Abstract

Despite the theoretical attention the phenomenon of scrambling in Japanese has received, very little is known about the functional nature of it in human discourse. Based on an analysis of natural texts, this study explores the linguistic properties of scrambled sentences and investigates why they occur in discourse.

An analysis of a variety of texts ranging from conversational to formal scripts revealed that scrambled constituents have two prominent properties: ‘heaviness’ (containing a subordinate or relative clause) and ‘referentiality’ to the immediately preceding context. These two properties account for almost all occurrences (95 percent) of scrambling, indicating that scrambled sentences are not a spontaneous variation of a canonical order, and that their distribution should be accounted for in a principled manner.

The study further investigates whether the motivation for scrambled sentences is discourse related, such as to provide useful cues to the listener/reader in order to optimize comprehension. The results showed that neither the principle of ‘given before new’ nor changes of topic directly and uniformly account for the occurrence of all scrambled sentences. The study concludes that comprehension-based perspectives are unable to account for the motivation of scrambled sentences and advocates investigation from production-based perspectives.

Keywords: word order; scrambling; discourse; production; Japanese.

1. Introduction

The phenomenon of ‘scrambling’, the mutation of the canonical SOV word order in Japanese, has been extensively investigated in a theoretical framework (e.g., Miyagawa 1997; Saito 1985, 1989; Tada 1993). In contrast to its theoretical aspects, little is known about the nature and
function of scrambling in natural discourse. How commonly it occurs, whether there are common linguistic features across all cases, what function it plays in a discourse, and why speakers/writers produce (write or utter) scrambled sentences are all yet to be investigated.

Investigation of the linguistic nature of scrambled sentences is not straightforward. Unlike sentences with a topicalized phrase, scrambled sentences do not strike one as semantically/pragmatically distinct from their canonical counterparts. Compare the canonical and topicalized sentences shown in (1). There is a sense of ‘exhaustive reading’ on John-ga in (1a) (Kuno 1973). In (1b), ringo ‘an apple’ is topicalized; the constituent is marked by wa and fronted to the sentence-initial position. By topicalization, the sense of exhaustive reading on John disappears; instead, the sense of ringo as an entity receiving contrastive reading and/or representing ‘given’ information or the information registered in the discourse, emerges (e.g., Kuno 1973).

(1) a. (Canonical sentence)
   
   John-ga ringo-o tabeta.
   
   ‘John ate an apple. [It is John who ate an apple.]’

b. (Topicalized sentence)
   
   Ringo-wa John-ga tabeta.
   
   ‘As for the apple [in contrast to other foods], John ate [it].’

Now compare the canonical sentence with the scrambled sentence shown in (2). Sentence (2a) is in canonical order. Sentence (2b) demonstrates the scrambling of an accusative case-marked direct object.

(2) a. (Canonical sentence)
   
   John-ga ringo-o tabeta.
   
   ‘John ate an apple.’

b. (Scrambled sentence)
   
   Ringo-o John-ga e tabeta.
   
   ‘John ate an apple.’

The difference in meaning between (2a) and (2b) is not clear, except that the emphatic stress on the scrambled constituent ringo-o results in some type of ‘focus’. Just how exactly the meaning of scrambled sentences differs from canonical ones is not easily recognized simply by reading the sentences.
Studies based on the meaning and syntactic well-formedness of scrambled sentences on a sentence level—although representing a valid method in a theoretical investigation—have limits in functional analyses because they do not incorporate factors from the discourse level or psychological aspects of the participants. Given the subtle difference in meaning between scrambled sentences and their canonical counterparts, it is imperative to observe the phenomenon of scrambling in a context that extends beyond the created sentence; one must observe scrambled sentences in natural texts and examine them from a multidisciplinary perspective.

The article is organized as follows. First, the method of data collection is summarized and the results are reported. Then sentence-internal properties of scrambling, such as characteristics of scrambled constituents, and sentence-external properties, such as the flow of information in the surrounding discourse are investigated. Finally, the study questions whether the occurrences of scrambled sentences are motivated by comprehension-based factors, in order to provide useful cues to the listener/reader so as to optimize comprehension of the sentence in a stretch of discourse.

2. Methods of text analysis

In order to observe naturally occurring scrambled sentences and their surrounding contexts, a text analysis was conducted. The current section summarizes the method and reports the frequency of scrambled sentences.

2.1. Articles used for the study and modes of production

Eighteen short articles, listed in Appendix A, were selected from several magazines ranging in formality: a working woman’s magazine (More), a magazine for married women (Lee), a young mother’s magazine (Petit Enfant), and a political magazine (Bungei Shunjuu). The texts ranged from one to fourteen pages, some including illustrations and pictures. The articles were selected from several magazines in order to examine a wide range of formality and types of speaking/writing, from a very colloquial style (transcription of a forum discussion) to highly literary writing (essay).

Rather than a corpus analysis of just one type of article, various types of articles with different levels of formality were examined in order to avoid results that reflect corpus-sensitive biases. As the sentences examined ranged from highly colloquial to formal, written language, the analysis does not differentiate between scrambling in spoken and written
language. This by no means indicates an assumption that scrambled sentences behave exactly the same in both spoken and written languages. However, despite the obvious difference in pace of production, the aspects common to spoken and written language, such as the maintenance of coherence and communication as a goal, justify this approach. Our goal in the current study is to examine the properties of scrambled sentences, regardless of production mode, as a point of departure. Of course, any claim that is specific to the mode of production must be independently investigated in future studies.

The articles were classified into four types according to the mode and formality of the language: (i) casual conversational or semiconversational script; (ii) casual writing (essay); (iii) formal conversational script (political roundtable discussion); and (iv) formal writing (political essay). The total number of sentences in each category is shown in Table 1.

2.2. Types of sentences examined

The phenomenon of scrambling occurs only in sentences in which the predicate classifies more than one argument. The analysis in this study focused on sentences with the most typical predicates of that type: transitive and ditransitive verbs. Sentences with a stative verb such as *aru* ‘there is’ and the copula, *da/desu*, were excluded from the study due to the fact that their canonical orders are unclear.

Table 2 summarizes the types of sentences targeted in the current study.

Among the 2,635 sentences in all articles, 798 sentences fit the categories in Table 2. Each sentence was tagged according to the following factors: word order (canonical or scrambled), type(s) and number of phonologically null arguments, and the presence of topicalization. The order of each sentence was evaluated based on overt constituents. A sentence with a phonologically null pronoun in the subject position was considered a canonically ordered sentence, as long as the rest of the sentence followed
the canonical order. A sentence with the agentive wa-marked topic at the beginning of a sentence was also considered canonical. All other types of topicalization, including the topicalization of dative and accusative-marked argument and adjunct phrases, were not counted as scrambled sentences. The word order in the subordinate clause was not examined.

The scrambled sentences were categorized into three groups based on the type of scrambling. The first type is the case where a constituent is fronted to the initial position of the matrix clause, ‘short-distance scrambling’. Sentences shown in (3b–e) represent various instances of short-distance scrambling, while (3a) represents the canonical order. All sentences have roughly equivalent meanings.

(3) a. *John-ga Mary-ni ringo-o ageta.*
   John-nom Mary-dat apple-acc gave

   Mary-dat John-nom - apple-acc gave

c. *Ringo-o John-ga Mary-ni e ageta.*
   apple-acc John-nom Mary-dat gave
Sentences (3b–c) demonstrate how one constituent may be displaced to the sentence-initial position. Theoretically, more than one constituent may undergo scrambling, as shown in (3d–e). As noted in Shibatani (1990), however, scrambling of more than one constituent results in a slightly awkward sentence.

The second type of noncanonical order is the permutation of two internal arguments, the indirect and direct objects, marked by the dative and accusative cases, respectively. A scrambled order among those arguments is referred to as ‘internal scrambling’.

