Examining the Suppression of Predictive Inferences Among Japanese EFL Readers

Shingo NAHATAME
Graduate School, University of Tsukuba

Abstract

Although making inferences (e.g., predictive inferences) during reading plays an important role in text comprehension, readers do not always make correct inferences and it is necessary to suppress the inferences when they are disconfirmed by the following context. However, the suppression of inferences has received little attention in previous research. Thus, the present study aimed to examine whether Japanese EFL readers suppress the inferential information which they activate during reading, focusing on predictive inferences. In addition, the effects of disconfirming the activated inferences on readers’ text representations were also investigated. In the experiment, 37 Japanese university students read several sets of short narratives and engaged in probe verification and cued recall tasks. The result of the verification task showed that the activation of predictive inferences remained immediately after they were disconfirmed by the context. Moreover, the results of the cued recall task revealed that the disconfirmed inferences were not deleted but were maintained in readers’ final text representations and that the disconfirmation had a negative impact on the comprehension of explicit text information. The present study suggests there is difficulty in suppressing the activated predictive inferences among EFL readers, and provides further understanding of the nature of inferential information in reading.

1. Review of Previous Studies

Successful text comprehension involves not only understanding information actually mentioned in the text, but also understanding information not explicitly described in the text. This is achieved by making inferences. Although inferences can be classified into various types (e.g., bridging and elaborative inferences; van Dijk & Kintsch, 1983), predictive inferences have received most attention from many researchers in the L1 reading area (e.g., Cook, Limber, & O’Brien, 2001; McKoon & Ratcliff, 1986). Predictive inferences are the most typical type of elaborative inference. Such inferences are not necessary for comprehension but embellish explicitly stated information and contribute to building richer text comprehension. Predictive inferences are implicit anticipations of the likely consequences of events or actions described in the passage. For example, when reading the sentence The director and cameraman were ready to
shoot close-ups when suddenly the actress fell from the 14th story (McKoon & Ratcliff, 1986), readers would predict the actress’s death as a consequence of the event described in the sentence (i.e., readers activate the predictive inference). A large body of research has examined under what conditions predictive inferences are activated during L1 reading, and has found that the activation is influenced by various factors, such as the contextual constraint of the passage (Cook et al., 2001; Klin, Guzmán, & Levine, 1999) and whether or not the inferences maintain the local coherence of the passage (Klin, Murray, Levin, & Guzmán, 1999). Although some limited studies have examined inference activation among EFL readers (e.g., Yoshida, 2003), few have investigated the activation of predictive inferences including the aforementioned factors suggested in L1 studies. Thus, Nahatame (2011) attempted to resolve this issue and found that Japanese EFL readers strongly activated predictive inferences during reading only when the context strongly suggested one outcome could be predicted and the inferences were necessary for preserving local coherence of the text. In addition, it was also revealed that the inferences strongly activated during reading were also maintained in readers’ text memory (i.e., encoding of activated inferences into readers’ text representations). This previous study enhanced our understanding of the conditions under which EFL readers strongly activate and encode predictive inferences. However, in real-life reading such as reading in the classroom, the predictions readers activate are not always consistent with the subsequent context. In the case of drawing incorrect predictions, it is necessary for readers to suppress (i.e., decrease the activation of) such predictions in order to achieve coherent text comprehension and efficiently use their attentional resources for text processing. Despite their significance, in previous studies much less attention has been paid to the suppression of predictive inferences than the activation of these inferences.

To date, several previous studies have examined the suppression of information explicitly written in the text (cf. inferential information) such as narrative characters’ names and goals (e.g., Gernsbacher, Robertson, Palladino, & Werner, 2004; Linderholm et al., 2004). These previous studies concluded that L1 readers maintain the activation of explicit text information during reading as long as it is relevant to their current comprehension, whereas they suppress the information once it is less relevant. Additionally, Gernsbacher et al. (2004) demonstrated that more skilled readers had better ability to suppress text information than less skilled readers.

