言語研究（*Gengo Kenkyu*）**160**: 123–153（2021）　　doi: 10.11435/gengo.160.0_123

# The Timing of Filler-Gap Dependency Formation in Second Language Comprehension

Itsuki Minemi　　Masataka Yano

*The University of Tokyo/ JSPS Research Fellow*　　*Tokyo Metropolitan University*

**Abstract:** This study investigated at what point in a sentence native Japanese speakers begin to build long-distance dependencies while reading English relative clauses (RCs) by conducting two self-paced reading experiments. Experiment 1 revealed that Japanese learners of English construct long-distance dependency immediately after reading a verb by demonstrating "plausibility mismatch effects" at the verb site. On the contrary, Experiment 2 did not find evidence that they form a long-distance dependency before a verb, that is, no "transitivity mismatch effect." On the basis of these results, this study proposes that Japanese learners of English initiate long-distance dependency formation immediately after encountering a verb, but it does not precede the appearance of a verb. This extends previous findings that it is difficult to generate predictions in the processing of non-native languages, to a structural processing level.*

**Key words:** Psycholinguistics, Sentence Comprehension, Non-Native Language, English, Filler-Gap Dependencies

## 1. Introduction

In native language (L1) comprehension, sentence processing operates incrementally in the sense that the language processor builds a syntactic representation immediately after perceiving new linguistic input (Altmann and Kamide 1999, Crain and Fodor 1985, Frazier and Clifton 1989, Kamide, Altmann, and Haywood 2003, Miyamoto 2002, Tanenhaus et al. 1995, Traxler and Pickering 1996). The present study uses the term INCREMENTAL to refer to a process in which the language processor incorporates bottom-up information from actual input into a mental representation without a substantial delay. Recent studies have further shown evidence for top-down, PREDICTIVE L1 processing, which refers to the processing of some information before encountering bottom-up evidence avail-

* We would like to thank Masakatsu Inoue for his statistical support, and Hajime Ono, Daichi Yasunaga, and Yuki Hirose for their invaluable advice. We are also grateful to two anonymous reviewers for their insightful comments and suggestions. This paper is based on a B.A. thesis submitted by the first author to the Department of Linguistics, Kyushu University. We are thankful to the lab members for their supports. This work was supported in part by JSPS KAKENHI Grant Number 18J21393 (PI: Itsuki Minemi).

able from input (Aoshima, Phillips, and Weinberg 2004, Arnold 2001, Kehler et al. 2007, Nakano, Felser, and Clahsen 2002, Omaki et al. 2015, Rohde and Horton 2014).

In non-native language (L2) comprehension, the extant literature shows that L2 learners build syntactic structures as incrementally as L1 speakers do in the processing of certain structural dependencies, including FILLER-GAP DEPENDENCIES; they form a structural dependency immediately after verb information becomes available (Cunnings 2017, Cunnings et al. 2010, Felser 2015, Felser et al. 2012, Kim, Baek, and Tremblay 2015, Omaki and Schulz 2011). In contrast, predictive processing is proposed to be more limited in L2 comprehension because L2 learners consume substantial resources during incremental processing and thus cannot allocate sufficient resources to predictive computation (Dussias et al. 2013, Grüter, Rohde, and Schafer 2014, 2017, Hopp 2013, Ito, Martin, and Nieuwland 2017, Kaan 2014, Kaan, Dallas, and Wijnen 2010, Mitsugi and Macwhinney 2016). However, only limited types of construction have been examined thus far, and no research has investigated whether L2 learners can build a filler-gap dependency as predictively as L1 speakers do. Therefore, it is unclear to what extent the hypothesis that predictive processing is restricted in online L2 comprehension can be generalized. Hence, the present study addresses this issue by investigating whether L2 learners of English form filler-gap dependencies predictively to the same extent as do native speakers of English. In the following sections, we first discuss the distinction between the incremental and predictive processes in the filler-gap dependency formation and then turn to the processing of filler-gap dependencies in L2 comprehension.

### 1.1. Incremental and predictive processing of filler-gap dependencies

Filler-gap dependency refers to a dependency between a dislocated element (a FILLER) and its canonical position (a GAP).[1] In a relative clause (RC) in English, as shown in (1), the head noun phrase (NP) *the laboratory* is located in the clause-initial position but needs to be interpreted in the object position of the preposition *to*, which is indicated with <GAP>, because the RC head NP cannot receive a thematic role in a nonargument position.

(1) That is the laboratory which Irene used a courier to deliver the samples to <GAP>.

In order to correctly interpret this type of sentence, the processor needs to identify a gap site and integrate a filler with that position (i.e., filler-gap depen-

[1] We note here that we use the terms "gap" or "filler-gap dependency" because they make descriptions or discussions on long-distance dependencies easier and simpler and not because the gaps are assumed to account for the processing of long-distance dependencies. Discussions in this paper can also be understood with non-gap analyses (e.g., Pickering 1993, Pickering and Barry 1991). We do not address the issue of whether the processor makes use of gaps to comprehend long-distance dependencies in the present study.

dency formation). Previous research has proposed two types of filler-gap dependency parsing mechanisms: FILLER-DRIVEN PARSING (the active filler hypothesis: Crain and Fodor 1985, Frazier and Clifton 1989) and VERB-DRIVEN PARSING (McElree and Griffith 1998, Pickering and Barry 1991). The filler-driven parsing hypothesis assumes that the sentence processor constructs filler-gap dependencies as soon as possible, that is, the processor begins searching for a gap position immediately after it detects the filler and postulates a gap at the earliest structurally eligible position without waiting for bottom-up supportive information such as verb transitivity. In this sense, filler-driven parsing is categorized as predictive processing. In contrast, the verb-driven parsing account assumes that the processor does not create dependencies until it obtains bottom-up information such as verb's information that encourages the gap-postulation process.

We next discuss the distinction between those two filler-gap dependency parsing hypotheses by observing how each parsing mechanism works in a phrase-by-phrase fashion when processing RC sentence (1). First, once the processor reaches the RC head NP *the laboratory* and the RC complementizer *which*, it recognizes that the sentence involves a filler-gap dependency and that a gap is needed in a later part of the sentence. This process is categorized as an incremental process because bottom-up structural information from the RC complementizer *which* confirms that the gap is required. On the other hand, postulating the gap into a specific position at this moment, such as in the RC subject position, is a predictive process because no bottom-up information confirms such structural analyses. In other words, the processor cannot determine a specific gap site based on only the bottom-up information that has been obtained so far. Predictively positing a specific gap at this moment requires a top-down process, such as that the processor aims to postulate a gap at the earliest structurally eligible position (i.e., filler-driven parsing). Indeed, the subject gap is predicted once it encounters the RC complementizer in L1 comprehension, as evidenced by a longer reading time (RT) at the RC subject NP *Irene* when the RC subject position is potentially available for the gap, as in (1), than when the RC subject is not available for the gap, as in *That is the laboratory to which Irene used a courier to deliver the samples* <*GAP*> (i.e., a SUBJECT FILLED-GAP EFFECT: Lee 2004).[2]

Then, the processor associates the NP *Irene* with the RC subject position (an incremental process), which means that the RC subject position is no longer avail-

[2] Some studies argue that L2 learners also exhibited the subject filled-gap effects and suggest that they adopt the predictive subject-gap postulation process as well as L1 speakers do (Aldwayan, Fiorentino, and Gabriele 2010, Canales 2012, Johnson, Fiorentino, and Gabriele 2016). However, those studies compared reading times at an embedded subject NP *Barbara* between a *wh*-sentence such as *My brother asked who Barbara will photograph us beside* <*GAP*> *at the graduation* and a yes-no question sentence such as *My brother asked if Barbara will photograph us beside Mom at the graduation*, so there remains a possibility that the lexical difference between *who* and *if* caused the longer RTs in the *wh*-sentence. Hence, whether the processor engages in the predictive subject-gap postulation process in L2 comprehension remains controversial.

able for the gap. If the parsing mechanism is filler-driven, it predicts a gap at the object position of an upcoming verb because that is the next earliest available gap position (i.e., the hyper-active gap filling hypothesis: Omaki et al. 2015). Omaki et al. (2015) tested this hypothesis using an eye-tracking experiment with the RC sentences in (2), in which they manipulated verb transitivity (intransitive (2a/b) vs. transitive (2c/d)) and islandhood (i.e., whether a sentence involves an RC island (2b/d) or not (2a/c); island domains were indicated with square brackets []). Because it is prohibited to postulate a gap within islands (Chomsky 1973, 1986, Ross 1967), the island condition serves as a control condition against the non-island conditions.

(2) a intransitive, non-island:
The book that the author chatted regularly about <GAP> was named for an explorer.
b intransitive, island:
The book that the author [who chatted regularly] saw <GAP> was named for an explorer.
c transitive, non-island:
The book that the author wrote regularly about <GAP> was named for an explorer.
d transitive, island:
The book that the author [who wrote regularly] saw <GAP> was named for an explorer.

The filler-driven hypothesis claims that the processor posits a gap at the direct object position of the verb before the input of the verb if the sentence does not include an island. This necessitates the anticipation of a transitive verb because only transitive verbs can take a direct object. Thus, the filler-driven account assumes that the processor predicts a transitive verb as well as a direct object gap. Specifically, this account expects an RT disruption at the intransitive verb *chatted* in the non-island condition (2a) (i.e., a TRANSITIVITY MISMATCH EFFECT) because the intransitive verb *chatted* violates the expectation. However, such an RT disruption is not expected in the transitive, non-island condition (2c) because the transitive verb *wrote* satisfies the expectation that the verb would be transitive and host a direct object gap. Therefore, the filler-driven hypothesis predicts a longer RT for intransitive verbs in the non-island condition (2a) than in the island condition (2b) but predicts no RT differences in transitive conditions, such as (2c) and (2d).

