling from other types of non-operator movements, for example, NP-movement and German scrambling. I suggested a condition on the application of Merge based on this unique property of Japanese (and Korean) scrambling.

Acknowledgements

References

Ausin, A., 1998. A minimalist approach to the remnant movement constraint. MS, University of Connecticut.

Belletti, A., Rizzi, L., 1988. Psych-verbs and Θ-theory. Natural Language and Linguistic Theory 6, 291–352.

Bošković, Ž., Takahashi, D., 1998. Scrambling and last resort. Linguistic Inquiry 29, 347–366.

Chomsky, N., 1981. Lectures on Government and Binding. Foris Publications, Dordrecht.

Chomsky, N., 1986a. Knowledge of Language: Its Nature, Origin, and Use. Praeger, New York.

Chomsky, N., 1986b. Barriers. MIT Press, Cambridge, MA.

Chomsky, N., 1993. A minimalist program for linguistic theory. In: Hale, K., Keyser, S.J. (Eds.), The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger. MIT Press, Cambridge, MA.

Chomsky, N., 1994. Bare Phrase Structure. MIT Occasional Papers in Linguistics No. 5.

Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, MA.

Chomsky, N., 1998. Minimalist Inquiries: The Framework. MIT Occasional Papers in Linguistics No. 15.

Chomsky, N., 1999. Derivation by Phase. MIT Occasional Papers in Linguistics No. 18.

Chomsky, N., Lasnik, H., 1977. Filters and control. Linguistic Inquiry 8, 425–504.

Dejima, M., 1999. A Study on Remnant-Scrambled Sentences in Japanese and their LF Representations. MA thesis, Nanzan University.

Diesing, M., 1992. Indefinites. MIT Press, Cambridge, MA.

Epstein, S.D., Groat, E., Kawashima, R., Kitahara, H., 1998. A Derivational Approach to Syntactic Relations. Oxford University Press, Oxford.

Fukui, N., 1986. A Theory of Category Projection and its Applications. PhD dissertation, MIT.

Grewendorf, G., Sabel, J., 1999. Scrambling in German and Japanese: adjunction versus multiple specifiers. Natural Language and Linguistic Theory 17, 1–65.

Harada, K., 1972. Constraints on WH-Q binding. Studies in Descriptive and Applied Linguistics 5, 180–206.

Hoji, H., 1985. Logical Form Constraints and Configurational Structures in Japanese. PhD dissertation, University of Washington.

Kawamura, T., 2001. A Feature-Checking Analysis of Japanese Scrambling. MA thesis, Nanzan University.

Kitahara, H., 1997. Elementary Operations and Optimal Derivations. MIT Press, Cambridge, MA.

Kitahara, H., 2000. Case and scrambling: a derivational view. MS, Keio University.

Kuno, M., 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: Poser, W. (Ed.), Papers from the Second International Workshop on Japanese Syntax. CSLI Publications, Stanford.

Lasnik, H., 1985. Illicit NP movement: locality conditions on chains? Linguistic Inquiry 16, 481–490.

Lasnik, H., 1999. Chains of arguments. In: Epstein, S.D., Hornstein, N. (Eds.), Working Minimalism. MIT Press, Cambridge, MA.

Lasnik, H., Saito, M., 1984. On the nature of proper government. Linguistic Inquiry 15, 235–289.

Lasnik, H., Saito, M., 1992. Move α. MIT Press, Cambridge, MA.

Lebeaux, D., 1988. Language Acquisition and the Form of the Grammar. PhD dissertation, University of Massachusetts, Amherst.

Lebeaux, D., 1998. Where does the binding theory apply? MS, NEC Research Center, Princeton.

Lee, R. K., 1994. Economy of representation. PhD dissertation, University of Connecticut.

Mahajan, A., 1990. The A/A-Bar Distinction and Movement Theory. PhD dissertation, MIT.

Miyagawa, S., 1997. Against optional scrambling. Linguistic Inquiry 28, 1–25.

Miyagawa, S., 2000. The EPP, scrambling, and *wh*-in situ. MS, MIT.

Müller, G., 1996. A constraint on remnant movement. Natural Language and Linguistic Theory 14, 355–407.

Murasugi, K., 1991. Noun Phrases in Japanese and English. PhD dissertation, University of Connecticut.

Nakamura, M., 1996. Sokubaku Kankei (Binding Relations). Hitsuji Shobo, Tokyo.

Nemoto, N., 1993. Chains and Case Positions: A Study from Scrambling in Japanese. PhD dissertation, University of Connecticut.

Riemsdijk, H. van, Williams, E., 1981. NP-structure. The Linguistic Review 1, 171–217.

Saito, M., 1985. Some Asymmetries in Japanese and their Theoretical Implications. PhD dissertation, MIT.

Saito, M., 1989. Scrambling as semantically vacuous A′-movement. In: Baltin, M., Kroch, A. (Eds.), Alternative Conceptions of Phrase Structure. University of Chicago Press, Chicago.

Saito, M., 1992. Long distance scrambling in Japanese. Journal of East Asian Linguistics 1, 69–118.

Saito, M., 1994. Improper adjunction. MIT Working Papers in Linguistics 24, 263–293.

Saito, M., 1995. Radical reconstruction. MS, University of Connecticut.

Saito, M., 1996. Barriers for A-adjunction. Japanese/Korean Linguistics 5, 1–36.

Saito, M., 2000. Japanese scrambling in a comparative perspective. MS, Nanzan University.

Saito, M., Fukui, N., 1998. Order in phrase structure and movement. Linguistic Inquiry 29, 439–474.

Saito, M., Hoshi, H., 2000. The Japanese light verb construction and the minimalist program. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.), Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik. MIT Press, Cambridge, MA.

Tada, H., 1990. Scrambling(s). MS, MIT.

Tada, H., 1993. A/A-bar partition in derivation. PhD dissertation, MIT.

Takano, Y., 1998. Object shift and scrambling. Natural Language and Linguistic Theory 16, 817–889.

Ura, H., 2000. Checking Theory and Grammatical Functions in Universal Grammar. Oxford University Press, Oxford.

Yatsushiro, K., 1999. Case Licensing and VP Structure. PhD dissertation, University of Connecticut.

Yoshimura, N., 1989. Parasitic pronouns. MS, University of Southern California.