In contrast to the suppression of explicit text information, the suppression of inferential information has been investigated in only a limited amount of research. Potts, Keenan, and Golding (1988), for example, tested the activation level of predictive inferences by measuring naming and lexical decision times for target words and comparing them between the predictive and disconfirming conditions. In the predictive condition, participants read the passages inducing targeted predictive inferences (e.g., No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. He had been feeling angry for weeks, but had refused to seek help; A targeted inference is break). On the other hand, the second sentence of the predictive passage was replaced by the one that precluded the targeted inferences (i.e., disconfirming
context) in the disconfirming condition (e.g., No longer able to control his anger, the husband threw the delicate porcelain case against the wall. By sheer luck the vase hit at such an angle that it was not damaged). Potts et al. found no significant differences in naming and lexical decision times between the two conditions, suggesting that readers maintain the activation of the predictive inference even after they receive information indicating that the inference is not appropriate. Contrary to these results, Iseki (2006) demonstrated that the activation of predictive inferences was decreased (i.e., suppressed) in the case of the disconfirming condition, using a task in which participants were required to judge whether or not the target sentence made sense. Given this discrepancy and the limited amount of research, there is no converging evidence regarding the suppression of inferential information during reading. In addition, since previous research has been conducted in the L1 context, it still remains unclear whether EFL readers can suppress predictive inferences which they activate during reading. Thus, the first goal of the present study is to examine the suppression of predictive inferences during EFL reading. Considering the effects of reading ability on suppression noted in Gernsbacher et al. (2004), the present study investigated inference suppression, including L2 reading proficiency as a variable.

In terms of the methodology, previous studies have examined the suppression of predictive inferences based on the analysis of response times obtained in the on-line tasks (Potts et al., 1988; Iseki, 2006). Although the on-line measurement can provide evidence for suppression immediately after the disconfirming context (i.e., during reading), it is impossible to further clarify whether or not readers delete the disconfirmed inferences from their final representations (i.e., after reading). Few studies have attempted to address this issue. Importantly, Zwaan and Radvansky (1998) noted that the content of the representations during reading is distinguished from the content of the final text representations after reading, and therefore the suppression and deletion need to be considered as different processes. Thus, the suppression and deletion of the inferences should be examined individually. Furthermore, in order to provide useful pedagogical suggestions, it is advisable to explore the effects of disconfirming the activated predictive inferences on readers’ explicit text comprehension. For these reasons, the second goal of this study is to investigate how the disconfirmation of the activated inferences affects readers’ final text representations (i.e., the deletion of the inferences and accuracy of the text comprehension) using an off-line measurement.

To achieve these goals, three research questions (RQs) were addressed in this study:

RQ1: Do EFL readers suppress the activated predictive inferences when the subsequent context disconfirms the inferences?
RQ2: Do EFL readers delete the disconfirmed inferences from their final text representations?
RQ3: How does the disconfirmation of the inferences influence EFL readers’ text comprehension?
2. Method

2.1 Participants
A total of 37 Japanese university students participated in this study. Their majors were varied, including practical engineering, economics, and so on. All the participants were native speakers of Japanese and had been studying English for more than six years.

2.2 Materials
A 54-item random cloze test was used as a L2 reading proficiency test. It was created using a 410-word text from Timed Readings, Book one (Spargo, 1989).

Sixteen narrative texts adopted from Virtue, van den Broek, and Linderholm’s (2006) study were used as experimental passages, with slight modifications. These were also used in Nahatame (2011). Each of the 16 texts consisted of four sentences and originally had two versions: predictive and neutral versions. Nahatame (2011) with Japanese EFL readers confirmed that targeted predictive inferences were strongly activated during reading in the predictive version of these texts. In the predictive text, targeted inferences were strongly constrained by the context and they were necessary for maintaining the local coherence of the texts. In the inference text in Table 1, for example, the fourth sentence of the text (i.e., the final sentence of the predictive text) strongly constrains the possible inference to be activated (i.e., steal). In addition, not only is the targeted inference the likely outcome of the described character’s actions, it is also the motivation for his doing so. Readers cannot achieve coherent text comprehension without this inference because it explains why the character took the action, and therefore it is required to preserve the local coherence of the passage. On the other hand, the neutral texts described different topics from predictive texts and did not promote predictive inferences.

