Incremental Processing of Gap-filler Dependencies: Evidence from the Processing of Subject and Object Clefts in Japanese

Barış Kahraman[1], Atsushi Sato[2], Hajime Ono[3], and Hiromu Sakai[2]
ÇANAKKALE ONSEKİZ MART UNIVERSITY[1], HIROSHIMA UNIVERSITY[2], and KINKI UNIVERSITY[3]

1. Introduction

In the field of sentence processing, there is a general consensus on the *incrementality* of sentence processing. The term incrementality is widely defined as immediate assignment of thematic roles before they are signaled unambiguously (e.g., Kamide et al. 2003). Following this definition, the present study regards incrementality as predicting and building a representation for upcoming elements of the sentence or clause, even before the exact input is encountered (Kamide 2008). However, it is still unknown to what extent the sentences are processed incrementally. Particularly, the processing of gap-filler dependencies in head-final languages casts serious doubt on the issue of incrementality. The present study uses clefts sentences in Japanese to deal with the issue of incrementality of gap-filler dependencies.

In the sentence processing terms, a filler refers to the moved argument in a

---

sentence, and a gap refers to the original position of the filler (e.g., Fodor 1989).

(1) Who do you like ____.

In example (1), who is a filler, and its original object position is a gap. The filler and the gap are found in various structures such as clefts, relative clauses (RCs), scrambled sentences, and WH-questions (e.g., Aoshima et al. 2004). It is assumed that in order to understand the meaning of these kinds of structures, the parser needs to link the filler and the gap (Fodor 1989).

Previous studies have shown that when the filler precedes the gap, the parser immediately postulates a gap and starts a forward search for that gap, as soon as the filler is identified (e.g., Aoshima et al. 2004, Stowe 1986). In other words, the filler-gap dependencies are established incrementally. However, in the case of structures like RCs in head-final languages, the gap precedes the filler.

(2a) SR: ____ Giin-o hihanshita kisha-i-wa wakakatta senator-ACC criticized reporter-TOP young

“The reporter who criticized the senator was young.”

(2b) OR: Giin-ga ____ hihanshita kisha-i-wa wakakatta senator-NOM criticized reporter-TOP young

“The reporter who the senator criticized was young.”

In example (2), kisha (reporter) is the filler, and the line shows the gap position. Only a few studies have attempted to examine the incrementality of gap-filler dependency processing (Chinese: Lin 2006, Korean: Kwon 2008, Turkish:
Kahraman et al. 2010). These studies have focused on the processing asymmetry between subject and object RCs (SRs / ORs respectively). It is well known that SRs are easier to process than ORs in many languages. It is assumed that this processing asymmetry reflects the relative ease / difficulty of establishing filler-gap dependencies (e.g., Gibson 1998, O’Grady 1997).²

In head-final languages, it is conceivable that the position where the processing asymmetry is observed between SRs and ORs indicates that the gap-filler dependencies are formed at that position (Kahraman et al. 2010, Lin 2006). However, the processing asymmetry between SRs and ORs was observed in different positions among head-final languages. Thus, results are not conclusive, and it remains unsettled whether the parser postulates a filler as soon as the gap-site is identified, and starts to construct a dependency prior to the filler’s appearance.

For example, in Chinese, Lin (2006) reported that SRs were read faster than ORs, and the processing asymmetry was observed at the RC-head (filler). Lin argued that the process of establishing a gap-filler dependency started when the filler is encountered. Therefore, it cannot be said that the gap-filler dependency was processed incrementally in Chinese. In Turkish, Kahraman et al. (2010) reported that SRs were easier to process than ORs, and the processing asymmetry was observed at the RC-verb before the filler was encountered. They argued that the Turkish parser starts to construct the gap-filler dependency before the filler appears, and hence the processing of gap-filler dependencies is incremental in Turkish. Similarly, Kwon (2008) showed that processing asymmetry between SRs and ORs was observed both at the RC-verb and the RC-head in Korean. The processing asymmetry observed at the RC-verb position indicates that the processing of gap-filler dependencies starts before a
filler appears in Korean as well. In Japanese, on the other hand, the processing asymmetry between SRs and ORs was always observed at the RC-head as in the case of Chinese (see (2)) (e.g., Ishizuka 2005, Ueno & Garnsey 2008). However, in Japanese, neither relative pronouns nor RC-markers are used. Therefore, RCs are structurally ambiguous between a matrix clause and a subordinate clause at the embedded verb, and the upcoming filler is not signaled at all until the filler appears. Thus, from the existing data, it is very hard to determine whether the parser starts constructing the gap-filler dependencies before the filler is encountered in Japanese. In order to provide convincing evidence from Japanese, we need to look into other constructions, in which the structural ambiguity is resolved before the filler appears; hence the existence of the filler is signaled. As we explain in the next section, cleft sentences would provide a good test case.

