Verb Subcategorization Information
During Sentence Comprehension by Japanese EFL Learners:
Evidence From Self-Paced Sentence Anomaly Task

Chieko NAGAI
Graduate School, Kobe University
Satoshi YABUUCHI
Kyoto Seika University
Ken-ichi HASHIMOTO
Kyoto University of Foreign Studies
Kosuke SUGAI
Kinki University
Hirokazu YOKOKAWA
Kobe University

Abstract

This study investigated the use of verb subcategorization information during sentence comprehension by Japanese EFL learners, using a self-paced reading task with the embedded anomaly technique. In order to observe when the syntactic structure was determined, filler-gap sentences were constructed as stimuli. Four types of verbs were used: simple transitive verbs, dative verbs, infinitive complement verbs and intransitive preference verbs. The plausibility of the direct object of the embedded verb was also manipulated. The results demonstrate that the high proficiency learners utilize verb subcategorization information of simple transitive verb; however they do not use it automatically as natives. The results also indicate that they cannot utilize that of other types of verbs. The low proficiency learners show the difficulty in processing the sentences with long-distance dependencies.

1. Introduction

Recent studies show that one of the main factors that causes L2 learners difficulties in fluent reading is lack of automaticity including lower-level processes like word recognition and parsing. We assume that such less-automated sentence processing may result from the learners’ lack of lexical information knowledge in their mental lexicon.

Based on this view, this study investigates lexical influence on parsing, focusing on how
Japanese EFL learners utilize verb subcategorization information while they comprehend a sentence. The final goal of the study is to apply the findings to the efficient reading instruction for Japanese EFL learners.

2. Literature Review

2.1 Filler-Gap Dependencies

In sentence processing studies, the sentences with long-distance dependencies ("filler-gap" sentences: FG) are often used as stimuli in order to investigate when the syntactic structure is determined (e.g., Tanenhaus et al., 1989). In a filler-gap sentence, the "filler" must be associated with a "gap" that occurs later in the sentence, at least in English. For example, in a sentence like "Which window did the boy break ___ at midnight?" the wh-questioned phrase which window is semantically interpreted as the direct object of the verb break. In this case, the filler is which window, and the gap is posited and filled after break. If a reader immediately accesses and uses the subcategorization information that the verb break takes a direct object, then the filler-gap association will take place at the main verb position.

2.2 Sentence Anomaly Task

Tanenhaus et al. (1989) conducted an experiment to investigate lexical influence on the parsing of sentences with long-distance dependencies. To determine when a gap was posited and filled, they used a plausibility judgment task that they refer to as the "embedded anomaly technique" in which participants pressed a "No" key if the sentence stopped making sense. They conducted a self-paced reading task with word-by-word presentation. They used two types of verbs: verbs that are typically used transitively and verbs that are typically used intransitively, and two types of fillers: fillers that were plausible as the object of the verb and fillers that were implausible objects. For example, in a sentence like "The sheriff wasn't sure which horse/rock the cowboy raced down on the hill." horse is a plausible object of the verb race, whereas rock is an implausible object. The results demonstrated that gaps were posited and filled immediately at the embedded verb for transitive preference verbs, even when the filler was implausible, whereas gaps were not initially posited following intransitive verbs. They also compared the verbs that are typically used with both an object and an infinitival complement (e.g., remind) with verbs that do not take infinitival complement (e.g., call). The result for the call-type verbs was that gaps were posited and filled at the embedded verbs, even when the filler was implausible. In contrast, there wasn't plausibility effect at the embedded verb in either the judgment or reading time for remind-type verbs, but object gaps were filled at complementizer to only when the filler was a plausible object. From these results, they demonstrate that verb-argument structure is used for filler-gap assignment and it interacts with plausibility in determining when a filler will be assigned to a possible gap. Boland et al. (1989) conducted experiments to further explore the use of verb
subcategorization in the gap-filling process. They examined the difference between simple transitive and infinitive complement verbs, and included a new type of verb, dative verbs. They also used a sentence anomaly task in a self-paced word-by-word reading task. The results showed that gaps were posited and filled at a word after embedded verbs in the sentences with dative verbs. This study, like Tanenhaus et al. (1989), suggests that verb subcategorization information is immediately accessed and used in parsing sentences with filler-gap dependencies in L1.