Constituents that are involved in scrambling are not limited to a single NP. A sentential complement may undergo short-distance scrambling, as in (5b), or internal scrambling, as in (5c).

Finally, a constituent in the subordinate clause may occasionally appear at the initial position of the matrix clause, as shown in (6). This type is referred to as ‘long-distance scrambling’.

In (6), the accusative-marked constituent ookina ie ‘a big house’, which belongs to the sentential complement of omou ‘think’, is scrambled to the sentence-initial position of the matrix clause.
Table 3. *Number and frequency of scrambled sentences in all texts*

<table>
<thead>
<tr>
<th>Type of word order</th>
<th>Number</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-distance</td>
<td>11</td>
<td>0.41</td>
</tr>
<tr>
<td>Internal</td>
<td>7</td>
<td>0.27</td>
</tr>
<tr>
<td>Long-distance</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19</strong></td>
<td><strong>0.72</strong></td>
</tr>
</tbody>
</table>

### 3. Frequency

Table 3 reports the number and frequency of scrambled sentences in all texts, categorized by type.

The most frequent type of scrambling was short-distance scrambling, followed by internal scrambling. Only one case of long-distance scrambling was observed in the texts. No sentences involved scrambling of more than one constituent. The frequency of scrambled sentences across the texts was low; only 19 scrambled sentences were observed in the 2,635 sentences, i.e., less than one percent of the sentences analyzed in the current study. Since sentences with an intransitive verb or with a copula were excluded from the study, the actual number of scrambled sentences may be higher than reported here. However, the low frequency of scrambled sentences observed in this study is in accord with previous corpus studies of written (Kokuritsu Kokugo Kenkyujo 1964) and spoken texts (Hinds 1983).

Table 4 summarizes the types, number, and frequency of scrambled sentences by text type.

Overall, the frequency of scrambled sentences is low regardless of the type of text, ranging from zero to one-and-a-half percent of all sentences found in each type of text. Among the types of texts, scrambling was most frequently observed in casual writing. The most infrequent context was formal writing, with no occurrences of scrambled sentences found. Because the number of scrambled sentences is small, however, it is not possible to statistically evaluate and compare the text types in which scrambled sentences most frequently occur.

### 4. Syntactic and pragmatic properties of scrambled constituents

Finding a linguistic property that is common across all occurrences of scrambled sentences would be a significant step towards revealing the functions of scrambling in the discourse and the reasons why speakers/writers produce such sentences. We begin by examining the nature of scrambled constituents. The results show that two distinct properties divide the observed scrambled sentences into two groups: the ‘heaviness’
of the constituent and reference to the preceding discourse. We examine each property in detail.

4.1. **Heaviness**

In a majority of the scrambled sentences (fourteen out of nineteen) a scrambled constituent is syntactically ‘heavy’, that is, it contains a subordinate clause (sentential complement or a relative clause) and typically consists of several content words (ranging from three to thirteen in number, with an average of 5.6). This tendency is reported in other text analyses (Dryer 1980 and Hawkins 1994) and also demonstrated experimentally (Yamashita and Chang 2001). Scrambling of heavy constituents can be observed in both short-distance scrambling, as in (7) and (8), and internal scrambling, as shown in (9).

(7) (from ‘How to live life positively’, see Appendix A)

*Demone,* [[dare-ni totte-mo, zinsee -ni oite mudana][however no matter who-to life -in meaningless koto-wa nani hitotu-mo nai][thing-top there is none -comp watasi-wa e omou kedo ne. I -top think but emph ‘You know, no matter who you are, there is nothing in life that is meaningless to you, I think.’]

(8) (from Senji Kuroi, ‘Kenzaburo Oe and the Nobel Prize’, see Appendix A)

[[Watasi-ga oboeteiru hanasi]-o kare-wa e sukkari I -nom remember story -acc he -top completely wasuretesimatte-iru-kamosirenai. has forgotten maybe ‘He may have forgotten completely the story that I remember.’

Table 4. **Number and percentage of scrambled sentences in the analyzed texts**

<table>
<thead>
<tr>
<th>Scrambling</th>
<th>Total</th>
<th>Percentage of all sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short-distance</td>
<td>Internal</td>
</tr>
<tr>
<td>1. Casual conversational/</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>semiconversational script</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Casual writing</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3. Formal conversational script</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4. Formal writing</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
The most frequent are the cases shown in (7), where the sentential complement of a bridge verb such as *omou* ‘think’ or *iu* ‘say’ is scrambled. The scrambled constituent in sentence (8) has a relative clause. Sentence (9) demonstrates internal scrambling, in which the accusative case-marked direct object with a subordinate clause precedes the dative case-marked indirect object.

A remarkable trend was that internal scrambling shows the strong influence of heaviness; among the six internally scrambled sentences, five contained heavy phrases.\(^7\) Interestingly, canonically ordered sentences with a dative-marked indirect object preceding an accusative-marked direct object showed a similar tendency; in canonical order, heavy constituents tended to precede light ones. Among eight canonical sentences with an overt indirect and a direct object, seven either had the heavy dative-marked indirect object preceding a light accusative-marked direct object or both constituents were ‘light’ (noncomplex). Only one contained a heavy accusative-marked phrase.

4.2. Referentiality

The other prominent characteristic of the scrambled constituents is that they make reference to the preceding utterance. Twenty-six percent, or five out of the nineteen scrambled sentences, contain a determiner or an anaphor referring to something directly mentioned in the preceding context or inferable from it. Almost all these constituents are light and therefore this type of scrambled sentence is in almost complementary distribution to those with heavy constituents.\(^8\) Here are some examples.\(^9\)

(10) (from ‘Roppongi in the sixties’; A is a producer, B a singer who comments in astonishment at A’s remark, and C the actress herself)
A: ‘I didn’t think that you [a female actress] were in such a turmoil at that time, because when I saw you at Ondiene [the name of a theatrical show], you told me, “I am the most ‘journalistic’ woman now’.’

B: ‘Sonna hureezu-o yoku 22,3-sai-no onnanoko-ga e tukatta ne. Such phrase-acc emp 22 or 3 year old girl-nom used emp “Such an [arrogant] word, a girl who is just twenty-two or three uses!”’

C: ‘I was young and ambitious.’

(11) (from ‘Men of my type’; the speaker is a model talking about Robin Williams)
‘Akarukute atatakai kanzi dasi, sono tatakasa -o cheerful and warm look and that warmthness-acc mawari -ni -mo e hurimaite kuresoo. surrounding people -to -also -sprinkle looks like “[He] looks like he is cheerful and warm, and [he] looks likely to give such benevolence to people around him.”’

(12) (from ‘Kenzaburo Oe and the Nobel Prize’; the writer grew up with Oe Kenzaburo, the Nobel Prize winner, looking back at his childhood)
‘Hirosi, my cousin, and Kenzaburo Oe were classmates in the French class at the university. At that time, I had failed the college entrance exam twice, and I was still preparing for another. I admired them, and I felt an inferiority complex even at times when the three of us were having a good time. The relationship between them and me was a little awkward because they knew I was somewhat jealous of them, and yet they were very considerate of it and treated me very nicely. Keredo, sonna zibun-no pozisyon-o, watasi-wa e sorehodo however such self-gen position-acc I -top that much iyada -to -wa omotte inakatta. hateful-comp -top did not think “However, I did not hate my position [lit. such position of myself] [among the three of us] that much.” Because I was not a college student [because I failed the entrance exam], I had no need to keep up with them.’

The scrambled constituents in (10) and (11) directly refer back to the words in the preceding context. The referent in (12) is not directly mentioned previously but is inferable from the previous sentence.
Almost all (95 percent) of the scrambled constituents have either the property of heaviness or referentiality. This indicates that a scrambled order is by no means a spontaneous, free variation of canonical word order; on the contrary, the distribution of the scrambled sentences should be accounted for in a principled manner. Nevertheless, the properties of scrambled constituents alone do not seem to straightforwardly provide an explanation for the function of scrambled sentences. Note that the two kinds of properties found here, heaviness and referentiality, are unrelated to each other. While heaviness is a syntactic property, referentiality is a pragmatic/discourse property. It would be possible to account for the two types of scrambling based on two independent functions, but a principle which accounts for all types of scrambling ought to be more parsimonious.

