Sources of Difficulty in Processing Scrambling in Japanese*

Edson T. Miyamoto  Shoichi Takahashi  
NAIST  MIT  

November 20, 2001

1 Introduction

Word order flexibility in Japanese has generated a vast literature in theoretical linguistics (Bošković and D. Takahashi 1998; Hale 1980; Hoji 1985; Miyagawa 1997; Saito 1985; Takano 1996; *inter alia*). However, there have been few studies investigating the processing of alternative word orders in this language, and their results have often yielded conflicting results. In particular, behavioural studies do not agree as to whether non-canonical word orders are more difficult to process than their canonical counterparts (Mazuka, Itoh, and Kondo 2001; Yamashita 1997).

We report a self-paced reading experiment (Experiment I) that taken together with a previous result (Mazuka, Itoh, and Kondo 2001) supports the view that the processing of non-canonical word orders requires longer reading times. In the second part of the paper, we discuss the possible sources of this difficulty.

We propose that the parsing of non-canonical word orders involves positing a gap required by the displaced element in the sentence.\(^1\) Creating such a gap leads to longer reading times especially when readers have limited memory resources available or when the gap is far away from its antecedent (following Gibson 1998; Just and Carpenter 1992).

However, in all experiments reported thus far in the literature, including Experiment I below, a number of alternative factors may have been responsible for the slow-down observed, rather than the proposed filler-gap association process. For example, the slow-down may have been detected because NPs with different case markers were compared; or perhaps because non-canonical word orders are less frequently encountered than their canonical counterparts; or because

---

\(^1\) A revised version of this paper is to appear in M. Nakayama (Ed.) *Sentence Processing in East Asian Languages*.

\(^1\) It will not be relevant for the present discussion whether the gap is encoded as an empty category (a trace left behind by movement; Chomsky 1981) or a slash feature (Gazdar, Klein, Pullum & Sag, 1985).
there is a preference for accusative NPs to be adjacent to the subcategorizing verb. A second type of experiment designed to counter those alternative explanations is described. We report some preliminary results of a self-paced reading experiment that support the gap hypothesis. Moreover, a corpus count is also reported suggesting that the behavioural data collected are not easily explained by frequency factors.

2 Scrambling in Japanese

The canonical word order in Japanese is SOV (subject-object-verb). However, constituents can be scrambled (i.e. moved leftward in the sentence) as long as the verb is the last element in the clause. For example, in (1a) (from Mazuka, Itoh, and Kondo 2001), brother-acc was scrambled to the beginning of the sentence and gap indicates its original position according to generative linguistics proposals. The canonical word order is shown in (1b).²

(1) a. Scrambled accusative NP
   ototo-o Mariko-ga gap, mita.
   brother-acc Mariko-nom saw
   ‘Mariko saw her brother.’

   b. Canonical word order
   Mariko-ga otooto-o mita.
   Mariko-nom brother-acc saw
   ‘Mariko saw her brother.’

As far as the processing of scrambled word orders is concerned, two questions are particularly relevant. The first question is whether scrambled word orders are more difficult to process than their canonical versions. The second question is whether a more complex structure is computed for scrambled sentences. The two questions are often assumed to be related in that difficulty is taken to be a sign of more complex structure building. In the first part of this paper, we will concentrate on the first question and will not concern ourselves with the possible factors that could lead to such a difficulty.

An earlier study did not observe any significant slow-downs in the processing of scrambled word orders (Yamashita 1997). More recently, however, Mazuka, Itoh, and Kondo (1998; 2001) detected longer reading times at Mariko-nom in the scrambled sentence (1a) above when it was compared to brother-acc in the canonical (1b), leading them to conclude that scrambled sentences must involve some extra processing cost.

One objection that may be raised is that NPs with nominative case may be used as clause boundary markers in Japanese (Inoue 1991; Miyamoto 2000). Consequently, the slow-down observed by Mazuka and colleagues is not necessarily related to extra costs incurred by the scrambling of the accusative NP.

²See, for example, Hoji 1985, for syntactic tests to determine which order is canonical.
The use of nominative NPs as clause boundary markers may be necessary because Japanese does not provide any other indicator to signal the beginning of an embedded clause; in particular, a complementizer when present comes at the end of the clause. Nominative NPs are good candidate markers because they cannot be scrambled (Saïto 1985); hence, their position is fixed in the initial segment of the clause. The result by Mazuka and colleagues, although suggestive, cannot be unambiguously interpreted without further investigation of the role of nominative NPs in parsing.

3 Experiment I: Scrambling and Memory Load

In this section, we provide further evidence that scrambling requires some extra effort to be processed.

It is possible that previous studies (e.g. Yamashita 1997) did not detect any significant differences in the reading times of scrambled sentences in relation to their canonical versions because the items used were relatively simple and may have been read rapidly in all conditions. Mazuka and colleagues provide some evidence in this direction as the slow-down in their experiment was more pronounced when center-embedded sentences were used (Mazuka, Itoh, and Kondo 2001).