Table 1 Sample of Experimental Passages Used in This Study

<table>
<thead>
<tr>
<th>Inference Text</th>
<th>Neutral Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Predictive and Disconfirming Text)</td>
<td>Brad and Fred had been working in New York City for six months. Brad’s parents were coming to stay with them this weekend. He was trying to think about what fun tourist spots they would visit. Brad and Fred decided to take them to the Museum of Art first.</td>
</tr>
<tr>
<td>Brad was looking for a present for his wife’s birthday. He wanted to find something special for her, but he couldn’t afford to buy anything nice. In the accessories department, he saw an expensive ruby ring sitting in a display. Seeing no salespeople or customers around, he quietly made his way to the display and took out his bag. Writing down the price of the ring on his memo pad, he promised to buy it someday.</td>
<td>Target word: <strong>steal</strong></td>
</tr>
</tbody>
</table>

Note. An italicized sentence in the inference text appeared only in the disconfirming text.
Each text had a corresponding target word and a yes-no comprehension question. Target words were the verbs which represented the inferential information suggested by the predictive text (e.g., steal). The comprehension questions focused on details of the passage that were not directly relevant to the targeted inference and were used to encourage participants’ careful reading and to assess their understanding of each passage.

Since one of the purposes of this study was to examine the suppression of the activated predictive inferences, another version of the inference text (i.e., disconfirming text) was created in addition to the predictive text. In the disconfirming texts, as in previous studies (e.g., Iseki, 2006), a single sentence was added after the final sentence of each predictive text. This extra sentence provided the information precluding the inference suggested by the predictive text. In Table 1, for example, the disconfirming sentence Writing down the price of the ring on his memo pad, he promised to buy it someday was added to the predictive text. The disconfirming sentences for each passage were created by the researcher, referring to the materials in previous studies. Considering the importance of the disconfirming sentences in this study, these sentences were assessed by two other raters using a 7-point scale in terms of their plausibility as sentences following the original text (i.e., predictive text). As a result, the sentences rated as less plausible were modified through discussions with raters. In addition, all disconfirming sentences were composed of between 15 and 20 simple words. Note that these processes ensured that the different features of disconfirming sentences such as the plausibility, length, and difficulty did not affect the results. All disconfirming sentences had been checked and corrected by a native English speaker.

In sum, three sets of 16 experimental passages were used in this study, each of which had two inference versions (i.e., predictive and disconfirming texts) and a neutral version. Table 1 shows sample experimental passages of each version. Four material sets (A-1, A-2, B-1, B-2) were created to counterbalance the passages across versions, and each of the four sets included 12 of 16 experimental passages. This confirmed that each passage appeared an equal number of times in the three conditions (i.e., predictive, disconfirming, and neutral conditions).

In addition to the experimental passages, 12 filler passages were used to balance the number of yes and no responses of the probe verification (see the procedure section). Therefore, a certain word included in each filler passage was used as the target word corresponding to the passages. The filler passages also consisted of four sentences but did not induce any inferences. All of the filler passages were included in each material set.

2.3 Procedure

Participants were tested individually. First of all, participants were given the cloze test and asked to finish it within 17 minutes. Afterwards, the experiment had the following two phases: (a) an on-line phase, in which participants performed the on-line task, and (b) an off-line phase, in which they performed the off-line task. The procedures in these two phases were based on those in Nahatane (2011) and were as described below.
a. *On-line phase*: The on-line phase employed the program SuperLab 4.5 on a computer. During this phase, participants were randomly assigned to one of the four material sets. First, they read the instructions about their task on the computer screen. Each trial began with the word *Ready?* appearing in the center of the screen. When ready to begin the trial, participants pushed the YES button on the Response Pad RB-730 and then the first sentence of the passage appeared. Participants read the passage at their own pace sentence by sentence, pushing the Yes button to bring on each successive sentence of the passage. After reading the final sentence of the passage, pressing the Yes button led to the appearance of a warning signal (XXX) for 750 ms in the center of the screen. The target word flanked by asterisks (e.g., **steal**) followed the warning signal. Then participants were required to verify whether the target word had appeared in the passage they had just read (i.e., a probe verification task). They had to do this as quickly and accurately as possible. They pressed the Yes or No button to indicate their responses, and the accuracy and latencies were recorded. After their responses, another signal (???) appeared for 500 ms, which was replaced by a comprehension question. Participants made their responses and received feedback about their accuracy. This trial was repeated for each of the 24 passages (i.e., 12 experimental and 12 filler passages), which were randomly presented to the participants. Before beginning the experiment, the participants were given five practice items and followed the same procedure described above.