In Omaki et al.'s (2015) eye-tracking experiment, a longer RT was observed for the intransitive verb in the non-island condition (2a) than in the island condition (2b), as expected by the filler-driven hypothesis. The RT increased because of the prediction error induced by the intransitive verb (the transitivity mismatch effect). Moreover, no RT contrast was observed in the transitive conditions. These results show that the processor predictively posits a direct object gap and anticipates the transitive verbs before encountering the verbs. Thus, the transitivity mismatch effect demonstrates that fillers, rather than verbs, drive gap prediction in the

L1 processing of filler-gap dependencies.

### 1.2. Filler-gap dependency formation in L2 comprehension

As mentioned above, it is generally assumed that predictive processing is limited in L2 comprehension (Grüter et al. 2014, 2017, Kaan 2014, Kaan et al. 2010). However, as to the processing of filler-gap dependencies in an L2, it is unclear whether the processor predictively builds dependencies as no previous research has investigated this issue. Although several studies have revealed that L2 learners posit an object gap at least as soon as encountering a transitive verb, this does not imply that they process filler-gap dependencies as predictively as native English speakers because the bottom-up subcategorization information of a verb is available when the verb is encountered (i.e., verb-driven parsing: Cunnings et al. 2010, 2010, Dallas, DeDe, and Nicol 2013, Dallas and Kaan 2008, Kim et al. 2015, Omaki and Schulz 2011, Williams 2006, Williams, Möbius, and Kim 2001).

For instance, Omaki and Schulz (2011) showed that Spanish learners of English create an object gap immediately after the verb appears by conducting a self-paced reading experiment with a plausibility mismatch paradigm (Traxler and Pickering 1996). They manipulated islandhood (non-island (3a/b) vs. island (3c/d)) and the potential semantic plausibility between the filler and verb by changing the filler (*city* for the implausible conditions (3a/b) and *book* for the plausible conditions (3c/d)). In the non-island, implausible condition (3a), for example, the RC head NP *city* was semantically implausible as the object of the verb *wrote* ('*the author wrote the city*').

(3) a non-island, implausible:
The city that the author wrote regularly about <GAP> was named for an explorer.
b non-island, plausible:
The book that the author wrote regularly about <GAP> was named for an explorer.
c island, implausible:
The city that the author [who wrote regularly] saw <GAP> was named for an explorer.
d island, plausible:
The book that the author [who wrote regularly] saw <GAP> was named for an explorer.

The results showed that an RT was longer at the adverb *regularly* (the post-critical spillover region) in the non-island, implausible condition (3a) than in the non-island, plausible condition (3b), which is most likely due to the semantic plausibility mismatch between the filler *the city* and the verb *wrote* (i.e., the PLAUSIBILITY MISMATCH EFFECT). On the other hand, in the island conditions (3c/d), there was no plausibility mismatch effect. These results suggest that Spanish learners of English form filler-gap dependencies at least as soon as they encounter the verb and compute semantic interpretation of the dependency, and RTs in the

spillover region therefore increased only when the dependency was semantically implausible. Similar results were found for L2 learners of English who have other L1 backgrounds (Cunnings et al. 2010 for Chinese-speaking learners, Felser et al. 2012 for German-speaking learners, Kim et al. 2015 for Spanish- and Korean-speaking learners).

Nonetheless, no study has examined the predictive filler-driven parsing process in L2 English. It is expected that L2 learners of English exhibit transitivity mismatch effects if they process English filler-gap dependencies predictively. On the other hand, taking into account that it is argued that predictive processing in L2 comprehension is more restricted than in L1 comprehension, it might be the case that L2 learners of English do not adopt predictive parsing while processing filler-gap dependencies.

The present study explores this issue by focusing on the processing of English filler-gap dependencies by Japanese speakers. Japanese is a head-final language, that is, a language in which a head element such as a verb appears late in sentences. Therefore, substantial memory resources are required for the processor to refrain from constructing a sentence structure until a head element appears during the processing of such languages (i.e., verb-driven parsing). In fact, it has been revealed that L1 speakers form filler-gap dependencies predictively before bottom-up verb information becomes available in Japanese (Aoshima et al. 2004, Nakano et al. 2002). Thus, it would be possible that Japanese speakers process English filler-gap dependencies predictively as well, given a potential positive influence from the L1 sentence comprehension strategy. This study can provide new insight into predictive processing in L2 comprehension by focusing on the relationship between predictive processing in language comprehension and L2 learners' processing strategy in their L1.

## 2. The present study

The present study aims to reveal the role of the prediction system in the human sentence comprehension mechanisms by looking into the predictive formation of filler-gap dependencies by Japanese learners of English. Although the previous findings indicate that the processor completes the filler-gap dependency formation at least immediately after a verb in L2 comprehension, whether it initiates the dependency formation even before a verb (i.e., filler-driven parsing) in L2 comprehension remains an open question. Given that some previous studies propose the hypothesis that predictive processing is generally restrictive in L2 comprehension relative to in L1 comprehension (Grüter et al. 2014, 2017, Kaan 2014, Kaan et al. 2010), it would be worth exploring the issue of whether the processor starts searching for a gap before bottom-up verb information becomes available in L2 comprehension as well in order to reveal a clearer picture of predictive processing in online sentence comprehension. Thus, the present study investigated whether Japanese learners of English initiate the filler-gap dependency formation in advance of the verb (filler-driven parsing) by conducting a self-paced reading experiment with a transitivity mismatch paradigm (Experiment 2). However, this

study first explored verb-driven parsing in the processing of filler-gap dependencies in English by Japanese-speaking learners using a plausibility mismatch paradigm (Experiment 1) before directly examining predictive object-gap postulation in the processing of filler-gap dependencies by Japanese learners of English as no research has examined whether Japanese learners of English complete the dependency construction at least immediately after the verb is encountered (i.e., verb-driven parsing). The predictive gap postulation can be regarded as "hyper-incremental" processing; therefore, no predictive processing is expected as far as the processor does not exhibit the incremental processing. In the following sections, we first describe the experimental methods that were used in both Experiments 1 and 2 because the two experiments were performed simultaneously in the same session.

### 2.1. Participants

Forty-seven Japanese learners of English, who were undergraduate or graduate students at Kyushu University, participated in the experiment. Their mean age was 21.96 (standard deviation (*SD*) = 1.80), and the mean age at which they were first exposed to English was 9.96 (*SD* = 3.09). Fourteen learners had studied in English-speaking countries for 6.11 months on average (*SD* = 6.24). Immediately after the reading experiment on the same day as the self-paced reading experiments, the L2 English proficiency levels of the participants were assessed using the Oxford Placement Test 2 (Allan 2004). The mean score was 144.62 (*SD* = 9.72), thus indicating that the participants were upper-intermediate level learners (corresponding to CEFR: B2, IELTS: 5.5, and TOEIC: 525–780). All were paid for their participation, and both experiments were approved by the ethics committee of the Department of Linguistics at Kyushu University.

### 2.2. Procedure

The self-paced reading experiments were conducted on the Linger software version 8.6.3, which was developed by Doug Rohde. The experimental stimuli were presented in a non-cumulative, word-by-word moving window presentation (Just, Carpenter, and Woolley 1982). First, a series of dashes was presented on the screen. Each time a participant pressed the spacebar, the dashes were replaced with words one by one from left to right. After each experimental sentence, another sentence was presented as a comprehension question to check if a participant had correctly understood the sentence. In the comprehension question, the participants were asked whether the event described in the question sentence was consistent with the experimental sentence. The feedback for these answers was given only for incorrect answers.

There were 32 sets of experimental target sentences for each experiment and 64 filler sentences. The filler sentences included 32 sentences with RCs and 32 without RCs to prevent the participants from adapting to the experimental sentences. In addition, the target sentences in each experiment served as the filler sentences for the other experiment, since the experiments were carried out in the

same experimental session simultaneously. The experimental sentences were distributed among four lists so that a participant read only one token from each set. The experiment lasted approximately an hour and a half.

## 3. Experiment 1

Experiment 1 examined whether Japanese learners of English posit an object gap at least immediately after reading an RC verb. We used an experimental paradigm similar to that of Traxler and Pickering (1996) and Omaki and Schulz (2011): a two-by-two factorial design with the manipulations of the sentence type (RC vs. non-RC) and the plausibility of the object-gap dependencies (implausible vs. plausible).[3] In the current study, the plausibility focused on the thematic relations among the first NPs, the second NPs, and the first verbs (the embedded verbs) in the RC condition.

### 3.1. Materials

Example sentences for the critical items are shown in (4) below, and all experimental items, including filler sentences, are available from the corresponding author's OSF webpage https://osf.io/rupd6/. The sentences in (4a) and (4b), which contain an RC, were our target sentences. In (4a), the filler NP, *parents*, is implausible as an object of the verb *feed*, which takes the second NP *infant* as its subject, i.e., *the infant feeds the parents* (thematic-role reversals; Chow et al. 2016). On the contrary, in (4b), the first NP *infant* is a plausible object of the verb, the subject of which is *parents* (i.e., *the parents feed the infant*).

Sentences (4c) and (4d) were control conditions because they did not include an RC. In the control conditions, a plausibility mismatch between the filler and the verb did not occur because they did not include a filler-gap dependency. Although Traxler and Pickering (1996) and Omaki and Schulz (2011) used RC islands to create control conditions, this study used a simple NP-of-NP construction primarily because the purpose of the experiment was to investigate the plausibility mismatch effects and not whether L2 learners respect island constraints. The essential difference between the RC and non-RC conditions was whether the sentences involved a filler-gap dependency, which was sufficient for the study purpose.

