Reality Exploration and Discovery: Pattern Interaction in Language & Life

Linda Uyechi and Lian Hee Wee (Eds.)

CENTER FOR THE STUDY
OF LANGUAGE
AND INFORMATION

1

Evidence for the Linearization-Based Theory of Long-Distance Scrambling in Japanese

Shûichi Yatabe

1.1 Introduction

On a purely descriptive level, it can be said that scrambling is clause-bound in some languages but not in others. Scrambling in German is more or less clause-bound; apparent exceptions arguably involve clause union and are to be dealt with in terms of function composition (in Combinatory Categorial Grammar parlance) or argument composition (in Head-Driven Phrase Structure Grammar parlance). Likewise, it has been noted that scrambling in Malayalam is clause-bound (Mohanan 1982). On the other hand, scrambling in languages like Japanese and Korean does not seem to be clause-bound at least at first blush. The Japanese sentence in (1) is one example involving long-distance scrambling, i.e. scrambling that seems not to be clause-bound.

(1)	?[Sono	dôbutsuen	ni],	Tarô	ga	panda	ga	iru	to	itta
	[that	zoo	DAT]	Tarô	NOM	panda	NOM	are	COMP	said

rashii no desu. <12, 17, 10, 1>
it seems

'It seems that Tarô said that there are pandas at that zoo.'

Reality Exploration and Discovery: Pattern Interaction in Language & Life.
Linda Uyechi and Lian Hee Wee (Eds.).

(The four numbers after (1) here as well as those after (10) below represent the result of questionnaire studies. See section 1.3 for details.)

There have been two fundamentally different theories of long-distance scrambling in languages like Japanese and Korean: the movement-based theory, according to which long-distance scrambling is an operation that induces filler-gap dependency (see Harada 1977 and Saito 1985), and the linearization-based theory, according to which long-distance scrambling alters the linear order of expressions without altering the syntactic structure involved (see Yatabe 1996 and Yatabe 2003). The movement-based theory views long-distance scrambling as something analogous to *wh*-movement in a language like English; the linearization-based theory, on the other hand, views long-distance scrambling as something analogous to extraposition out of NP in a language like English.

The purpose of this paper is to present some evidence for the linearization-based theory of long-distance scrambling in Japanese. First, in Section 1.2, the outline of the linearization-based theory will be described and its predictions will be delineated. Then, in Section 1.3, the result of several questionnaire studies regarding those predictions will be presented and its implications will be discussed.[1]

1.2 Extraposition and long-distance scrambling

The linearization-based theory of long-distance scrambling is based on the view that long-distance scrambling in a language like Japanese is analogous to extraposition out of NP in languages like English and German, and it builds on Kathol and Pollard's analysis of extraposition in German (Kathol and Pollard 1995). Thus, in this section, Kathol and Pollard's theory of extraposition will be described in subsection 1.2.1, before the linearization-based theory of long-distance scrambling in Japanese is presented in subsection 1.2.2.

1.2.1 The Kathol-Pollard theory of extraposition

Extraposition out of NP in English has the following three properties, which set it apart from canonical instances of filler-gap dependency. First, a string α can be extraposed out of a phrase β only if α constitutes the right periphery of β, as shown by (2). (In this subsection, the words

[1] I would like to thank Mohanan for various memorable remarks that have provided me with food for thought over the years. On one occasion, he told me that you need to have a lot of "ammunition" when you publish a paper which argues for an unconventional view such as the view that long-distance scrambling in Japanese does not give rise to filler-gap dependency. In writing this paper, I hope I am being a good rebel, packing a reasonable amount of ammunition.

that have been extraposed are shown in italics.)

(2) *It appears I have given the assignment to a fool after all(,) *complete and utter*. (from Stucky 1987)
(Cf. There may be somebody here *more technically oriented than I*.)

Second, when two or more constituents are extraposed out of a phrase, the linear order between those constituents must be preserved, as shown by (3).

(3) (from Dowty 1996)

a. Can you give me the names of any newcomers as soon as possible *from Finland who may have programming experience*? (Cf. newcomers from Finland who may have programming experience)

b. *Can you give me the names of any newcomers as soon as possible *who may have programming experience from Finland*? (Cf. *newcomers who may have programming experience from Finland)

And third, extraposition violates some of the island conditions that constrain canonical instances of filler-gap dependency such as topicalization. For instance, extraposition out of a subject NP always violates the condition that bans extraction out of a grammatical subject. Likewise, a sentence like (4) indicates that extraposition is also allowed to violate the Coordinate Structure Constraint.