The verification task, which is the most significant task in the on-line phase, assessed the activation of inferential information represented by target words. This task has been often used for measuring inference activation during reading in prior research (e.g., Fincher-Kiefer, 1995, 1996; McKoon & Ratcliff, 1986). McKoon and Ratcliff (1986) assumed that if readers activated the inferential information represented by the target word, then a correct response would be slow. Such a delay in verifying a target word is caused by confusion about the source of the memory for the target word; that is, at the time of the verification, readers experience difficulty in distinguishing inferential information they activate during reading from the information explicitly presented in the text (Fincher-Kiefer, 1995, 1996).

b. *Off-line phase*: After completing the on-line phase, the participants engaged in a reading task and a recall task. Each participant was given a booklet consisting of 12 experimental passages that they had not read in the on-line phase, and was instructed to read it within 10 minutes. Each of the participants read an equal number of passages in each of the three versions, and each version of each passage was presented to one fourth of the participants. After reading the passages, the participants were provided with the first sentence of each passage. Then they were required to recall and write down as much about the passages as possible in Japanese, with the first sentence as a cue (i.e., a cued recall task). Participants were given an unlimited amount of time to complete the recall task.

Prior research (e.g., Klin, Murray et al., 1999; Murray & Burke, 2003) used the cued recall
task to examine whether or not readers maintain the activated inferences in their text memory. If the inference is maintained and becomes part of readers’ final text representations, then readers should have produced the inferential information in their recall protocols. On the other hand, if the activated inference is lost or deleted from the representations, there should be very few inference intrusions in recall.

2.4 Scoring and Data Analysis

The cloze test was marked using the acceptable-word scoring method. Thirty percent of the data were randomly selected from the pool of the data and scored by the two raters separately, resulting in an agreement ratio of 91.36%; disagreements were resolved through discussion among the raters. The remaining data were scored by one of the raters (i.e., the researcher) alone. The test indicated high reliability (Cronbach’s $\alpha = .84$) after excluding five low discrimination items.

Regarding the analysis of the verification task, participants’ data from a passage were eliminated when they mistakenly skipped a sentence of the passage. This was because it was possible that their verification accuracy and latencies might have been affected by not reading a certain sentence. In addition, verification latencies which were 2.5 standard deviations beyond the mean for a participant were excluded, along with those less than 100 ms. These processes led to the removal of 2.08% of the data. Analysis was conducted on the verification latency for correct responses according to prior research (e.g., McKoon & Ratcliff, 1986). Further analysis of the verification data is noted in detail in the result section (3.2.2).

As in previous studies, recall protocols were scored in terms of two types of information: text and inferential information. Text information refers to the information explicitly described in the passages, that is, the three common sentences included in both the predictive and the disconfirming texts, except for the first sentence used as a cue. For the scoring of text information, the experimental passages were parsed into a set of idea units (IUs) by the two raters based on Ikeno’s (1996) criteria, resulting in high inter-rater reliability, $r = .95$; any disagreements were resolved through discussion. In contrast, inferential information was the concepts represented by the target words (e.g., steal), which were not actually mentioned in the text. When the target word was included in participants’ recall protocols (e.g., …Seeing no people around, he approached the display and stole the ring; translated from Japanese), it was scored as the production of inferential information. As in the scoring of the cloze test, 30% of recall protocols were marked by the two raters separately in terms of text and inferential information, resulting in an agreement ratio of 92.09%. After resolving disagreements, the remaining data were scored by one rater. Neutral texts were not included in the recall analysis because they had different text content from the inference texts and therefore it was difficult to compare the protocols between neutral and inference texts.

