Given adult native speakers’ knowledge of grammar, two immediate questions arise in psycholinguistics. The first is how children acquire such knowledge. The second question is how adults use this knowledge when processing utterances in their native language. The present paper investigates the latter question in the word-by-word processing of scrambling in Japanese.∗

1 Introduction

When reading a sentence, people do not have access to all its words at once, instead they process each word as it is heard or read off the input string. However, people do not wait for the last word to come in to start understanding the meaning of the sentence (Marslen-Wilson & Tyler 1980, inter alia). Sentences are parsed incrementally as words are associated with each other immediately as soon as they are detected, and partial interpretations are built. For example, when presented with the fragment “Mary gave…” , an English speaker associates Mary as the subject of gave before the rest of the sentence becomes available.

In this paper, we investigate two types of questions in the processing of scrambling in Japanese. First, how is the grammar used when sentences are processed? Second, is the word-by-word parsing process the same for all human languages, or are language-specific parameterizations necessary?

1.1 The grammar and incremental processing

Even complex grammatical constraints such as subjacency have a local effect in sentence processing in the sense that readers (we will restrict the present discussion to reading) apply such constraints as early as possible rather than wait till the end of the sentence to filter out ungrammatical alternatives. For example, in the processing of fronted wh-phrases in English, it has been shown that readers try to create a gap for the

∗The experimental results reported in this paper were supported by Kanda University of International Studies and its Grant-in-Aid for COE Research (08CE1001) from the Japanese Ministry of Education, Science and Culture.
wh-phrase as soon as grammatically permissible. This strategy can lead to mistakes as illustrated by the filled-gap effect (FGE) in the following sentence (adapted from Crain & Fodor 1985).

(1) Who \text{ did the children force } \wedge \text{ us to sing the songs for } \text{ gap } \text{ yesterday?}

As soon as the word who is read it is clear that a corresponding gap has to be available in the input sentence, and readers attempt to insert such a gap at the earliest grammatical position available. For example, readers posit a gap at the position indicated with the wedge sign ‘\(\wedge\)’ in (1), so that the wh-phrase is initially interpreted as the direct object of force. When the next word us is read, it becomes apparent that the gap cannot be in this position and a FGE takes place as attested by slow reading times observed at this point compared to a control sentence (Crain & Fodor 1985; also Stowe 1986, for related results). Readers have to resume searching for the gap position, which is eventually found immediately after the preposition for. One important aspect of the FGE is that it indicates that readers are not just attaching each incoming word into the partial tree structure being built, but that they are also predicting how the sentence is going to continue by building portions of the tree structure (in the present case, by inserting the gap) without waiting for confirmation (i.e., without waiting to see if an overt direct object will follow).

Another related finding is that no FGE is observed if the potential position for the gap is inside an island (Stowe 1986), suggesting that readers take such island constraints in consideration in order to determine where to posit a gap.

In summary, the grammar actively determines the flow of the parsing process. Readers use grammatical knowledge (e.g. where a wh-gap can be inserted) not only to associate a word to the partial tree structure, but also to determine the possible continuations for the sentence fragment read.

### 1.2 Beyond the grammar: Parsing and parameterizations

Assuming that it is possible to separate the process of parsing sentences into two components — namely a grammar and a parser (i.e., an algorithm that uses a grammar to process sentences) — it is an open question whether the algorithm has to be parameterized for each language. Clearly, grammars are different for different languages in the sense that they vary in word order, they allow wh-phrases to be pronounced in different positions, and so forth. Thus, it is conceivable that a specific parser has to be available for each (type of) language. Alternatively, parameterizations may be restricted to the grammars and the parsing algorithm is such that it can handle the grammar of any natural language. Note that there is no straightforward way of showing that the parsers for different languages are the same. Thus, until evidence to the contrary becomes available, the working assumption should be that they are the same.

In the following, we will assume that this is in fact the case and that a single parsing algorithm handles languages as different as English and Japanese. Any potential differences observed in the processing of those languages will have to be explained based on the interaction of the respective grammars with the parsing algorithm. The goal of such a research project is to explain the processing of different constructions across various languages using the same type of processing mechanism.
We will argue that the processing of scrambled constituents in Japanese takes place in a manner similar to the processing of fronted wh-phrases in languages such as English. The proposal has the advantage of using mechanisms that have been previously investigated in English, therefore their properties are well-understood and they make clear predictions in the present case as long as we consider some relevant differences between fronted wh-phrases and scrambling.