In the present study, using subject and object clefts (SCs / OCs respectively) in Japanese, we examined whether the Japanese parser incrementally starts to construct the gap-filler dependency before the filler is encountered. The results of the self-paced reading experiment showed that OCs were read faster than SCs, and the processing asymmetry was observed only at the embedded verb position. The results indicated that the Japanese parser starts to construct the gap-filler dependency incrementally in cleft sentences.

2. Cleft Sentences in Japanese

Unlike RCs, the main function of clefts is focus (Mihara & Hiraiwa 2006). However, they share many syntactic characteristics with RCs. Particularly their word orders are basically identical in Japanese. As seen in example (3), *no*-*wa* is attached to the embedded verb in clefts. The particle *no* is taken as a complementizer, and *wa* is a topic marker (Hiraiwa & Ishihara 2002). In the
present study, we assume that the complex of the particle *no-wa* is a reliable cleft marker for the parser.

(3a) SC:  ____ i Giin-o hihanshita-no-wa kisha-i -da.
  senator-ACC criticized-C-TOP reporter-COP
  “It was the reporter who criticized the senator.”

(3b) OC:  Giin-ga  ____ i hihanshita-no-wa kisha-i -da.
  senator-NOM criticized-C-TOP reporter-COP
  “It was the reporter who the senator criticized.”

Due to the use of *no-wa*, the Japanese parser should be able to resolve the structural ambiguity at the embedded verb, and predict the existence of upcoming filler. If the parser immediately postulates a filler and incrementally starts a forward search for it from the embedded verb position, the processing asymmetry between SCs and OCs would be observed at that position. Previous studies in head-final languages argued that, the depth of embedding of a gap (i.e. *structural distance* (O’Grady 1997) can account for the processing asymmetry between SRs and ORs. If the depth of embedding of the gap is the decisive factor for the processing difficulty of gap-filler dependencies, SCs should be processed more easily than OCs, because the gap is more deeply embedded in OCs than SCs. To the best of our knowledge, no study has reported how Japanese clefts are processed in real time. In order to test our predictions, we conducted two self-paced reading experiments.

3. Experiments

3.1 Experiment 1

We first wanted to make sure that our experimental materials are reliable. For
this purpose we prepared RCs, and compared the processing of SRs and ORs.

**Method**

**Experimental material:** We prepared 30 sets of SCs and OCs as target sentences, and 66 filler sentences. The semantic plausibility between the agent and the theme was confirmed in a norming study.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR: Kyonen sobo ga inaka de kahoosita tooi shinsek-i wa jiko de nakunatta last year grandma NOM village LOC nursed distant relative TOP accident LOC died “The distant relative who the grandma last year died in an accident”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR: Kyonen sobo ga inaka de kahoosita tooi shinsek-i wa jiko de nakunatta last year grandma ACC village LOC nursed distant relative TOP accident LOC died “The distant relative who the grandma last year died in an accident”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Experimental sentences are identical except for the second region. In the SR condition, an accusative NP, and in the OR condition, a nominative NP comes after a time adverb. A locative adverb then follows these nouns. We used these adverbs in the target sentences because they may appear at the focus position in Japanese clefts, and we wanted to eliminate a possible expectation for adverbs at the focus position. Regarding the canonical order of locative adverbs, Tamaoka et al. (2004) argued that native speakers of Japanese prefer to place locative adverbs immediately after the nominative NP. We thus inserted a locative adverb after nominative NP in the OC condition. In the SC condition, for the sake of the unification, a locative adverb was inserted after the accusative NP as well. Our prediction for the results is as follows. If our experimental material is reliable, SRs should be read faster than ORs, and the processing asymmetry should be observed at the RC-head (e.g., Ishizuka 2005).
**Participants and procedure:** Thirty-six undergraduate or graduate students at Hiroshima University participated in Experiment 1. They were all native speakers of Japanese. The experimental materials were divided into 2 lists by a Latin Square design, and each participant was assigned to either one of the lists. Experimental sentences were presented on the computer screen in a random order, using Linger 2.94 (by Douglas Rohde) as a word-by-word, non-cumulative self-paced reading task. After reading a sentence, participants answered a yes-no comprehension question. Before the experiment, the participants were given a practice set of 6 sentences, and encouraged to read sentences as naturally as possible.

