The Distance Effect in the Processing of English Oblique RCs*

Chae-Eun Kim
(Korea University)


This study investigates whether native speakers of English indicate any difficulty in processing two types of oblique relative clauses, both involving a directional prepositional phrase. Previous work on sentence processing with native English speakers has shown that subject RCs were read faster than direct object RCs in a self-paced reading task. However, the previous research is problematic due to uncontrolled animacy of heads, comparison between canonical and non-canonical word-order structures, and the frequency factor. Therefore, the present study tests oblique RCs with different lengths of filler-gap dependency in order to separate the effects of distance between the filler and the gap in English relative clauses from the effects of other factors. The study examines the comprehension accuracy and total reading times of 40 adult native speakers who read stimuli with English oblique relative clauses in an on-line reading task. Overall, the results show that the participants had lower accuracy scores on the oblique RCs with a long filler-gap dependency (FGD) than on the oblique RCs with a short FGD. In addition, they had more difficulty processing the oblique RCs with a long FGD than oblique RCs with a short FGD, and this difficulty increased in the relative clause and main verb regions, as indicated by longer reading times. The results support an effect of the length of the filler-gap dependency on processing.

Keywords: relative clauses (RCs), L1 sentence processing, filler-gap dependency, oblique

1. Introduction

Relative clause (RC) structure is one of the syntactic phenomena that

* I am very grateful to the three anonymous reviewers for their helpful comments and suggestions. All remaining errors are my own.
have received much interest in the psycholinguistics literature, because RCs have long-distance dependencies between the head and the gap. Previous studies on sentence processing of RCs have usually compared two types of RCs, subject and direct object RCs, as in (1) and (2).

(1) Subject RC:
   The reporter [that _ attacked the senator] disliked the editor.

(2) Direct object RCs:
   The reporter [that the senator attacked _] disliked the editor.

This popular comparison, which has been employed in first/second language acquisition and psycholinguistics, has confirmed that subject RCs have processing advantages over direct object RCs (Warren & Maratsos 1978; Flynn & Lust 1980; Ford 1989; King & Just, 1991; Gibson 1998; Traxler et al. 2002; Gennari & MacDonald 2008). Even when the two types of RCs involve identical lexical items in the structures, direct object RCs are much more complex to process than subject RCs. Hence, the complexity of direct object RCs has been explained by the properties of its syntax. However, Kidd et al. (2007) report that direct object RCs in actual speech typically have inanimate head nouns, as in (3).

(3) Direct object RC with an inanimate head noun:
   The report [that the senator submitted _] surprised the governor.

According to Kidd et al. (2007), it is very rare to find direct object RCs with an animate head as in (2) in either child speech or child-directed speech. However, the previous studies did not separate animacy of the head nouns from syntactic complexity, which raises the question of how to design a better test of the processing difficulty of certain types of RCs.

---

1) These studies have tested whether native and second language speakers of English process relative clauses in the way predicted by Keenan and Comrie’s (1977) Noun Phrase Accessibility Hierarchy (NPAH). If the same hierarchy (subject > direct object > indirect object > …) governs relativization, then subject RCs are less marked typologically than other types of RCs (i.e., those to the right in the hierarchy).
If the reported difficulty of direct object RCs might be partly due to an animacy effect, we do not yet know to what extent the difficulty is due to the properties of its syntax.

The present study investigates RC processing in English and examines how the factor of syntactic complexity affects the comprehension of RC construction in English. The next section discusses previous studies on relative clause processing that compare two types of English oblique RCs. In addition, factors that have been shown to influence the computation of various types of relative clause will be discussed, with a focus on the effects of syntactic complexity on sentence processing. The rationale behind the present study, which seeks to resolve several problems in previous research, is explained. In section 3, this study's method is explained in detail. Analysis of the data and discussion of the results follow in section 4. In section 5, the implications that can be drawn from the results of this study are discussed. Section 6 concludes the paper and proposes possible follow-up studies.

