Task effects on reading implausible sentences in an L2:
Evidence from self-paced reading*

Jung Hyun Lim** · Hee-Don Ahn***
(Konkuk University)

Jung Hyun Lim · Hee-Don Ahn. Task effects on reading implausible sentences: Evidence from self-paced reading. *Language and Linguistics 66, 00–00.* This study investigates how L1 Korean speakers of English respond to semantic information in syntactically unambiguous sentences, manipulating plausibility of the sentence. Task effects are also examined by using different types of comprehension questions. The experiment in a self-paced reading paradigm found that L2 speakers displayed a sensitivity to plausibility information from the beginning of the sentence, and that the sensitivity was influenced by the comprehension question types. The study provides behavioral data of L2 learners' use of plausibility information and emphasizes the effect of tasks on L2 reading. The study further suggests that L2 processing is strategic and goal-dependent, supporting for the good-enough language processing in L2 processing mechanism.

Keywords: L2 sentence comprehension, task effects, self-paced reading task

* This work was supported by the KU Research Professor Program of Konkuk University. We thank three anonymous reviewers for useful comments.
** First author  *** Corresponding author
1. Introduction

When one reads a sentence, the reader needs to know words, grammar, its context, and/or background knowledge to make the sentence understandable and comprehensible. Our ability to read a sentence is amazingly rapid considering that understanding a sentence requires the complicated process to put together lexical, syntactic, semantic, and pragmatic information. One of the important questions in psycholinguistics is how the language parser integrates syntactic and semantic information during sentence reading. In the past decades, investigation of reading garden-path sentences such as a sentence in (1) below (from Christianson, Hollingworth, Halliwell, & Ferreira, 2001) has been of interest in both native and non-native language processing research, because it gives us valuable information on how people comprehend and reanalyze the syntactically ambiguous sentences.

(1) While the man hunted the deer ran into the woods.

Previous studies examining analyses of garden-path sentences by both native (henceforth, L1) and non-native (henceforth, L2) speakers found that an ambiguous sentence such as (1) was more likely to be misinterpreted than an unambiguous sentence such as ‘while the man hunted the pheasant, the deer ran into the woods’. To be more specific, participants tended to misanalyze ‘the deer’ as a direct object of the verb ‘hunt’, resulting in difficulties to process the verb ‘ran’, which is syntactically inconsistent with the initial direct-object analysis (Christianson, Williams, Zacks, & Ferreira, 2006; Juffs & Harrington, 1996; Roberts & Felser, 2011).

Another line of research to study the integration of syntactic and semantic information has examined how the parser deals with the information when syntactic and semantic cues are in conflict. Existing studies with L1 English native speakers suggested that language processing can sometimes be ‘good-enough’ in reaching at the final interpretation of a sentence, especially
when the outcomes of syntactic and semantic routes are different (Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007). In addition, Swets, Desmet, Clifton, and Ferreira (2008) found that the depth of language processing can also vary depending on task demands at hands, supporting for the 'good-enough' language processing.

To pursue the question of how syntactic and semantic information are integrated, plausibility information has usually been manipulated along with syntactic structures. Previous studies have focused on how semantic information influences structural rules in parsing strategies (Pickering & Traxler, 1998; Trueswell, Tanenhaus, & Garnsey, 1994). There has been no doubt that plausibility has an effect in reading a sentence, but the issue is at what stage the information is used and how both syntactic and semantic information are interacted in processing the local/global ambiguities of a sentence.

Although there are many studies to manipulate both semantic and syntactic information in looking at both L1 and L2 processing, there are no studies so far to examine the effect of plausibility on L2 reading by teasing apart from syntactic structure manipulation. The present study examines how plausibility affects L2 speakers' real-time reading comprehension, without any complexity of syntactic structures. This study also investigates how the effect of plausibility is influenced by task demands manipulating types of comprehension questions, which has never been questioned before. The next section reviews the importance of plausibility in on-line reading comprehension, followed by the literature review on the task effects on language processing. Then, the present study is introduced describing method and results. The results are discussed in the light of L2 sentence comprehension and in the frame of the good-enough language processing.