(4) ?She was comparing Kyoto University today *and Tokyo University*. (from Yatabe 2003)

This example is a result of dislocating the conjunct *and Tokyo University* out of the coordinate structure *Kyoto University and Tokyo University*, and it is nevertheless only awkward, unlike an example like (5), which is completely unacceptable.

(5) *[And John]$_i$, I saw [Bill t_i]

The theory of discontinuous constituency proposed in Kathol and Pollard 1995 offers one way to explain why extraposition has these properties.[2] In Kathol and Pollard's theory, the portion of a syntactic structure that determines grammatical dependency relations is represented by means of an unordered tree, that is, a tree with no specifications as to the ordering of its constituents. The information concerning the linear order between the constituents is contained in what are called

[2]Due to space limitation, other theories of discontinuous constituency, such as those that make use of LFG mechanisms, will not be considered in this article.

$$\begin{bmatrix} \text{S} \\ \text{DOM} \left\langle \begin{bmatrix} \textit{the man} \\ \text{NP[nom]} \end{bmatrix}, \begin{bmatrix} \textit{bought} \\ \text{V} \end{bmatrix}, \begin{bmatrix} \textit{it} \\ \text{NP[acc]} \end{bmatrix} \right\rangle \end{bmatrix}$$

$$\begin{bmatrix} \text{VP} \\ \text{DOM} \left\langle \begin{bmatrix} \textit{bought} \\ \text{V} \end{bmatrix}, \begin{bmatrix} \textit{it} \\ \text{NP[acc]} \end{bmatrix} \right\rangle \end{bmatrix} \quad \begin{bmatrix} \text{NP[nom]} \\ \text{DOM} \left\langle \begin{bmatrix} \textit{the} \\ \text{Det} \end{bmatrix}, \begin{bmatrix} \textit{man} \\ \text{N}' \end{bmatrix} \right\rangle \end{bmatrix}$$

FIGURE 1 Total compaction of an NP

order domains, each of which is associated with a node in the unordered tree. An order domain is a list of *domain objects*, and is given as the value of the DOM feature. A domain object is very much like a sign; unlike a sign, however, it does not carry any information as to its internal morphosyntactic structure.

Let us take a concrete example. Figure 1 shows part of the structure assigned to the English sentence *The man bought it.* What is shown in this figure is an unordered tree; there is no linear precedence relation between the VP node and the NP node. The order domain (i.e. the DOM value) of the VP node consists of two domain objects, one that is pronounced *bought*, and the other one that is pronounced *it*. The order between these two domain objects *is* significant; it indicates that this VP is to be pronounced *bought it*, rather than *it bought*. Likewise, the order domain of the NP node tells us that this NP is to be pronounced *the man*, and the order domain of the S node tells us that the S node is to be pronounced *The man bought it.*

Let us take a closer look and see how the order domain of the S node is related to the order domains of the NP node and the VP node in Figure 1. The two domain objects in the order domain of the VP node are both integrated, unaltered, into the order domain of the S node.[3] Notice that the domain object that is pronounced *bought* precedes the domain object that is pronounced *it* in the order domain of the S as well as in the order domain of the VP. This is a consequence of the constraint given in (6) (see Kathol 1995).

(6) The Persistence Constraint:
Any ordering relation that holds between domain objects α and β in one order domain must also hold between α and β in all other order domains that α and β are members of.

[3]The exposition here employs procedural terms in order to enhance readability, although a purely constraint-based theoretical framework is presupposed throughout.

Next, let us see how the order domain of the NP is related to the order domain of the S in Figure 1. The order domain of the NP node contains two domain objects, but this NP node contributes to the order domain of the S node only one domain object, which is pronounced *the man*. What is at work here is an operation called *total compaction*. (7) illustrates the way the total compaction operation takes a sign and turns it into a single domain object.

(7) $$\begin{bmatrix} \alpha_0 \\ \textsc{dom} \left\langle \begin{bmatrix} \beta_1 \\ \alpha_1 \end{bmatrix}, \ldots, \begin{bmatrix} \beta_n \\ \alpha_n \end{bmatrix} \right\rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \beta_1 \circ \cdots \circ \beta_n \\ \alpha_0 \end{bmatrix}$$

What is shown on the left of the arrow is the input to the operation; the input is a sign. The first line of a sign ("α_0" in this case) is its SYNSEM value; the second line ("DOM ...") shows what its order domain looks like. On the right of the arrow is shown the output of the operation; the output is a domain object. The first line of a domain object ("$\beta_1 \circ \cdots \circ \beta_n$" in this case) is its PHON value. (The small circle is an operator that concatenates strings.) The second line of a domain object ("α_0" in this case) is its SYNSEM value.