We conducted an experiment in which the memory load of the task was increased by using complex sentences which required close attention from the participants. The basic construction used involves ditransitive verbs such as give (‘ageru’) and introduce (‘shookai-suru’). The canonical order is shown in (2a) below. In (2b), the accusative argument was scrambled to the position prior to the dative argument.

(2) a. Canonical: Subject NP-dat NP-acc ditransitive-verb

b. Scrambled: Subject NP1-acc NP-dat gap1 ditransitive-verb

This particular construction was chosen because its scrambled version is commonly used by native speakers; therefore, familiarity issues are less likely to be a concern in this case (see Experiment III below for further discussion on frequency factors). Moreover, it has been suggested that both word orders in (2) are canonical (Miyagawa 1997), despite the fact that (2b) has been traditionally assumed to involve scrambling (Hoji 1985; Takano 1996; inter alia). Therefore, processing results may be helpful in resolving the controversy at the syntactic level. Another advantage of this construction is that scrambling occurs inside the VP; thus, differences in reading times cannot be caused by the detection of clause boundaries.

Both clauses in (2) were further embedded as indicated below.

(3) a. Canonical: [Subject [NP:ec]\textsubscript{r} NP-dat \textsubscript{f} NP-acc \textsubscript{r} ditransitive-verb]\textsubscript{f} \textsubscript{r} NP\textsubscript{f}\textsubscript{r} verb\textsubscript{comp} NP-nom report-verb.
b. Scrambled:  

\[ \text{Subject } [e^j_{> RC} \ NP_{i\text{-acc}} \ NP_{i\text{-dat}} \ \text{gap}_i] \]

\[ \text{ditransitive-verb} \ NP_{j\text{-acc}} \ \text{verb-\text{comp}} \ NP_{\text{nom}} \ \text{report-verb}. \]

The ditransitive clause is a relative clause modifying an accusative NP (relative clauses precede the modified head in Japanese), which in turn is part of a sentential complement. The empty category \( e^j \) indicates that the subject position inside the ditransitive clause was relativized. The main clause is headed by a report-verb (see the Materials section below for an example sentence used). Fillers of comparable complexity were included. Complex items were used in order to force participants to read carefully, and intently store words in working memory.

We adopt an incremental processing model according to which case-marked NPs are associated within a partial representation of the sentence before the subcategorizing verb is read (see Miyamoto 2000, for discussion and experimental evidence on the use of case markers in this process). Under this model, readers are able to detect a non-canonical word order even without a verb, and as a consequence, in the present experiment, the crucial region for comparison is the object NP in italics that precedes the ditransitive verb in each sentence. In (3b), NP-\text{dat} is the first word that unambiguously indicates that the accusative NP was scrambled. Before that point, readers could assume that “Subject NP-\text{acc}” is the initial segment of a simple transitive sentence which does not involve scrambling and only requires a transitive verb to be completed. The italicized NP-\text{acc} in (3a) provides the baseline reading time.

3.1 Method

3.1.1 Participants

Thirty-eight native speakers of Japanese participated in the study for \$20 each. They had all gone to the U.S. as adults and were residents of the Boston area at the time of the experiment.

3.1.2 Materials

Twenty-five sentence pairs like the following were constructed (see Appendix A for the complete list of items used).

(4) a. Ofisu-de shokuin-ga kakaricho-ni ocha-o dasita
    office-nom employee-nom manager-dat tea-acc served
    josei-o teineini hometa-to Aiharasan-ga hanasiteita,
    woman-acc politely praised-comp Aihara-nom said
b. Ofisu-de shokuin-ga ocha-o kakarichoo-ni dasita
office-nom employee-nom tea-acc manager-dat served
josei-o teineini hometa-to Aiharasan-ga hansiteita.
woman-acc politely praised-comp Aihara-nom said

'At the office, Aihara said that the employee politely praised
the woman who had served tea to the manager.'

As previously observed, the regions of interest are the NPs in italics. In
(4b), manager-dat is the earliest point in which it is clear that the NP tea-acc
was scrambled.

The experiment included another three conditions which were unrelated
to the present claims. Five lists were created by distributing the twenty-five stimuli
in a Latin Square design. Each participant saw exactly one of the lists intermixed
with 53 unrelated foil items in pseudo-random order. After each sentence, partici-
pants answered a yes/no comprehension question presented on a new screen.

3.1.3 Procedure
The experiment was conducted on a Power Macintosh 7500/100 running PsyScope
(Cohen, MacWhinney, Flatt, and Provost 1993) with a button-box. Participants
were timed in a phrase-by-phrase self-paced non-cumulative moving-window
reading task (Just, Carpenter and Woolley 1982).\(^3\) The segmentation in the
sentences in (4) (represented by intervening spaces) was the actual segmen-
tation used in the self-paced reading presentation. Sentences were shown using
Japanese characters (as in Appendix A) with the uniform-width font Osaka
Toohaba 14. Stimuli initially appeared as dots with intervening spaces indicat-
ing the segments, and participants pressed the leftmost button of the button-box
to reveal each subsequent region of the sentence and cause all other regions to
revert to dots. At the end of each sentence, a yes/no question appeared on a new
screen, which participants answered by pressing one of the two rightmost but-
tons. No feedback was given. Corresponding data points were eliminated from
further analyses if the participant did not answer the comprehension question
correctly.