2. Background
2.1. Effects of plausibility on reading sentences

Many studies with L1 English native speakers have shown that reading about an implausible event in a sentence occurs disruption during real-time comprehension reading times (Cook & Myers, 2004; Patson & Warren, 2009; Rayner, Warren, Juhasz, & Liversedge, 2004) and also yields inaccurate interpretations (Christianson et al., 2001; Ferreira, 2003). One of the reasons of studying the effect of plausibility on time course of on-line reading comprehension is that it broadens our understanding on whether and how semantic information affects sentence parsing and at what stage of processing it has an effect along with integration of syntactic information.

Two types of plausibility manipulation have usually been conducted to examine the effect of plausibility in sentence processing. First, a word is manipulated to make the sentence pragmatically anomalous as in 'John used a pump to inflate the large carrots for dinner' or implausible as in 'John used an axe to chop the large carrots for dinner' (from Rayner et al., 2004). Second, sometimes the whole event of the sentence can be manipulated in an implausible way such as 'The man bit the dog' (from Ferreira, 2003).

Using the former kind of plausibility manipulation, some studies have reported the immediate effect of plausibility on sentence parsing (Ni et al., 1998; Thornton & MacDonald, 2003). For example, Ni et al. (1998) examined participants' eye-movements in reading the sentences as in (2).

(2) a. It seems that the cats won't usually eat the food we put on the porch. (control condition)
b. It seems that the cats won't usually bake the food we put on the porch. (pragmatic anomaly condition)
c. It seems that the cats won't usually eating the food we put on the porch. (syntactic anomaly condition)
Results showed that participants took longer times in reading the syntactic anomaly condition (2c) than the other two. In contrast, the residual reading times, which consider each participant's differences and length of each stimuli sentence, showed the longest in the pragmatic anomaly condition (2b). Although the study reported the earlier effect of plausibility, the results were quite inconsistent depending on the reading time measures, yielding a totally different pattern.

More recently, Rayner et al. (2004) pointed out that the results of Ni et al. (1996) study contained a problematic variability in their data, which might have come from the different word lengths in the critical region across the experimental sentences. Rayner et al. (2004) thus conducted eye-tracking experiments in which plausibility was more systematically manipulated using identical syntactic structure across different conditions. They manipulated a noun phrase (NP) as in appropriate or inappropriate thematic-role filler for an accompanying verb, and also the degree of the implausibility in the context. The sample experimental sentences are presented as in (3).

(3) a. John used a knife to chop the large carrots for dinner.  
   (plausible)  
   b. John used an axe to chop the large carrots for dinner.  
   (implausible)  
   c. John used a pump to inflate the large carrots for dinner.  
   (anomalous)

Results showed that readers immediately detected the anomaly of the sentence when a word was inappropriate theme for the following verb such as (3c). Meanwhile, the implausible sentence (3b) had a little bit delayed effect of implausibility on reading the target noun region, that is 'the large carrots' in the
example sentence above. Rayner et al. (2004) concluded that the disruption causing from the implausibility in the time course of reading is varied as the degree of the violation's severity, such that the unlikely theme as in (3c) had a delayed effect than the inappropriate condition (3b).

While some studies reported the earlier effect of plausibility on sentence parsing (Ni et al., 1998; Rayner et al., 2004; Thornton & MacDonald, 2003) as reviewed above, other studies have shown that plausibility rarely involves in initial parsing (Clifton, Traxler, Mohamed, Williams, Morris, & Rayner, 2003; Pickering & Traxler, 1998). As such, there still exist the remaining issue of timing when the plausibility plays a role in the parsing; that is, the stage of initial parsing decision or later reanalysis.

The whole event of the sentence has also been manipulated to examine how semantic information interacts with structural information in sentence processing. Previous studies to probe this research question have found that even native speakers sometimes misinterpreted a syntactically unambiguous sentence when the event of a sentence was unlikely to happen in the real world or inconsistent with the common world knowledge. As a fundamental study, Ferreira (2003) examined how native speakers of English interpret unambiguous active and passive structures manipulating plausibility of the events in the sentences. The sample sentences are describe below in (4).

(4) a. The patient was treated by the doctor. (passive-plausible)
   b. The doctor was treated by the patient. (passive-implausible)
   c. It was the doctor that treated the patient. (subject cleft-plausible)
   d. The patient treated the doctor. (active-implausible)

The study reported that passive structures yielded 80% accuracy in implausible version and 90% accuracy in plausible version, indicating larger inaccuracy rates in implausible than plausible sentences. The
active structures, however, resulted in similar accuracy rates in both implausible and plausible sentences: 95% and 98%, respectively. This difference was statistically significant, showing a reliable interaction between syntactic structure and plausibility.