5. Discourse principles and comprehension-based motivations for scrambling

Examination of scrambled constituents on a sentence-internal level alone imposes limits on the investigation of the function of scrambled sentences because it largely neglects the relationship between the scrambled sentence and surrounding discourse. We therefore shift the scope of investigation from a sentence-internal to a discourse level to see whether these seemingly unrelated sentences share a common denominator in the discourse. Two discourse principles that have been claimed to convey information efficiently to the listener/reader are examined: ‘given before new’ and ‘change of topic’. In doing so we also explore whether scrambled sentences are produced to maximize efficiency in communication, that is, whether they are the most optimal form for conveying information in each discourse.

5.1. Given before new

The correlation between word order and informational status has been noted by many scholars. In particular, constituents representing information which is already known to the participants in a discourse (‘given’ information) are claimed to generally precede those representing information not yet known (‘new’ information). We will call this principle ‘given before new’ (e.g., Clark and Clark 1977; Halliday 1967; Prince 1980, 1992). Generally, starting a sentence with a constituent representing given information aids comprehension because it gives the listener/reader a base of information to build on. Topicalization in Japanese, the displacement of a wa-marked constituent to the sentence-initial position, has been claimed to reflect this principle and represent
given information before new (e.g., Kuno 1973; Hinds 1987; Inoue 1980; Maynard 1982).

Notice that the phenomenon of scrambling shares a characteristic with topicalization: it displaces a constituent to the left, typically to the sentence-initial position. We will therefore examine the information status of the scrambled constituent to see whether scrambled sentences occur out of the same motivation as topicalization, that is, so as to place the constituent representing given information before new material.

The typology employed in the study is the one proposed in Prince (1992). Prince distinguishes entities representing given/old information from those representing new information using the speaker’s assumption of what the hearer knows and what the discourse records. Information is ‘discourse-old’ if it represents an entity evoked in the previous discourse, while it is ‘hearer-old’ if the speaker assumes that the hearer knows the entity. In English, a generic NP such as the sun or the White House is hearer-old in a conversation among adults, even if it has not been mentioned in the discourse, because the speaker assumes that the hearer knows these entities. We consider both discourse-old and hearer-old elements as given information.

The examination of scrambled constituents and postponed constituents revealed that the given before new principle does not uniformly account for the distribution of scrambled sentences. Seven out of the nineteen sentences found in the texts indeed represent given information. They are discourse-new but hearer-old, or simply discourse-old; they are either generic or become known to the hearer as a result of inference, even though they had not been directly mentioned in the previous discourse. Examples (10) to (12) shown earlier are in accord with this principle. The scrambled constituents sonna hureezu-o ‘such a word [phrase]’, sono atatakasa-o ‘such benevolence’, and sonna zibun-no pozisyon-o ‘such position of myself’ in (10) to (12) respectively, are all discourse-old because they became known to the listener through the conversation.

However, in many cases the scrambled constituents represent new information. Typically, those scrambled constituents are sentential complements representing the speaker’s opinion. These scrambled constituents represent information that is neither generic nor evoked in the preceding discourse.

(13) (from ‘How to live life positively'; the speaker is a writer being interviewed)

‘My work [as a scenario writer] benefits from a variety of experiences in life. In that sense, I feel lucky that I went through
many jobs, which gave me opportunities to see many things in society.

Demone, dare-ni totte-mo, zinsee-ni oite mudana
however no matter who-dat life -in meaningless
koto-wa nani hitotu-mo nai -to
thing-top there is none -comp
watasi-wa e omou kedo ne.
I -top think but emph

“You know, no matter who you are, there is nothing in life that is meaningless to you, I think.”

In (13), the speaker states her opinion on the value of experiencing a variety of things in life after stating that she went through many jobs before she became a writer. Although the scrambled constituent, dare-ni totte-mo, zinsee-ni oite mudana koto-wa nani hitotu-mo nai-to ‘no matter who you are, there is nothing in life that is meaningless to you’ is related to her previous statement, it is an addition of new information to the discourse, namely, her statement of how she values various work experiences. Its content is neither evoked by the discourse nor assumed by the speaker to be known to the listener. A similar example is shown in (14).

(14) (from ‘Roppongi in the sixties’; A is a producer, B an actress, C a singer; they are talking about the sixties when they were young)
A: ‘I must have been a high school student. Mariko [speaker B] must have been a young junior high student, rather than an actress.’
B: ‘I bet I was only fifteen.’
C: ‘Sorede ore, oboeteru n da kedo, and so I remember ext cop but
Utino oyazi-to ohukuro-ga Iikurakata-mati-ni
my old man-and mom-nom Iikurakata-town-at
resutoran-o hazimeru n da -to
restaurant-acc start ext cop-comp
Syoo-tyan-ga e itta n da yo.
Sho-chan-nom said ext cop emp
‘and so, I remember, “My old man and mom will start [running] a restaurant at Iikurakata-matani”, Sho-chan said.’
So we went there. And we found out that the day [we went] was the opening day.’

The case of internal scrambling poses a further problem. In this type of scrambling, an accusative case-marked direct object is not fronted to the sentence-initial position but to the position between the subject/topic and
dative-marked indirect object. Because the scrambled constituents are not at the sentence-initial position, the information status of scrambled constituents is irrelevant to the aim of starting a sentence with given information. Examples (15) and (16) demonstrate this.

(15) (from ‘Ijime [Bullying]’; the speaker is a mother whose daughter is being bullied by another girl)

‘[Upon finding out that my daughter was being bullied by a girl] I decided to talk to another mother in the same kindergarten to observe how things are. But the situation seemed to remain the same.

Demo, uti-no ko-wa, K-ko chan ni izimereteiru koto-o but my-gen child-top K-ko -by being bullied fact-acc itido-mo watasi-ni e hanasi-masen. once even I -dat talk neg

“My child would never tell me that she was being bullied by K-ko, not even once.”

I myself did not want to admit that my child was being bullied, so I pretended that I did not notice. But I am so worried that my daughter, who cannot express her emotional pain, is hurt inside very much.’

(16) (from Senji Kuroi, ‘Kenzaburo Oe and the Nobel Prize’; the speaker is a writer who grew up with Oe Kenzaburo, the Nobel Prize winner, talking about the impact of the Nobel Prize on Oe Kenzaburo’s reputation in overseas)

‘[The fact that China had not shown a strong interest in Oe Kenzaburo’s literature] had made me reconsider the significance which Japanese literature holds in the literature of the world.

Nooberusyoo-no zyusyoo-wa, Ooe Kenzaburoo-si-no Nobel Prize-gen receiving-top Oe Kenzaburo-Mr.-gen bungaku-ga sono hukami -to hirosa-ni oite ikkoku- no work-nom its depth and width in one country-gen iki-o koeru mono de aru koto-o sekai -ni e meezi-sita. range-acc exceed thing cop fact-acc world-dat showed

“[Mr. Oe’s] receiving of the Nobel Prize clearly demonstrated to the world that Mr. Oe Kenzaburo’s literature exceeds, in its depth and width, the [standard of] one country.”

But much is left as to the question that how Mr. Oe’s literature has been read and how it will affect people’s minds.’

Both sentences (15) and (16) start with a wa-marked topic representing given information. Notice that these sentences start with given
information due to the fact that the first word is a wa-marked topic; the scrambled internal arguments are irrelevant to the ‘given’ status of the beginning of the sentence.