(4) a RC, implausible:
The$_1$ parents$_2$ who$_3$ the$_4$ infant$_5$ did$_6$ not$_7$ feed$_8$ a$_9$ nutritious$_{10}$ dish$_{11}$ lost$_{12}$ weight$_{13}$ rapidly.$_{14}$
b RC, plausible:

[3] Unlike Traxler and Pickering (1996) and Omaki and Schulz (2011), who manipulated semantic plausibility by changing filler nouns, we manipulated it by reversing two arguments. Considering that it is well-known that the semantic relatedness between content words in a sentence has a significant influence on the processing of a word, we aimed to uniquely interpret the implausibility effect as reflecting the implausibility of the semantic meaning of filler-gap dependencies in the implausible condition and not merely as a priming effect between the content words in the plausible condition (Chow et al. 2016).

The$_1$ infant$_2$ who$_3$ the$_4$ parents$_5$ did$_6$ not$_7$ feed$_8$ a$_9$ nutritious$_{10}$ dish$_{11}$ didn't$_{12}$ drink$_{13}$ milk.$_{14}$

c non-RC, implausible:
The$_1$ infant$_2$ of$_3$ the$_4$ parents$_5$ did$_6$ not$_7$ feed$_8$ a$_9$ nutritious$_{10}$ dish$_{11}$ to$_{12}$ his$_{13}$ dogs.$_{14}$

d non-RC, plausible:
The$_1$ parents$_2$ of$_3$ the$_4$ infant$_5$ did$_6$ not$_7$ feed$_8$ a$_9$ nutritious$_{10}$ dish$_{11}$ to$_{12}$ their$_{13}$ pets.$_{14}$

One may wonder whether the RC sentences are grammatical because some theoretical literature argues that relativization of the indirect object NP such as *parents* in a double-object construction is grammatically prohibited in English (e.g., Hudson 1992, Larson 1988),[4] while other studies do not treat such sentences as completely ungrammatical sentences (e.g., Barss and Lasnik 1986, Jackendoff and Culicover 1971). Thus, to validate that those RC sentences that contain relativization of the indirect object in a double object construction like (4a/b) can be used as experimental sentences, we conducted an acceptability judgment with native speakers of English.[5] The judgment included ungrammatical sentences as well as the target sentences above to test whether those RC sentences show higher acceptability that ungrammatical sentences. The RC, implausible condition (4a) did not exhibit higher acceptability than the ungrammatical filler sentences ($p$ = 0.942; RC, implausible condition: $M$ = 3.743, $SE$ = 0.236; ungrammatical sentences: $M$ = 3.814, $SE$ = 0.231). Critically, however, the RC, plausible condition

[4] We are grateful to the anonymous reviewers for pointing out this issue.

[5] The acceptability judgment survey included the same 32 sets of target RC sentences like (4) and 32 filler sentences. Half of the filler sentences were ungrammatical due to number agreement mismatch such as **The athlete go to a bar with his friend.*, contained strict subcategorization violations such as **The model told his friends at his birthday party last night.*, or contained island constraint violations such as **The lady* [*that the artist inspires the designer* [*who lives with* <*GAP*>]] *works at a famous restaurant.* Thirty-five native speakers of English were recruited using Amazon Mechanical Turk (https://www.mturk.com/) and Prolific (https://www.prolific.co/). Acceptability of each sentence was rated with a seven-point scale (1 as "unacceptable" and 7 as "acceptable"). The acceptability data were $z$-transformed and analyzed using LME models (Baayen et al. 2008, Bates et al. 2015) with dummy contrast coding (ungrammatical sentences as the baseline condition). Data from one participant were excluded before the statistical analysis because the mean of their $z$-transformed acceptability of the ungrammatical filler sentences was higher than 2 *SD*s from the by-participants mean, which may indicate that s/he did not correctly judge sentence acceptability, as ungrammatical sentences should not be rated with such high acceptability; we also checked that the acceptability patterns did not change substantially because of the data exclusion. While acceptability of the non-RC sentences was not statistically analyzed as it is not relevant to the issue of whether the target RC sentences are acceptable, their mean acceptability was higher than that of the ungrammatical filler sentences (the non-RC, implausible condition (4c): $M$ = 5.147, $SE$ = 0.185; the non-RC, plausible condition (4d): $M$ = 5.085, $SE$ = 0.191).

(4b) showed higher acceptability than the ungrammatical sentences ($p$ = 0.011; RC, plausible condition: $M$ = 4.515, $SE$ = 0.236). These results indicate that the RC sentences with extraction of the indirect object in a double object construction are not completely ungrammatical sentences, at least as far as the plausible ones, as some previous studies have argued (Barss and Lasnik 1986, Jackendoff and Culicover 1971). The acceptability of the implausible RC sentences would be lowered due simply to their implausibility. Thus, the present study used those RC sentences in the reading time experiment.

To diagnose the plausibility mismatch effects, the critical region (Region 8 *feed*) was the first verb region. Regions 1, 3, 4, 6, 7, 9, 10, and 11 were lexically matched for all conditions. In Regions 2 and 5, two different NPs (*parents* and *infant*) were used among the conditions to generate the plausibility differences. Considering that the words used in Regions 12–14 were entirely different, no statistical analyses were reported for these regions. In comprehension question sentences, we tested whether the participants understood sentences correctly by focusing on the interpretation of the matrix clause. For example, the comprehension sentence for (4a) was *The parents lost weight.*, for which participants were expected to answer "yes" since the event that was described by the comprehension sentence was consistent with that of the target sentence. All of the comprehension question sentences are available from the OSF page given above.

### 3.2. Prediction

If Japanese learners posit an object gap at least as soon as the verb is encountered, a plausibility mismatch effect will be observed at the first verb position, Region 8 *feed*, in the RC, implausible condition (4a). Specifically, the RT for the verb would be longer in the RC, implausible condition (4a) than in the plausible condition (4b) due to the implausible thematic relationship between the arguments and the verb. By contrast, this RT difference would not be expected in the non-RC conditions since no plausibility mismatch effect is expected in these conditions due to the lack of filler-gap dependencies.[6] Statistically, an interaction between plausibility and sentence type is expected to be significant in Region 8 because this would reflect the RT disruption in the RC, implausible condition.

### 3.3. Data analysis

To estimate the effects of plausibility and sentence type on the responses to the

[6] An anonymous reviewer noted that the sentence in the non-RC, implausible condition also contains a semantic implausibility because the NP *the infant* seems implausible as a subject of the verb *feed*. We acknowledge this potential implausibility due to the implausible subject NP. Nevertheless, it would be expected that the RC, implausible condition (4a) should induce a larger plausibility mismatch effect than the non-RC, implausible condition (4c) because the plausibility mismatch effect due to the thematic-role reversal, such as *the infant feeds the parent* vs. *the parents feed the infant*, should occur only in the RC, implausible condition (4a). Therefore, it would not be the case that the implausible subject NP cancels out the target thematic-role reversal plausibility mismatch effects.

comprehension task and the RTs, response data were submitted to logistic mixed-effects regression model analyses (Jaeger 2008), and RT data were submitted to linear mixed-effects (LME) model analyses (Baayen, Davidson, and Bates 2008). Model selection was performed using a "parsimonious" approach (Bates et al. 2015). The maximum models included plausibility, sentence type, their interaction term, and item order as the fixed effects and random intercepts and slopes for the participants and item sets.

Each condition was coded using dummy coding depending on SENTENCE TYPE (RC vs. non-RC) and PLAUSIBILITY (implausible vs. plausible): the implausible and RC conditions were coded "0," and the plausible and RC conditions were coded "1." These values were centered when the actual models were constructed. Pairwise comparisons were performed with the non-centered codes to calculate the simple slopes for each factor. The pairwise comparisons for plausibility in the non-RC condition and sentence type in the plausible condition were executed using reverse coding values: "0" for the plausible and non-RC conditions and "1" for the implausible and RC conditions.

The models were implemented in R version 3.5.3 (R Core Team 2019) with the "lme4" package version 1.1.21 (Bates, Mächler, et al. 2015). *P*-values were approximated with the "lmerTest" package version 3.1.0 (Kuznetsova, Brockhoff, and Christensen 2017).

### 3.4. Results

#### 3.4.1. Comprehension accuracy

The mean accuracy rate for the comprehension tasks was 81.986% (*SE* = 1.551). Table 1 and Table 2 show the mean in each condition and the results of the logistic regression model analysis, respectively. As shown in Table 2, a main effect of sentence type, an interaction between plausibility and sentence type, and a main effect of item order were found to be significant. Paired comparisons revealed significant simple effects of plausibility of opposite direction in the RC condition ($B$ = 1.303, $SE$ = 0.441, $z$ = 2.955, $p$ = 0.003) and the non-RC condition ($B$ = −1.158, $SE$ = 0.439, $z$ = −2.641, $p$ = 0.008) and indicated a significant effect of sentence type in the plausible condition ($B$ = −1.936, $SE$ = 0.431, $z$ = −4.494, $p$ < 0.001) but not in the implausible condition ($B$ = 0.526, $SE$ = 0.439, $z$ = 1.197, $p$ = 0.231).