The results of the study were inconsistent with previous studies which mostly assumed that the language processor operates in a complete, detailed, and accurate way to generate sentence meaning (Frazier & Clifton, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994). Based on Ferreira (2003)'s study and the other studies reporting misinterpretations using active/passive sentences and garden-path sentences (Christianson et al., 2001; 2006), Ferreira and her colleagues proposed a 'good-enough' (GE) approach to language processing and comprehension (Ferreira et al., 2002). According to the GE language processing, we process linguistic input through two routes, a syntactic route which deals with grammar, and a heuristic route which processes semantic plausibility or a noun-verb-noun (NVN) word order template. When we comprehend a sentence, the representations from the two routes usually appear in reconcile, resulting in a faithful and accurate interpretation. However, the output of each route is sometimes not reconciled especially when the representation from one route appears to be good-enough for comprehension, yielding an inaccurate final interpretation. This is the case when the representation from syntactic route is not consistent with the general knowledge; that is, something that is not likely to happen in the real world, as in a sentence 'the dog was bitten by the man'.

Similarly, the good-enough or shallow language processing has also been found in L2 speakers (Jacobs & Felser, in press; Lim & Christianson, 2013b; Roberts & Felser, 2011). Lim and Christianson (2013b) examined reading patterns of L1 Korean speakers of learning English using a subject and object relative clause construction in both plausible and implausible events, and compared them with that of native English speakers. The study also manipulated task demands on sentence processing in two reading sessions: comprehension and translation. Results showed that L2 learners were more influenced by plausibility information
than native speakers, although both groups equally displayed incremental processing of syntactic information. The effect of task types was also found: L2 learners were involved in deeper processing when they were asked to read for translation than comprehension. The study suggested that although both native and L2 processing computed both syntactic and semantic information, L2 speakers were more sensitive to semantic information, which was discussed under the good-enough language processing approach.

Roberts and Felser (2011) also examined the influence of plausibility information using locally ambiguous garden-path sentences in both native and non-native processing. The study manipulated the plausibility of the direct object of the immediately preceding verb in the sentences containing temporary subject-object ambiguities such as *The inspector warned the boss (crimes) would destroy very many lives* and *While the band played the song (beer) pleased all the customers*. The L1 Greek participants were more dependent on plausibility information than native speakers, showing the shallow and good-enough processing in L2 processing mechanism. The study also found that L2 speakers failed to reanalyze the sentences during on-line comprehension, compared to native speakers, depending on the degree of ambiguity of the garden-path sentences.

As reviewed above, the effect of plausibility on sentence processing seems obvious in both native and non-native processing mechanism. Both native and non-native speakers appears to be influenced by pragmatic plausibility information either at earlier or later stage in parsing, and sometimes the plausibility overrides the representations of the syntactic route. The good-enough language processing is also supported by the findings that readers sometimes employ different strategies on reading depending on task at hands. The next section describes the task effects on language processing related to the GE language processing approach.

2.2. Task effects on sentence comprehension
A recent study in L1 processing reported that reading goals and task demands can influence processing routines (Swets et al., 2008). Swets et al. (2008) examined whether language processing is affected by certain types of tasks and sentences. There has been an issue of whether globally ambiguous sentences are more difficult to process than unambiguous sentences. Some studies showed disadvantages of syntactic ambiguity inferred from longer reading times (Clifton, Staub, & Rayner, 2007) while other studies reported advantages of syntactically ambiguous sentences (van Gompel, Pickering, Pearson, & Leversedge, 2005). Swets et al. questioned whether different types of comprehension questions can affect reading times on ambiguity resolutions in the relative clauses. The study employed three types of relative clauses which were a) fully ambiguous, b) disambiguated in favor of high (N1) attachment, and c) disambiguated in favor of low (N2) attachment. The experimental sentences were followed by three types of comprehension questions: a) questions about the interpretations of the relative clause, b) questions about the superficial understanding of the sentence, and c) questions about the superficial understanding on an occasional basis. The study revealed the advantage of ambiguities, showing faster reading times on the condition where readers were asked about superficial understanding of a sentence. Participants seemed to put much less effort to resolve ambiguity when the question asked at a superficial level. Importantly, when the question asked about the relative clause and its attachment, participants took significantly longer reading times than the superficial question conditions, and spent more times in the ambiguous sentences than in the disambiguated sentences. The authors interpreted the results that language processing is not always complete, and ambiguous sentence can sometimes be unresolved especially when there is no demands to disambiguate it. This study showed the importance of task demands and reading goals on language processing which can modulate reading strategies and the depth of language processing. Also, the authors stated that their results can be the evidence of the good-enough processing in language comprehension.