**Results and Discussion:** Statistical analyses for reading times were conducted only on sentences wherein comprehension questions were answered correctly. The accuracy rate for the comprehension questions was 90% for SRs and 87% for ORs. This difference was significant by only item analysis \([F_1 (1,35) = 0.84, p = .37; F_2 (1,29) = 4.49, p < .04]\), indicating that ORs were slightly harder to comprehend than SRs. Accuracy rates for both participants and items were much higher than the chance level. Therefore we included all participants and items in the statistical analysis, but we discarded the reading time data shorter than 250 milliseconds (ms) or longer than 3 SD point. This trimming process affected 2.3% of the reading time data. Mean reading times for each region in two conditions are shown in Fig. 1. The main region (word) of interest is the RC-head (region 6).

The results of ANOVA for repeated measures showed that reading time difference was significant by subject analysis at the RC-head (region 6) \([F_1 (1,35) = 4.85, p < .03; F_2 (1,29) = 3.84, p = .06]\). In other regions, there was no significant difference \((F's < 2.6)\). This demonstrates that the head noun of SRs
was read faster than that of ORs.

![Fig.1 Mean reading times for SRs and ORs](image)

In Experiment 1, we showed that SRs are easier to process than ORs, and the processing asymmetry was observed at the RC-head (filler). This result replicates the previous studies in Japanese (e.g., Ishizuka 2005, Ueno & Garnsey 2008). The results showed that our experimental material is reliable to test the processing of cleft sentences. Therefore, we converted all RCs into cleft sentences, and compared to reading times of SCs and OCs in Experiment 2.

3.2 Experiment 2

The aim was to test whether there is a processing asymmetry between SCs and OCs in Japanese. If so, where is the processing asymmetry observed?

**Method**

**Experimental material**: We used 30 sets of SCs and OCs which are converted from RCs in Experiment 1, with 80 filler sentences. These sentences were divided into two lists by a Latin Square design. Our predictions are as follows. The gap is always more deeply embedded in OCs than SCs. If the depth of embedding of a gap is the decisive factor for the difficulty of forming gap-filler dependencies, SCs should be easier to process than OCs. Moreover, if the Japanese parser starts constructing the gap-filler filler dependency before the
filler is encountered, the processing asymmetry would be observed at the embedded verb position, where the structural ambiguity is resolved.

### Table 2. Target sentences in Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SC:</strong></td>
<td>Kyonen sobo-o inaka-de kaihoosi-nowa- toi shiniski-dato haha-ga iiita Last year grandma-ACC village-LOC nursed-NOWA distant relative-COPC mother-NOM Said</td>
<td>“My mother said that it was the distant relative who nursed the grandma last year.”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OC:</strong></td>
<td>Kyonen sobo-ga inaka-de kaihoosi-nowa- toi shiniski-dato haha-ga iiita Last year grandma-NOM village-LOC nursed-NOWA distant relative-COPC mother-NOM Said</td>
<td>“My mother said that it was the distant relative who the grandma nursed last year.”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Participants and procedure:** Twenty-six undergraduate or graduate students at Hiroshima University were participated in Experiment 2. They were all native speakers of Japanese, and none of them participated in Experiment 1 or the norming study. The procedure was identical to Experiment 1.

**Results and Discussion:** The accuracy rate of the comprehension questions was 84% for SCs and 85% for OCs. This difference was not significant. Statistical analyses for reading times were conducted only on sentences wherein comprehension questions were answered correctly. Accuracy rates for both participants and items were much higher than the chance level. Therefore we included all participants and items in the statistical analysis, but we discarded the reading time data shorter than 250 milliseconds (ms) or longer than 3 SD point. This trimming process affected 3.9% of the data. Mean reading times for each region are shown in Fig. 2. The main region of our interest is the embedded verb (region 4) and the focus position where the filler appears (region 6).

The results of ANOVA for repeated measures showed that reading times difference was significant at the embedded verb position (region 4) \([F_1 (1,25) = 9.76, p < .005; F_2 (1,29) = 5.12, p < .03]\). In other regions there was no
significant difference ($F_s < 2$). This demonstrates that the embedded verb in the OC condition was read faster than the SC condition.

![Fig.2 Mean reading times for SCs and OCs](image)

The results show that OCs are easier to process than SCs in Japanese. This result is inconsistent with previous RC processing studies in Japanese (e.g., Ishizuka 2005, Ueno & Garnsey 2008). Therefore we cannot explain this result by the depth of embedding of a gap (O’Grady 1997). On the other hand, the fact that the processing asymmetry was observed only at the embedded verb position is consistent with our prediction regarding the incrementality of gap-filler dependencies. In the next section, we will discuss the results of two experiments, considering the position where the processing asymmetry was observed, and the processing ease of clefts and RCs.