Table 1. Mean accuracy rates for the comprehension task in Experiment 1.

| | Mean Accuracy | *SE* |
|---|---|---|
| (4a) RC, implausible | 82.979% | 2.198 |
| (4b) RC, plausible | 88.830% | 2.192 |
| (4c) non-RC, implausible | 81.663% | 2.737 |
| (4d) non-RC, plausible | 74.468% | 2.793 |

Table 2. Summary of fixed effect estimates of the logistic mixed-effects regression model of the comprehension task response data in Experiment 1.

| | B | *SE* | *z* | *p* | |
|---|---|---|---|---|---|
| (intercept) | 2.463 | 0.234 | 10.509 | <0.001 | * |
| plausibility | 0.073 | 0.045 | 1.632 | 0.103 | |
| sentence type | −0.705 | 0.045 | −15.801 | <0.001 | * |
| sentence type×plausibility | −2.461 | 0.880 | −2.797 | 0.005 | * |
| item order | 0.217 | 0.023 | 9.441 | <0.001 | * |

*: $p < 0.050$

Final Model: glmer (response data ~ sentence type + plausibility + sentence type×plausibility + (1 + sentence type×plausibility | participant) + (1 + sentence type×plausibility | set) + item order, family = "binomial")

### 3.4.2. Reading time

The following were removed from the analyses: trials for which the participants answered the comprehension task incorrectly (18.020%), RT data points that were extraordinarily short or long (e.g., shorter than 100 milliseconds (ms) or longer than 6000 ms for Region 8; histograms of raw RT data in each region are available on the corresponding author's OSF page, the URL of which is given above), and residuals that exceeded 3 *SD*s from the mean RT estimated by the final models (2.532%) (Arai and Roland, 2016, Baayen 2008, Baayen and Milin, 2010). Therefore, 20.552% of the data were excluded. Figure 1 presents the mean RT for each condition per region. The regions of interest are from Region 8 *feed*. Except for Regions 12–14, the statistical analysis results are reported for regions in which

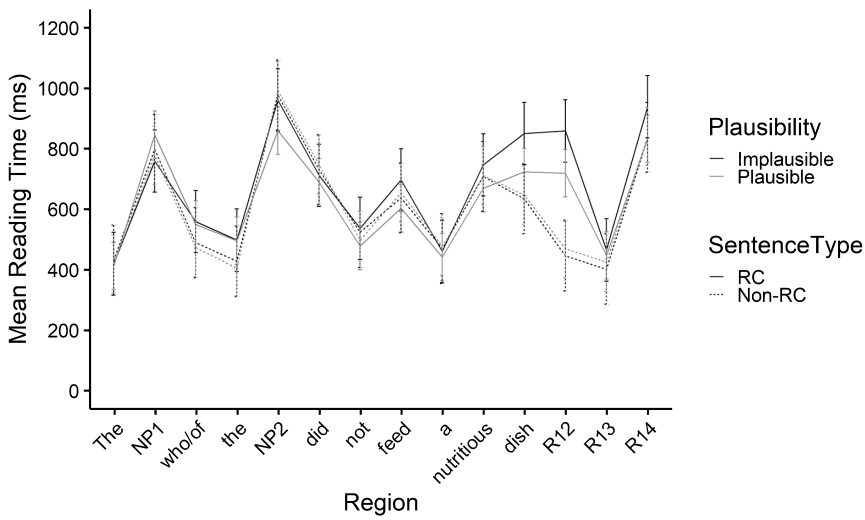


Figure 1. Mean RT for each condition per region in Experiment 1. Error bars indicate 95% confidence interval (CI) of the mean.

significant effects were observed. Other regions showed no significant effects (all *p*s > 0.100).

In all regions, a main effect of item order was significant, indicating that the RTs became shorter as the experiment proceeded (all *p*s < 0.050). In advance of the critical region (Region 8 *feed*), significant effects were found in Regions 2, 3, 4, and 7. Region 2 *NP1* exhibited a significant interaction of sentence type and plausibility (B = −91.681, *SE* = 45.392, *t* = −2.020, *p* = 0.044) and in Region 5 *NP2*, the interaction was marginally significant (B = 130.646, *SE* = 71.435, *t* = 1.829, *p* = 0.068). These interaction effects are likely due simply to the word difference among the conditions as different words, *parents* or *infant*, were used in those regions across the conditions (e.g., for Region 2, *parents* was used in the RC, implausible (4a) and the non-RC, plausible (4d) conditions). An effect of sentence type was found to be significant in Region 3 *who* or *of* (*B* = −75.164, *SE* = 13.782, *t* = −5.454, *p* < 0.001) and Region 4 *the* (*B* = −84.128, *SE* = 11.998, *t* = −7.012, *p* < 0.001), in which an RT was longer in the RC conditions than in the non-RC conditions. An effect of plausibility was significant in Region 7 (*B* = −39.926, *SE* = 17.567, *t* = −2.279, *p* = 0.023), reflecting a longer RT in the implausible conditions than in the plausible conditions.

Table 3 shows the results for the critical region (Region 8 *feed*). This table shows that the effect of plausibility was marginally significant. Crucially, an interaction between sentence type and plausibility was significant because of an RT slowdown in the RC, implausible condition (4a), as illustrated in Figure 2. Pairwise comparisons revealed a significant simple effect of plausibility in the RC conditions, reflecting the longer RT in the implausible condition (4a) than in the plausible condition (4b) (*B* = −91.043, *SE* = 33.040, *t* = −2.756, *p* = 0.007). However, no such effect was found in the non-RC conditions (*B* = 12.321, *SE* = 34.823, *t* = 0.354, *p* = 0.724). Furthermore, the paired comparisons revealed a marginally significant simple effect of sentence type in the implausible conditions due to an RT slowdown in the RC condition (4a) (*B* = −61.723, *SE* = 33.782, *t* = −1.827, *p* = 0.071).

Table 3. Summary of fixed effect estimates of the LME model of the RT data in Region 8 *feed* in Experiment 1.

| | B | *SE* | *t* | *p* | |
|---|---|---|---|---|---|
| (intercept) | 648.962 | 34.517 | 18.801 | <0.001 | * |
| sentence type | −10.267 | 22.730 | −0.452 | 0.652 | |
| plausibility | −41.925 | 22.620 | −1.853 | 0.065 | † |
| sentence type×plausibility | 103.391 | 50.655 | 2.041 | 0.049 | * |
| item order | −68.892 | 12.014 | −5.734 | <0.001 | * |

†: *p* < 0.100, *: *p* < 0.050, ||: no correlation was calculated with random factors

Final Model: lmer (RT ~ sentence type + plausibility + sentence type×plausibility + (1 | participant) + (1 + sentence type×plausibility || set) + item order)

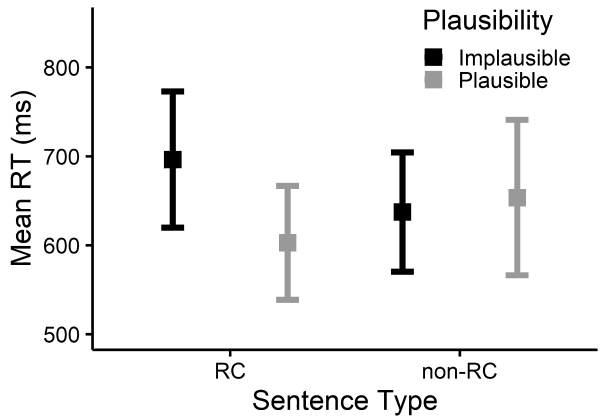


Figure 2. Mean RT for each condition in Region 8 *feed* in Experiment 1. Error bars indicate 95% CI of the mean.

In the post-critical regions, as detailed in Table 4 and Figure 3, Region 11 *dish* showed significant effects of sentence type and plausibility and a significant interaction between them. Pairwise comparisons found a significant simple effect of plausibility in the RC conditions (B = −121.560, *SE* = 33.090, $t$ = −3.674, $p$ < 0.001), with a longer RT in the implausible condition (4a) than in the plausible condition (4b), and a significant simple effect of sentence type in both the implausible and plausible conditions, reflecting an RT slowdown in the RC conditions (the implausible conditions: B = −229.293, *SE* = 33.886, $t$ = −6.767, $p$ < 0.001; the plausible conditions: B = −94.292, *SE* = 33.512, $t$ = −2.814, $p$ = 0.005).

Table 4. Summary of fixed effect estimates of the LME model of the RT data in Region 11 *dish* in Experiment 1.

| | B | *SE* | *t* | *p* | |
|---|---|---|---|---|---|
| (intercept) | 717.146 | 34.517 | 18.801 | <0.001 | * |
| sentence type | −161.848 | 23.883 | −6.777 | <0.001 | * |
| plausibility | −56.874 | 23.883 | −2.385 | 0.017 | * |
| sentence type×plausibility | 135.000 | 47.550 | 2.839 | 0.005 | * |
| item order | −32.713 | 47.550 | −2.592 | 0.009 | * |

*: $p$ < 0.050

Final Model: lmer (RT ~ sentence type + plausibility + sentence type×plausibility + (1 | participant) + (1 | set) + item order)

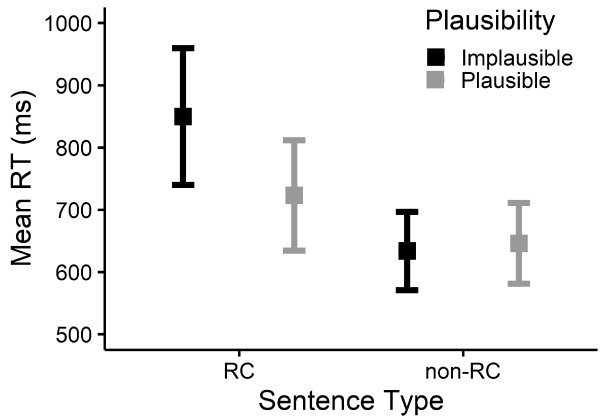


Figure 3. Mean RT for each condition in Region 11 *dish* in Experiment 1. Error bars indicate 95% CI of the mean.

### 3.5. Discussion

The purpose of Experiment 1 was to test the following hypothesis: when processing filler-gap dependencies, Japanese learners of L2 English would postulate an object gap immediately after the sentence processor finds a verb in Region 8. If this were the case, plausibility mismatch effects would be observed in the verb region. Specifically, there would be a longer RT for the verb in the RC, implausible condition than in the RC, plausible condition, and there would be no such RT collapse in the non-RC conditions.

The results of Experiment 1 were consistent with the hypothesis. The LME models revealed significant interaction between sentence type and plausibility in Region 8 *feed*, and the pairwise comparisons confirmed that the interaction was induced by an RT slowdown because of the plausibility mismatch effects in the RC, implausible condition (4a) and the longer RT in the RC, implausible condition (4a) than in the RC, plausible (4b) and the Non-RC, implausible conditions (4c). Given that only the implausible condition that included thematic-role reversal manipulation exhibited the RT slowdown in the verb region, the plausibility mismatch effect in the region should be accounted for by the implausibility induced by the thematic-role reversal.