2 The incremental processing of scrambling

We propose that Japanese readers insert the gap for a scrambled constituent as soon as grammatically possible. Consider how the word-by-word processing of the following fragment proceeds within our proposal.

(2) Ueitoresu-wa kokku-o
    waitress-top cook-acc

Up to this point, a Japanese reader can assume that this is the beginning of a transitive construction in which a transitive verb (e.g. saw) will follow. However, when a dative NP is read next, it is clear that the NP sequence is actually the beginning of a ditransitive clause in which the accusative NP precedes the dative NP. Assuming that the dative-accusative order is canonical (Hoji 1985), readers must posit a gap for the accusative NP, and the representation at this point would be as follows.

(3) Ueitoresu-wa kokku,-o rejigakari-ni (gap_i)
    waitress-top cook-acc cashier-dat

We assume that the verb need not be present in order for the gap to be posited, following an incremental processing model in which a partial representation of the sentence is created based on the case markers of the NPs (see Miyamoto 2002, and references therein).

Note that there is a crucial difference between the search for gaps in fronted wh constructions and in scrambling. In the former, as soon as a wh-phrase is detected, the reader knows that a gap is required. In scrambling, however, it is not always immediately clear that a constituent was scrambled. For example, in (3), it is not clear that the accusative NP was scrambled until the dative NP is detected. Therefore, in wh processing, there is the realization that a gap is necessary, and at a later point a gap is inserted. In scrambling, in contrast, the two processes occur simultaneously. The point in which it becomes apparent that a gap is necessary is often a position where a gap can be inserted. We will not explore the possible consequences of such a difference in this paper, but this is the type of question that can be explored within this proposal.

According to previous processing results, the creation of a gap should have a number of behavioral effects. Below we provide the background for two such predictions and then report the results of a self-paced reading experiment and a probe recognition experiment.
2.1 Measurable effects in the processing of gaps

Two types of effects have been observed in gap processing. First, the farther the gap is from its antecedent, the harder, and consequently the longer it takes to process the sentence (Gibson 1998; Just & Carpenter 1992).

Second, as is the case with pronouns, gaps have been shown to prime their antecedent, in other words, gaps facilitate the recollection of their antecedent (Bever & McElree 1988; MacDonald 1989). For example, consider the two English sentences below (from Bever & McElree 1988).

(4) a. Passive construction
   The *astute* lawyer, who faced the female judge was suspected ⟨gap⟩ constantly.

   b. Adjectival construction
   The *astute* lawyer who faced the female judge was suspicious constantly.

After reading each sentence, native speakers of English were faster to determine that the probe word *astute* had appeared in (4a) in comparison to its gap-less counterpart (4b). The assumption is that the gap primes the entire NP *the astute lawyer*, therefore even portions of it, such as the adjective *astute*, are more quickly recognized.

If the processing of scrambling in Japanese involves the insertion of a gap, it should present effects similar to the ones just discussed. First, a gap should take longer to be processed if it is farther from the scrambled constituent. Second, a probe should be recognized faster if the constituent it is a part of is scrambled rather than in-situ (cf. Nakayama 1995, who reports experimental results arguing against such a facilitation; but see Miyamoto & Takahashi 2002, for a different interpretation of Nakayama’s results). Note that the first prediction, which is tested in Experiment 1, implies longer reading times at the gap position, whereas the second prediction, tested in Experiment 2, involves faster recognition times of a probe presented after the sentence is read.

3 Experiment I — Self-paced reading

In this experiment, we measured reading times per word in order to show that longer latencies are observed if the gap is farther from its antecedent, namely the scrambled constituent (see Miyamoto & Takahashi 2002, for an overview of previous reading time results and discussion of various factors that need to be considered in this type of experiment).

3.1 Method

3.1.1 Participants

Thirty-two native speakers of Japanese, undergraduates at Kanda University of International Studies in the Kanto area of Japan, participated in the study.
3.1.2 Materials

Twenty sets with four sentences each were constructed (see (5) for an example set).