The principle of given before new does not directly account for all cases of scrambled sentences. Some cases of scrambling start the sentence with new information, while in some cases the scrambling is irrelevant to the ‘given’ status of the beginning of a sentence. This also demonstrates that the functional nature of scrambling is not identical to that of topicalization in Japanese, despite the similarity in their constructions.

5.2. Change of topic

Previous analyses of English and other languages have demonstrated a correlation between syntactic representations and the continuity of discourse topics. Morphologically complex sentences, including sentences with marked word order, are claimed to be associated with a change in topic (e.g., Givón 1979, 1983; Fletcher 1985; Fox 1985). Listeners/readers use the marked structure as a cue in detecting the change of topic. Scrambling is clearly a marked structure because of its deviation from canonical order; therefore, we now investigate whether scrambled sentences are uttered to cue a change of topic.

The continuity of topic can be defined using the content of the sentence and the one immediately preceding it (Fletcher 1985): if the topic of a scrambled sentence is the same as the one immediately preceding it, the topic is classified as ‘continuing’, whereas if the topic of a scrambled sentence is different from the immediately preceding sentence, it is classified as ‘changed’.

Table 5 summarizes the number and percentage of scrambled sentences which continue a topic and which change the previous topic.

As shown in Table 5, most scrambled sentences (89 percent) continue an ongoing topic. The sentences in (10) to (13) and (15) to (16) all fall into this category. Two sentences appear at the change of topic. It does not appear, however, as if the scrambling is the main signal of the

| Table 5. Number and percentage of scrambled sentences in continuing topic and at the change of topic |
|---------------------------------------------------------------|-----------------|-----------------|
| Scrambled sentences continuing the topic                        | 17              | 89              |
| Scrambled sentences at the change of topic                      | 2               | 11              |
| Total                                                           | 19              | 100             |
change in these two cases. As shown in (17) and (18), the sentences include a conjunction or phrase that aggressively indicates the shift in the topic.

(17) (from Sadaharu Oh, ‘Parting with Giants’; the speaker is a former coach of the Yomiuri Giants, a professional baseball team in Japan) ‘... I was very selfish in terms of my life with baseball. I never thought of taking my children out on weekends. I know some baseball players wake up early in the morning to see their children off to school and go back to sleep, but I never did anything that would negatively affect my training in baseball. I never consulted my wife about my move in career [changing teams] either, because she said, “Do whatever you want”. I feel a little sorry for her, though, because my second daughter got married recently and she [my wife] must be a little lonely.

_Tonikaku, katu yorokobi-o sensyu-ni e osiete agetai desune._

At any rate _win joy -acc players-dat _teach for them cop emp

“At any rate, I want to show the players the joy of winning.”

I don’t want to compare them with the Giants players, but I cannot help it.’

(18) (from ‘Roppongi in the sixties’; A is a producer, B an actress, C a singer; they are talking about the sixties when they were young)

A: ‘I must have been a high school student. Mariko [speaker B] must have been a young junior high student, rather than an actress.’

B: ‘I bet I was only fifteen.’

C: ‘_Sorede ore, oboeteru n da kedo,_ and so _I remember ext cop but

_Uti-no oyazi -to ohukuro -ga Iikurakatamati -ni my _old man-and mom -nom Iikurakata town-at

_resutoran-o_ hazimeru n _da -to restaurant-acc start _ext cop-comp

_Syoo-tyan-ga e itta n _da _yo._ Sho-chan-nom said ext cop emp

‘And so, I remember, “My old man and mom will start [running] a restaurant at Iikurakatamati,”’ Sho-chan said.’

So we went there. And we found out that the day [we went] was the opening day.’

The phrases, _tonikaku ‘at any rate’ or sorede ore, oboeteru n da kedo … ‘and so, I remember …’_ alone suffice to signal the change in topic. Even if the occurrence of scrambling at the change of topic in these two cases
may not be purely coincidental, it does not appear to be the dominant force in signaling the change.

5.3. *Are scrambled sentences uttered to optimize comprehension?*

So far we have examined scrambled sentences in a surrounding discourse in light of two powerful discourse principles: given before new, and change in topic. Sentences that follow these principles are considered to facilitate comprehension. The given information provides the listener/reader with a basis to build on. Seeing/hearing a marked word order can signal a change in topic to the listener/reader. Contrary to expectations, scrambled sentences appear to occur independent of these principles.

The lack of correlation between scrambled sentences and the two principles may indicate that we have not yet found the right measure to capture all occurrences of scrambling, that is, how scrambling contributes to comprehension. However, it also suggests the possibility that scrambled sentences are produced for reasons other than facilitating comprehension. In order to examine whether or not scrambled sentences are indeed produced for comprehension-based reasons, we now examine whether the scrambled sentences are always the form that conveys information most efficiently. If a scrambled sentence is indeed chosen to optimize comprehension, ‘unscrambling’, or replacing it with the canonical counterpart while keeping the original context, should make the sentence degraded or difficult to understand. If, on the other hand, replacing it with a canonical sentence does not affect comprehension, or even improves comprehension, then it is likely that scrambled sentences are motivated by reasons other than the optimization of comprehension.

When put back in canonical order, all cases of scrambled order with a referential phrase seem to shift the priority of the information conveyed and therefore convey information less efficiently. As contrasted below, the canonical counterparts are not as natural as the original scrambled sentence, even though both are grammatical. The hash mark (#) indicates increased difficulty in interpreting the sentence or a noticeable change in meaning from the original sentence in the original context.

(19) (from ‘Yomiuri’s proposal for the Constitution’; the speaker is agreeing with another who has just been pointing out that it is unfair that congressmen limit the right to appeal to the Constitutional Court to cabinet members and congressmen)

a. (Scrambled order)

Kokkaiginnattigga zibuntatini tugonowarui koto o
congressmen-nom oneselfs-dat inconvenient matter-acc
‘Do you think that the congressmen bother to appeal to the Constitutional Court [about] something that is inconvenient to themselves [if the matter were judged]?’

b. (Canonical order)

\[
\text{Kokkaigintati-ga kenpoosaibansyo -ni}\ 
\text{congressmen-nom the Constitutional Court-dat}\ 
\text{zibuntati-ni\ \ tugoo-no warui koto -o}\ 
\text{oneselves-dat inconvenient matter -acc}\ 
\text{wazawaza moositatemasu ka.}\ 
\text{bother to appeal Q}\ 
\]

‘Do you think that the congressmen bother to appeal to the Constitutional Court [about] something that is inconvenient to themselves [if the matter were judged]?’

In sentence (19), the antecedent of zibuntati ‘themselves’ in the scrambled order is easily accessed because there is only one NP, kokkaigintati ‘the congressman’, that immediately precedes the anaphor. In contrast, the antecedent is accessed less directly in the canonical order, because another constituent, kenpoosaibansyo-ni ‘by the Constitutional Court’, intervenes between the anaphor and its antecedent. Therefore the listener/reader of the sentence must go back two constituents to determine which one is the antecedent.

Not only must one go back to the beginning of the sentence and recall the first word upon encountering the anaphor, but as a false candidate for the antecedent the second constituent also momentarily distracts the hearer from processing the anaphoric expression. Although the Constitutional Court is not a typical agentive noun, it may be interpreted as the group of people who work there, as in the case of Kuni-wa kenpoosaibansyo-ni sekinin-o oraseta ‘the government let the Constitutional Court assume its responsibility’. Therefore, both the nominative-marked constituent, kokkaigintati ‘the congressmen’, and the dative-marked constituent, kenpoosaibansyo ‘the Constitutional Court’, may potentially serve as the antecedent, and some inference about agentivity and number must be made before one interprets the intended antecedent.\(^{10}\) Other scrambled sentences with referential phrases demonstrate the same point.

Unlike the example in (19), however, many canonical sentences are just as informative as, or even more informative than, the scrambled sentences
in the original context. For instance, in example (20), whether ‘the pictures with the crest of the paulownia leaf’ comes before or after ‘each reporter’ does not change the ease of comprehension.