Not many studies have yet considered task effects on L2 processing despite of
possible important implications on L2 reading processing and comprehension. Recently, several studies started to report the findings that secondary tasks, such as an explicit grammaticality judgment task, a memory task, and a translation task, increased L2 learner’s attention to syntactic information (Jackson & Bobb, 2009; Leeser, Brandl, & Whiteglass, 2011; Lim & Christianson, 2013b; Williams, 2006). For example, as mentioned in the above section, Lim and Christianson (2013b) reported the effect of task demands on L2 sentence processing. In their study, translation was used as a type of reading task, which requires additional cognitive demands in language processing. L2 participants were asked to read for comprehension in one session and for translation in the other session. Results revealed that L2 learners were better able to use the syntactic information, and their reading time patterns were more similar to native processing in the translation session. The study emphasized the effect of task types and reading goals on L2 processing and further suggested that L2 processing is also strategic and goal-dependent just as L1 processing, supporting for the good-enough language processing approach.

As reviewed thus far, the use of plausibility information has been investigated with much interests because it provides information about whether plausibility plays a role in the initial parsing or later analysis and how plausibility information interacts with syntactic information in the context of sentence parsing. Most of the studies examining the use of plausibility information have manipulated syntactic structures as well (Ferreira, 2003; Lim & Christianson, 2013b) or have used syntactically ambiguous sentences (Jacob & Felser, in press; Roberts & Felser, 2011; Swets et al., 2008). In order to control for any confound effects of different syntactic structures, Rayner et al. (2004) and Patson and Warren (2009) systematically manipulated plausibility using identical syntactic structures. These two studies manipulated an
individual word to differentiate the severity of plausibility in the context. Yet, there are lack of studies that systematically manipulate plausibility of the whole events of a sentence controlling for syntactic structures in both L1 and L2 processing research. To fill the gap in the literature, the present study targets L2 learner’s use of semantic information without any burdens of processing syntactic complexity. Also, as reviewed above, it seems that task types apparently affect language processing and reading strategies. Therefore, it would be another interesting question to ask how task demands modulate processing of plausibility in L2 sentence comprehension.

3. The present study

The present study thus aims to answer to the research questions of a) at what stage of the L2 sentence parsing plausibility has an effect, and b) how different types of task influence processing of plausibility information. As previous studies reported L2 learners’ large sensitivity to semantic information (Jacob & Felser, in press; Lim & Christianson, 2013a, b; Roberts & Felser, 2011), we predict that L2 speakers would respond to the implausibility of the sentence from the early stage of the sentence. In addition, as task demands have been found to have an effect on reading strategies in both L1 and L2 (Lim & Christianson, 2013b; Swets et al., 2008), L2 speakers are expected to show different reading times depending on the task types and to reveal possible interaction between plausibility and task types. Specifically, there would be longer reading times when the task requires more processing demands. In addition, we would carefully make predictions that when the question type requires more attention and focus, participants would show higher levels of accuracy in the comprehension questions. Since there are no previous studies to examine both plausibility and task
demands, it is hard to make specific predictions regarding the interactions, but any findings would add valuable information to second language processing literature.

3.1. Participants

Thirty-six undergraduate students at a large university located in Seoul, Korea participated in this study. They received 10,000 Won (KRW) for their participation in the experiment. All were native Korean speakers who started receiving English education at schools after their puberty.