The experimental trials were preceded by one screen of instructions and eight
practice trials. All sentences were presented on a single line. The experiment took
participants approximately 20 minutes.

3.1.4 Analysis
We analyzed comprehension question response accuracy and reading times.
Analyses were conducted on residual reading times per region (Ferreira and
Clifton 1986), derived by subtracting from raw reading times each participant’s

\(^3\)Each phrase was constituted of a content word plus functional particles such as a case
marker, a postposition, a complementizer.
predicted time to read regions of the same length (measured in number of characters), which in turn was calculated from a linear regression equation across all of a participant’s sentences in the experiment. The residual reading times were trimmed so that data points beyond three standard deviations from the relevant condition × region cell mean were discarded, corresponding to less than 2.5% of the total data. The means and analyses presented below are based on the trimmed residual reading times. Analyses conducted on raw reading times yielded a similar pattern of results as the residual reading times analyses reported. We report the residual reading times because the number of characters was not controlled for in the critical region.

3.2 Results

3.2.1 Comprehension Question Accuracy

Comprehension question performance was low as participants correctly answered 61% of the canonical word order items and 65% of the scrambled items. The difference was not statistically significant (Fs < 1).

3.2.2 Reading Times

At the critical region, the argument prior to the ditransitive verb, the scrambled condition was significantly slower than the canonical condition (185.10 msec and -207.2 msec respectively in the residual reading time analysis; 1588.68 msec and 1215.29 msec in the raw reading time analysis) in the analysis by subjects (F(1,37) = 7.48; P < 0.05) as well as by items (F(1,23) = 5.41; P < 0.05). The two conditions did not differ significantly at the ditransitive verb (Fs < 1). No significant differences were observed at any of the other regions.

Because of the low comprehension performance, a further analysis was conducted in which the data from the ten participants with the worst comprehension performance were eliminated. In this analysis, the correct response rates for the scrambled condition (73%) and for the canonical condition (69%) were not statistically different (Fs < 1). The scrambled condition was marginally slower than the canonical condition at the region prior to the ditransitive verb (analysis by subjects: F(1,27) = 3.79, P = 0.062; analysis by items: F(1,22) = 3.34, P = 0.081). There was no significant difference at the ditransitive verb (Fs < 1.8) or at any of the other regions.

3.3 Discussion

The present result provides evidence that scrambling leads to longer reading times as previously observed by Mazuka, Itoh, and Kondo (1998; 2001). Because scrambling in the present case occurs inside the VP, the slow-down observed cannot be related to clause boundary detection.

---

4One item was eliminated in the analysis by items because it had no data points for the canonical condition as participants answered the comprehension question incorrectly.
However, there are a number of alternative factors that could be used to explain the slow-down in this experiment. In the next section, we discuss some possible alternatives.

Of particular interest in Experiment I is the fact that the slow-down occurs at the NP preceding the ditransitive verb, but not at the verb itself (see Mazuka, Itoh, and Kondo 2001, for a similar pattern of results; and the discussion below on the direct association hypothesis; Pickering and Barry 1991).

4 Sources of Difficulty in Scrambling

Although Experiment I provides evidence that scrambled word orders involve more processing difficulty than their respective canonical word orders, the source of such a difficulty is not entirely clear. In this section, we will argue that the difficulty can be explained based on the assumptions below.

(I) Readers compute gaps for scrambled constituents in sentences with non-canonical word order.

(II) The positioning of a gap leads to processing slow-downs.

According to these assumptions, readers should slow down whenever it is necessary for a gap to be posited in a scrambled sentence. Furthermore, given proposals in the language processing literature related to working memory constraints (Gibson 1998; Just and Carpenter 1992), we also adopt the following.

(III) The slow-down in (II) is affected by the amount of working memory that readers have available during parsing and by the distance between the gap and its antecedent.

Assumption (III) can be justified given that the processing of a gap necessarily requires maintaining and eventually finding a discourse entity (the antecedent for the gap) in working memory and, therefore, the less working memory available the harder the retrieval of the relevant entity, thus, leading to slower processing times. The distance component of (III) is supported by evidence from a self-paced reading experiment using the following sentences in English (Gibson and Warren 1998). According to Gibson and Warren, the verb annoyed was read faster in (5a) than in (5b) because the intermediate trace (after concluded) in the former sentence facilitated the processing of the trace being posited at annoyed. Under one possible interpretation, the trace at annoyed was processed faster in (5a) because its antecedent had been reactivated more recently at the intermediate trace.

(5) a. The victim [who, the counselor concluded \(_{CP} \text{gap}_i \) that the critical comment had annoyed \(_{\text{gap}_i} \) ] has switched to another psychoanalyst.

    b. The victim [who, the counselor’s conclusion about the critical comment had annoyed \(_{\text{gap}_i} \) ] has switched to another psychoanalyst.
4.1 The Incremental Processing of Scrambling

In the following, we discuss how assumptions (I) to (III) above apply to the processing of scrambled ditransitive constructions in Japanese such as the one schematically represented below.