(20) (from ‘Being a Prime Minister’; the speaker is a former journalist who specialized in Prime Ministers, who is looking back on the old days)

‘I remember Prime Minister Sato well . . . . We went on a picnic, spending all day, along with reporters. He seemed to enjoy eating lunch with people around him, including reporters.

a. (Scrambled order)

Zenbu syasin-o totte kurete,
all pictures-acc take for us
kaeruto kiri no ha no naikaku-no
on returning paulownia-gen leave-gen Cabinet-gen
monsyoo-o ouinsita karaa syasin-o
crest -acc stamped color picture-acc
kisya hitorihitori-ni e purezento-site kureta.
reporter each -dat presented for them

“He [the Prime Minister of Japan] took pictures of everyone, and when [we] returned, [he] gave pictures with the crest of the paulownia leaf, the Cabinet crest, to each reporter.”
Those pictures of former Prime Minister Sato and me remind me, even now, of the good old days.’

b. (Canonical order)

Zenbu syasin-o totte kurete, kaeruto
all pictures-acc take for us on returning
[kisya hitorihitori]-ni
reporter each -dat
kiri -no ha -no naikaku-no monsyoo-o
paulownia-gen leave-gen Cabinet-gen crest -acc
ouinsita karaa syasin -o purezento-site kureta.
stamped color picture-acc gave as a present

‘He [the Prime Minister of Japan] took pictures of everyone, and when [we] returned, [he] gave each reporter the pictures with the crest of the paulownia leaf, the Cabinet crest.’

The following example demonstrates that in some cases a sentence is even better comprehended when put back into canonical order. The scrambled order in (21) temporarily leads the listener to interpret the referent of uti-no oyazi ‘my old man’ as speaker C’s father. This interpretation holds until the listener hears the matrix subject Syoo-tyan because until then it is the only possible referent. In contrast, in the canonical sentence, the
referent of ‘my old man’ is clear because the nominative-marked subject Syoo-tyan, which strongly signals the possible referent of ‘my old man’, immediately precedes the word.

(21) (from ‘Roppongi in the sixties’; A is a producer, B an actress, C a singer; they are talking about the sixties when they were young)
A: ‘I must have been a high school student. Mariko [speaker B] must have been a young junior high student, rather than an actress.’
B: ‘I bet I was only fifteen.’

a. (Scrambled order)
C: ‘Sorede ore, oboeterun da kedo, and so I remember ext cop but Uti-no oyazi-to ohukuro-ga Iikurakata -mati-ni my old man-and mom-nom Iikurakata -town-at resutoran-o hazimeru n da-to restaurant-acc start ext cop-comp Syoo-tyan-ga e itta n da yo. Sho-chan-nom said ext cop emp ‘And so, I remember, “My old man and mom will start [running] a restaurant at Iikurakatamati,” Sho-chan said.’ So we went there. And we found out that the day [we went] was the opening day.’

b. (Canonical order)

Most scrambled sentences with heavy phrases do not show a change in effectiveness or become more informative when they are unscrambled. Thus, the examination revealed that scrambled sentences are not always the most efficient sentence form for conveying information in a given context. In some cases scrambled sentences are not interchangeable with canonically ordered sentences without compromising the efficiency of comprehension. In many cases, however, putting sentences back into canonical order does not make a difference to the discourse;
at times it even facilitates comprehension. This constitutes a further argument against the claim that scrambled sentences are produced for comprehension-based reasons.

6. Possibility of production-based motivations for scrambling

We have observed that scrambled sentences have systematic characteristics, such as the heaviness and the referentiality of constituents that are scrambled, which suggests that they are not random variants of canonical order. Yet we have not seen that they are produced to optimize comprehension. Why do speakers/writers produce scrambled sentences? A possible alternative to a comprehension-based explanation is a production-based explanation, incorporating factors which influence various aspects of the cognitive processes involved in a speaker’s/writer’s generation of a sentence.

In order to claim that an utterance is made to maximize the listener’s/ reader’s comprehension one must assume that the speaker/writer follows the cooperative principle (Grice 1975). However, uttering a sentence in a timely fashion in speech involves cognitive processes that take place at an amazingly high speed; according to Levelt (1989; cited by Bock 1991), speakers select a word about once every 400 milliseconds from a vocabulary of approximately 30,000 words. Multiple levels of cognitive activity, including retrieval of lexical items from the mental lexicon while maintaining coherence by constantly updating the conversational record, must take place in a conversation. Even though the demands of time may not be as pressing when writing as they are when speaking, one must keep track of previous discourse and maintain cohesion while verbalizing concepts. In such a demanding task, it is quite possible that some sentences are not produced for the sake of optimal comprehension.

Indeed, a number of psycholinguistic studies in English give experimental evidence which suggests that the choice of syntactic structures is speaker-based (Arnold 1998; Stallings et al. 1998; Ferreira and Dell 1996; Wasow 1997a, 1997b). They demonstrate that factors such as efficiency in the speaker’s working memory during the processes of utterance or psychological availability of a word over others affect word order in generating a sentence. Wasow (1997a, 1997b), for instance, cites the case of a baseball announcer who always utters a sentence with a heavy NP shift whenever he announces a new player coming to bat.

(22) That will bring to the plate Barry Bonds.

Wasow diagnoses that the announcer uses such an order instead of the canonical order *That will bring Barry Bonds to the plate* in order to gain a
second to check the name of the next player. Wasow (1997a) further conducted a corpus analysis that demonstrates that, whenever possible, speakers choose a word order that delays their commitment to one structure over others and therefore maintains the flexibility of an utterance for longer.

Furthermore, in an experimental study Ferreira and Dell (1996) demonstrate that speakers use a form that is easier to produce, even though it may not be the optimal way of conveying information. It is well known that the ‘temporal ambiguity’ of a sentence, or the possibility that the sentence may be interpreted in more than one way, adds a source of difficulty in comprehension (e.g., Berwick and Fong 1995). One example is how a sentence ‘unfolds’ after the verb know in English. In a sentence fragment such as The coach knew ..., the sentence may end with a direct object, as in The coach knew you, or a subordinate clause, as in The coach knew you missed a practice yesterday. Observe the representative sentences used in Ferreira and Dell (1996):

(23) a. The coach knew [that] you missed a practice yesterday.
    b. The coach knew [that] I missed a practice yesterday.

In (23a), the nominative case is not overtly reflected in you. Therefore without the overt complementizer that, the phrase the coach knew you is temporarily ambiguous. In contrast, the structure in (23b) is not ambiguous regardless of the overt complementizer because the nominative case is reflected in the subject of the sentential complement, the pronoun I. The pronoun unambiguously signals that it is the subject of the sentential complement. Ferreira and Dell (1996) hypothesized that if the subjects were to choose a sentence to avoid ambiguity and therefore facilitate the listener’s comprehension, they would overtly utter the complementizer that in (23a) more than (23b). In the experiment the subjects were given sentences with the complementizer half of the time and without it half of the time, in both types of sentences, and told to remember the sentence. After a while, the cue the coach told was given, and the subjects said the sentence aloud. Despite the presence of temporal ambiguity and the fact that the complementizer that aids comprehension in (23a), the subjects did not produce more sentences with that in (23a) than in (23b). In another experiment, Ferreira and Dell (1996) observed that speakers uttered a sentence in the form in which it is easier to say. They showed that when a word is easier to retrieve because it has already been mentioned in the sentence, speakers omitted an optional word before it in order to utter the word sooner. The results of both experiments indicate that speakers do not always choose a sentence that optimizes the listener’s comprehension.

Note that both the heavy NP shift and the deletion of the complementizer shown earlier do not involve a semantic change from
the other possible structure. Scrambling in Japanese is analogous to these English constructions in the sense that the construction itself does not change the semantic properties of canonical sentences. It is quite possible that scrambled sentences are also produced out of production-based motivations.