The domain object that is created by totally compacting a sign X is placed in the order domain of the mother of X. In Figure 1, the domain object that is created by totally compacting the subject NP has been placed in the order domain of the S.

The order between the domain object that comes from the subject NP and the domain objects that come from the VP is determined by a linear precedence statement that states that a V must follow its subject in English. Although domain objects coming from two or more daughter nodes can be stringed together in any order as long as they do not violate any constraints explicitly stated in the grammar, the order of the three domain objects is completely determined in this case, due to the Persistence Constraint and the linear precedence statement concerning subject NPs.

So far, we have seen two processes whereby the order domain of a given node is integrated into that of its mother. First, a node can be totally compacted. Second, a node may undergo no compaction whatsoever. Henceforth the latter situation is going to be described by saying that the node in question has been *liberated*. The VP in Figure 1 has been liberated.

There is a third process allowed by the theory: a given node can be *partially compacted*. Partial compaction takes a sign and turns it into one or more domain objects, as opposed to total compaction, which always produces a single domain object. (As will become clear shortly,

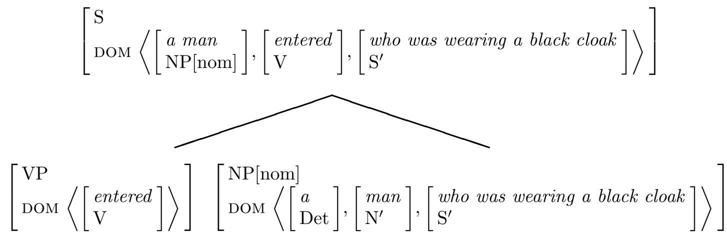

FIGURE 2 Partial compaction of an NP

total compaction can be seen as a special case of partial compaction.) (8) and (9) illustrate the way the partial compaction operation takes a sign and turns it into one or more domain objects, which are to be placed in the order domain of the mother of that sign.

(8) Partial compaction for head-first languages:

$$\begin{bmatrix} \alpha_0 \\ \textsc{dom} \left\langle \begin{bmatrix} \beta_1 \\ \alpha_1 \end{bmatrix}, \ldots, \begin{bmatrix} \beta_n \\ \alpha_n \end{bmatrix} \right\rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \beta_1 \circ \cdots \circ \beta_i \\ \alpha_0 \end{bmatrix}, \begin{bmatrix} \beta_{i+1} \\ \alpha_{i+1} \end{bmatrix}, \ldots, \begin{bmatrix} \beta_n \\ \alpha_n \end{bmatrix}$$

(9) Partial compaction for head-last languages:

$$\begin{bmatrix} \alpha_0 \\ \textsc{dom} \left\langle \begin{bmatrix} \beta_1 \\ \alpha_1 \end{bmatrix}, \ldots, \begin{bmatrix} \beta_n \\ \alpha_n \end{bmatrix} \right\rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \beta_1 \\ \alpha_1 \end{bmatrix}, \ldots, \begin{bmatrix} \beta_{i-1} \\ \alpha_{i-1} \end{bmatrix}, \begin{bmatrix} \beta_i \circ \cdots \circ \beta_n \\ \alpha_0 \end{bmatrix}$$

In (8), the DOM value of the sign that is fed to the operation as the input has n domain objects in it. Of those domain objects, the first (i.e. leftmost) i domain objects are bundled together and turned into a single domain object, while the remaining domain objects, if any, are left out of the bundle and continue to be separate domain objects. (9) is a mirror image of (8). Roughly speaking, partial compaction of a sign α is achieved by first obtaining a sign α' by deleting a certain number of domain objects in the DOM value of α (the rightmost $n - i$ domain objects in the case of (8) and the leftmost $i - 1$ domain objects in the case of (9)) and then totally compacting α'.

Various types of extraposition constructions result when an expression is partially compacted and surfaces as a discontinuous constituent. Figure 2 shows how the English extraposition construction can be generated via partial compaction. Here, the subject NP has been partially compacted. The relative clause has been left out of the bundle and appears in the sentence-final position. What puts the relative clause in this particular position is an English-particular linear precedence statement, which is not formulated in this paper.

In this theory, the fact that only things that constitute the right edge of a phrase can be extraposed out of it is a consequence of the Persistence Constraint.[4] The fact that the linear order between two or more extraposed constituents must be preserved is also a consequence of the Persistence Constraint. And the fact that extraposition violates some of the island conditions is unsurprising, as extraposition is assumed not to induce filler-gap dependency.