Although there was a concern that the implausible relationship between the subject NP *the infant* and the verb *feed* alone might induce plausibility in the verb region, the RC, implausible condition (4a) elicited a longer RT than did the non-RC counterpart (4c), which also contained *the infant* as the subject and *feed* as the verb. Therefore, the RT slowdown in the RC, implausible condition (4a) should be due to the thematic-role reversal implausibility. Thus, the plausibility mismatch effects found in the verb region suggest that the processor incrementally posits an object gap as soon as it processes the verb during the comprehension of English filler-gap dependencies by Japanese learners (i.e., incremental verb-driven parsing).

This finding is consistent with the previous findings that L2 learners of English complete the filler-gap dependency formation at least immediately once verb information is available (Cunnings et al. 2010, Felser et al. 2012, Kim et al. 2015, Omaki and Schulz 2011).

Moreover, the sensitivity to plausibility mismatch due to thematic-role reversal indicates that Japanese learners of English compute the interpretation of the RC as soon as they reach the RC verb, although not all of the arguments are revealed at this moment. At the verb, it turns out that the filler, *the parents* or *the infant*, is the indirect object of the verb *feed*. However, no direct object NP appears at that moment. Nevertheless, the processor completes the filler-gap dependency formation by postulating the gap in the indirect object position, further initiates the interpretation process, and detects the plausibility mismatch induced by the thematic-role reversal. Thus, the plausibility mismatch effect in Region 8 suggests that Japanese learners of English form filler-gap dependencies at least as soon as the verb is encountered.

In addition, the significant effect of sentence type in Region 11 *dish* is also likely due to the bottom-up semantic computation process. In this region, the direct object of the double object construction *dish* is unfolded in the RC conditions, and every element of the double object RC then becomes available. For example, in (4a), the subject is *the infant*, the verb is *feed*, the indirect object is *the parents*, and the direct object is *a nutritious dish*. The processor executes semantic computation of the RC again using the bottom-up information. In the RC, implausible condition (4a), the interpretation turns out to be ultimately implausible, and RTs thus slowed down in this condition compared to the RC, plausible condition (4b).

The implausibility of the final interpretation of the RC is likely to be attributed to the comprehension question accuracy. The results of the comprehension question demonstrate that the participants correctly understood the target sentences in most cases; they correctly answered the comprehension question in more than 80% of the target trials on average. However, the logistic mixed-effects model exhibited a significant interaction of sentence type and plausibility on the proportion of correct answers. The participants gave wrong answers more in the RC, implausible condition (4a) than in the RC, plausible condition (4c), whereas the mean accuracy was lower in the non-RC, plausible condition (4d) than in the non-RC, implausible condition (4c). The low accuracy in the RC, implausible condition (4a) is likely to reflect the implausibility of the RC sentences. Owing to using a thematic-role reversal design to manipulate the semantic plausibility of filler-gap dependencies, the final interpretation of the RC became implausible in the sense that the NP *the infant* is a thematically unnatural subject of the verb *feed*, which takes the NP *the parents* as the object. The participants were sensitive to this implausibility and therefore sometimes failed to correctly compute the interpretation of the entire sentence in the RC, implausible condition (4a), although the comprehension question never asked about the interpretation of the RC. Nevertheless, there was no difference in the mean accuracy rate between the RC,

implausible (4a) and the non-RC, implausible (4c) conditions, and both conditions showed a relatively high accuracy rate (over 80%). Given this, the implausibility of the final interpretation of the entire sentence might not influence the participants' processing performance drastically.

Overall, Experiment 1 demonstrates that Japanese learners of English postulate an object gap and compute the plausibility of the RC at least immediately after reading the verb. However, it is still unclear whether they perform this process without depending on bottom-up verb information even before such information is available (i.e., predictive filler-driven parsing) or whether it is dependent on bottom-up information (i.e., incremental verb-driven parsing). Therefore, Experiment 2 examined this issue by conducting another self-paced reading experiment with a transitivity mismatch paradigm.

## 4. Experiment 2

Experiment 2 sought to explore whether Japanese learners of English adopt filler-driven parsing, that is, whether they postulate an object gap before confirming that a verb is transitive. Experiment 1 found that an object gap is postulated at least after the processor encounters a verb, suggesting that the processor posits it at least as soon as it receives the verb information. On the other hand, the filler-driven hypothesis predicts that the search for gap positions starts immediately after the processor detects a filler. Therefore, this hypothesis was tested on the L2 English processing by Japanese learners using a transitivity mismatch paradigm similar to that used in Omaki et al. (2015).

### 4.1. Materials

An example set of experimental sentences is shown in (5), in which two factors were crossed: transitivity (intransitive vs. transitive) × sentence type (RC vs. non-RC). The verbs were intransitive in the intransitive conditions (5a) and (5b). By contrast, the verbs were transitive in the transitive conditions (5c) and (5d). Thus, a direct object position was available in the transitive, RC condition (5c). However, in the intransitive, RC condition (5a), a gap was not able to be placed at the direct object positions because intransitive verbs do not hold a direct object position.

(5) a intransitive, RC:
The$_1$ colleague$_2$ who$_3$ the$_4$ worker$_5$ did$_6$ not$_7$ talk$_8$ or$_9$ debate$_{10}$ about$_{11}$ dropped$_{12}$ a$_{13}$ newspaper.$_{14}$

b intransitive, non-RC:
The$_1$ colleague$_2$ of$_3$ the$_4$ worker$_5$ did$_6$ not$_7$ talk$_8$ or$_9$ debate$_{10}$ about$_{11}$ his$_{12}$ own$_{13}$ brothers.$_{14}$

c transitive, RC:
The$_1$ colleague$_2$ who$_3$ the$_4$ worker$_5$ did$_6$ not$_7$ respect$_8$ or$_9$ scorn$_{10}$ sometimes$_{11}$ drove$_{12}$ a$_{13}$ car.$_{14}$

d transitive, non-RC:
The$_1$ colleague$_2$ of$_3$ the$_4$ worker$_5$ did$_6$ not$_7$ respect$_8$ or$_9$ scorn$_{10}$ his$_{11}$ own$_{12}$

elder$_{13}$ brothers.$_{14}$

### 4.2. Prediction

If the processing of English filler-gap dependencies by Japanese learners is driven by fillers, it is expected that the sentence processor would initiate a search for a gap immediately after finding a relative pronoun. Therefore, the processor would create the direct object gap before the verb in the RC conditions because the earliest gap position after the subject position is the direct object position. In other words, it is predicted that the processor would anticipate a transitive verb. If this is the case, an RT for the intransitive verb *talk* in the RC condition (5a) is expected to be longer than in the non-RC condition (5b) (i.e., transitivity mismatch effects), but there would be no such RT difference in the transitive conditions (5c) and (5d). Statistically, an interaction between transitivity and sentence type is expected to be significant.

By contrast, if Japanese learners' filler-gap dependency construction is triggered by the verbs, no transitivity mismatch effects would be found in the intransitive conditions because there is no bottom-up motivation to anticipate a transitive verb. Only after the verb is processed would the Japanese learners decide a gap position using bottom-up verb information. Therefore, if verbs trigger filler-gap dependency formation, there would be no transitivity mismatch effects, and there would further be a longer RT in the RC conditions than in the non-RC conditions in both the transitive and intransitive conditions because of the processing costs associated with bottom-up gap postulation (i.e., filler-gap dependency formation). That is, in sentences with filler-gap dependency, it is expected that the RC verb would trigger the dependency formation, resulting in a higher processing load than in sentences without a filler-gap dependency. An effect of sentence type is expected to be significant in the statistical analysis if the dependency formation is driven by verbs.

### 4.3. Data analysis

The same data analysis procedure was used as in Experiment 1. Response data from the comprehension question tasks were submitted to logistic mixed-effects regression models, and RT data were submitted to LME models, in which transitivity, sentence type, their interaction, and item order were the fixed effects. Each condition was coded "0" or "1" depending on transitivity (intransitive vs. transitive) and sentence type (RC vs. non-RC), with "0" being given for the intransitive and RC conditions. The coding values used for the analysis were centered.

### 4.4. Results

#### 4.4.1. Comprehension accuracy

In Experiment 2, the mean rate of comprehension accuracy was 76.495% (*SE* = 1.495). Table 5 summarizes the mean for each condition, and Table 6 shows that the logistic regression model exhibited no significant effects. The results indicate that the participants paid attention to the experiment well, although the mean

accuracy was lower than that of Experiment 1. No further discussion is given on the results of the comprehension question accuracy as no significant effect was found across the conditions.

Table 5. Mean accuracy rates for the comprehension task in Experiment 2.

| | Mean Accuracy | *SE* |
|---|---|---|
| intransitive, RC | 75.272% | 2.815 |
| intransitive, non-RC | 77.989% | 2.272 |
| transitive, RC | 77.446% | 2.566 |
| transitive, non-RC | 75.272% | 2.074 |

Table 6. Summary of fixed effect estimates of the logistic mixed-effects regression model of the comprehension task response data in Experiment 2.

| | B | *SE* | *z* | *p* |
|---|---|---|---|---|
| (intercept) | 1.714 | 0.236 | 7.267 | <0.001 * |
| transitivity | −0.026 | 0.039 | −0.683 | 0.495 |
| sentence type | 0.050 | 0.039 | 1.304 | 0.192 |
| transitivity×sentence type | −0.394 | 0.595 | −0.663 | 0.508 |
| item order | 0.003 | 0.021 | 0.140 | 0.889 |

*: $p < 0.050$

Final Model: glmer (response data ~ transitivity + sentence type + transitivity×sentence type + (1 + transitivity×sentence type | participant) + (1 + transitivity× sentence type | set) + item order, family = "binomial")

### 4.4.2. Reading time

The followings were removed from the analyses: trials for which the participants answered the comprehension task incorrectly (23.493%), RT data points that were extraordinarily short or long (e.g., shorter than 100 ms or longer than 4000 ms for Region 8; distribution histograms of the raw RT data were also available on the OSF page), and residuals that exceeded 3 *SD*s from the mean RT estimated by the final models (2.660%) (Arai and Roland 2016, Baayen 2008, Baayen and Milin 2010). Therefore, 26.153% of the data were excluded via the data screening.