Although exact characterization of production-based motivations is beyond the scope of the current study, the current data suggests possible future directions. One possibility is that scrambled sentences are produced in order to improve efficiency in the speaker’s production. We have observed that heavy constituents tend to be scrambled. Shifting a heavy constituent, either to the sentence-initial position or the sentence-final position, has been argued to affect the working-memory associated with production because it usually enables the speaker to produce one clause at a time. Producing one clause at a time is assumed to be lighter in memory load than two clauses because when working with two clauses, the elements of one clause must be kept in a buffer while the second clause is processed (e.g., Hawkins 1994; Gibson 1998).

The other possibility is that the prominence of the semantic content of a constituent makes the speaker say it earlier in the sentence. It has been experimentally demonstrated that speakers choose a sentence which starts with a psychologically accessible constituent, that is, they tend to say the arguments that attract their attention first in a sentence (e.g., Levelt 1989; Osgood 1980). A scrambled constituent with a referential phrase is highly available in the speaker’s mind because it has been mentioned in the previous discourse or is inferable from it. Therefore the speaker may opt to start a sentence with such a phrase, which results in a scrambled sentence. Heavy constituents may attract the speaker’s attention by virtue of containing more words and therefore being semantically more complex than light ones.

When examining production-based motivations including working memory and the prominence of semantic content, it is important to find a principle that accounts for all cases of scrambling, including both heaviness and referentiality. Furthermore, in addition to text analysis, such a study should be corroborated by production experiments asking subjects to produce sentences while the experimenter systematically controls each factor.

7. Conclusions and suggestions for future study

This study has reported on the linguistic properties of scrambled sentences found in natural texts and discussed possible speaker/writer motivations for producing them. The examination of scrambled sentences at both the
sentence and the discourse level has revealed little indication that they are produced for comprehension-based reasons. It was suggested that they may be generated out of production-based considerations.

These observations question the function of word order in a language with greater flexibility in word order such as Japanese. It is certainly striking that scrambled sentences, which have a word order so explicitly different from the canonical order, demonstrate little contribution to comprehension; this finding presents quite a contrast to the function of topicalized sentences, which unarguably play an active role in discourse. However, considering the relative prominence of the structure of scrambling, it may not be surprising that scrambled sentences do not play a major role in discourse. Unlike topicalization in Japanese, scrambling does not involve any special morphological marking; scrambling simply involves constituents that are placed out of their canonical order. In Japanese such displacement of arguments alone may not be a cue strong enough to contribute to comprehension in discourse. Notice that the realization of the so-called canonical order in actual texts in Japanese, that is, the appearance of a sequence of constituents marked by *ga*, *ni*, and *o*, is not as strict as in English. Because Japanese has null pronouns: even canonically ordered sentences in Japanese can start out with constituents marked other than *ga*, as in (24), or appear to miss arguments, as shown in (25a–b).

(24) *Ringo-o tabeta.*
    apple-acc ate
    ‘[He] ate an apple.’

    John-nom Mary-dat gave
    ‘John gave [it] to Mary.’

b. *John-ga ringo-o ageta.*
    John-nom apple-acc gave
    ‘John gave an apple [to him/her/them].’

As shown, even many canonical sentences in Japanese either start out with constituents that are marked as being in other than nominative case, or appear out of order because of phonologically null arguments. Therefore scrambled sentences, which also start out with constituents marked as being in other than nominative case, or which place constituents out of order, may not serve as a strong cue for comprehension. This is in strong contrast to the case of inversions in English. In English, a language with relatively strict word order, an inversion or leftward
displacement of a constituent plays a major role in discourse (e.g., Birner 1994; Ward 1988). The contrast of the function of leftward movement in English and scrambling in Japanese may suggest that cues helping comprehension in a language vary depending on the linguistic nature of the particular language. One possibility is that the cues helping comprehension in a language vary according to the syntactic nature of each language. In a language with a strict word order, a change in word order can be highly productive as a cue in discourse and aid in comprehension. On the other hand, in a language with a fairly flexible word order, morphological marking may play a greater role in discourse than the word order. A further cross-linguistic examination that compares the functions of canonically ordered sentences, sentences with marked structure, and their interaction with morphology is necessary.

Appendix A: List of articles analyzed in the study

Casual conversational or semiconversational scripts

Multiple authors, ‘Roppongi in the sixties’. In Bungei-shunjuu, December 1994, 328–336.
— ‘How to live positively’. In MORE, December 1994, 201–207.
— ‘Skin and vitamin C’. In MORE, December 1994, 325.

Casual writing

Multiple authors, ‘Ijime’. In Petit Enfant, May 1994, 43–47.
Formal conversational script

Multiple authors, ‘Being a Prime Minister’. In Bungei-shunjuu, December 1994, 192–198.
—‘Yomiuri’s proposal for the constitution’. In Bungei-shunjuu, December 1994, 123–130.

Formal writing


Appendix B: Scrambled sentences with relevant preceding/subsequent context

(A scrambled phrase is indicated by underlining, and its original position is indicated by e. The preceding and subsequent context is given in English translation by the author.)

(1) (from ‘Roppongi in the sixties’; A is a producer, B an actress, C a singer; they are talking about the sixties when they were young)

A: ‘I must have been a high school student. Mariko [speaker B] must have been a young junior high student, rather than an actress.’

B: ‘I bet I was only fifteen.’

C: ‘Sorede ore, oboeterun kedo, and so I remember ext cop but Uti-no oyazi-to ohukuro-ga likurakata-mati-ni my old man-and mom-nom Iikurakata-town-at resutoran-o hazimeru n da -to restaurant-acc start ext cop-comp Syoo-tyan-ga e itta n da yo. Sho-chan-nom said ext cop emp ‘And so, I remember, ‘My old man and mom will start [running] a restaurant at Iikurakatamati,’ Sho-chan said.’ So we went there. And we found out that the day [we went] was the opening day.’

(2) (from ‘Roppongi in the sixties’; A is a producer, B a singer who comments in astonishment at A’s remark, and C the actress herself)

A: ‘I didn’t think that you [a female actress] were in such a turmoil at that time, because when I saw you at Ondiene [the name of a theatrical show], you told me, “I am the most ‘journalistic’ woman now”.’
B: ‘Sonna huree-zu-o yoku 22,3-sai-no onnanoko-ga e tukatta ne. Such an [arrogant] word, a girl who is just twenty-two or three uses!’

C: ‘I was young and ambitious.’

(3) (from ‘Children with no nationality’; A is speaking about the court ruling in which the Japanese nationality was granted to a child when the Japanese government could not show who the mother of the child is, B is the interviewer)

A: ‘... At first, the ruling was that unless the government can prove someone to be his mother little Andele is granted Japanese nationality. However, at the appeal court, they say, ‘If you want the [Japanese] nationality, prove by yourself that your mother is unknown, and we don’t care how difficult it is for you [to do it]’ to the three-year-old child.’

B: ‘That is ridiculous.’

(4) (from ‘Living in our thirties’; the speaker is a singer)

‘After a four-year break, I tried to keep my workload light. I was busy with my private life, and such a pace was just fine with me. When it comes to life, the question is what the highest priority is at each moment. I decide confidently-can’
“You can never do everything at once; if I think one thing is important, then I cut off other things [so that I may concentrate on one thing at a time].”
I am pretty good at prioritizing things. I may not look that way, but I am in control of my life pretty well.’

(5) (from ‘Living in our thirties’; the speaker is an actress being interviewed about her life and children)
‘I suppose that if I put my children in a private school, [because there are no entrance exams in junior high school and high school] my children could have had an easier life.
Demo, sonotudo zyuken -to iu haadoru-o norikoete iku but each time entrance exam called hurdle-acc overcome go hoo -ga, mi-no aru direction-nom fulfilling zinsee-ni naru -to watasi-wa e omou no.
life -becomes -comp I -top think ep
“I think life becomes more fulfilling if one overcomes a hurdle of an entrance exam each time [one enters school].”
So I talked with all four children and made them go through entrance exams each time. I think I did well and worked hard on this matter.’