The putative fact noted in Guéron 1980 that extraposition sometimes has an effect on semantic scope relations remains unexplained in Kathol and Pollard's theory. In Yatabe 2001, however, a variant of their theory is proposed which provides an explanation for Guéron's observation.

1.2.2 The linearization-based theory of long-distance scrambling

The linearization-based theory of long-distance scrambling is based on the idea that long-distance scrambling in a language like Japanese is the mirror image of extraposition in a language like English. This idea is not as far-fetched as it might appear at first blush. First, examples like (10), whose resemblance to familiar English extraposition constructions is evident, lend plausibility to the view that the Japanese equivalent of extraposition shifts expressions to the left rather than to the right. Note that this example involves leftward dislocation of the phrase *Yoshino san kara no* out of an NP meaning 'the last letter from Yoshino san'.

(10) [Yoshino san kara no], tabun kore ga saigo no tegami ni
[Yoshino from GEN] probably this NOM last GEN letter DAT
naru n darô na. <13, 13, 2, 0>
become NML I suspect
'This will probably be the last letter from Yoshino san, I suspect.'

Second, the fact that long-distance scrambling in Japanese can dislocate expressions out of an S is not a problem for the proposed theory, as extraposition in English is likewise capable of dislocating an expression out of an S, as in the following sentence. Notice that this example involves extraposition of the innermost S out of the second innermost S.

(11) I had hoped that it was true for many years that Rosa Luxemburg had actually defected to Iceland. (from Gazdar 1981)

[4] The definition of partial compaction alone does not ensure this consequence. Without the Persistence Constraint, it might be possible to extrapose β_{i+1} in (8) without also extraposing β_n, for instance.

Finally, the fact that long-distance scrambling in Japanese is arguably unbounded and not subject to a constraint analogous to the so-called Right-Roof Constraint[5] does not pose a problem for the linearization-based theory either, because constraints on extraposition are known to vary from language to language. For instance, Müller 2004 shows that, unlike English, German allows relative clauses and complement clauses to be extraposed over an indefinite number of NP boundaries. It will therefore not be particularly surprising if the Japanese counterpart of extraposition turns out to be capable of dislocating an expression over an indefinite number of S boundaries.

The theory of long-distance scrambling proposed in Yatabe 1996 is based on the assumption that the compaction operation is applied in accordance with the constraints given in (12).[6]

(12) a. In a head-complement structure whose head is verbal, the head is liberated and the non-head is partially compacted.

b. In a head-adjunct structure whose head is verbal, the head and the adjunct are both partially compacted.

c. In a headed structure whose head is nominal and whose non-head is not a marker, the head is totally compacted[7] and the non-head is partially compacted.

d. In a head-marker structure, the head and the marker are both liberated.

e. In a coordinate structure, each of the conjuncts is totally compacted.

Let us also assume that complementizers and case particles are markers and thus subject to (12d) in Japanese. Given these assumptions, sentences involving long-distance scrambling are automatically generated, unless we posit in the grammar of Japanese a linear precedence statement that specifically prohibits long-distance scrambling. In what follows, it will be assumed, for the sake of simplicity, that Japanese syntax has only two linear precedence statements: one that requires heads to follow their dependents (i.e. complements and adjuncts), and another one that requires markers to follow what they mark.

[5] It is in fact quite difficult to demonstrate that long-distance scrambling in Japanese is not bounded, but see Yatabe 1996 for an attested example in which an expression seems to have been long-distance-scrambled over two S boundaries.

[6] Another assumption is that head-complement structures and head-adjunct structures in Japanese are strictly binary. See Yatabe 1996 for some more detail.

[7] Without this requirement, it would be possible in a language like Japanese to extrapose Y in a structure like $[_{\mathrm{N}'_1}\ X\ [_{\mathrm{N}'_2}\ Y\ Z]]$ out of N'_1 by 'first' extraposing it to the left of X, despite the fact that Y does not constitute the left edge of N'_1.