4. General Discussion and Conclusions

In Japanese, RCs and clefts were processed differently in two respects. First, the two constructions differed with respect to the position where the processing asymmetry was observed. In the case of RCs the processing asymmetry was observed only at the RC-head, whereas it was the embedded verb in clefts.
Previous studies argued that the processing asymmetry reflects the ease / difficulty of a dependency formation between the filler and the gap (Gibson 1998, O’Grady 1997). Our results showed that the gap-filler dependency was easier to establish in SRs than ORs, and this process took place after the filler is encountered. In clefts, on the other hand, it was easier to establish the gap-filler dependency in OCs than SCs, and the dependency formation started at the embedded verb position. Since the Japanese parser starts constructing the dependency before the filler is encountered, it can be said that the gap-filler dependencies are processed incrementally in Japanese clefts as in the case of Korean and Turkish RCs (Kahraman et al. 2010, Kwon 2008).

In Turkish, Kahraman et al. (2010) pointed out that RC-markers which are attached to the embedded verb play an important role on the incremental processing of gap-filler dependencies. They argued that the Turkish parser immediately uses information from RC-markers to resolve the structural ambiguity and make predictions for the upcoming constituents. In the present study, the only difference between RCs and clefts was the use of cleft marker no-wa. Due to the use of no-wa, the processing of clefts considerably differed from RCs. This suggests that the Japanese parser also immediately utilizes information from the cleft marker, and predicts the upcoming filler. Therefore, the cleft marker in Japanese plays an important role, as RC-markers in Turkish does, to process sentences.

The second difference between the RCs and clefts was the processing ease. As we pointed out above, SRs are easier to process than ORs. This result is in line with previous studies in various languages and Japanese. In the case of clefts, OCs are easier to process than SCs. This result is different from studies in English clefts (e.g., Gordon et al. 2001), and studies in Japanese RCs (e.g.,
Ishizuka 2005). This difference suggests that the processing ease of RCs and clefts in Japanese cannot be explained by the same structural factors such as the depth of embedding of a gap (O’Grady 1997) or linear distance (Gibson 1998) at the same time. Why then cleft sentences are processed differently from RCs in Japanese? Regarding this question, we would like to point out three possibilities.

In English, Reali & Christiansen (2007) showed that the frequency of RC types is reflected in their processing ease. In Japanese, the distribution of RCs and clefts might be different. In other words, the frequency of SRs might be higher than ORs, and OCs might be higher than SCs. If this kind of distributional information is crucial in sentence processing, the processing ease of RCs and clefts might turn out to be different. As a second possibility, the expectation for upcoming constituents at the embedded verb position might be different in RCs and clefts (e.g. Hale 2006). For example, the choice of the element at the focus position in OCs is quite narrow, compared with SCs. In OCs, it is quite likely that the object NP will appear at the focus position. On the other hand, in SCs, not only the subject NP but also other elements such as -kara, -tame, (‘that is why...’) might have also been expected, and the calculation of those possible constituents might have caused longer reading time in SCs at the embedded verb position. In the case of RCs, the expectation for the upcoming constituents might be same for SRs and ORs, and due to the depth of embedding of a gap, SRs would have been processed easily than ORs at the head-noun. Third possibility is the discourse function of RCs and clefts (e.g., Roland 2009). Roland pointed out that the discourse functions of SRs and ORs are different. In Japanese, SCs and OCs might be used for different purposes in the discourse, and this difference might have somehow affected our results. However, at this moment we cannot distinguish among these possibilities. We leave this issue for
future studies.

Overall, our study suggests that, the processing of gap-filler dependencies is incremental in Japanese clefts. Moreover, the cleft marker plays an important role on the incremental processing of gap-filler dependencies as in Turkish RCs.

Notes
*This research was supported by (1) Grant-in-Aid for Scientific Research (B) “Neurocognitive basis for language learning through the processing of input and output (PI: Hiromu Sakai, #20320060)” by JSPS; (2) Grant-in-Aid for Scientific Research (B) (PI: Kentaro Nakatani, #21320083) by JSPS; (3) Grant-in Aid for Young Scientists from the MEXT (PI: Hajime Ono, # 21720152). We would like to thank Tsutomo Sakamoto for sharing his RCs with us, Rosalyn Chiu, Marshall Higa for their help, Zhong Chen, Kazumitsu Chujo, Yukiko Hatasa, Ruiko Ohama, Hiroyuki Shirakawa and Mikihiro Tanaka for their constructive comments. All remaining errors and shortcomings are of course our own.

1We make a distinction between “filler-gap dependencies” and “gap-filler dependencies”. The former is used for the cases when the filler precedes the gap, and the latter is used when the gap precedes the filler.

2 Many studies showed that factors such as animacy, frequency, predictability have an impact on the processing of RCs. Nevertheless, this does not mean that filler-gap or gap-filler dependencies are unrelated to processing of RCs. We assume that these factors affect the ease of establishing a dependency.

References


(heiwaeiyu@gmail.com)