The region-by-region mean RTs for each condition are illustrated in Figure 4. The regions of interest were Region 8, in which the verbs appeared, and Region 9 because the effects in Region 8 could spill over. The regions of interest or those that showed a significant effect are reported, except for Regions 11–14 because lexically different words were used across the conditions in these regions. Other effects were not significant (all *p*s > 0.100).

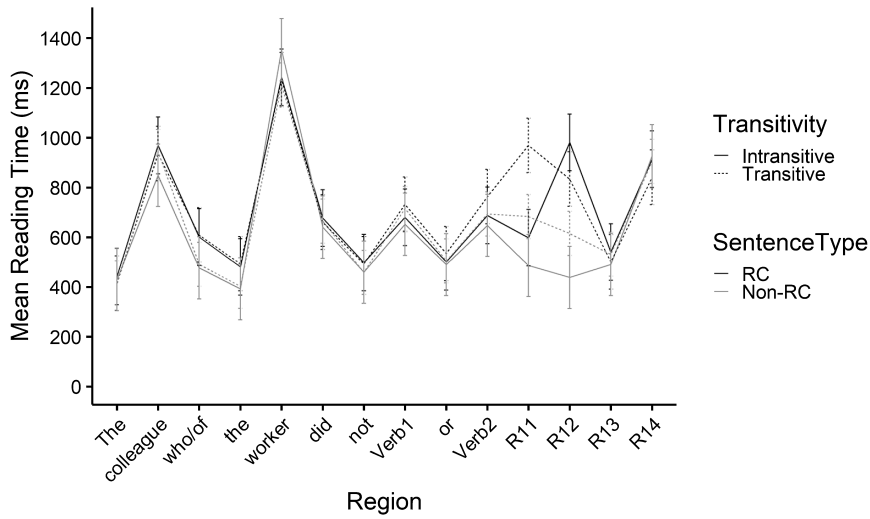


Figure 4. Mean RT for each condition per region in Experiment 2. Error bars indicate 95% confidence interval (CI) of the mean.

An effect of item order was significant in all regions (all *p*s < 0.050). Prior to the critical regions, a significant effect of transitivity was found in Region 1 *The*, with a longer RT in the intransitive conditions ($p < 0.050$), and an effect of sentence type was found to be significant in Region 3 *who* or *of*, Region 4 *the*, and Region 7 *not*, reflecting a longer RT in the RC conditions than in the non-RC condition (*p*s < 0.050). An interaction of these two factors was marginally significant in Region 2 *colleague* ($p = 0.092$).

The critical region (Region 8 *talk* or *respect*) exhibited a significant effect of transitivity owing to a longer RT in the transitive conditions than in the intransitive conditions, as seen in Table 7 and Figure 5. However, neither an effect of sentence type nor an interaction of transitivity and sentence type was significant.

Table 7. Summary of fixed effect estimates of the LME model of the RT data in Region 8 *talk* or *respect* in Experiment 2.

| | B | *SE* | *t* | *p* |
|---|---|---|---|---|
| (intercept) | 699.580 | 38.863 | 18.001 | <0.001 * |
| transitivity | 51.301 | 24.931 | 2.058 | 0.040 * |
| sentence type | −21.450 | 25.029 | −0.857 | 0.392 |
| transitivity×sentence type | 26.958 | 49.923 | 0.540 | 0.589 |
| item order | −65.450 | 13.241 | −4.943 | <0.001 * |

*: $p < 0.050$

Final Model: lmer (RT ~ transitivity + sentence type + transitivity×sentence type + (1 | participant) + (1 | set) + item order)

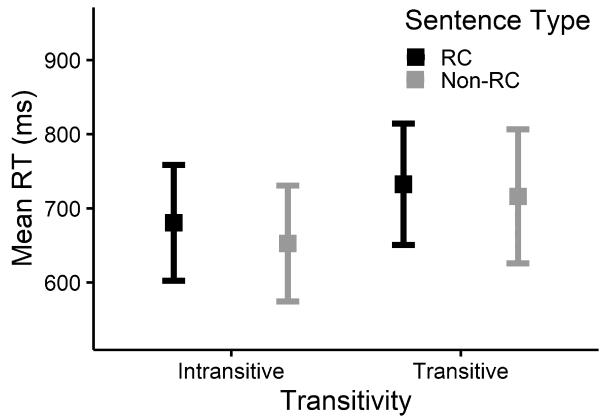


Figure 5. Mean RT for each condition in Region 8 *talk* or *respect* in Experiment 2. Error bars indicate 95% CI of the mean.

In the spillover region, Region 9 *or*, the model revealed marginally significant effects of transitivity and sentence type, although their interaction was not significant, as shown in Table 8. The effect of transitivity reflected the same RT pattern as observed in Region 8, that is, there was a longer RT in the transitive conditions than in the intransitive conditions. The effect of sentence type was evidenced by an RT slowdown in the RC conditions relative to an RT in the non-RC conditions. The RT pattern for each condition is shown in Figure 6. Region 10 *debate* or *scorn* showed a significant effect of sentence type (B = −53.807, *SE* = 22.340, *t* = −2.409, *p* = 0.016) due to a longer RT in the RC conditions than in the non-RC conditions.

Table 8. Summary of fixed effect estimates of the LME model of the RT data in Region 9 *or* in Experiment 2.

| | B | *SE* | *t* | *p* | |
|---|---|---|---|---|---|
| (intercept) | 509.322 | 15.261 | 33.375 | <0.001 | * |
| transitivity | 24.027 | 12.944 | 1.856 | 0.064 | † |
| sentence type | −25.316 | 12.963 | −1.953 | 0.051 | † |
| transitivity×sentence type | −4.031 | 25.878 | −0.156 | 0.876 | |
| item order | −35.125 | 6.806 | −5.161 | <0.001 | * |

†: $p < 0.100$, *: $p < 0.050$

Final Model: lmer (RT ~ transitivity + sentence type + transitivity×sentence type + (1 | participant) + (1 | set) + item order)

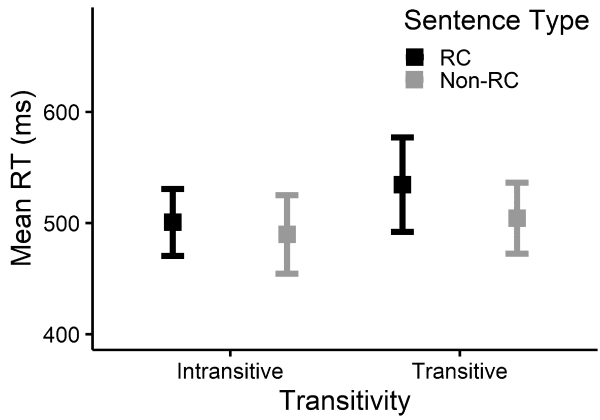


Figure 6. Mean RT for each condition in Region 9 *or* in Experiment 2. Error bars indicate 95% CI of the mean.

### 4.5. Discussion

Experiment 2 examined whether Japanese learners of English process filler-gap dependencies predictively by positing a gap before bottom-up verb information becomes available (i.e., predictive filler-driven parsing). In other words, Experiment 2 investigated whether Japanese learners postulate a direct object gap in advance of input of the actual verb and, as a consequence, predict a transitive verb during the processing of English filler-gap dependencies. If the dependency formation were driven by fillers, a transitivity mismatch effect should occur in the verb region; there would be a longer RT in the intransitive, RC condition (5a) than in the intransitive, non-RC condition (5b) and no such differences in the transitive conditions. However, no transitivity mismatch effect was observed in Experiment 2. The RT for the intransitive verb *talk* in the RC condition (5a) was not longer than in the non-RC condition (5b). Therefore, there was no supporting evidence for filler-driven parsing in the processing of filler-gap dependencies by Japanese learners of English.

Instead, the effect of sentence type was significant in the spillover regions, Regions 9 and 10, in which the RT was longer in the RC conditions than in the non-RC conditions. This finding indicated that the Japanese learners of English experienced more difficulty processing the verbs in the RC conditions than those in the non-RC conditions. This result is in line with the verb-driven hypothesis because it expects that the sentence processor initiates filler-gap dependency formation after the verb is encountered. The RT slowdown in the RC conditions could reflect a processing load associated with dependency formation based on bottom-up verb information. If this is the case, the results from Experiment 2 support the verb-driven analysis.

In the transitive, RC condition, the longer RT in this condition can be attributed to an additional processing load associated with the verb-driven object

gap postulation. That is, it is likely to be the case that the RT became slower in the transitive, RC condition because the sentence processor initiated object gap postulation as soon as it confirmed that the transitive verb could hold an object, resulting in a greater processing load than in the transitive, non-RC condition. The intransitive verbs also provide the processor rich bottom-up information about the gap position because most intransitive verbs used in the experiment took a prepositional phrase (PP) complement (e.g., the intransitive verb *talk* in (5) took a complement PP with a preposition *about*). Given this, the lexical information from the intransitive verbs was sufficient to lead the processor to posit a gap within the PP. The longer RTs for Regions 9 and 10 in the intransitive, RC conditions may be a consequence of this verb-driven PP-gap postulation.