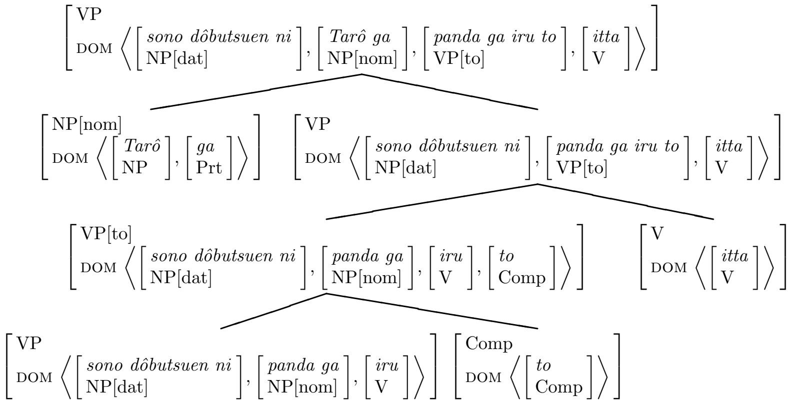

FIGURE 3 Long-distance scrambling via partial compaction

Figure 3 shows how part of sentence (1) is generated in this theory. At the bottom of the unordered tree shown in Figure 3, there is a head-marker structure with a VP functioning as a head and a complementizer functioning as a marker. (Note that what is called a VP here is in fact a complete clause; the term *VP* is used here to refer to all phrases in Japanese that are headed by a verb.) The VP and the complementizer are both liberated, and the order domain of the VP[to] ends up containing four domain objects. When this VP[to] then combines with the V *itta* to produce a VP, the V is liberated while the VP[to] is partially compacted; the first element in the order domain of the VP[to] escapes the compaction and continues to be a separate domain object. The topmost local subtree in the figure is a head-complement structure that combines this VP and a subject NP. The VP is liberated, and the subject NP is partially (in fact totally) compacted. The domain object that comes from the subject is allowed to be anywhere in the order domain of the topmost VP, as long as it does not follow the domain object corresponding to the V *itta*. Hence it is perfectly legitimate for that domain object to be immediately preceded by the domain object corresponding to the phrase *sono dôbutsuen ni*, which is syntactically a part of the embedded clause.

This theory of long-distance scrambling makes the following predictions. First, it predicts that a string α can be long-distance-scrambled out of a phrase β only if α constitutes the left periphery of β. The second prediction is that, when two or more constituents are long-distance-scrambled out of a phrase, the linear order between those constituents must be preserved. And the third prediction is that long-distance scrambling should be allowed to violate some of the island conditions that constrain canonical instances of filler-gap dependency such as topicalization in English. Specifically, long-distance scrambling is predicted to be capable of dislocating a conjunct, in violation of the Coordinate Structure Constraint (see (4)).

1.3 Questionnaire studies

In this section, the results of several questionnaire studies whose purpose was to test the three predictions of the linearization-based theory of long-distance scrambling will be presented in the first subsection, and the implications of those results will be discussed in the next subsection. The respondents in these studies, conducted in 2006 and 2007, consisted of students at the University of Tokyo who were not linguists, and they were compensated for their time. When the relative acceptability of two or more examples was of interest, the order between those examples was randomized for each respondent. In what follows, the four figures immediately following example sentences show the number of speakers who found the sentence 'completely natural under the intended reading', 'slightly unnatural under the intended reading', 'considerably unnatural under the intended reading', and 'completely impossible under the intended reading', respectively. The *average rating* for a linguistic material L, which will be represented as $r(L)$, is defined here as $(1a+2b+3c+4d)/(a+b+c+d)$, when the questionnaire result for L is $<a, b, c, d>$. A linguistic material L that is associated with a questionnaire result is shown here with no diacritic if $1 \leq r(L) < 2$, with '?' if $2 \leq r(L) < 2.5$, with '??' if $2.5 \leq r(L) < 3$, with '?*' if $3 \leq r(L) < 3.5$, and with '*' if $3.5 \leq r(L) \leq 4$.[8]

1.3.1 Questionnaire results

(13) and (14) form a minimal pair that tests the first prediction of the linearization-based theory, i.e. the prediction that an expression α can be scrambled out of another expression β only if α constitutes the left edge of β. (14) is a sentence obtained by exchanging the two expressions in (13) that both modify the noun *kutsu*. The relative clause *sono toki o-shiro ni ita*, which is long-distance-scrambled to the sentence-initial position in both sentences, originates at the left edge of the NP meaning 'the glass shoe of the person called Cinderella, who was at the castle at the time' in (13) (as shown in (15a)), but not in (14) (as shown in (15b)).

[8]The notion of average rating is only intended as an expedient; the way it is defined and used here is arbitrary to a certain extent.