(6) \textit{NP-nom \textit{NP-acc} NP-dat \ldots}

In the incremental parsing model adopted here, Japanese speakers create a partial representation for NPs based on various sources of information such as plausibility, animacy and, in particular, case marking. Initially, when reading the sequence in (6), there is a preference to associate the first two NPs in a simple transitive construction (see Miyamoto 2000, for experimental evidence). However, as the dative NP is read, there is a preference for a ditransitive clause to be built (Yamashita 1997). Moreover, it is also clear at this point that the accusative NP was scrambled within the ditransitive clause, and its gap can be posited immediately after the dative NP. In this respect, we follow previous findings in the processing of filler-gap dependencies in assuming that gaps are inserted as early as possible in the input string (e.g., Frazier and Clifton 1989, and references therein). In short, even before the verb is read, there is enough information for a gap to be inserted after the dative NP in (6).

In the processing of fronted \textit{wh}-phrases in English, filler-gap dependencies involve two separate processes: detecting a \textit{filler} (the \textit{wh}-phrase) and, at a later point, creating a related gap. In contrast, in Japanese scrambling, the two processes often occur simultaneously because it is not immediately apparent that a constituent was displaced. Thus, in (6) the fact that \textit{NP-acc} was scrambled only becomes clear when \textit{NP-dat} is read, at which point the gap is also posited.

According to assumption (II), there should be a slow-down at the indirect object \textit{NP-dat} in order for the gap to be inserted. Moreover, following (III), such a delay should be longer if the distance between the accusative NP and the position of the gap (in other words the position of the dative NP) is larger.

The adoption of (I) to (III) provides an account of Japanese scrambling which is compatible not only with prevailing theoretical linguistics proposals for the phenomenon (e.g., Hoji 1985), but also with previous behavioural literature according to which the processing of gaps have a number of observable effects during on-line processing (Bever and McElree 1988; Gibson and Warren 1998; MacDonald 1989; \textit{inter alia}).

In what follows, we discuss three alternative factors that could also explain the slow-down observed in Experiment I. As should be clear from the discussion, this inability to determine the source of the difficulty more precisely is not a problem exclusive to Experiment I, but rather it may be unavoidable whenever a scrambled sentence is directly compared to its corresponding canonical version.

4.2 A Case Marker Processing Hierarchy

In experiments comparing canonical and scrambled word orders directly (such as Experiment I, as well as Mazuka, Itoh, and Kondo 2001; Yamashita 1997),
one unavoidable design restriction is that NPs with different case markers have to be compared. Consequently, it could be suggested that the reading time differences at those points result from intrinsic properties of the case markers being processed, and not from scrambling itself.

In more general terms, there may be a hierarchy according to which some case markers may require longer processing time than others. The nominative *ga* seems particularly marked in a number of ways as it involves complex discourse settings (see, for example, Kuno 1973, for the exhaustive reading of the *ga* marker) and idiosyncratic syntactic properties (e.g. nominative subjects cannot be scrambled; Saito 1985). Consequently, among the various case markers, the nominative *ga* may take the longest to be processed.

The difference between the dative *ni* and the accusative *o* markers is more subtle, without one being necessarily more salient than the other. But if importance of a constituent for the interpretation of a clause correlates with processing slow-down, then an accusative marked NP should require longer reading times than a dative NP (see Sadakane and Koizumi 1995, for a hierarchy based on affectedness also Babayonshev and Gibson 1995; Chomsky 1981, for differences between structural and inherent case).

If confirmed by empirical evidence, such a case marker hierarchy may pose problems to scrambling experiment in which NPs with distinct case markers are compared. For example, in Mazuka and colleagues’ experiment, the slow-down observed at the nominative NP in the scrambled condition (in comparison to an accusative NP in the canonical condition) may have been caused by the processing of the nominative marker and not by scrambling.

In Experiment I above, the dative marker in the scrambled condition was read more slowly than the accusative marker in the canonical condition, whereas the case marking hierarchy would predict the opposite result. Thus, case marking differences may be less of an issue in this case. However, independent results investigating the processing time of the various case markers would be desirable.

### 4.3 Frequency

A second factor that requires consideration is the relative frequency of the constructions compared. If less frequent constructions require longer reading times to be processed, then any slow-down observed in scrambled sentences can be readily explained because these constructions are less common than their respective canonical versions (see the corpus count in Experiment III confirming this tendency).

But note that even if the correlation between frequency and reading time is shown to hold, frequency in itself is not an explanation. It would be desirable to provide independent motivation (e.g. cognitive or historical factors) that led such frequency patterns to emerge in the first place. Moreover, there is evidence suggesting that frequency in corpora do not always match behavioural preferences in Dutch (Mitchell and Frenck-Mestre 1998) and in English (Gibson, Schütze, and Salomon 1996).
Nevertheless, considering the importance of frequency in lexical access and several models integrating it to sentence-level processing (e.g., Elman 1991; MacDonald, Pearlmutter, and Seidenberg 1994), frequency factors should be carefully considered when interpreting experiments investigating alternative word orders.

4.4 The Adjacency Constraint on Direct Objects

A third factor that cannot be easily dismissed in scrambling experiments is the preference for direct objects to be placed adjacent to their subcategorizing verbs. This preference is strictly observed in English as shown by the following sentences (from Stowell 1981).