3.2. Materials

The materials were constructed with 48 experimental and 96 filler sentences, adapted from both Ferreira (2003) and Lim & Christianson (2013b). All experimental items consisted of a noun phrase (e.g. the dog), a relative pronoun (that), a subject relative clause (e.g., bit the man), and a matrix verb phrase (e.g., was in the yard). The experiment employed a 2 (plausibility) X 2 (question type) mixed design, resulting in four lists. The experimental design was adapted from Swets et al. (2008). Plausibility is a within-subject factor and question type is a between-subject factor. Plausibility was manipulated only within the relative clause such as 'the dog that bit the man' (plausible) vs. 'the man that bit the dog' (implausible). The semantic properties were all normed by English native speakers in the previous studies, through a survey which asked people to rate each sentence on a scale from 1 ("implausible") to 7 ("extremely likely") (for more information about the norming survey, see Lim & Christianson (2013b)). Comprehension question type was manipulated between-participants in order to see the effect of question types on processing plausibility. There are two types
of questions: a relative clause question and a superficial question. One type of question asked about the content included in the relative clause (e.g., Did the dog bite the man?), and the other type of question probed about participant's understanding of the sentence at a very superficial level (e.g., Was something in the yard?). Below are examples of the four conditions.

(5) a. The dog that bit the man was in the yard. / Did the dog bite the man? (plausible/RC-question)
   b. The dog that bit the man was in the yard. / Was something in the yard? (plausible/superficial-question)
   c. The man that bit the dog was in the yard. / Did the dog bite the man? (implausible/RC-question)
   d. The man that bit the dog was in the yard. / Was something in the pool? (implausible/superficial-question)

Half of the questions for each condition had affirmative answers and the other half of them required negative responses. The yes/no answers were counterbalanced across the lists to control for the type of answer.

Filler sentences were identical for all lists, except the question types. The question types were modified to match the degree of the depth of the questions to the questions of experimental sentences for each condition. For the two lists with the RC-questions, the questions for filler items were created to require more attentions. For the other two lists with the superficial questions, the questions were made to be more general than the RC-question conditions. The experimental sentences were distributed across four lists and the four conditions for each item rotated across lists. Accordingly, participants read each experimental item for only one condition in the within-participants design.
3.3. Procedures

The experiment was run in a quiet room, and each participant sat in front of a computer programmed with E-prime professional software version 2.0. Korean native speakers were asked to read each English sentence appeared on the monitor in a self-paced, word-by-word, non-cumulative reading paradigm. After reading each sentence, a comprehension question appeared. Participants were asked to answer to a question by pressing yes or no button on the keyboard. A short practice session was conducted before they started a main experiment to get participants used to the procedure. There was a break session after they read half of the items in each list. After finishing the main experiment, they completed a cloze test that was used in Lim & Christianson (2013b), and a questionnaire that asks for language background information. The cloze test consisted of a short narrative story (adapted from American Kernel Lessons: Advanced Students' Book, 1981) and every fifth to seventh word was replaced with a blank with three words that participants can choose from. There were total 40 blanks and each blank received 1 point if it was correct. The mean score of the participants in this study was 30.8 out of 40. A procedure lasted about approximately 45 minutes to 1 hour depending on participant’s pace.

3.4. Results

Comprehension accuracy was overall 90% of the time on the filler sentences. Reading times analyses were conducted on all experimental sentences. Reading times greater than two standard deviations from the mean were replaced with the mean of the particular sentence in the same condition. Three critical regions of interests were defined as in: the relative clause region consisting of the words within the relative clause, the matrix verb region, and the post-verb region which consisted of the rest of the sentences after the matrix verb. Reading times of each word within the relative clause and the post-verb region were summed to
create the single dependent variable. An example of the three critical regions is as below (6).

(6) The dog that **bit** the man **was** in the yard.

Reading times were analyzed using a linear mixed-effects model (Baayen, Davidson, & Bates, 2008). Using a stepwise selection procedure, factors failing to reach at least \( p < .01 \) were excluded from the final model. Cloze scores were added as a factor to examine if reading times differed in response to L2 proficiency, but it did not reach at a significant level in any of the reading times analyses in all critical regions. The possible reasons for the absence of the effect of proficiency will be discussed in the Discussion section later. Thus, effects of plausibility, question type, and the interaction between plausibility and question type were included as predictors in the final, best-fitted model. Table 1 presents the grand mean reading times for the three critical regions in the four conditions and Table 2 shows the results of the linear mixed-effects model.