(13) ??[Sono toki o-shiro ni ita], kore ga, [Shinderera to yû]
[that time castle DAT was] this NOM [Cinderella COMP call]
hito no garasu no kutsu na n desu. <2, 9, 5, 6>
person GEN glass GEN shoe COP I say politely
'This is the glass shoe of the person called Cinderella, who was at the castle at the time.'

(14) ?*[Sono toki o-shiro ni ita], kore ga, garasu no, [Shinderera
[that time castle DAT was] this NOM glass GEN [Cinderella
to yû] hito no kutsu na n desu. <0, 4, 6, 12>
COMP call] person GEN shoe COP I say politely

(15) a. [[[sono toki o-shiro ni ita] [Shinderera to yû] hito]
[[[that time castle DAT was] [Cinderella COMP call] person]
no] garasu no kutsu
GEN] glass GEN shoe
'the glass shoe of the person called Cinderella, who was at the castle at the time'

b. garasu no [[[sono toki o-shiro ni ita] [Shinderera to
glass GEN [[[that time castle DAT was] [Cinderella COP
yû] hito] no] kutsu
call] person] GEN] shoe

The Wilcoxon signed-rank test showed that the rating for (13) was significantly higher than the rating for (14) ($T = 7$, $n = 15$, $p < 0.001$), as predicted by the theory.

(16) and (17) are another minimal pair that tests the prediction that an expression α can be scrambled out of another expression β only if α constitutes the left edge of β. As shown in (18), the NP *ude ni* must be at the left edge of the phrase whose literal meaning is 'to give a twist to one's arms', in order for the phrase to have the idiomatic meaning 'to make extra efforts'. Therefore it can be said that the scrambled phrase originates at the left edge of the embedded clause in (16) but not in (17), as long as the embedded clause is to be interpreted in its idiomatic sense.

(16) ?*[Ude ni] aitsu ga yori o kaketai tte itteta yo.
[arm DAT] that guy NOM twist ACC want to give COMP was saying
<0, 7, 10, 13>
'That guy was saying that he wants to make extra efforts.'
('That guy was saying that he wants to give a twist to his arms.' (Lit.))

(17) *[Yori o] aitsu ga ude ni kaketai tte itteta yo.
[twist ACC] that guy NOM arm DAT want to give COMP was saying
<0, 0, 5, 25>

(18) a. Ude ni yori o kaketai. <20, 8, 1, 1>
arm DAT twist ACC want to give
'(I) want to make extra efforts.'
('(I) want to give a twist to my arms.' (Lit.))

b. *Yori o ude ni kaketai. <0, 0, 9, 21>
twist ACC arm DAT want to give

Again, as predicted by the proposed theory, (16) was found to be significantly better than (17) ($T = 0$, $n = 15$, $p < 0.0001$).

Next, (19) and (20) form a minimal pair that tests the second prediction of the linearization-based theory, i.e. the prediction that, when two or more constituents are long-distance-scrambled out of a phrase, the linear order between those constituents must be preserved. In both sentences, two phrases (*Ken no koto o* and *hannin to*) have been long-distance-scrambled out of the embedded clause. As shown in (21) and (22), the phrase *Ken no koto o* must precede the phrase *hannin to*, when the two phrases are co-arguments of the verb *shinjikonde*; (21) was indeed found to be significantly better than (22) ($T = 0$, $n = 14$, $p < 0.0001$). Therefore it can be said that (19) preserves the pre-scrambling order between the two scrambled phrases and that (20) does not.

(19) ??[Ken no koto o] [hannin to], Tanaka ga keisatsu ga
[Ken GEN matter ACC] [culprit COMP] Tanaka NOM police NOM
shinjikonderu tte itteta. <1, 8, 4, 3>
believe COMP said
'Tanaka said that the police firmly believed Ken to be the culprit.'

(20) *[Hannin to] [Ken no koto o], Tanaka ga keisatsu ga
[culprit COMP] [Ken GEN matter ACC] Tanaka NOM police NOM
shinjikonderu tte itteta. <0, 1, 4, 11>
believe COMP said

(21) [Ken no koto o] [hannin to] keisatsu ga shinjikonde
[Ken GEN matter ACC] [culprit COMP] police NOM believe-GER
iru. <11, 5, 0, 0>
be-PRES
'The police firmly believe Ken to be the culprit.'

(22)?*[Hannin to] [Ken no koto o] keisatsu ga shinjikonde
[culprit COMP] [Ken GEN matter ACC] police NOM believe-GER
iru. <0, 4, 3, 9>
be-PRES

As predicted by the theory, (19) was found to be significantly better than (20) ($T = 8$, $n = 14$, $p < 0.01$).