(7) a. Paul quickly opened the door.
    b. *Paul opened quickly the door.

Sentence (7b) is ungrammatical because the adverb *quickly* intervenes between the verb and the direct object. In Japanese, the adjacency constraint is also observed, but less directly because of word order flexibility. For example, in informal conversations, the accusative marker can be omitted from direct objects, but, in this situation, grammaticality is degraded if the direct object is not immediately followed by the verb (Saito 1985).

If the adjacency constraint is also a factor during parsing (as intuitive judgements suggest), then it should not be surprising that a slow-down is observed immediately after a scrambled accusative NP because the following constituent is not going to be the expected transitive verb.

In the next section, we describe a different experimental design which avoids the confounding factors discussed above, with the added advantage of testing a proposed property of gaps, namely, that they take longer to process when they are far away from their antecedents.

5 Distance between Gap and Antecedent

In order to eliminate the alternative factors that could have influenced the result of Experiment I, we devised the experimental design in (8). The basic construction is the simple ditransitive construction in (2). In the present case, however, a long intervening phrase (an XP for short) is included before the two object NPs (as schematically shown in (8ab)) or between them as in (8c). In (8ab), the accusative NP is in its canonical position; whereas in (8c), it was scrambled over the dative NP (the gaps indicate the original position of the accusative NP).

(8) a. Canonical (preposed XP):
   Subject XP NP-dat NP-acc ditran-verb.
b. Short scrambling (preposed XP):
   Subject XP NP, -acc NP-dat gap, ditran-verb.

c. Canonical (intervening XP):
   Subject NP-dat XP NP-acc ditran-verb.

d. Long scrambling (intervening XP):
   Subject NP, -acc XP NP-dat gap, ditran-verb.

As before, the region of interest is the argument prior to the ditransitive verb because this is where readers may be positing a gap for the accusative NP in the two scrambled conditions. If assumption (II) above is correct, there should be a slow-down in this region in (8b) when compared to the same region in (8a). However, because (8ab) are relatively simple sentences, the memory load may be small and the positing of the gap in (8b) may be easily overridden by other factors. In particular, given that the case markers are different, they could potentially wash out any differences related to the positing of the gap.

Instead, consider (8d) in relation to (8b). Participants should be positing a gap in both sentences. But if the distance between the gap to its antecedent is a factor, then the slow-down should be more pronounced in (8d) than in (8b). One advantage in this procedure is that the NPs compared have the same dative case marker. Hence, one of the issues raised in the previous section is addressed.

One problem in that comparison is that, intuitively, an XP is more awkward when it intervenes between the two object NPs, independent of scrambling. Thus, (8d) may be more difficult than (8b) not only because of the distance between gap and antecedent, but also because of the intervening XP. In order to obtain a baseline for the intervening-XP effect alone, we included the canonical conditions with a preposed XP (8a) and with an intervening XP (8c). In this type of experiment, then, we would like to find out whether there is an interaction between the position of the XP (preceding or intervening between the object NPs) and the position of the accusative NP (scrambled or canonical).

The adjacency constraint, which posed a problem in Experiment I, becomes less of an issue in the present design. Because the accusative NP is not immediately followed by the verb in either of (8b) or (8d), the processing of the dative NPs in these two conditions should be equally affected by the adjacency constraint.

Another advantage of the present design is that it may resolve the frequency bias discussed in the previous section as well. An intervening XP is awkward regardless of the order in which the object NPs appear. Thus, the frequency of canonical sentences with intervening XP such as (8c) should be low; in which case, a frequency-based model would be less likely to account for an interaction in reading times (see Experiment III for further discussion and relevant corpus counts).

As in Experiment I, the positing of a gap in the present construction occurs before the relevant predicate is read; thus, we should also be able to investigate processing proposals in which gaps are not adopted (e.g., Pickering and
Barry 1991), According to Pickering and Barry, gaps are not computed by readers, and any effects should only be detectable when a constituent (an argument NP, for example) and its predicate are associated. Thus, in the present case, their model predicts no differences when the argument NPs are read, but only at the ditransitive verb. On the other hand, if the positing of gaps does lead to slow-down as it is assumed here, then the delay should be observed before the verb is read.

5.1 Experiment II: Intervening Adverbs — A Preliminary Report

We are in the process of conducting a series of experiments using the basic paradigm depicted in (8) with various types of intervening XPs. Among the alternatives considered are adverbs (e.g. locatives) and relative clauses. Preliminary analyses of a version of the paradigm using a long adverbial phrase (locatives and temporal) as XPs have yielded promising results. The self-paced reading set-up was the same as the one used in Experiment I. Thirty-two native speakers of Japanese, who were students or staff members at Kanda University of International Studies, participated in the experiment. The comprehension performance in the four conditions (between 87.5% and 91.9%) did not differ significantly. In the reading time analyses, the predicted interaction was observed at the object NP preceding the ditransitive verb (F(1,31) = 7.95, P < 0.01 in the analysis by subjects; F(1,19) = 5.03, P < 0.05 in the analysis by items). The two canonical conditions did not differ significantly in this region (Ps > 0.2); whereas the long scrambling condition was significantly slower than the short scrambling condition (P < 0.05). The four conditions did not differ significantly at the ditransitive verb (Fs < 1).