Finally, any data from one participant with more than 28% errors on the comprehension questions was excluded, on the basis of prior research (e.g., Klin, Guzmán et al., 1999). The analysis was therefore based on the data from the remaining 36 participants.
3. Results

3.1 Cloze Test

Based on the scores on the cloze test, participants were divided into two proficiency groups, upper \((n = 18, M = 36.50, SD = 2.23)\) and lower groups \((n = 18, M = 25.41, SD = 5.51)\). Then a 2 (Proficiency: upper, lower) \times 4\) (Material sets: A-1, A-2, B-1, B-2) analysis of variance (ANOVA) was conducted on the scores. The results showed that the main effect of proficiency was significant, \(F(1, 28) = 55.11, p < .001, \eta^2_p = .66\), but neither the main effect of the material sets nor the interaction between proficiency and material sets was significant, \(F(3, 28) = 0.92, p = .446, \eta^2_p = .09; F(3, 28) = 0.14, p = .938, \eta^2_p = .01\). This result means that participants in the upper group had L2 reading proficiency significantly higher than those in the lower group and that there was no significant difference in readers' proficiency among the four material set conditions. Therefore, L2 reading proficiency but not the material set conditions was included as a variable in the analysis of verification and recall tasks.

3.2 On-Line Task

3.2.1 Comprehension Questions

The participants correctly answered, on average, 91.67% of the comprehension questions for experimental passages. A repeated-measures ANOVA with three reading conditions as a within-participants variable found no significant difference in correct answer rates among conditions (Predictive \([M = .92, SD = .14]\), Disconfirming \([M = .92, SD = .13]\), and Neutral \([M = .91, SD = .15]\)), \(F(2, 70) = 0.08, p = .920, \eta^2_p = .00\). This means that participants comprehended the passages equally well in each of the three reading conditions when performing the on-line task. Therefore, it is important to note that if a difference in the performance of the verification task was obtained among reading conditions, it cannot be interpreted as being the result of the difference in passage comprehension among conditions.

3.2.2 Verification Task

The mean verification latency for correct responses as a function of reading conditions and proficiency groups is reported in the Appendix. In order to compare the strength of the predictive inference activation among conditions, inference activation scores were obtained (see Graesser, Wiemer-Hastings, & Wiemer-Hastings, 2001; Virtue et al., 2006). These scores were measured by subtracting the verification latency for correct responses in a neutral condition from those in each of the two inference conditions (i.e., predictive and disconfirming conditions). For example, if the mean verification latency is 1000 ms in the neutral condition and 1200 ms in the predictive condition, the activation score in the predictive condition is 200 ms. The score would be significantly greater than zero if inferences were activated during reading. Additionally, the more strongly the inferences were activated, the greater the score would be.

Table 2 and Figure 1 show the mean inference activation scores in each inference condition.
and proficiency group. Prior to the test of the suppression of activated inferences, it was necessary to verify that readers had actually activated the targeted inferences during reading. Therefore, one-tailed paired t tests were conducted on the inference activation scores in the predictive condition to determine whether they were significantly above zero for each level of different proficiency groups. The results showed that the scores were significantly greater than zero for both upper and lower proficiency groups, t (17) = 2.60, p = .010, d = 0.60; t (17) = 2.08, p = .027, d = 0.48. Thus, it was confirmed that targeted inferences were activated during reading in both proficiency groups.

Table 2 Activation Scores (ms) in Each Inference Condition and Proficiency Group

<table>
<thead>
<tr>
<th>Proficiency groups</th>
<th>Predictive condition</th>
<th>Disconfirming condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Upper</td>
<td>178.05</td>
<td>290.11</td>
</tr>
<tr>
<td>Lower</td>
<td>144.86</td>
<td>295.24</td>
</tr>
</tbody>
</table>

Then, in order to examine the suppression of activated inferences (RQ1), the inference activation scores in the predictive condition were compared to those in the disconfirming condition. If the disconfirming text serves to decrease the activation of the inference (i.e., suppression), then the activation scores in the disconfirming condition should have been lower than those in the predictive condition. If, on the other hand, the inference remains activated even after it is disconfirmed by the subsequent context, then the activation scores in the disconfirming condition should have been the same as those in the predictive condition. The mean inference activation scores were submitted to a 2 (Condition: predictive, disconfirming) x 2 (Proficiency: upper, lower) mixed ANOVA with condition as a within-participants variable, and proficiency as a between-participants variable. The results showed that neither the main effects of condition nor proficiency approached significance, F (1, 34) = 0.75, p = .392, ηp² = .02; F (1, 34) = 0.72, p = .401, ηp² = .02. In addition, the interaction between condition and proficiency was not significant, F (1, 34) = 0.42, p = .521, ηp² = .01. Although the activation scores of the disconfirming condition seemed slightly higher than those of the predictive condition, this difference was not significant. These results indicated that the strength of inference activation was not different between the predictive and disconfirming conditions, regardless of readers’ L2 proficiency; in other words, the inferential information was not suppressed, and remained active immediately after the disconfirming context.
3.3 Off-Line Task (Cued Recall Task)