Finally, (23) and (24) test whether Japanese scrambling obeys the Coordinate Structure Constraint. (23) involves scrambling of a conjunct, whereas (24) involves scrambling of part of a conjunct out of that conjunct.[9]

(23) ?[Kyôdai to], kanojo ga [Tôdai to] o kurabeteru
[Kyoto U. and] she NOM [Tokyo U. and] ACC is comparing
tte, shitteta? <2, 14, 9, 3>
COMP did you know
'Did you know that she is comparing Kyoto University and Tokyo University?'

(24) *[Tanaka sensei no], kore wa, [saisho no chosho to]
[Prof. Tanaka GEN] these TOP [first GEN book and]
[Suzuki sensei no saigo no chosho] da . <0, 0, 6, 6>
[Prof. Suzuki GEN last GEN book] are
'These are Prof. Tanaka's first book and Prof. Suzuki's last book.'

As indicated above, the average rating for (23) was in the '?' range, and the average rating for (24) was in the '*' range.

1.3.2 Discussion

The results of the questionnaire studies invariably confirm the predictions of the linearization-based theory of long-distance scrambling. (23) is predicted to be grammatical, because the embedded clause in the sentence can be assigned the structure shown in Figure 4. On the other hand, (24) is predicted to be ungrammatical; no well-formed structure can be assigned to it, due to the requirement, stated in (12e), that a conjunct must be totally compacted.

On the other hand, the questionnaire results above do not make sense in the standard, movement-based theory of long-distance scrambling. In the movement-based theory, there is no grammar-internal reason why long-distance scrambling should be possible only from the left edge of

[9] (23) becomes structurally ambiguous when the second occurrence of *to* is omitted. In order to ensure that the respondents notice the presence of the second *to* in this example, a variant of the sentence without the second occurrence of *to* was also included in the questionnaire.

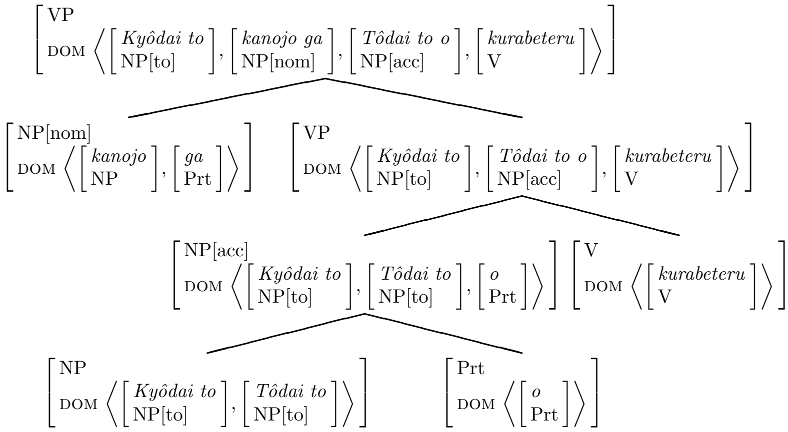

FIGURE 4 Scrambling of a conjunct

an expression, although there could be a processing-related reason why that must be so. Likewise, in the standard theory, there is no grammar-internal reason why the linear order between two scrambled expressions must be preserved, although, again, there could be a processing-related reason why that must be so. The fact that long-distance scrambling in Japanese does not obey the Coordinate Structure Constraint (CSC), as shown by (23), is particularly problematic for the movement-based theory; the theory seems to make it impossible to explain how learners could come to know that dislocation of a conjunct is allowed in Japanese but not in the case of leftward movement in English.

The awkwardness of (13), (16), and (19) is not predicted by the proposed theory, but it is not predicted by the movement-based theory either. It appears that both theories need to be augmented with "soft constraints" (see Bresnan and Nikitina's contribution to this volume) against dislocating an expression across more than one NP boundary, dislocating an idiom chunk, and dislocating more than one constituent out of a single clause, although it is not clear at this point whether such constraints would have to be viewed as part of the grammar.

As has been noted by Harada 1977, scrambling in Japanese does appear to obey the Complex NP Constraint, just like topicalization and relativization in English. This, however, does not mean that the movement-based theory of scrambling is on the right track; there are reasons to believe that the Complex NP Constraint is not a syntactic constraint and hence is irrelevant in the present context (see Kluender 1998). In fact, there are attested examples in which long-distance scrambling violates the Complex NP Constraint. One example is sentence (21) in Yatabe 1996, and another example is the following sentence,[10] in which the dative argument of the verb *fitto shita* 'fit' has been long-distance-scrambled out of a relative clause.