6 Experiment III: Corpus Frequency

One recurring problem in the study of scrambled word orders is that in general they occur with less frequency than their corresponding canonical versions. In the previous section, however, we suggested that constructions with intervening adverbial phrases (AdvPs for short; such as (8cd)) seem to have low frequency regardless of whether the object NPs are scrambled or not. Thus, if the two intervening-AdvP conditions have similarly low frequencies, then frequency models may not easily account for the interaction observed between scrambling and intervening AdvPs in Experiment II.

More specifically, consider a model according to which frequent constructions are read faster than similar but less frequent constructions. For a study involving constructions such as the ones in Experiment II, a natural prediction of the frequency model would be that an interaction effect observed in the behavioural data should also be observable in the frequency of the constructions in an appropriate corpus. In the preliminary results of Experiment II, we observed that the scrambled condition with intervening AdvP ((8d); condition D for short)

12
is harder to process than the scrambled condition with preposed AdvP (condition B). On the other hand, the canonical conditions with preposed AdvP (condition A) and with intervening AdvP (condition C) do not differ significantly. Therefore, frequency counts should also yield a similar pattern of results such that \( \text{freq}(B)/\text{freq}(D) > \text{freq}(A)/\text{freq}(C) \), where \( \text{freq}(X) \) is the frequency of construction \( X \) in a given corpus.

A corpus count was conducted in order to determine the frequency of the relevant constructions.

6.1 Procedure

A frequency count was conducted on sentences from the Mainichi Shinbun (a general daily newspaper) between 1991 and 1995, and from the Nippon Keizai Shinbun (a daily financial newspaper) between 1990 and 1996. The sentences were automatically parsed using SLUNG (Mitsuishi, Torisawa, and Tsujii 1998). Only sentences with 50 or less Japanese characters were included because of accuracy constraints of the SLUNG parser.

From the partially bracketed constituents produced by SLUNG, we obtained 215,999 sentences that contained both a \( ni \) marked NP and an \( o \) marked NP in the same clause. Next, sentences containing the verbs used in Experiment II were extracted (see Appendix B for the complete list of verbs used), resulting in 5,102 sentences. These sentences were manually checked by a native speaker to remove irrelevant cases (e.g., collocations, temporals marked with the \( ni \) postposition, constituents other than an AdvP intervening between the two object NPs). The remaining 1,834 sentences were counted and classified according to whether the accusative NP was scrambled and whether an AdvP intervened between the two object NPs.

6.2 Results

We obtained 1,834 sentences which contained two arguments (a dative NP and an accusative NP) and a ditransitive verb in the same clause. Among those sentences, 1,202 (65.5%) were sentences with adjacent object NPs in canonical order (condition A). In 565 sentences (30.8%), the accusative NP was scrambled to the position immediately before the dative NP (condition B). In 46 sentences (2.5%), the object NPs were in canonical word order with an intervening AdvP (condition C). And in the remaining 21 sentences (1.2%), the accusative NP was scrambled with an intervening AdvP between the two object NPs (condition D).

The frequency of condition A is approximately 26 times the frequency of condition C. A similar proportion is observed between the other two conditions as the frequency of condition B is approximately 27 times the frequency of condition D.

---

*We would like to thank Hiroshi Kanayama for providing the bracketed sentences.*
6.3 Discussion

The results confirm the generalization that canonical word orders are more frequently used than scrambled word orders. Even with a commonly used construction such as NP-acc NP-dat (condition B), its canonical version NP-dat NP-acc (condition A) is approximately twice as frequent.

As expected, very few sentences (3.7% overall) presented an intervening AdvP between the two object NPs. Moreover, given that the relative frequencies of adjacent and non-adjacent object NPs do not interact with scrambling, it indicates that frequency cannot easily explain the interaction observed in the behaviour study. As observed earlier, a direct conversion from reading time patterns onto frequency counts, should have yielded a larger ratio between conditions B and D when compared to conditions A and C.

An objection could be raised in relation to the selection of the sentences used in the count. For better comparison and in order to minimize manual labour, we restricted our counts to sentences that contained the ditransitive verbs used in the items of Experiment II. However, because the slow-down in Experiment II occurred before the verb was seen, clearly the verb itself does not play a role in the slow-down. Thus, it could be argued that, also in the corpus count, the verbs are not relevant and should not be used as a criterion to select sentences; instead, all sentences with a ni marked NP and an o marked NP, regardless of the accompanying verb, should have been included in the count. With this larger sample, it is possible that the results could have reached the interaction levels that frequency models would predict. Nevertheless, the present results are promising in that they suggest that the constructions we are utilizing are less likely to be influenced by frequency factors than the usual constructions used in scrambling experiments.