(5) a. Scrambled/Adjacent
   Ueitoreisu-wa doogu-ga okareteiru sooko-de kokku-o rejigakari-ni (gap) shookaishita sooda.
   cashir-dat introduced seems

b. Scrambled/Far
   Ueitoreisu-wa kokku-o doogu-ga okareteiru sooko-de
   cashir-dat introduced seems
   rejigakari-ni (gap) shookaishita sooda.
   waitress-top cook-acc tools-nom stored storage-loc

   rejigakari-ni (gap) shookaishita sooda.
   cook-acc introduced seems

   ‘The waitress seems to have introduced the cook to the cashier in the storage room where the tools are stored.’

In the first two conditions (the two Scrambled conditions), the accusative NP cook was scrambled, whereas in the last two conditions (the Canonical conditions) it remains in its canonical position immediately before the verb. In the Adjacent conditions (5a,c), the two object NPs (cashier-dat and cook-acc) are adjacent to each other; in contrast, in the Far conditions (5b,d), an adjunct phrase (a locative or a temporal) intervenes between the two object NPs.

The region preceding the verb introduced is the crucial region for comparison. In both Scrambled conditions, a gap has to be posited for the scrambled accusative NP as soon as cashier-dat is read. However, because the distance between the gap being created and its antecedent is greater in the Scrambled/Far condition, we predict that participants will take longer to create the gap when reading cashier-dat in this condition in comparison to the Scrambled/Adjacent condition. In contrast, no such a difference should be observed at the equivalent position, cook-acc, in the two Canonical conditions.

Four lists were created by distributing the test stimuli in a Latin Square design. Each participant saw exactly one of the lists intermixed with 50 unrelated foil items in pseudo-random order so that at least one foil item intervened between two test items.
3.1.3 Procedure

The experiment was conducted on a Power Macintosh running PsyScope (Cohen, MacWhinney, Flatt, & Provost 1993) with a button-box. Participants were timed in a phrase-by-phrase self-paced non-cumulative moving-window reading task (Just, Carpenter & Woolley 1982). Each phrase included a content word and a functional particle such as a case marker, a postposition, a complementizer. The segmentation in the sentences in (5) (indicated with intervening spaces) was the actual segmentation used in the self-paced reading presentation. Sentences were shown using Japanese characters with the uniform-width font Osaka Toohaba 14. Stimuli initially appeared as dots with intervening spaces indicating the segments, and participants pressed the leftmost button of the button-box to reveal each subsequent region of the sentence and cause the present region to revert to dots. At the end of each sentence, a yes/no question appeared on a new screen; participants answered by pressing one of the two rightmost buttons of the button-box, and then auditory feedback was provided. Data points were eliminated from the reading time analyses if the participant did not answer the corresponding 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 15 minutes.

3.2 Results

3.2.1 Response accuracy

The average correct responses to the comprehension questions did not differ for the four conditions (Scrambled/Adjacent, 87.5%; Scrambled/Far, 88.7%; Canonical/Adjacent, 90.6%; and Canonical/Far, 91.9%; F's < 1).

3.2.2 Reading times

The analyses of the reading time results for the critical region (the object NP preceding the main verb) were as follows. There was an interaction between position of the adjunct and word order (F1(1,31) = 7.95, P < 0.01; F2(1,19) = 5.03, P < 0.05). The Canonical/Far (787 ms) and the Canonical/Adjacent (842 ms) conditions did not differ (F1(1,31) = 1.54, P = 0.22; F2(1,19) = 1.16, P = 0.29); whereas the Scrambled/Far condition (903 ms) was slower than the Scrambled/Adjacent condition (763 ms; F1(1,31) = 5.84, P < 0.05; F2(1,19) = 5.04, P < 0.05).

There was an unexpected slow-down in the region preceding the critical region. The Canonical/Far condition (1055 ms) was slower than each of the other three conditions (Canonical/Adjacent 818 ms; Scrambled/Far 863 ms; Scrambled/Adjacent 847 ms; all Ps < 0.05); whereas the other three conditions did not differ from each other (all F's < 1).
3.3 Discussion

As predicted, the Scrambled/Far condition was slower than the Scrambled/Adjacent condition in the critical region. This is the region that makes clear that the accusative NP *cook* was scrambled, and a gap is posited for this NP. Because of the intervening adjunct, the distance between the gap and its antecedent is greater in the Scrambled/Far condition, and consequently we observe longer reading times for this condition. The Canonical conditions, in contrast, do no differ because there is no gap being posited at this point and, therefore the position of the adjunct phrase is irrelevant.