Table 3 displays the mean proportion of targeted inferences and text information recalled by the upper and lower proficiency groups. Inferential information in the predictive condition was recalled to the same degree as the data from Nahatame (2011), who indicated the strong encoding of activated inferences, \( t(63) = 1.48, p = .145, d = 0.37 \). Thus, readers in the present study also maintained activated inferences in their final text representations when they read the predictive passages.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Inference</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td>Predictive condition</td>
<td>Upper</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>.19</td>
</tr>
<tr>
<td>Disconfirming condition</td>
<td>Upper</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>.14</td>
</tr>
</tbody>
</table>

To explore RQ2, a 2 (Condition: predictive, disconfirming) x 2 (Proficiency: upper, lower) mixed ANOVA was conducted on the proportion of inferential information included in participants’ protocols, and similar results were found to those obtained in the analysis of the verification task. Neither the main effects of condition nor proficiency were significant, \( F(1, 34) = 1.05, p = .312, \eta_p^2 = .03; F(1, 34) = 2.92, p = .097, \eta_p^2 = .08 \). Additionally, the interaction between condition and proficiency did not approach significance, \( F(1, 34) = 0.05, p = .827, \eta_p^2 = .00 \). Thus, it was revealed that the inferential information readers activated during reading was recalled in the disconfirming condition to the same degree as in the predictive condition, irrespective of readers’ L2 proficiency; that is, the activated inferences were not deleted but maintained in readers’ final text representations even when passages included the disconfirmation.

Then, to answer RQ3, a 2 (Condition: predictive, disconfirming) x 2 (Proficiency: upper, lower) mixed ANOVA was also conducted on the proportion of text information recalled in participants’ protocols. The result showed that the main effects of condition and proficiency were statistically significant, \( F(1, 34) = 4.12, p = .050, \eta_p^2 = .11; F(1, 34) = 4.98, p = .032, \eta_p^2 = .13 \). Readers in the upper proficiency group significantly recalled more text information than those in the lower proficiency group, which was consistent with the results of the cloze test reported in the previous section 3.1. More importantly, it is interesting to note that readers recalled less text information in the disconfirming condition than in the predictive condition. However, the interaction between condition and proficiency was not significant, \( F(1, 34) = 0.23, p = .635, \eta_p^2 = .01 \), suggesting that the difference between predictive and disconfirming conditions was consistent among different levels of proficiency groups.
4. Discussion

Suppression of the predictive inferences during reading (RQ1)

The first goal of this study was to examine whether EFL readers suppress the activated predictive inferences when the subsequent context disconfirms it. The result of the verification task analysis found no significant difference in the activation scores between the predictive and disconfirming conditions, suggesting that the inclusion of the disconfirming context did not serve to suppress the activated inferences. This result was consistent between upper and lower proficiency readers. Although the difference was not significant, the activation scores in the disconfirming condition were slightly higher than those in the predictive condition. This, one may argue, indicates the possibility that the obtained results reflect just the difference in the memory load between the two conditions, but not the difference in the strength of inference activation. That is, readers might need a slightly longer time to verify the target words in the disconfirming condition because of the increased memory load, which could be caused by the larger number of sentences in the disconfirming text compared to the predictive text (i.e., five vs. four sentences). Potts et al. (1988), however, also showed a relative facilitation of inference activation in the disconfirming condition (although no significant difference to the predictive condition was found) as in this study, using (a) materials in which the number of sentences was not different between the predictive and disconfirming texts, and (b) the naming task for target words, which was less likely to be affected by the memory load. Thus, the present results would not reflect the different memory load, but demonstrate that EFL readers do not successfully suppress the predictive inferences immediately after the information that disconfirms the inferences (the answer to RQ1).