Table 1. Grand mean reading times (in milliseconds) of three critical regions in four conditions

<table>
<thead>
<tr>
<th></th>
<th>Plausible-RC</th>
<th>Plausible-All</th>
<th>Implausible-RC</th>
<th>Implausible-All</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>1969.94</td>
<td>1899.65</td>
<td>2051.12</td>
<td>2034.48</td>
</tr>
<tr>
<td>Matrix verb</td>
<td>508.80</td>
<td>488.54</td>
<td>605.35</td>
<td>571.68</td>
</tr>
<tr>
<td>Post-verb</td>
<td>1476.93</td>
<td>1314.39</td>
<td>1471.06</td>
<td>1438.70</td>
</tr>
</tbody>
</table>

Table 2. Fixed effects in the linear mixed-effects model of reading times

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>SE</th>
<th>t-value</th>
<th>( p )</th>
</tr>
</thead>
</table>

- 15 -
<table>
<thead>
<tr>
<th>Region</th>
<th>Intercept (Plausibility)</th>
<th>88.17</th>
<th>22.56</th>
<th>(\langle .01)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RC</strong></td>
<td>Intercept</td>
<td>1998.80</td>
<td>88.17</td>
<td>22.56 (\langle .01)</td>
</tr>
<tr>
<td>Plausibility</td>
<td>54.00</td>
<td>15.17</td>
<td>3.47</td>
<td>(\langle .01)</td>
</tr>
<tr>
<td><strong>Matrix Verb</strong></td>
<td>Intercept</td>
<td>543.59</td>
<td>17.56</td>
<td>30.95 (\langle .01)</td>
</tr>
<tr>
<td>Plausibility</td>
<td>44.92</td>
<td>6.81</td>
<td>6.60</td>
<td>(\langle .01)</td>
</tr>
<tr>
<td><strong>Post-verb region</strong></td>
<td>Intercept</td>
<td>1425.27</td>
<td>80.77</td>
<td>17.65 (\langle .01)</td>
</tr>
<tr>
<td>Plausibility</td>
<td>29.61</td>
<td>10.74</td>
<td>2.76</td>
<td>(\langle .01)</td>
</tr>
<tr>
<td><strong>Plau \times Q-types</strong></td>
<td>32.55</td>
<td>10.74</td>
<td>3.03</td>
<td>(\langle .01)</td>
</tr>
</tbody>
</table>

Note. \(p \langle .05\); Plau = Plausibility; SE = standard error

In the RC region and at the matrix verb, participants took longer times in reading implausible sentences than plausible sentences regardless of the question types, yielding significant effects of plausibility in both regions \(p \langle .01\). In both regions, there were no main effects of question types and interaction between two predictors. In the post-verb region, there was a main effect of plausibility, indicating that implausible conditions took longer to be read than plausible conditions. An interaction between plausibility and question types yielded a significant effect, indicating that reading times significantly differed depending on whether (im)plausible sentences were followed by RC-related questions or superficial questions. In other words, when people read implausible sentences, question types did not differentiate reading times, whereas when they read plausible sentences, they significantly sped up when the question types were superficial.

In order to examine how plausibility and question types affect readers' interpretations of sentences and reading strategies, times to answer to comprehension questions were analyzed using a linear
mixed-effects model. Table 3 shows how much times readers took to answer the questions in the four conditions. Main effects of both plausibility and question types were obtained ($t = 3.38$, $p < .01$ and $t = -3.78$, $p < .01$, respectively). Readers took significantly longer times to answer to comprehension questions when sentences were written with implausible than plausible events. Also, when they were asked about the content of the whole sentence at a superficial level, they took much less time to answer the questions. This finding suggests that people change their reading strategies depending on the task at their hands. In other words, readers seemed to involve in deeper processing of the sentence to answer the questions that asked about more specific contents, i.e., relative clauses, indicated by the longer times in the RC-question conditions.

Table 3. Question-Answering times in the four conditions (in milliseconds)

<table>
<thead>
<tr>
<th>Conditions</th>
<th>pl-RC</th>
<th>pl-all</th>
<th>impl-RC</th>
<th>impl-all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3730.88</td>
<td>2536.97</td>
<td>4207.60</td>
<td>2795.53</td>
</tr>
</tbody>
</table>

Note. pl=plausible; impl=implausible

Finally, comprehension accuracy was analyzed for all experimental sentences to investigate how the two factors have influences on their accuracy performance. A logit mixed model was used to analyze accuracy data because the data contain binary distribution (Jaeger, 2008). Table 4 shows the results of the logit mixed model, and Figure 1 illustrates the proportion of accuracy in the four conditions for better understanding of the data.