In an L2 setting, Williams (2006) examined the degree to which natives and non-natives interpret English sentences incrementally using stop making sense tasks. He required participants to read sentences word by word to make on-line plausibility judgment. The results showed that both natives and non-natives processed sentences incrementally and they were essentially similar with regard to lexical processing. However, semantic processing of L2 was less efficient than for L1 and appeared to be task-and person-independent. Yabuuchi et al. (2001) investigated how Japanese EFL learners use verb subcategorization during sentence processing using filler-gap sentences and a sentence anomaly task. They found that Japanese EFL learners have a similar tendency to natives’ with regard to processing sentences including verbs typically used transitively, but not for other types of verbs; verbs that typically take infinitive complement and dative verbs.

2.3 Residual Problems

This study partly replicates Yabuuchi et al. (2001), and attempts to further explore the use of verb subcategorization in the gap-filling process of Japanese EFL learners. However, there were several problems in the previous study. First, the words used in the experimental sentences were not sufficiently controlled. They should be carefully controlled with length, frequency and familiarity. Second, the participants’ English proficiency was not appropriately assessed. The proficiency levels should be more clearly distinguished. Third, some inappropriate sentences were included in the stimuli. Therefore, we conducted an experiment with improving these problems in the study.

3. Experiment

3.1 Research Questions

In order to examine how Japanese EFL learners use verb subcategorization information during sentence comprehension, two research questions have guided the study: 1. How do Japanese EFL learners use subcategorization information in processing filler-gap sentences on line? 2. How does the degree of proficiency affect processing filler-gap sentences?

3.2 Method

3.2.1 Participants

Thirty-five undergraduate and graduate students learning English as a foreign language
participated in the experiment. They were divided into three groups, high (Group A \( n = 14 \)), intermediate (Group B \( n = 8 \)), and low (Group C \( n = 13 \)) proficiency groups according to SLEP Test (Secondary Level English Proficiency Test) reading section score. The score range of the Group A, Group B, and Group C is 61-73, 51-60, 27-48, and the mean score and (standard deviation) are 66 (4.4), 55 (2.7), 39 (6.8) respectively (Full score = 75).

3.2.2 Material

The verbs were selected based on the verb norming studies that were tested in a sentence generation task to assess verb’s preference (e.g., Traxler, 2005). Four types of verbs were used: (1) simple transitive verbs (e.g., call), (2) dative verbs (e.g., buy), (3) verbs that are typically followed by infinitive complement (e.g., advise), (4) intransitive preference verbs (e.g., hurry). Eight verbs of each type (over 60% preference) were chosen and placed in embedded questions in which the questioned noun phrase was the object of the verb in filler-gap sentences. Two types of fillers were used: fillers that were plausible as the object of the verb, and fillers that were implausible objects. In addition, for each verb, we constructed the baseline (BL) condition in which there is not a filler-gap dependency. Thirty-two sentences were constructed as experimental sentences using four types of verbs, and another 32 sentences were constructed as distracters. Four sets of experimental sentences were prepared to counterbalance the design. Thus, 160 sentences were made overall (each participant read 64 sentences). All the words used in experimental sentences were more than 5.0 familiarity rate on a 7 point scale (7 = highly familiar; Yokokawa [ed.], 2006). The number of syllables of wh-questioned noun phrases was also controlled.

3.2.3 Procedure

A self-paced phrase-by-phrase reading task was conducted using the psychological experiment software SuperLab Pro. All phrases appeared in the center of the computer screen. Before the experiments started, the instructions were given, and a practice session was conducted including four trials. Each trial began with the mark “+++++” on the screen. When participants pressed the spacebar for the first time, the mark changed to the first phrase in the sentence. Participants were instructed to continue pressing the spacebar as long as the sentence continued to make sense. When they thought the sentence had stopped making sense, they were to press a different key. This was called a “no-response.” All times between presentations of a phrase were measured to millisecond accuracy by the software. Thus, we had two ways to observe implausibility: observing increased reading times and an increased no-response rate at a particular position. The experiment was approximately 20 minutes long including instructions and trial tests.

3.2.4 Analysis

We focused on comparison of reading times and no-response rate at the critical regions between ambiguous plausible sentences and ambiguous implausible sentences. To answer the
research questions mentioned in 3.1, the data of each type of verb and proficiency group were separately analyzed using two-way ANOVAs: plausibility (plausible: implausible) x ambiguity (FG: BL). In order to make proficiency difference clear, the results of Group B were not analyzed.