These results suggest that readers have some difficulty in the suppression of activated predictive inferences. Contrary to the assumption that more skilled readers have efficient suppression mechanisms (Gernsbacher et al., 2004), activated inferences were not suppressed by either upper or lower proficiency readers in this study. This raises the possibility that the difficulty of suppression shown in this study could have been caused by text features rather than insufficient reading ability. In the disconfirming texts of this study, there were always four sentences (including five to 10 IUs) that induced the predictive inference and only one sentence (including three to four IUs) that disconfirmed the inference. Given this imbalance, it is possible that the disconfirming context consisting of a single sentence was not enough for readers to fully suppress the activated inferences. In addition, the disconfirming context suggested that the inference was not appropriate, but neither explicitly eliminated it nor encouraged readers to generate an alternative inference instead of the disconfirmed one (this issue is discussed further in the discussion of RQ3). Therefore, the quality of the disconfirming context as well as the quantity could be a factor affecting the difficulty of suppression. This assumption is based on the findings of several previous studies (e.g., Guéraud, Harmon, & Peracchi, 2005; O’Brien, Rizzella, Albrecht, & Halleran, 1998), which demonstrated that readers’ comprehension difficulty caused by the
existence of the inconsistent text information was decreased if there was a sufficient amount and quality of context qualifying the inconsistency. Thus, it is likely that the inconsistency between the activated predictive inferences and actual text information can also be resolved with a quantitatively and qualitatively sufficient disconfirming context. More specifically, the suppression of inferences could be facilitated if readers receive more information in several sentences which disconfirms the inferences, or the subsequent context clearly disconfirms them and induces other inferences.

Deletion of the predictive inferences after reading (RQ2)

The second goal of this study was to examine the effects of disconfirming the activated inferences on EFL readers’ final text representations. The present results found that the inferential information was recalled in the disconfirming condition to the same degree as in the predictive condition and this was consistent between different proficiency groups. Thus, the predictive inferences disconfirmed by the subsequent context were not deleted but were maintained in readers’ final text representations. Combining this result with findings concerning the suppression, it can be said that the inferences which were activated but not suppressed during reading remained active in readers’ final representations.

However, note that this result does not mean readers could not correctly comprehend the disconfirming context and found the activated predictive inferences to be what actually happened in the narrative. A closer look at the participants’ protocols in the disconfirming condition indicated that readers often recalled inferential information with disconfirming information and comprehended that the inference did not in fact happen. For instance, some participants who read the disconfirming passage in Table 1 produced inferential information in the following way: *He saw no people around and tried to steal the ring, but he wrote down its price and promised to buy it someday* (translated from Japanese). Although it may sound somewhat inconsistent or unnatural because the inference was recalled with information disconfirming it, it is interesting that readers nevertheless maintained both inferential and disconfirming information in their comprehension. Therefore, the answer to RQ2 could be that EFL readers do not completely delete the disconfirmed inferences from their final representations; they maintain these inferences with additional information. The reason why readers maintained both inferential and disconfirming information together is discussed in the following section.

Effects of disconfirming the activated inferences on text comprehension (RQ3)

The results of the recall task also demonstrated that readers had less text comprehension in the disconfirming condition than in the predictive condition; that is, the disconfirmation of the activated predictive inferences had a negative impact on readers’ understanding of explicit text information (the answer to RQ3). This finding can be explained by two possibilities.
One possible explanation is that readers failed to establish the coherence of the passages due to the disconfirmation of the activated inferences. In this study, targeted inferences explain the reason for the character’s action, and therefore are activated to preserve the local coherence of the passage. Although these inferences are disconfirmed by the following context in the disconfirming condition, readers maintained the activated inferential information with the disconfirmation in their final representations, as noted above. This result indicates readers understood that the inferences might be incorrect but this did not result in the generation of an alternative inference explaining why the character took the action (e.g., He looked around because he did not want to be seen making a note of the price of the ring.). Thus, readers still maintained the disconfirmed inferences in their final representations, which led to somewhat inconsistent comprehension, as shown by the example in the discussion of RQ2. In sum, it would be difficult for readers to achieve coherent comprehension of the passages in the disconfirming condition because the activated inferences were disconfirmed and they could not generate the alternative explanatory inferences. As a result, readers produced less text information in the disconfirming condition than in the predictive condition. As Graesser et al. (1994) highlighted the significance of generating inferences that explain the character’s action in narrative comprehension, it is likely that such inferences were also important for constructing coherent comprehension in this study.