The long reading times at the adjunct head (*storage-loc*) in the Canonical/Far condition had not been predicted. There are two explanations that can be considered at this point, but we will leave further investigation for future research. One possibility is that the slow-down at this point indicates that there is a gap being created for the dative NP *cashier*. In other words, the order in which the adjunct precedes the dative NP is canonical. A second possible explanation is that readers were expecting an accusative NP to follow the dative NP (given that most verbs used in the experiment were ditransitives), and they were surprised when the adjunct head was detected instead.

4 Experiment II — Probe recognition

In this experiment, we measured the reaction time to recognize a probe word in order to verify if there is any facilitation after a scrambled sentence (see Miyamoto & Takahashi 2002, for further discussion and related results).

4.1 Method

4.1.1 Participants

Thirty-five students and administrative staff from Kanda University of International Studies participated in the experiment. None of them had taken part in the previous experiment.

4.1.2 Materials

Twenty sets of items with two conditions each were constructed based on items used by Nakayama (1995).1 The set in (6) is an example. The two conditions have exactly the same content words in the same linear order. The only difference is in the position of the accusative and nominative case markers in italics.

---

1We would like to thank Mineharu Nakayama for providing the original items of his experiment and for helpful discussion.
(6)  a. Canonical condition
Gakkoo-de mondai-o dashita kooshi-ga mukuchina
school-loc question-acc asked lecturer-nom quiet
gakusei-o mita.
student-acc saw
‘The lecturer who asked the question at school saw the quiet student.’

b. Scrambled condition
[Gakkoo-de mondai-o dashita kooshi]-o mukuchina
school-loc question-acc asked lecturer-acc quiet
gakusei-ga \( \langle \text{gap} \rangle \) mita.
student-nom saw
‘The quiet student saw the lecturer who asked the question at school.’

The sentences differ in their meanings but there does not seem be an *a priori* advantage of one meaning over the other in relation to probe recognition (in this respect, the present design is similar to that of Bever & McElree 1988, given that the agent and patient in their sentences in (4) are different as well).

The second word in each condition (*question* in the sentences above) was used in the probe recognition task. Two lists were created by distributing the test stimuli in a Latin Square design. Each participant saw exactly one of the lists intermixed with 40 unrelated foil items in pseudo-random order. Answers for the probe recognition task and for the comprehension question were counterbalanced for each list.

### 4.1.3 Procedure

The self-paced reading setup was the same as the one used for Experiment 1. Moreover, immediately after the last region of the sentence was read, a probe word surrounded by underscores was presented on a new screen and participants had to decide whether it had appeared in the sentence by pressing one of the two rightmost buttons of the button-box. Next, participants answered a yes/no comprehension question which appeared on a new screen. Auditory feedback was provided for the comprehension question only. Data points for a sentence were included in the reading time analyses and in the probe recognition latency analyses only if the participant gave correct answers for the probe recognition and the comprehension tasks.

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 15 minutes.

### 4.2 Results

#### 4.2.1 Response accuracy and reading times

Probe recognition accuracy did not differ between the canonical condition (93.1%) and the scrambled condition (95.1%; \( F_s < 1.6 \)) Response accuracy to the comprehension questions did not differ either (canonical, 78.2%; scrambled, 79.1; \( F_s < 1 \)).

The reading time analyses per region revealed no reliable differences.
4.2.2 Probe recognition

Participants were faster to recognize the probe after reading sentences in the scrambled condition (888 ms) than in the canonical condition (931 ms; $F_1(1,34) = 4.72, P < 0.05; F_2(1,19) = 12.51, P < 0.01$).

4.3 Discussion

The faster reaction time to recognize the probe after the scrambled condition suggests that the NP containing the probe was primed at the gap. Hence, the result supports the proposal that a gap is being created for the scrambled constituent (see Miyamoto & Takahashi 2002, for other related results and a discussion of Nakayama 1995, who did not detect reactivation in scrambled sentences; see also Nakano, Felser, & Clahsen 2002, for reactivation in long distance scrambling as measured with a cross-modal priming task).