Alternatively, there remains another possibility that is consistent with the verb-driven parsing hypothesis, that is, the RT slowdowns in the RC conditions in Regions 9 and 10 might be due to a filled-gap effect.[7] In both the transitive, RC and intransitive, RC conditions, an appearance of the disjunction *or* is not consistent with the expected sentence structures based on the verb-driven bottom-up processing because both the transitive and intransitive verbs do not generate an expectation of the disjunction. As a result, the disjunction *or* could cause additional processing costs due to the mismatch between the already constructed structural representation and the actual input (i.e., the filled-gap effect). Regardless of whether those effects were due to the gap postulation process or the filled-gap effect, the RT slowdowns in Regions 9 and 10 support the bottom-up verb-driven parsing hypothesis that Japanese learners of English initiate the filler-gap dependency formation only after the verb is encountered because no transitivity mismatch effect was observed.

Nonetheless, there might remain a concern that there were no transitivity mismatch effects simply because the Japanese learners merely misinterpreted the intransitive verbs as transitive. However, the intransitive verbs used in Experiment 2 were not so difficult that the Japanese learners, whose English proficiency level was upper-intermediate, were not able to precisely process those intransitive verbs. Six of the eight intransitive verbs in the experiment (*complain*, *debate*, *worry*, *respond*, *smile*, and *agree*) are listed in the New JACET 8000 (JACET Basic Words Revision Committee 2016) as words to be learned during high school, and the other two intransitive verbs (*chat* and *bow*) are listed as required words for the university entrance exam. Therefore, the learners may not have incorrectly processed these intransitive verbs as transitive ones. Hence, the lack of transitivity mismatch effects in Experiment 2 cannot be attributed to just misinterpretation of the verbs by Japanese learners.

Overall, Experiment 2 demonstrates that Japanese learners of English are unlikely to form filler-gap dependencies in advance of the verb. Specifically, given

[7] We would like to express our appreciation for a comment from an anonymous reviewer on this possibility.

that they might postulate a gap only after reading the verb, they do not predict gap positions. These results seem to suggest that the processing of English filler-gap dependencies by Japanese learners is driven by verbs rather than fillers, that is, incremental verb-driven parsing. This finding is consistent with findings in previous studies that argue that L2 learners have difficulty generating expectations in general (Dussias et al. 2013, Grüter et al. 2014, 2017, Hopp 2013, Ito et al. 2017, Kaan 2014, Kaan et al. 2010, Mitsugi and Macwhinney 2016). These studies explored the predictions in L2 processing at the discourse, semantic, or lexical levels. The present study extended these implications to syntactic predictions in the L2 processing of filler-gap dependencies.

## 5. General discussion

The present study examined whether filler-gap dependency formations are triggered by fillers or verbs (filler-driven parsing vs. verb-driven parsing) in Japanese learners' processing of English filler-gap dependencies. Experiment 1 first examined whether Japanese learners of English create filler-gap dependencies as soon as the verb is encountered. Previous studies on L2 comprehension of English filler-gap dependencies have found that L2 learners complete dependency formation immediately after the verb. However, no research has demonstrated that Japanese learners use the same parsing mechanism. Therefore, Experiment 1 addressed this issue by using a plausibility mismatch paradigm. The results showed reliable plausibility mismatch effects in a verb region, with an RT disruption at the verb in implausible RC sentences, such as *the parents who the infant did not feed a nutritious dish lost weight rapidly*. This provides evidence of the hypothesis that Japanese learners of English accomplish the filler-gap dependency formation at least immediately after the verb. However, it remained unclear whether the dependency construction by Japanese learners is driven by fillers or verbs because the results of Experiment 1 could be explained with either the filler- or verb-driven parsing hypotheses, considering that both hypotheses assume that the processor posits a gap at least after the verb appears. Therefore, to obtain greater clarity, Experiment 2 directly tested these two hypotheses by focusing on a transitivity mismatch effect.

Experiment 2 investigated whether Japanese learners of English predictively postulate gaps even before confirming bottom-up verb information. The filler-driven parsing hypothesis expects that intransitive verbs induce an RT disruption due to a transitivity mismatch effect in a filler-gap dependency environment if Japanese learners predict a transitive verb as a consequence of the preverbal object gap postulation. However, no such effects were found in the current experiment. Instead, an RT for the verbs was slower in RC sentences than in non-RC sentences regardless of verb transitivity. This result is most likely to suggest that Japanese learners of English initiate the filler-gap dependency formation as soon as bottom-up verb information becomes available, and the RTs for both transitive and intransitive verbs therefore increased, reflecting processing costs associated with the verb-driven gap postulation process. Thus, Japanese learners are not

likely to postulate an object gap before reading the verb or predict verb transitivity. Experiment 2 supports the verb-driven parsing hypothesis that Japanese learners of English initiate filler-gap dependency building only after the verb is encountered.

The present study offers an implication for psycholinguistic research that aims to reveal prediction mechanisms in the human sentence comprehension system in the sense that it extends the generality of the hypothesis that predictive processes are more restricted in L2 comprehension compared to L1 comprehension (e.g., Grüter et al. 2017, Kaan 2014) to the processing of a structurally complex sentence, filler-gap dependency. In other words, the present finding is inconsistent with the argument that L2 learners process filler-gap dependencies with the same time-course fashion as do native speakers (Cunnings 2017, Felser 2015). However, further investigation is needed considering several limitations of the present study.

Although the present study indicates the possibility that L2 learners do not process filler-gap dependencies as predictively as do native speakers of English, no implication is given as to the issue of what induces L2 learners' difficulty of generating predictions during the processing of filler-gap dependencies. Previous studies have suggested that either L2 proficiency (Dussias et al. 2013, Hopp 2013) or cognitive resources (Grüter et al. 2014, 2017) might cause difficulty with predictive processing in L2 sentence comprehension. This suggestion would warrant future research that examines a relationship between those factors and predictive filler-driven parsing in L2 comprehension by conducting experiments with L2 learners of English who have various individual differences. Furthermore, exploration of the influence of processing strategy in L1 to predictive processing in L2 would be necessitated. The present study only focused on Japanese learners of English, whose L1 has linguistically different properties from English in terms of canonical word order and wh-movement. It would be worth investigating whether L2 learners of English whose L1 is typologically closer to English predictively process English filler-gap dependencies.

Another possible factor for non-predictive L2 processing in a filler-gap dependency environment is related to structural revision (Jacob and Felser 2016, Jessen and Felser 2019, Roberts and Felser 2011). The filler-driven parsing hypothesis proposes that the sentence processor begins dependency formation as soon as it detects a filler. Assuming that the first available gap position is the subject position in most cases, the filler-driven hypothesis implies that structural revision is required when the subject position is filled with a lexical NP, such as in our experimental sentences, if the processor predictively postulates the subject gap. The processor is then unable to predict a direct object gap unless it correctly revises the initial subject-gap analysis immediately after the appearance of the actual subject NP. However, it has been argued that it is difficult for L2 learners to reanalyze a sentence structure that was built during the initial processing (Jacob and Felser 2016, Jessen and Felser 2019, Roberts and Felser 2011), thus indicating that revision difficulty in L2 comprehension may result in a weak or no preverbal prediction for a direct object gap or verb transitivity.

It may be possible to eliminate the potential effects of revision difficulty on the predictive filler-gap processing in L2 English by using sentences with *whom*, such as *the colleague whom the worker talked about yesterday quit his job*. The relative pronoun *whom* could rule out the possibility of positing a gap at the subject position and would allow for the testing of preverbal predictive processing by avoiding the influence of revision failure at the subject position. Another possible direction could be an exploration of preverbal subject-gap prediction in L2 English. Considering that the subject filled-gap effect also provides evidence for the filler-driven dependency formation, it would be worth investigating the effects in regard to the subject NPs. Although several previous studies aimed to test subject filled-gap effects in L2 English (Aldwayan, Fiorentino, and Gabriele 2010, Canales 2012, Johnson, Fiorentino, and Gabriele 2016), those studies seem to involve a methodological problem in that there remains a concern that the effects that they interpreted as the subject filled-gap effects were merely due to spillover effects of lexical difference in a previous region. They compared RTs of subject NPs like *Barbara* in sentences such as *My brother asked who Barbara will photograph us beside at the graduation* vs. *My brother asked if Barbara will photograph us beside Mom at the graduation*. Therefore, the word previous to the critical region was not lexically controlled, that is, *who* vs. *if*. Hence, further research is needed to reveal a more precise relationship between the revision process and the preverbal object prediction in L2 comprehension.

Another concern of the present study is the effect of participants' L2 proficiency level. The participants in this study are primarily at the upper-intermediate level of English proficiency, and Experiment 2 did not show evidence that they process English filler-gap dependencies predictively to the same extent as do native speakers of English. However, L2 learners of English who have higher proficiency may engage in predictive processing. In fact, previous research has argued that the more proficient L2 learners become, the more predictive their processes are (Dussias et al. 2013, Hopp 2013). Thus, the current implication is restrictive in that the results infer that Japanese learners of English, at least whose English proficiency is at the upper-intermediate level, do not form English filler-gap dependencies as predictively as do native speakers of English. Further research is needed to evaluate the influence of L2 proficiency on predictive processing in L2 comprehension.

Moreover, the present study has a methodological limitation. This study used a self-paced reading method, in which participants were asked to read sentences phrase-by-phrase by pressing a button to unfold the next phrase. This method is useful for research that focuses on predictive processing in the sense that it potentially encourages them to predict the upcoming phrases by prohibiting them from seeing adjacent phrases parafoveally. However, the button pressing is a secondary task that would impose some cognitive resource requirement upon participants. Taking into account that limitation of cognitive capacity in L2 comprehension is a potential factor of L2 learners' difficulty in generating predictions (Grüter et al. 2014, 2017), other experimental methodology should be considered that allows

participants to read sentences in a more natural way such as an eye-tracking reading experiment.

## 6. Conclusion

This study investigated the timing of filler-gap dependency formations in the processing of L2 English by Japanese learners, conducting two self-paced reading experiments. The results from Experiment 1 revealed that Japanese learners of English postulate an object gap immediately after reading the verb. However, Experiment 2 did not present any reliable evidence that they predict an object gap prior to the appearance of the actual verb. The results of Experiment 2 are consistent with the verb-driven parsing hypothesis that the sentence processor initiates gap-filling processes only after encountering the verb. Therefore, this study proposes that filler-gap dependency formations are triggered by bottom-up verb information in online L2 comprehension. However, further investigation is required to reveal what drives L2 learners' difficulty in predicting gap positions without bottom-up information.