(6) (from ‘Living in our thirties’; the speaker is a hair and makeup artist)
‘In your thirties, you must take care of yourself so that you stay an attractive woman. Fixing yourself up is one of the ways to make yourself attractive.
Zyosee-wa soko-ni iru dake-de hanayagu women-top there exit only-with make place lively sonzei da si, sore -wa syakaitekina existence cop and it -top social yakuwari na n da -to boku-wa e omoimasu. role cop ep cop -comp I -top think
“I think that women make surroundings lively with their mere presence, and it is the social role [assigned to] them.”
For example, if a child’s mother who is all fixed up shows up to the school open house, that alone makes the child happy and proud of his mother.’

(7) (from ‘Men of My Type’; the speaker is a model, talking about Robin Williams)
Akarukute atatakai kanzi dasi, cheerful and warm look and
sono atatakasa -o mawari -ni -mo e
that warmness -acc surrounding people -to -also
hurimaite kuresoo.
sprinkle looks like
‘[He] looks like he is cheerful and warm, and [he] looks likely to give such benevolence to people around him.’
He also looks like he is a positive thinker.’

(8) (from ‘How to live life positively’; the speaker is a writer being interviewed
‘My work [as a scenario writer] benefits from a variety of experiences in life. In that sense, I feel lucky that I went through many jobs, which gave me opportunities to see many things in the society.
Demone, dare-ni totte-mo, zinsee-ni oite mudana koto-wa
but no matter who-to life-in meaningless thing-top
nani hitottu-mo nai-to
there is none -comp
watasi-wa e omou kedo ne.
I-top think but emph
‘You know, no matter who you are, there is nothing in life that is meaningless to you, I think.’
Because when a person experiences something, he is sure to learn something from it. In that sense, nothing is a negative experience in one’s life.’

(9) (from ‘Ijime [Bullying]’, the speaker is a mother talking to another mother who has a little daughter, X-ko-chan, of the same age as her own daughter)
‘My daughter is very shy and cannot express her feelings well.
Itido X-ko-chan-to asobi-tai -tte, watasi-ni-wa e
once X ko-chan-with play-want-comp I -dat-top
itte ta no.
said emph
‘[She] told me that [she] wants to play with X-chan sometime.’
Can you and your daughter come to my house and visit over tea sometime?’

(10) (from ‘Ijime [Bullying]’; the speaker is a mother whose daughter is being bullied by another girl)
‘[Upon finding out that my daughter was being bullied by a girl] I decided to talk to another mother in the same kindergarten to observe how things are. But the situation seemed to remain the same.
Demo, uti-no ko-wa, K-ko chan ni izimerareteiru koto-o but my-gen child-top K-ko -by being bullied fact-acc itido-mo watasi-ni e hanasi-masen.

once even I -dat talk neg
‘My child would never tell me that she was being bullied by K-ko, not even once.’
I myself did not want to admit that my child was being bullied, so I pretended that I did not notice. But I am so worried that my daughter, who cannot express her emotional pain, is hurt inside very much.’

(11) (from Senji Kuroi, ‘Kenzaburo Oe and the Nobel Prize’; the speaker is a writer who grew up with Oe Kenzaburo, the Nobel Prize winner, talking about the impact of the Nobel Prize on Oe Kenzaburo’s reputation in overseas)
‘[The fact that China had not shown a strong interest in Oe Kenzaburo’s literature] had made me reconsider the significance which Japanese literature holds in the literature of the world.

Nooberusyoo-no zyusyoo-wa, Ooe Kenzaburoo-si-no Nobel Prize-gen receiving-top Oe Kenzaburoo-Mr.-gen bungaku-ga sono hukami-to work-nom its depth and hirosa -ni oite ikkoku- no iki-o koeru mono width in one country-gen range-acc exceed thing de aru koto-o sekai -ni e meezi-sita.
cop fact-acc world-dat showed
‘[Mr. Oe’s] receiving of the Nobel Prize clearly demonstrated to the world that Mr. Oe Kenzaburo’s literature exceeds, in its depth and width, the [standard of] one country.’
But much is left as to the question that how Mr. Oe’s literature has been read and how it will affect people’s minds.’

(12) (from ‘Kenzaburo Oe and the Nobel Prize’; the speaker is a writer who grew up with Oe Kenzaburo, the Nobel Prize winner, looking back at his childhood)
‘We thought that his name was Kensanroo Oe, instead of Kenzaburoo Oe. This is because one time, he himself [Kenzaburo Oe] told us so, with a serious face.
Sono syooko ni, -to kare-wa e tuzuketa.
its evidence cop-comp he-top continued
‘For the evidence [of it], he continued.’
My brother’s name is See‘yon’roo, instead of ‘Seesiroo’. At school, people call us brothers ‘Kenzaburoo’ and ‘Seesiroo’, but on the
governmental record, [we are] supposed to be [“Kensanroo” and “Seeyonroo”].

(13) (from ‘Kenzaburo Oe and the Nobel Prize’; the speaker is a writer who grew up with Oe Kenzaburo, the Nobel Prize winner, looking back at his childhood)

‘Hirosi, my cousin, and Kenzaburo Oe were classmates in the French class at the university. At that time, I had failed the college entrance exam twice, and I was still preparing for another. I admired them, and I felt an inferiority complex even at times when the three of us were having a good time. The relationship between them and me was a little awkward because they knew I was somewhat jealous of them, and yet they were very considerate of it and treated me very nicely.

However, I did not hate my position [lit. such position of myself] [among the three of us] that much.

Because I was not a college student [because I failed the entrance exam], I had no need to keep up with them.’

(14) (from ‘Kenzaburo Oe and the Nobel Prize’; the speaker is a writer who grew up with Oe Kenzaburo, the Nobel Prize winner, looking back at his childhood)

‘I don’t claim that we [I and Oe Kenzaburo] are the closest friends. He may have forgotten completely the story that I remember.’

But I wonder if he still remembers the story which starts as ‘I am a blue beard. I have blue hair all over my body.’ Hiroshi and I always laughed to death while we read what he wrote.’

(15) (from Sadaharu Oh, ‘Parting with Giants’; the speaker is a former coach of the Yomiuri Giants, a professional baseball team in Japan)

‘... I was very selfish in terms of my life with baseball. I never thought of taking my children out on weekends. I know some baseball players wake up early in the morning to see their children off to school and go back to sleep, but I never did anything that would negatively affect my training in baseball. I never consulted
my wife about my move in career [changing teams] either, because she said, ‘Do whatever you want’. I feel a little sorry for her, though, because my second daughter got married recently and she [my wife] must be a little lonely.

*At any rate, I want to show the players the joy of winning.*

I don’t want to compare them with the Giants players, but I cannot help it.’

(16) (from ‘Being a Prime Minister’; the speaker is former secretary to Prime Minister Fukuda, looking back on the old days)

‘Former Prime Minister Tanaka, whom I had known for a long time [but was the head of political group against Prime Minister Fukuda], said to me, ‘Hey, Nishimura, if things get sour, feel free to come to me’ in front of Mr. Fukuda.

But he [Mr. Tanaka] does not care about such things; he just says whatever he wants to say.’

(17) (from ‘Being a Prime Minister’; the speaker is a former journalist who specialized in Prime Ministers, who is looking back on the old days)

‘I remember Prime Minister Sato well … . We went on a picnic, spending all day, along with reporters. He seemed to enjoy eating lunch with people around him, including reporters.

I was told Fukuda-Mr.-somewhat he-top

Kakuee-with connected -pres -comp

omovareysainaika-to kotti-wa e hiyahiya sita.

could be judged -comp I -top worried

‘When I was told such a thing by him, I worried that Mr. Fukuda would wonder whether I am in a loop with Mr. Kakuee [Tanaka].’