2. Processing of Relative Clauses

2.1. Length of the filler-gap dependency

As pointed out, previous research on the processing of English relative clauses has mainly focused on comparison of subject and direct object relative clauses. The subject RC over direct object RC advantage has been reported in child language acquisition studies (Flynn & Lust 1980; O'Grady et al. 2003; Friedmann & Novogrodsky 2004), as well as in psycholinguistics (Warren & Maratsos 1978; King & Just 1991; Gibson 1998; Traxler et al. 2002; Gennari & MacDonald 2008). The processing difficulty of RCs is related to the length of the 'filler-gap dependency'; that is, the distance between the filler and gap determine the greater ease or difficulty of processing various types of RCs (Hawkins 1999). In particular, the distance factor (O'Grady 2011) predicts that English direct object RCs (4b) are processed with more difficulty than their subject counterparts (4a). Gibson (1998, 2000) proposes that the amount of lexical material
between the head and the gap, such as NPs and verbs, influences the difficulty of processing. This happens because the head noun and the gap in the unfolding relative clauses are closer in (4a) than in (4b).

(4) a. Subject RC
   the reporter [that _ attacked the senator]
   |____0____|

b. Direct object RC
   the reporter [that the senator attacked _]
   |______________1_______2____|

In particular, Gibson’s (1998) Dependency Locality Theory (DLT) explains that dependencies between the head and the gap are constrained by both processing storage and integration resources. No materials intervene between the filler and the gap in the subject RC (4a), thus requiring minimal effort from a reader to find the filler-gap dependency (Hawkins, 2004; O’Grady 2011). In contrast, direct object RCs draw on more integration resources because they have two intervening items of lexical material, thus requiring a longer time to connect the filler with the empty gap; the new discourse material (e.g., the senator and attacked) between the filler and the gap (4b) creates a burden on working memory before the dependency can be resolved. To describe this phenomenon, O’Grady (2011) proposes the distance factor in (5).

(5) The distance factor (O’Grady 2011, p. 22)
   The difficulty of processing a relative clause increases with the length of the filler-gap dependency.

O’Grady (1997, p. 179) previously proposed that the structural (hierarchical) distance between the filler and the gap would be the main source of processing difficulty in English RCs. When a gap is deeply embedded inside a structure, readers find its processing burdensome. This paper takes his idea of the (linear) distance as a factor in syntactic complexity to explain the processing of English RCs with filler-gap dependencies.
2.2. Animacy influence

Some research has proposed that the posited difficulty of direct object RCs arises from previous works’ use of stimuli with head noun animacy issues, as in (6a) (Traxler et al. 2002; Mak et al. 2002; Warren & Gibson 2002; Grodner & Gibson 2005; Reali & Christiansen 2007). These studies have claimed that the vast majority of direct object RCs in actual speech have inanimate head nouns, as in (6b).

(6) a. Direct object RCs with animate head:
the reporter [that the senator attacked _]
b. Direct object RCs with inanimate head:
the report [that the senator submitted _]

In addition, the claim that the processing difficulty of the filler-gap dependencies in direct object RCs can be minimized when the head noun is inanimate has been supported by various studies (Traxler et al. 2002; Mak et al. 2002; Warren & Gibson 2002; Grodner & Gibson 2005; Reali & Christiansen 2007). Aissen (2003), also, argues that subjects generally take animate NPs and direct objects take inanimate NPs in sentences. Altogether, the issue of the animacy of head nouns in direct object RCs is clearly one of the obstacles to finding the true source of difficulty of processing sentences with direct object RCs.

For this reason, Kim (2013, 2015) and Kim and O’Grady (2015) tested different types of relative clauses, subject RCs and indirect object RCs, to examine the filler-gap dependencies in complex RC patterns with items that controlled the animacy of the head nouns, as exemplified in (7a) and (7b).

(7) a. Subject RCs: the musician [that _ sent the book to the director]
b. Indirect object RCs: the musician [that the director sent the book to _]

Subject and indirect object relatives with an animate head require the same level of syntactic processing because animate nouns (the musician) are good candidates for both agents and experiencers, and they are also
semantically reversible (a musician can send a book to a director, and also a director can send a book to a musician). Therefore, these studies have teased the length effect apart from the animacy issue, as either or both factors may have an effect on the processing of RCs.

In Kim’s (2015) work, forty-two native speakers of English showed no difference in comprehension scores for these two types of RCs; however, their total reading times with indirect object RCs were longer than with subject RCs. In sum, Kim’s results support the suggestion that the length of the filler-gap dependency is strongly relevant to sentence processing, proving that readers found dependencies in indirect object RCs more difficult to compute.