4. Results and Discussion

4.1 Simple Transitive Verb

Figure 1 indicates mean reading time and no-response rate of each region of sentences with simple transitive verbs for Group A. There was a reliable interaction between plausibility and ambiguity at the embedded verb position (P5). Increased reading time was observed at the position of implausible FG sentences. They were read 459 ms slower than plausible sentences, which led to a main effect of plausibility, $F(1, 51) = 7.424, p < .01, \eta^2 = .12$. This suggests that the gaps were posited and filled at the embedded verb. No-response rate started to increase at the embedded verb position (P5) of FG sentences, but there was no significant plausibility effect. Increased no-response rate (35%) was also observed at a word after the verb (P 6), suggesting the gap-filling was delayed compared to natives. Overall no-response rate for implausible FG sentences was 68%, which was lower than that of natives in the previous studies. From these results, it seems that high proficiency learners tend to process FG sentences using subcategorization of simple transitive verbs; however not automatically as natives.

Figure 2 indicates mean reading time and no-response rate of each region of sentences with simple transitive verbs for Group C. Increased reading time and no-response rate already appeared at P3 position (which) in FG sentences. There was no significant plausibility effect either on reading time or no-response rate at the embedded verb of FG sentences. Overall no-response rate for implausible FG sentences was 65%, which was similar to the result of Group A. However, no-response rate was 42% for plausible FG sentences, which was higher than Group A. Taken together, it seems to be quite difficult for low proficiency learners to process sentences with complex structure such as FG sentences.

*Figure 1. Mean reading time (ms, left) and cumulative no-response rate (%) of simple transitive verb for Group A.*
Figure 2. Mean reading time (ms, left) and cumulative no-response rate (%) right) of simple transitive verb for Group C.

[FG] The man wasn't sure/ which horse (*song) / the boy / watched / this morning.
[BL] The man wasn't sure / whether / the boy / watched / the horse (*song) / this morning.

Note. FG Pla = plausible filler-gap sentence, FG Imp = implausible filler-gap sentence, P 1 = phrase one, ◆ = critical region (gap-filling position in the L1 studies), asterisk = implausible object of embedded verb, and underlined = critical region.

4.2 Dative Verb

Figure 3 indicates mean reading time and no-response rate of each region of sentences with dative verbs for Group A. Although increased reading time was observed at the critical region (P6, to the student) in implausible FG sentences, there was no significant difference between plausible and implausible FG sentences. However, there was a reliable interaction between plausibility and ambiguity before the critical region. The significant plausibility effect was found a word before the critical region (P5: taught), which was similar to the result in sentences with simple transitive verbs, $F = (1, 55) = 5.701, p = .0205, \eta^2 = .09$. In addition, increased no-response rate appeared at the embedded verb and became the highest in the following region, which was also quite similar to the tendency in sentences with simple transitive verbs. These suggest that high proficiency learners tend to regard the embedded verb as a simple transitive verb, and gaps seem to be posited and filled at the embedded verb. Overall no-response rate for implausible FG sentences was 61%, which was lower than that of natives in the previous studies. These results show that even high proficiency learners cannot use subcategorization information of dative verbs efficiently in processing FG sentences.

Figure 4 indicates reading time and no-response rate of each region of sentences with dative verbs for Group C. Reading time and no-response rate increased at P3 position in both FG sentences and baseline, suggesting the difficulty in processing complex sentences. There was no significant plausibility effect either on reading time or no-response rate at the critical region (P6) of FG sentences. Overall no-response rate for implausible FG sentences was 54%, which was far lower than that of natives in the previous studies.
Figure 3. Mean reading time (ms, left) left, and cumulative no-response rate (%) right of dative verb for Group A.

Figure 4. Mean reading time (ms, left) and cumulative no-response rate (%) right of dative verb for Group C.

[FG] The teacher was sure which news (*boy) the coach told to the student at school.
[BL] The teacher was sure whether the coach told the news (*boy) to the student at school.

4.3 Infinitive Complement Verb

Figure 5 indicates mean reading time and no-response rate of each region of sentences with infinitive complement verbs for Group A. Reading time and no-response rate increased at a word before to infinitive (P6), and no significant plausibility effect was observed either on reading time or no-response rate at P7, indicating gaps were not posited and filled at the position of to infinitive of FG sentences. These results suggest that high proficiency learners cannot use subcategorization information of infinitive complement verbs.

Figure 6 indicates mean reading time and no-response rate of each region of sentences with infinitive complement verbs for Group C. Again, increased reading time and no-response rate appeared at P3 and overall no-response rate for implausible FG sentences was low (58%), suggesting the difficulty in processing FG sentences. No significant plausibility effect was observed either on reading time or no-response rate at the position of to infinitive of FG sentences. This sentence type has the longest distance between the filler and the gap of all, which may be one of the factors that causes the difficulty in processing FG dependencies.
Figure 5. Mean reading time (ms, left) and cumulative no-response rate (%) right) of infinitive complement verb for Group A.