5 General discussion

At first sight it may seem hopeless to try to use mechanisms proposed for English to process Japanese because of the structural differences between the two languages. Word order would seem a good case in point. Because word order is relatively rigid in English, it does not require the kind of processing mechanism that seems necessary for scrambling. Thus, one could have concluded that the parser of Japanese must be different from that of English since it would require properties absent in the latter. The present research suggests that this is fact not the case, and that it is possible to handle both types of languages with a single parser as long as we are able to discern the crucial aspects of the phenomena at hand. Even though ordinary word order in English does not seem to provide a clue as to how scrambling is processed, there are nevertheless constituents (namely wh-phrases) that are pronounced far from their in-situ positions in this language. This is the commonality between fronted wh-phrases and scrambled constituents that allows us to propose a single mechanism to process them both.

The present results provide two types of evidence supporting the parallel advocated between the processing of fronted wh-phrases and the processing of scrambling. In both constructions, readers attempt to insert a gap for a prior constituent (a wh-phrase or a scrambled constituent) as early as grammatically permissible. The proposal allows us to explain the processing of scrambling based on a mechanism that has been characterized independently. Moreover, because the two constructions are not identical (e.g., wh-phrases trigger the search for a gap immediately, whereas in scrambling the search process is delayed), it should be possible to investigate more detailed aspects of the processing mechanism involved and how it is affected by idiosyncratic properties of the two constructions. Note that in this case the goal is not to look for differences between the two languages in order to propose a parameterization, but rather the intent is to look for differences between two types of constructions and investigate the kind of mechanism that can parse them both. In this sense, the fact that the constructions may come from different languages is irrelevant for the most part because we are not
interested in characterizing the parser for a given language as opposed to another (type of) language. But rather, we want to determine the types of requirements that sentences with a given type of configuration demand from the parser. Whether the configuration in question does not occur in a certain language is not particularly relevant. However, if a construction with the crucial properties is available in a language, then we should predict that the same processing principles apply.

For example, a language X in which all constituents are always pronounced in-situ is not particularly helpful in investigating the processing mechanism assumed for fronted wh-phrases and scrambling in this paper. But we do not want to claim as a consequence that the processing of X is different from the processing of English and Japanese. It is just that it so happens that X does not have the crucial constructions in which the relevant property of the parser can be studied. In contrast, given a language Y with a construction C in which a constituent is pronounced far from its in-situ position, we make the prediction that C is processed in the same way as fronted wh-phrases in English and scrambling in Japanese.

This immediately raises questions as to whether long-distance scrambling (i.e., cases in which a constituent is scrambled to a position outside its original clause) can be handled by the present proposal. In the sentences discussed in this paper, scrambling was VP internal (in Experiment 1) or to the periphery of the clause (in Experiment 2). In principle, we assume that all types of scrambling should be processed in the same way, whether this is the case is an empirical question that is just beginning to be investigated (see Nakano, Felser, & Clahsen 2002, for some relevant data).
References


Miyamoto
Nara Institute of Science and Technology
Graduate School of Information Science
8916-5 Takayama, Ikoma, Nara 630-0101, JAPAN
http://cl.aist-nara.ac.jp/~etm
[etm@is.aist-nara.ac.jp]

Takahashi
Massachusetts Institute of Technology
Department of Linguistics and Philosophy E39-245
77 Massachusetts Avenue Cambridge, MA 02139, USA
[s_t@mit.edu]
### Contents

1 Introduction ........................................ 1  
   1.1 The grammar and incremental processing ............ 1  
   1.2 Beyond the grammar: Parsing and parameterizations ... 2  

2 The incremental processing of scrambling ............. 3  
   2.1 Measurable effects in the processing of gaps ......... 4  

3 Experiment I — Self-paced reading ....................... 4  
   3.1 Method ............................................ 4  
      3.1.1 Participants .................................... 4  
      3.1.2 Materials .................................... 5  
      3.1.3 Procedure .................................... 6  
   3.2 Results ............................................ 6  
      3.2.1 Response accuracy ............................ 6  
      3.2.2 Reading times ................................ 6  
   3.3 Discussion ......................................... 7  

4 Experiment II — Probe recognition ....................... 7  
   4.1 Method ............................................ 7  
      4.1.1 Participants .................................... 7  
      4.1.2 Materials .................................... 7  
      4.1.3 Procedure .................................... 8  
   4.2 Results ............................................ 8  
      4.2.1 Response accuracy and reading times .......... 8  
      4.2.2 Probe recognition ............................ 9  
   4.3 Discussion ......................................... 9  

5 General discussion ..................................... 9