7 General Discussion

Experiment I provides evidence that scrambled word orders lead to longer reading times when compared to canonical word orders. Contrary to previous studies, we used a structure in which scrambling occurs inside the VP; therefore, the slow-down observed cannot be reduced to the processing of clause boundaries. Moreover, because the case markers on the NPs compared are not as salient as the nominative marker, the comparison may have avoided confounds at the case marking level as well. In addition, because the scrambled construction (the accusative-dative order) is commonly used by native speakers, familiarity issues may have been minimized (although Experiment III suggests that even in this case the canonical dative-accusative order can be twice as frequent). Based on Experiment I and the result reported by Mazuka, Itoh, and Kondo (1998; 2001), the first part of the paper concluded that scrambling in Japanese involves extra difficulty (contra Yamashita 1997).

This conclusion naturally raises the question as to what leads to the extra difficulty in scrambled word-orders. We argued that the kind of experimental
design used thus far, in which the canonical and scrambled sentences are directly compared, involves various confounding factors and is consequently ill-suited to provide a clear answer. We then proposed a different experimental design based on the assumption that the processing of a gap for a scrambled constituent is responsible for the slow-down observed in Experiment I. The design of Experiment II has several advantages over the traditional experiments in this area. One of them is illustrated in Experiment III, in which results indicate that frequency factors are less of an issue.

The results obtained suggest that readers slow down as soon as it is clear that a constituent was scrambled. The slow-down is more noticeable when the task requires close attention and, presumably, more working memory, or when the gap being posited is far away from its antecedent. The gap-processing hypothesis advocated here is compatible with studies in English showing that, first, there is reactivation of an entity when a related gap is processed (Bever and McElree 1988; MacDonald 1989), and second, the distance to the antecedent influences the time necessary to process the gap (Gibson and Warren 1998). Moreover, our proposal favours syntactic models in which the NP---NP order involves scrambling of the accusative NP (Hojjii 1985; Takano 1996), and it is not easily accommodated if this word order is assumed to be canonical (Miyagawa 1997).

7.1 Probe Recognition Experiments

Contrary to our conclusions, Sakamoto (2001) detected no measurable effects of scrambling in a series of experiments using a probe recognition task. In this type of experiment, an antecedent that was reactivated more recently (e.g., at a related pronoun or gap) is predicted to be recognized faster (see Bever and McElree 1988, for discussion). However, in Sakamoto’s experiments, no such advantage should be expected in the first place because the antecedent is reactivated equally recently — at the PRO — in all sentences.

In related work, Nakayama (1995) found that recognition of a probe was significantly slower when its antecedent was part of a scrambled constituent. This is the opposite of the expected result if gap processing were to facilitate probe recognition. However, when independent factors such as linear position of the scrambled constituent are controlled for, recognition turns out to be faster after the scrambled sentence in comparison to its canonical version, thus supporting the gap processing hypothesis (Miyamoto and Takahashi 2000; see Nakano, Felser, and Claesen, 2000, for a similar conclusion for long-distance scrambling).

7.2 Alternatives to Gap Processing

The present discussion has centered on linguistics approaches that encode gaps (as traces or as slash features); but it should be noted that syntactic proposals which do not assume gaps (e.g., Ates and Steedman 1982) cannot be ruled out without more careful scrutiny. It is necessary to consider how the latter type
of theories explains syntactic phenomena related to scrambling and how the corresponding mechanisms (that are posited instead of gaps) can be translated into a model of on-line processing. We mentioned in passing one such a model which attempts to provide an account of behavioural phenomena using categorial grammar (Pickering and Barry 1991; Pickering 1993; and for criticisms of their model see Gibson and Hickok 1993; Gorrell 1993). Pickering and Barry correctly observe that most behavioural experiments in English use sentences in which the gap occurs soon after the verb. Therefore, they argue that those results can be accounted for if one assumes that the association of a verb to a phrase requires the reactivation of the entities in the phrase. In this case, the processes that were claimed to occur at the gap take place at the verb itself. Although this proposal can account for various aspects of the English results, it requires further assumptions in order to explain the findings in Japanese scrambling, for which we have observed that the slow-down occurs at an NP, even before the verb is processed. The fact that Pickering and Barry’s model cannot readily account for the behavioural data in Japanese scrambling, is not necessarily an argument against categorial grammars.

8 Conclusion

The present paper reported a self-paced reading experiment as evidence for difficulty in the processing of scrambling in Japanese. It also discussed ways of further investigating the sources of this difficulty and suggested that the preliminary results of a self-paced reading experiment and a corpus count support the view that gaps for displaced constituents are posited during the on-line processing of scrambled sentences.

Acknowledgments

We would like to thank the audiences at the Keio Minimalist Syntax group (Keio University) and at Tsukiji Laboratory (University of Tokyo) for helpful comments on earlier presentations of this work. We would also like to thank Reiko Mazuka and Kenji Itoh for discussions clarifying various points in their studies and in the present paper, and Mineharu Nakayama for detailed comments on previous versions of this paper. All remaining errors are our own.

Experiment I was partially funded by the JST/MIT joint international research project ‘Mind/Articulation’ while the first author was a graduate student at MIT.

Experiment II was supported by a Grant-in-Aid for COE Research (08CE1001) from the Japanese Ministry of Education.

CD-ROMs of the Mainichi Shinbun and Nippon Keizai Shinbun were used in order to obtain the sentences for Experiment III.
References


18


Appendix A.