Figure 6. Mean reading time (ms, left) and cumulative no-response rate (%) right) of infinitive complement verb for Group C.

[FG] The teacher / was sure / which party (*woman) / the coach / permitted / the student / to attend / at school.

[BL] The teacher / was sure / whether / the coach / permitted / the student / to attend / the party (*woman) / at school.

4.4 Intransitive Preference Verb

Figure 7 indicates mean reading time and no-response rate of each region of sentences with intransitive preference verbs for Group A. Although reading time increased at the embedded verb of implausible FG sentences, there was no significant difference between plausible and implausible FG sentences. This indicates that the gaps were not posited and filled at the embedded verb. Increased no-response rate was observed at the embedded verb in both plausible and implausible FG sentences, suggesting high proficiency learners regarded the embedded verb as an intransitive verb. Increased no-response rate was also observed at a word after the embedded verb, suggesting the gap-filling was delayed. These results show that high proficiency learners have at least the knowledge of transitive and intransitive verbs.

Figure 8 indicates mean reading time and no-response rate of each region of sentences with intransitive preference verbs for Group C. Again, reading time and no-response rate increased at P3 position in both FG sentences and baseline, suggesting the difficulty in processing complex sentences.
Figure 7. Mean reading time (ms, left) and cumulative no-response rate (%) of intransitive preference verb for Group A.

Figure 8. Mean reading time (ms, left) and cumulative no-response rate (%) of intransitive preference verb for Group C.

[FG] The man / wasn’t / sure / which horse (*rock) / the boy / raced / this morning.
[BL] The man / wasn’t sure / whether / the boy / raced / the horse (*rock) / this morning.

5. Concluding Remarks and Further Studies

In the study, we examined how Japanese EFL learners utilize verb subcategorization information during sentence comprehension, using four types of verbs. The results show that they seem to be able to use subcategorization information of simple transitive verbs; however they cannot use that of other types. As Jurafsky (2003) argues, it may reflect frequency of syntactic frame, because simple transitive is the most frequent verb frame for Japanese EFL learners. Although high proficiency learners seem to store lexical information in their mental lexicon as well as natives, they cannot use it efficiently during sentence processing. Low proficiency learners may have insufficient lexical representation in their mental lexicon, so that they cannot use it for sentence processing. Taken together, we can hardly say that Japanese EFL learners, even high proficiency learners, can determine the syntactic structure automatically. The results also reveal that low proficiency learners do not seem to have subcategorization representation itself and it is quite difficult for them to process sentences with complex structure like filler-gap dependencies.

For further study, it seems to be necessary to investigate what kind of tasks or activities are efficient for improving representation of subcategorization information in mental lexicon and for promoting the ability to use it.
Acknowledgements

This study was partially supported by a JSPS Grant-in Aid for Scientific Research (A). We wish to express our gratitude to the teachers and students who have cooperated with the research.

References


Appendix

| Mean Reading Time (ms) and SD, and No-Response Rate (%) of FG Sentence |
|---|---|---|---|---|---|---|---|---|
| P3: RT(ms) | No(%) | P4: RT(ms) | No(%) | P5: RT(ms) | No(%) | P6: RT(ms) | No(%) | P7: RT(ms) | No(%) |
| Transitive Pla High | 1211(435) | 0 | 1207(701) | 14.29 | 1001(336) | 0 | 1198(561) | 14.29 |
| Low | 1768(668) | 11.54 | 1613(818) | 19.23 | 1034(263) | 0 | 1257(656) | 11.54 |
| Imp High | 1143(472) | 0 | 1210(493) | 0 | 1460(763) | 32.14 | 1305(768) | 35.71 |
| Low | 1823(724) | 15.38 | 1428(734) | 26.92 | 1196(795) | 7.69 | 1153(802) | 15.38 |
| Dative Pla High | 973(309) | 0 | 1232(657) | 10.71 | 1051(618) | 21.43 | 1269(598) | 25.00 | 1027(561) | 7.14 |
| Low | 1839(820) | 11.54 | 1638(862) | 3.85 | 1188(641) | 3.85 | 1224(398) | 23.08 | 1051(509) | 7.69 |
| Imp High | 1127(472) | 0 | 1155(475) | 10.71 | 1277(679) | 32.14 | 1450(274) | 32.14 | 1381(661) | 14.29 |
| Low | 1817(843) | 11.54 | 1324(590) | 26.92 | 1288(561) | 7.69 | 1395(448) | 15.38 | 1189(370) | 7.69 |