(25) Dakara, hontô ni, [[sô iu Nihon no wakamono no kenzen na,
so really [[such Japanese young people's wholesome

nan to iu ka, kibô to iu mono ni] wareware ga fitto shita
what shall I say thing called hope DAT] we NOM fit

seisaku o dashite iku to iu koto ga] hijô ni daiji da
policy ACC produce NML NOM] very important

to iu fû ni omotte iru tokoro desu.
that think

'So, really, I think it is very important that we should produce policies that fit Japan's young people's such wholesome hope.'

On the other hand, the constraint that bans movement of a conjunct seems to be a purely syntactic one. Some authors have argued that the part of the CSC that prohibits extraction out of a conjunct should be taken to be non-syntactic, but even those authors regard that part of the CSC that prohibits extraction of a whole conjunct as a syntactic constraint (Lakoff 1986, p. 161; Kehler 1996, p. 220).

1.4 Conclusion

In this article, the results of some questionnaire studies regarding long-distance scrambling in Japanese were presented, and it was argued that they provide some reasons to favor the linearization-based theory over the movement-based theory. The findings presented here indicate that there are at least two distinct kinds of dislocation in natural language syntax and that an analysis in terms of complex domain formation involving partial compaction is an adequate description of one of them.

References

Dowty, David R. 1996. Toward a minimalist theory of syntactic structure. In H. Bunt and A. van Horck, eds., *Discontinuous Constituency*, pages 11–62. Berlin: Mouton de Gruyter.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12:155–184.

Guéron, Jacqueline. 1980. On the syntax and semantics of PP extraposition. *Linguistic Inquiry* 11:637–678.

[10] http://www.mhlw.go.jp/kaiken/daijin/2007/02/k0206.html

Harada, S.-I. 1977. Nihongo ni "henkei" wa hitsuyô da. *Gengo* 6:88–95 (in no. 11) and 96–103 (in no. 12).

Kathol, Andreas. 1995. *Linearization-Based German Syntax*. Ph.D. thesis, Ohio State University, Columbus.

Kathol, Andreas and Carl Pollard. 1995. Extraposition via complex domain formation. In *33rd Annual Meeting of the Association for Computational Linguistics*, pages 174–180. San Francisco: Morgan Kaufmann.

Kehler, Andrew. 1996. Coherence and the coordinate structure constraint. In J. Johnson, M. L. Juge, and J. L. Moxley, eds., *BLS 22*, pages 220–231. Berkeley: Berkeley Linguistics Society.

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In P. Culicover and L. McNally, eds., *The Limits of Syntax (Syntax and Semantics 29)*, pages 241–279. San Diego: Academic Press.

Lakoff, George. 1986. Frame semantic control of the coordinate structure constraint. In A. M. Farley, P. T. Farley, and K.-E. McCullough, eds., *CLS 22 Part 2: Papers from the Parasession on Pragmatics and Grammatical Theory*, pages 152–167. Chicago: Chicago Linguistic Society.

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In J. Bresnan, ed., *The Mental Representation of Grammatical Relations*, pages 504–589. Cambridge, Massachusetts: MIT Press.

Müller, Stefan. 2004. Complex NPs, subjacency, and extraposition. *Snippets* 8:10–11.

Saito, Mamoru. 1985. *Some Asymmetries in Japanese and Their Theoretical Implications*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Stucky, Susan U. 1987. Configurational variation in English: A study of extraposition and related matters. In G. J. Huck and A. E. Ojeda, eds., *Discontinuous Constituency*, pages 377–404. Orlando: Academic Press.

Yatabe, Shûichi. 1996. Long-distance scrambling via partial compaction. In M. Koizumi, M. Oishi, and U. Sauerland, eds., *Formal Approaches to Japanese Linguistics 2 (MIT Working Papers in Linguistics 29)*, pages 303–317. Cambridge, Mass.: MITWPL.

Yatabe, Shûichi. 2001. The syntax and semantics of left-node raising in Japanese. In D. Flickinger and A. Kathol, eds., *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, pages 325–344. Stanford: CSLI. http://cslipublications.stanford.edu/HPSG/.

Yatabe, Shûichi. 2003. Does scrambling in Japanese obey the Coordinate Structure Constraint? In *Nihon Gengogakkai Dai-126-kai Taikai Yokôshû (Proceedings of the 126th Meeting of the Linguistic Society of Japan)*, pages 262–267. Kyoto: Linguistic Society of Japan.