Another possible explanation could be that disconfirming the activated inferences might alter a reader’s perceptual importance of text information. As mentioned earlier, the activated inferences were closely related to the character’s action. Thus, readers in the disconfirming condition found the character’s action less important when such inferences were disconfirmed, which caused less production of text information describing the character’s action. In other words, the readers’ focus of attention was directed to the disconfirming information from the preceding information; however, the disconfirming information was not scored in the prior recall analysis. In fact, further analysis showed that in the disconfirming condition readers recalled significantly more information in the disconfirming sentence ($M = .46$, $SD = .27$) than information in the three sentences preceding it ($M = .32$, $SD = .17$), $F (1, 34) = 15.70$, $p < .001$, $\eta^2_p = .32$.

In this study, it was difficult to distinguish between the two possibilities noted above, but they do not necessarily conflict with each other. It is probable that readers focused on the only disconfirming sentence because they could not establish the coherence between this sentence and the preceding sentence. Nevertheless, further research is needed to identify the cause of the negative effects on text comprehension of disconfirming the activated inferences.

5. Concluding Remarks

The primary goal of the present study was to examine the suppression of activated predictive inferences among EFL readers. The results of the verification task showed that the activation of inferential information remained immediately after the context disconfirming the
information. Although several previous studies have explored inference activation among EFL readers (e.g., Nahatame, 2011; Yoshida, 2003), few studies have focused on the suppression of activated inferences. Therefore, the present study represents a further step towards understanding the nature of inferential information during EFL reading. Successful text processing requires readers to allocate attentional resources flexibly and efficiently to the construction of the text representations, including the suppression of irrelevant information (Germbacher et al., 2004; Horiba, 2000). The present study suggests that EFL readers have some difficulty suppressing activated predictive inferences without sufficient information which disconfirms the inferences. Thus, it is implied that classroom teachers should not only encourage their students to actively make predictions during reading but also help them effectively suppress the predictions when they are not supported by the following context. For example, through questioning and activities, teachers should make students aware of the gap between their predictions and information actually described in the text, which could facilitate the suppression of the predictions and allow smooth processing of the following text. Additionally, the present results are suggestive for text writers. Although reading stories with unpredicted events or surprising endings can be enjoyable for students, it is suggested that writers should include sufficient information in these texts so that students can fully comprehend and enjoy reading such stories.

The present study also examined the effects of disconfirming activated inferences on readers’ final text representations. The results of the cued recall task revealed that the disconfirmed inferences are maintained, with some qualifications, in readers’ final representations and that the disconfirmation negatively affects readers’ understanding of the explicit text information. Although the negative effect of the disconfirmation on text comprehension is an interesting finding of this study, it should be interpreted with care because the cause of this negative effect was not specified in this study. Thus, further investigations are required before educational applications of this finding can be suggested.

Finally, it is important to note again that this is an initial study empirically evaluating the suppression of predictive inferences among EFL readers. Therefore, it suggests some significant directions for future research to further explore the suppression of predictive inferences. First, future research should examine under what conditions EFL readers can efficiently suppress predictive inferences. Such research should include consideration of some variables, such as the quantity and quality of the disconfirming context, types of instructions, and reading goals. This approach will provide more specific implications for teacher support in facilitating students’ effective suppression of predictive inferences. Furthermore, note that the standard deviations of the activation scores obtained in the present study were relatively large. Thus, there remains the possibility that learners’ individual differences apart from the L2 reading proficiency (e.g., working memory capacity) affect the suppression, and future research should confirm this. Although these points need to be clarified, the present study provided unique findings concerning
the suppression of predictive inferences; a process which plays a significant role in efficient reading and coherent text comprehension.

Acknowledgements

I would like to express my profound gratitude toward Professor Yuji USHIRO for his sound guidance. I would also like to thank three anonymous reviewers for their helpful comments to improve this paper.

References


### Appendix

#### Mean Verification Latency (ms) in Each Reading Condition and Proficiency Group

<table>
<thead>
<tr>
<th>Proficiency groups</th>
<th>Predictive condition</th>
<th>Disconfirming condition</th>
<th>Neutral condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Upper</td>
<td>1224.03</td>
<td>368.91</td>
<td>1347.29</td>
</tr>
<tr>
<td>Lower</td>
<td>1304.56</td>
<td>398.60</td>
<td>1322.38</td>
</tr>
</tbody>
</table>