Table 4. Fixed-effects in the logit mixed model of comprehension question accuracy
<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>SE</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.23</td>
<td>0.21</td>
<td>10.52</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Plausibility</td>
<td>-0.67</td>
<td>0.08</td>
<td>-8.21</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Q-type</td>
<td>0.54</td>
<td>0.20</td>
<td>2.67</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Plau × Q-types</td>
<td>0.27</td>
<td>0.08</td>
<td>3.26</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Note. p < .05; Plau = Plausibility; SE = standard error

Figure 1. Comprehension question accuracy in the four conditions (pl=plausible; impl=implausible)

As seen in Table 4, the significant main effects of plausibility and question types were obtained: in other words, readers’ comprehension accuracy significantly dropped in the implausible sentences than in the plausible sentences, and in the RC-related questions than in the superficial questions, respectively. An interaction between two factors also reached at the significance level (p < .01). This interaction showed that when sentences were described in plausible events, the readers’ comprehension performance did not differ regardless of the question
types, whereas when implausible sentences were followed by RC-related questions, participants' accuracy was reliably lower than the sentences followed by superficial questions. Even though they took longer times to read the implausible sentences throughout the sentences, their interpretations about the implausible sentences were still significantly incorrect than those about the plausible sentences.

4. Discussion

The current study set out to investigate L2 learners' use of plausibility information during comprehension as a function of task types. In a self-paced reading paradigm, the Korean speakers of English read both plausible and implausible sentences and answered to two types of comprehension questions, the question about the specific content in the relative clause, and the question about the whole sentence at a superficial level. The reading times for the three critical regions (RC, main verb, post-verb) were found to be longer in the implausible conditions, regardless of the question types, as expected. An interaction between plausibility and comprehension question types was obtained in the post-verb region. Results from the analysis of question-answering times also indicated that L2 speakers spent much longer times to answer to the questions when the sentences were implausible, and when comprehension question was about the relative clauses. Lastly, comprehension accuracy data showed that the L2 participants were accurate the most in the plausible sentences with the superficial questions, and the least in the implausible sentences with the relative clause questions, indicated by the interaction between the two factors.

Based on the longer reading times in the implausible sentences from where the plausibility information begins, L2 speakers seemed to make use of plausibility information at the very early stage of the sentence parsing. This result is consistent with previous studies (Lim & Christianson, 2013; Roberts & Felser, 2011) which suggest that L2 speakers were heavily influenced by semantic information during on-line comprehension. After reading the first noun phrase
(NP) in the sentence, for example, 'the man', the L2 readers were assumed to expect what to come next more or less to make the sentence plausible and comprehendible. Accordingly, when they read the relative clause that included the implausible event such as 'bit the dog', they noticed something strange and quite inconsistent with the general knowledge in the world, their processing times from the RC region significantly increased and the longer reading times lingered to the post-verb region.

An interesting finding was that plausibility and question type interacted in the post-verb region. The L2 readers spent longer times to read implausible sentences than plausible sentences only when the comprehension questions asked about the superficial content of the sentences, not when the question was about the relative clauses. In other words, the L2 participants were influenced by the comprehension question types. This is exactly in line with the results of Swets et al. (2008)'s study with monolingual speakers. It appears that the disruption by the implausibility occurred in the RC region was effective until the end of the sentence only when the question was about the RC itself. That is, the comprehenders realized that they would have to answer to the content in the RC region in the RC question conditions, such that their reading times did not significantly differ regardless of plausibility. When the question was about the superficial content, however, the implausibility seemed to be faded away in the reader's mind as the sentence moved along until the verb. By the time when the readers were toward the end of the sentence, they reminded of the implausibility of the sentence, which occurred in the RC region, resulting in much longer reading times than in the plausible condition.

This result suggests that the L2 participants changed their reading strategies depending on the question types that they were asked, just as the L1 participants in the study of Swets et al. (2008). In the superficial question group, they kept seeing a superficial question, they might have read the sentences with less focus. On the other hand, in the RC question group, since they were asked to answer about the content in the RC region, their attention to the implausibility
information might have been elevated throughout the sentence. Indeed, previous research suggests that attention and memory for certain representations of either lexical, semantic, or syntactic levels are very likely to be affected by experimental tasks (Townsend & Bever, 1991; Townsend, Hoover, & Bever, 2000). The current finding can be another evidence of the susceptibility of attention and memory as a function of different types of tasks, emphasizing the task effects on language processing.