## References

Aldwayan, Saad, Robert Fiorentino, and Alison Gabriele (2010) Evidence of syntactic constraints in the processing of wh-movement: A study of Najdi Arabic learners of English. In: Bill VanPatten and Jill Jegerski (eds.) *Research in second language processing and parsing*, 65–86. Amsterdam: John Benjamins Publishing Company.

Allan, D. (2004) *Oxford placement test 2*. Oxford: Oxford University Press.

Altmann, Gerry T.M. and Yuki Kamide (1999) Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73: 247–264.

Aoshima, Sachiko, Colin Phillips, and Amy Weinberg (2004) Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51: 23–54.

Arai, Manabu and Douglas Roland (2016) Gengo rikai kenkyuu ni okeru gankyuu undoo deeta oyobi yomizikan deeta no tookei kaiseki [Statistical analysis of eye-movement data and reading time data in language comprehension research]. *Tookei Suuri* [*Proceedings of the Institute of Statistical Mathematics*] 64: 201–231.

Arnold, Jennifer E. (2001) The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 31: 137–162.

Baayen, R. Harald (2008) *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Douglas Davidson, and Douglas Bates (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412.

Baayen, R. Harald and Petar Milin (2010) Analyzing reaction times. *International Journal of Psychological Research* 3: 12–28.

Barss, Andrew and Howard Lasnik (1986) A note on anaphora and double objects. *Linguistic Inquiry* 17: 347–354.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth, and Harald Baayen (2015) Parsimonious mixed models. *Available from ArXiv:1506.04967v2 (Stat.ME)*.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker (2015) Fitting linear

mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48.
Canales, Alonso Jose (2012) *Online processing of wh-dependencies in English by native speakers of Spanish*. Doctoral dissertation, University of Kansas University of Kansas.
Chomsky, Noam (1973) Conditions on transformations. In: Stephen R. Anderson and Paul Kiparsky (eds.) *A Festschrift for Morris Halle*, 232–286. New York: Holt: Rinehart & Winston.
Chomsky, Noam (1986) *Barriers*. Cambridge, MA: MIT Press.
Chow, Wing-Yee, Cybelle Smith, Ellen Lau, and Colin Phillips (2016) A "bag-of-arguments" mechanism for initial verb predictions. *Language, Cognition and Neuroscience* 31: 577–596.
Crain, Stephen and Janet Dean Fodor (1985) How can grammars help parsers? In: David R. Dowty, Lauri Karttunen, and Arnold M. Zwicky (eds.) *Natural Language Parsing*, 94–128. Cambridge: Cambridge University Press.
Cunnings, Ian (2017) Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition* 20: 659–678.
Cunnings, Ian, Claire Batterham, Claudia Felser, and Harald Clahsen (2010) Constraints on L2 learners' processing of wh-dependencies: Evidence from eye movements. In: Bill VanPatten and Jill Jegerski (eds.) *Language Acquisition and Language Disorders*, 87–110. Amsterdam: John Benjamins Publishing Company.
Dallas, Andrea, Gayle DeDe, and Janet Nicol (2013) An event-related potential (ERP) investigation of filler-gap processing in native and second language speakers: ERP study of L2 filler-gap processing. *Language Learning* 63: 766–799.
Dallas, Andrea and Edith Kaan (2008) Second language processing of filler-gap dependencies by late learners. *Language and Linguistics Compass* 2: 372–388.
Dussias, Paola E., Jorge R. Valdés Kroff, Rosa E. Guzzardo Tamargo, and Chip Gerfen (2013) When gender and looking go ahead in hand: Grammatical gender processing In L2 Spanish. *Studies in Second Language Acquisition* 35: 353–387.
Felser, Claudia (2015) Native vs. non-native processing of discontinuous dependencies. *Second Language* 14: 5–19.
Felser, Claudia, Ian Cunnings, Claire Batterham, and Harald Clahsen (2012) The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition* 34: 67–98.
Frazier, Lyn and Charles Clifton (1989) Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4: 93–126.
Grüter, Theres, Hannah Rohde, and Amy J. Schafer (2014) The role of discourse-level expectations in non-native speakers' referential choices. In: Will Orman and Matthew James Valleau (eds.) *Proceedings of the Boston University Conference on Language Development*, 179–191. Somerville, MA: Cascadilla Press.
Grüter, Theres, Hannah Rohde, and Amy J. Schafer (2017) Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism* 7: 199–229.
Hopp, Holger (2013) Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29: 33–56.
Hudson, Richard (1992) So-called 'double objects' and grammatical relations. *Language* 68: 251.
Ito, Aine, Andrea E. Martin, and Mante S. Nieuwland (2017) On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43: 635–652.

JACET Basic Words Revision Committee (2016) *The new JACET list of 8000 basic words*. Tokyo: Kirihara Shoten.

Jackendoff, Ray S. and Peter Culicover (1971) A reconsideration of dative movements. *Foundations of Language* 7: 397–412.

Jacob, Gunnar and Claudia Felser (2016) Reanalysis and semantic persistence in native and non-native garden-path recovery. *Quarterly Journal of Experimental Psychology* 69: 907–925.

Jaeger, T. Florian (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59: 434–446.

Jessen, Anna and Claudia Felser (2019) Reanalysing object gaps during non-native sentence processing: Evidence from ERPs. *Second Language Research* 35: 285–300.

Johnson, Adrienne, Robert Fiorentino, and Alison Gabriele (2016) Syntactic constraints and individual differences in native and non-native processing of wh-movement. *Frontiers in Psychology* 7: 549.

Just, Marcel A., Patricia A. Carpenter, and Jacqueline D. Woolley (1982) Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111: 228–238.

Kaan, Edith (2014) Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism* 4: 257–282.

Kaan, Edith, Andrea Dallas, and Frank Wijnen (2010) Syntactic predictions in second-language sentence processing. In: C. Jan-Wouter Zwart and Mark de Vries (eds.) *Structure preserved. Festschrift in the honor of Jan Koster*, 207–214. Amsterdam: John Benjamins Publishing Company.

Kamide, Yuki, Gerry T. M. Altmann, and Sarah L. Haywood (2003) The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49: 133–156.

Kehler, Andrew, Laura Kertz, Hannah Rohde, and Jeffrey L. Elman (2007) Coherence and coreference revisited. *Journal of Semantics* 25: 1–44.

Kim, Eunah, Soondo Baek, and Annie Tremblay (2015) The role of island constraints in second language sentence processing. *Language Acquisition* 22: 384–416.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13): 1–26.

Larson, Richard K. (1988) On the double object construction. *Linguistic Inquiry* 19: 335–391.

Lee, Ming-Wei (2004) Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33: 51–73.

McElree, Brian and Teresa Griffith (1998) Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *Journal of Memory and Language* 24: 432–460.

Mitsugi, Sanako and Brian Macwhinney (2016) The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition* 19: 19–35.

Miyamoto, Edson T. (2002) Case markers as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research* 31: 307–347.

Nakano, Yoko, Claudia Felser, and Harald Clahsen (2002) Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research* 31: 531–571.

Omaki, Akira, Ellen F. Lau, Imogen Davidson White, Myles L. Dakan, Aaron Apple, and

Colin Phillips (2015) Hyper-active gap filling. *Frontiers in Psychology* 6: 384.
Omaki, Akira and Barbara Schulz (2011) Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition* 33: 563–588.
Pickering, Martin (1993) Direct association and sentence processing: A reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes* 8: 163–196.
Pickering, Martin and Guy Barry (1991) Sentence processing without empty categories. *Language and Cognitive Processes* 6: 229–259.
R Core Team (2019) R: A language and environment for statistical computing Version 3.5.3 Vienna, Austria: R Foundation for Statistical Computing.
Roberts, Leah and Claudia Felser (2011) Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics* 32: 299–331.
Rohde, Hannah and William S. Horton (2014) Anticipatory looks reveal expectations about discourse relations. *Cognition* 133: 667–691.
Ross, John Robert (1967) *Constraints on variables in syntax*. Doctoral dissertation MIT.
Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and Julie C. Sedivy (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217): 1632–1634.
Traxler, Matthew J. and Martin J. Pickering (1996) Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35: 454–475.
Williams, John N. (2006) Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition* 9: 71–88.
Williams, John N., Peter Möbius, and Choonkyong Kim (2001) Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics* 22: 509–540.

*Author's contact information*:
Itsuki Minemi
College of Sport and Health Science, Ritsumeikan University
e-mail: minemi[at]fc.ritsumei.ac.jp

[Received 18 July 2019;
Accepted 9 February 2021]

## 【要　旨】

### 第二言語理解におけるフィラー・ギャップ依存関係形成のタイミング

峰見　一輝
東京大学／
日本学術振興会　特別研究員

矢野　雅貴
東京都立大学

本研究は，自己ペース読文実験を用いて，日本語母語話者が英語の関係節を読む際に，文中のどの時点で長距離依存関係を構築し始めるかを検証した。実験 1 では，関係節主要部と動詞の意味的関係が不自然である場合に，動詞位置で読み時間の増大が確認された。この結果は，日本語母語話者が少なくとも動詞を読んだ直後に，長距離依存関係を構築していることを示唆している。さらに，もし関係節動詞が出現する前に，予測的に依存関係が構築されているなら，動詞が自動詞であった場合，予測との不一致により読み時間が増大するはずである。しかし実験 2 では，そのような結果は観察されなかった。つまり，日本語母語話者は，動詞を読む前に予測的に依存関係を構築しているわけではない可能性が示唆された。本研究は，非母語における予測的処理は限定的であるという先行研究の知見が，長距離依存関係という複雑な構造処理のレベルにまで該当することを示している。