But he [Mr. Tanaka] does not care about such things; he just says whatever he wants to say.’

‘I remember Prime Minister Sato well … . We went on a picnic, spending all day, along with reporters. He seemed to enjoy eating lunch with people around him, including reporters.

Zenbu syasin-acc take for us

Kaeruto kiri no ha no naikaku-no monsyoo-o

on returning paulownia-gen leave-gen Cabinet-gen crest -acc

Ouinsita karaa syasin-acc

stamped color picture-acc

Kisya hitorihitori-ni e purezento-site kureta.

reporter each -dat- presented for them
“He [the Prime Minister of Japan] took pictures of everyone, and when [we] returned, [he] gave pictures with the crest of the paulownia leaf, the Cabinet crest, to each reporter.”

Those pictures of former Prime Minister Sato and me remind me, even now, of the good old days.’

(18) (from ‘Yomiuri’s proposal for the Constitution’; A is a female lawyer, criticizing the newly proposed Japanese Constitution by a publisher, B is a law professor):

A: ‘This proposal leaves many points unclear. For example, ‘The Preservation of Peace’ in Chapter 3 is one of them.’

B: ‘This is indeed a problem.’

A: ‘Motomoto genkoo-no kenpoo-wa, originally currently enforced constitution-top Makkaasaa-sooan-o nihonzin-ga e tekitooni MacArthur note-acc Japanese-nom roughly yakusita dake no mono na no desu. translated only thing cop ext cop “Basically, the current Japanese constitution is nothing but a Japanese translation of the MacArthur note.”

Because the existing Constitution is a direct translation of English, there are many points whose meaning is not totally clear.’

(19) (from ‘Yomiuri’s proposal for the Constitution’; A is a law professor, criticizing the Japanese congressmen, B is a lawyer)

A: ‘It is unfair that congressmen limit the right to appeal to the Constitutional Court to only the Cabinet members and congressmen.’

B: ‘Kokkaigimntati-ga zibuntati-ni tugoo-no warui koto-o congressmen-nom oneselfs-dat inconvenient matter-acc kenpooasibansyo -ni e wazawaza moositatemasu ka. the Constitutional Court-dat bother to appeal Q “Do you think that the congressmen bother to appeal to the Constitutional Court [about] something that is inconvenient to themselves [if the matter were judged]?”’

A: ‘For instance, I believe that the laws of the protection of political parties are unconstitutional, but I bet no congressman would bring this issue to the Constitutional Court.’
Notes

1. The current study was supported by the University of Illinois Research Board. A portion of this article was presented at the 11th Annual CUNY conference on Human Sentence Processing, in New Brunswick, NJ, City University of New York, and University of Illinois. I thank Yasuko Suzuki for data collection and discussion. I also thank Yuki Hirose, J. J. Nakayama, Yoshi Kitagawa, Keiko Uehara, Janet Fodor, Di Bradley, and Jennifer Venditti for comments on an earlier version of the article. All shortcomings are, of course, mine.

2. Throughout, the paper uses the following notations:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom</td>
<td>nominative case marker</td>
</tr>
<tr>
<td>acc</td>
<td>accusative case marker</td>
</tr>
<tr>
<td>dat</td>
<td>dative case marker</td>
</tr>
<tr>
<td>gen</td>
<td>genitive case marker</td>
</tr>
<tr>
<td>top</td>
<td>topic marker</td>
</tr>
<tr>
<td>comp</td>
<td>complementizer</td>
</tr>
<tr>
<td>cop</td>
<td>copula</td>
</tr>
<tr>
<td>emph</td>
<td>emphasis marker</td>
</tr>
<tr>
<td>ext</td>
<td>extended predicate, which expresses the speaker’s empathetic emotion</td>
</tr>
<tr>
<td>neg</td>
<td>negative marker</td>
</tr>
<tr>
<td>Q</td>
<td>question marker</td>
</tr>
</tbody>
</table>

In all examples the scrambled constituent is underlined, and the position where the constituent would be observed in the canonical sentence is marked with $e$.

3. The current article employs the terminology of ‘scrambling’ without an assumption that such order is derived from the canonical order. Recent developments in syntactic analysis in the minimalist framework distinguish the categorical status of constituents and the derivations of various word orders in Japanese. For instance, the noncanonical order ga-o-ni may involve syntactic movement motivated by the focus, while those that are not may be base generated (Miyagawa 1997). While the findings in the present study are in general compatible with such analysis, it is not possible to examine all cases of noncanonical sentences found in the current texts in light of the focus movement because naturally occurring sentences do not always include elements which unambiguously signal the presence of focus.

4. The complete list of scrambled sentences observed in the current study is presented in Appendix B.

5. I define heaviness in terms of structural complexity. A constituent is heavy if it includes a subordinate clause such as a relative clause or a sentential complement. Because the heavy constituent includes a clause, it is also generally long, in the sense of the number of words included. The sentences which share this property are numbers (1), (3) to (6), (8), (9) to (11), (14) to (17), and (19) in Appendix B.

6. Throughout the article, the best effort to represent the subtle changes in interpretation introduced by scrambling is made in the English translations. However, this is merely an approximation. Moreover, sometimes scrambled sentences are translated in the same way as their canonical counterparts simply because there are no English equivalents which correctly represent the Japanese sentences.

7. Yamashita and Chang (2001) experimentally demonstrated that heavy accusative or dative marked arguments are significantly more likely to be scrambled. In particular, it observed a robust effect of weight on the internal scrambling of accusative-marked argument.
8. Sentences (2), (7), (12), (13), and (19) in Appendix B share this property.

9. In describing conversation, the outline of the preceding and subsequent context is given in English when relevant. The same speaker continues his or her turn regardless of the language shown, i.e., English or Japanese transcription, unless another speaker's utterance is shown to interrupt.

10. The canonical sentence sounds awkward partially because, for some unknown reason, two dative-marked constituents appear consecutively (Keiko Uehara, personal communication). The sentence improves somewhat by inserting an adverb as shown here.

(i)  (Canonical order)

\[ ?Kokkaigiinntati-ga kenpoosaibansyo -ni [hidoku zibuntati -ni tugoo-no congressmen-nom the Constitutional Court-dat terribly oneselfs-dat warui koto -o wazawaza moositatemasu ka. inconvenient matter -acc bother to appeal Q ‘Do you think that the congressmen bother to appeal to the Constitutional Court something that is terribly inconvenient [if the matter is judged] to themselves?’\]

The difficulty of accessing the antecedent, however, still remains, in comparison to the scrambled order.

(ii)  (Scrambled order)

\[ Kokkaigiinntati-ga [hidoku zibuntati -ni tugoo-no warui koto] -o congressmen-nom terribly oneselfs-dat inconvenient matter-acc kenpoosaibansyo -ni e wazawaza moositatemasu ka. the Constitutional Court-dat bother to appeal Q\]

11. This, of course, does not apply to the heavy NP shift in English, which, as shown earlier, does not seem to serve as a strong cue. Note that the heavy NP shift is different from the leftward movement of constituents shown in Ward (1988) in the sense that it occurs in a highly predictive environment associated with particular predicates. It also does not occur in a sentence-initial position, which seems to serve as the strong cue in the discourse.

References


Hiroko Yamashita received her Ph.D. in Japanese Linguistics from the Ohio State University (1994). She is an assistant professor at the Rochester Institute of Technology where she has been teaching since 2001. Her main research areas are the comprehension (processing) and production of Japanese and Korean, and the perception of Chinese characters, Kanas, and Hangul. Her major publications include ‘The effects of word-order and case marking information on the processing of Japanese’ (1997) and ‘The use of verb information in a pro-drop language’ (1995). Address for correspondence: Dept. of Language and Literature, 92 Lomb Memorial Drive, Rochester, NY 14623, USA <hyamash@mail.rit.edu>.