The following are the items used in Experiment I. The slashes indicate the segmentations used in the self-paced reading presentation. Both conditions are presented for item 1. For all remaining items, the scrambled condition (B) can be obtained from the canonical condition (A) by exchanging the order of the third and fourth regions.

1A. Canonical:
オフィスで / 職員が / 係長に / お茶を / 出した / 女性を / 丁寧に / 誓めたと / 相原さんが / 話していた。

1B. Scrambled:
相原さんが / 話していた / 丁寧に / 誓めたと / 女性を / 出した / 係長に / お茶を / 職員が / オフィスで /
1B. Scrambled:
オフィスで / 職員が / お茶を / 係長に / 出した / 女性を / 丁寧に / 誓めたと / 相原さんが / 話していた。

2A. バーで / 訪問客が / やくざに / 花びんを / 見せた / ホステスを / ひそかに / 殺したと / 井原さんが / 言った。

3A. 教室で / 先生が / 大学生に / 公式を / 説明した / 研究者を / 簡単に / 招待したと / 上野さんが / 言った。

4A. 喫茶店で / 不良が / 友人に / 後輩を / 紹介した / オーナーを / 思いきり / なくしたと / 入江さんが / 濡らした。

5A. 京都で / 少年が / 係員に / 荷物を / 渡した / 駅員を / 徹底的に / 捕していたと / 太陽さんが / 聞いた。

6A. 食堂で / 運転手が / おばさんを / 定食を / 注文した / 若者達を / すぐ / 告訴したと / 川村さんが / 言った。

7A. 田舎で / 作家が / アシスタントに / カメラを / 預かった / 写真家を / すぐ / 音にしたと / 川村さんが / 言った。

8A. 駐車場で / 主人が / マネージャーに / 車を / まかせた / 有名を / 堂々と / 説得したと / 木村さんが / おっしゃった。

9A. 自宅で / お兄さんが / 友人に / ビデオを / 貸した / 仲間を / ためらいがちに / だまして / けい子さんが / 濡らした。

10A. 下町で / おばあさんが / 戸屋に / 借金を / 払った / 酒屋を / 簡単に / 追い出したと / 阪本さんが / 言った。

11A. カラオケで / 歌手が / オーナーに / 歌を / 挙げた / お客を / 優しく / 座らせたと / 田中さんが / 言った。

12A. 大阪で / 女優が / 監督に / 手袋を / 投げた / 観客を / 一生懸命 / 思い続けていたと / 中村さんが / 聞いた。

13A. 駅で / 婦人が / 長男に / かばんを / 預けた / 友達を / 静かに / 瞑めていたと / 平野さんが / 話していた。

14A. 会議で / 社長が / 部長に / 新製品を / 預けた / 副社長を / 怒って / やめさせたと / 木村さんが / つぶやいた。

15A. 農家で / 老人が / 一人娘に / 土地を / 残した / 親戚を / すぐに / 呼び出したと / 松本さんが / しゃべった。

16A. デパートで / 店員が / 女の子に / 洋服を / 見せていた / 高校生を / 緊急 / 見たと / 総務さんが / 頼んだ。

17A. 運動会で / 役員が / 光母さんが / お菓子を / 焼いた / 子供達を / 大声で / 呼んだと / 木村さんが / 言っていた。

18A. 広島で / 政治家が / 校長に / ピアノを / 預けた / 市長を / 無理に / 推薦したと / 山本さんが / 濡らした。

19A. バーで / 音楽家が / 蕨子に / 花束を / あげた / 知人を / ささげなく / 呼んだと / 長谷川さんが / 聞いた。

20A. 長崎で / 少女が / お姉さんに / おみやげを / 頼んだ / 親友を / 楽しみに / からかって / 木村さんが / 叫んでいた。

21A. レストランで / 男性が / 女人に / 指輪を / くれた / 売り子を / すぐに / 驚か / 通報したと / レポーターが / 聞いた。

22A. 公園で / 男の子が / 両親に / キャンディを / ねだった / 仲良しを / 元気よく / 呼び出したと / 山下さんが / 言った。
23A. 研究室で / 教授が / 先輩に / 壺を / 作った / 後輩を / 周到に / おとし入れたと / 秘書が / 溜らした。
24A. 横浜で / 女優が / 母親に / 車を / 買った / 男性を / 親切に / 世話したと / 宮田さんが / 話した。
25A. 大学で / 教授が / 学生に / 古文書を / 貸した / 図書館司書を / 急いで / 呼び出したと / 植田さんが / しゃべった。

Appendix B

The following are the ditransitive verbs that were used in Experiments II and III.
推薦する ‘recommend’
渡す ‘hand over’
紹介する ‘introduce’
注文する ‘order’ (food/drinks)
預ける ‘entrust’
届ける ‘deliver’
賃ぐ ‘give’ (as a tribute)
説明する ‘explain’
貸す ‘lend’
残す ‘leave/hand down’
あげる/上げる ‘give’
差し上げる ‘give’ (respectfully)
押し付ける ‘force upon’
出す ‘put out/serve’
勧める ‘suggest’