The question-answering times in this study were also telling about the effect of task on reading. The finding that the L2 speakers took much longer times to answer to the RC questions can be the evidence of the L2 learners’ strategic reading depending on the question types. If the types of questions did not have any influence on reading, participants would have shown the similar times to answer the two types of questions, which did not happen in this study.

Furthermore, the comprehension accuracy data in this study provides important information on L2 processing mechanism. Our L2 participants were significantly more accurate in the plausible sentences than the implausible sentences, indicated by the main effect of plausibility. The main effect of question types showed more accuracy rates in the superficial questions than the RC questions, which was opposite to our prediction. We expected that if the task requires more demands, the accuracy levels would be higher. This finding was unexpected, but is well-explained by the interaction between plausibility and question types. The interaction between plausibility and question types indicated that the differences in accuracy between plausible and implausible conditions dependent on the question types were significantly large (see Figure 1 above). The L2 participants were the most inaccurate in the implausible sentences with the RC questions. It seems that the L2 learners’ tendency to rely on the world knowledge is such strong that even when the task required the specific information on the content in the RC, their final interpretation appeared to be already overridden by the representation from the semantic/heuristic route. Therefore, the L2
speakers comprehended the sentence in the way that they thought it to be 'good-enough' for comprehension, which resulted in a final interpretation which was an incomplete and not faithful to the text, when the sentence was implausible. This result is compatible with the L1 studies suggesting the good-enough processing approach in language comprehension (Christianson et al., 2001; Ferreira & Patson, 2007) and supports that L2 processing can also be accounted under the GE processing (Lim & Christianson, 2013b).

Even though they spent longer times in reading the implausible sentences, they still failed to reach at the final interpretations correctly. This provides an important caveat to the future research in language processing literature. Most of the studies measured reading times in critical regions of interests and inferred from the reading times to discuss processing difficulties. However, as this study showed, even if participants spent longer reading times in the critical region, they still might get the final interpretation which is not faithful to the text. Thus, it needs to be carefully dealt with reading time measures when interpreting its processing difficulties in language comprehension.

Finally, as mentioned above, we added the L2 learners' cloze test scores as a predictor when analyzing the reading times and comprehension accuracy data to examine the effect of L2 proficiency on processing plausibility information in the different types of tasks. The L2 proficiency did not reach at the significance level statistically in any of the data, however. We carefully speculate two possible reasons for the absence of L2 proficiency. First, it might be due to the nature of the cloze test. As described in the Procedure section, the cloze test used in the current study included each blank with three options to choose, instead of blanks that participants need to fill out. Lim and Christianson (2013a) used the same cloze test of the current study but with a fill-in-the-blank manner. The cloze test in their study was better able to discriminate low-proficiency learners from high-proficiency learners than the one used in the current study. However, given
that Lim & Christianson (2013a)'s study was conducted in the U.S.A., and that the subject pools are different from their study and ours, we cannot surely say that the different method of the cloze test was attributed to the absence of proficiency effect in the present study. Second, most of the participants in this study were undergraduate students at a large university in Seoul. There is a possibility that this group is so homogeneous that we could not observe any proficiency effect in our experiment. Although the current study did not find L2 proficiency effect, it would be interesting to further research on whether and how L2 proficiency influences on processing plausibility information and different types of questions, by recruiting clearly lower-proficiency and higher-proficiency learners.

5. Conclusion

The present study investigated how L1 Korean speakers of English make use of semantic information in on-line reading without manipulation of syntactic structures and how different types of comprehension questions can affect their reading patterns in the time course of processing. In consistent with previous studies, L2 learners made use of semantic information at the early stage of processing. It was also found that L2 learners modulated their reading strategies depending on the comprehension question types, and that their final interpretations on the implausible sentences were sometimes based on the good-enough representations, supporting for the good-enough language processing approach. Task effects observed in this study can have several pedagogical implications after all: for example, teachers can modulate their language task materials depending on teaching focus for language learners. This study adds new experimental data to L2 processing research emphasizing effects of task demands on real-time sentence comprehension.
References


- 24 -


Compass 1:71-83.