### 2.3. Controlling word-order canonicity

Thus far, we have discussed how, because the traditional comparison (subject RC vs. direct object RC) was problematic, some studies have attempted to control the animacy factor by using subject and indirect object relatives (Kim 2013, 2015; Kim & O’Grady 2015). Although the results clearly showed that subject RCs were easier to produce and comprehend than indirect object RCs, which appears to support a distance effect, an issue still arises with respect to word order. Kim and O’Grady (2015) pointed out that their comparison (subject vs. indirect object RCs) is still troublesome because while subject relatives follow English canonical word order, indirect object relatives do not, as in (8).

\[ \begin{align*}
(8) & \quad \text{a. Subject RCs: the musician [that } _\_ \text{ sent the book to the director]} \\
& \quad \quad S \quad V \quad DO \quad IO \\
& \quad \quad \text{b. Indirect object RCs: the musician [that the director sent the book to } _\_ \text{]} \\
& \quad \quad IO \quad S \quad V \quad DO
\end{align*} \]

As in (8a) the subject RC follows a canonical SVX-like word order whereas the indirect object RC displays a non-canonical pattern. This fact implies a possible alternative explanation for the difficulty of indirect object RCs. To disentangle word-order canonicity from the distance factor, Kim (in progress) compares the difficulty of direct object RCs with the difficulty of their oblique counterparts, as in (9).
There are two advantages to this comparison. First, oblique relatives, like direct object relatives, usually have inanimate heads — effectively controlling the animacy factor. Second, relative clauses with two inanimate arguments cannot be comprehended based on real-world relations alone: actual syntactic processing is required to interpret the book in (9b) as the theme (or patient)/direct object and the carton as the locative/oblique. Third, because both patterns follow non-canonical word order, the proposed comparison allows a more direct test of the effect of length on filler-gap dependencies. If in fact a longer dependency increases processing cost, oblique relatives should prove to be more difficult than their direct object counterparts.

Overall, these results indicate that the participants comprehended direct object RCs better. Comprehension accuracy in the direct object RC and oblique RC conditions was 83% and 56%, respectively, and differed significantly across conditions. Total RT data analyses demonstrate that the participants actually read the oblique relatives much more slowly than they read the direct object relatives. There is a robust gap effect on sentence processing, establishing that the readers had greater difficulty in the oblique relative than the direct object relative condition. In sum, based on both comprehension scores and reading time measurements, the oblique relatives were more difficult than the direct relatives.

2.4. Frequency effect

However, even if some previous studies support the length effect (Kim 2013, 2015, in progress), a new problem arises, because the predicted results might be attributed to frequency: subject RCs are more frequent
than indirect object RCs, and direct object RCs are more common than oblique RCs. The present study aims to test a purer distance effect by controlling the possible roles of factors such as animacy, word order, and frequency, which may affect RC processing. In order to control for frequency, two types of oblique RCs are compared because both involve a directional prepositional phrase but they differ in the length of the filler-gap dependency (FGD), as (10) shows (see Reali & Christiansen, 2007).

(10) a. Oblique RC with short FGD
   the car [that the boy walked to _ from the church]

b. Oblique RC with long FGD
   the church [that the boy walked to the car from _]

Though oblique RCs of both types are low in frequency in some corpora (Hale, 2006, p. 658), they have similar word-order canonicity and both contain inanimate nouns, but the filler-gap dependency (FGD) is shorter in (10a) than in (10b).

The distance hypothesis predicts that oblique RCs with a short FGD, such as (11a), will be easier to process than oblique RCs with a long FGD, such as (11b).

(11) a. Oblique RC with short FGD:
   the car [that the boy walked to _ from the church]
   1  2

b. Oblique RC with long FGD:
   the church [that the boy walked to the car from _]
   1  2  3

Two referential elements (the boy, walked) intervene between the filler (the car) and the gap in (11a); but in (11b), three referential elements (the boy, walked, the car) intervene between the filler (the church) and the gap. Thus, this hypothesis predicts that in English, oblique RCs with a longer FGD will be more difficult for readers than oblique RCs with a shorter FGD.
2.5. Research question

The research question motivating the present study is as follows:

Do native English speakers process oblique RCs with a short FGD more easily than oblique RCs with a long FGD when the animacy of the head nouns, word-order canonicity, and frequency are all controlled? Does the factor of distance affect readers' sentence processing?

The prediction of the study is that, with animacy, canonicity, and frequency controlled, a greater length of filler-gap dependency in an oblique relative clause should cause the readers more difficulty in processing than a shorter length of FGD in an oblique relative clause.

3. Method

In this experiment, participants read sentences containing relative clauses. Half of the sentences contained an oblique relative clause with a short FGD and half contained an oblique relative clause with a long FGD. In both types of sentences, the sentential subjects were animate and human (e.g., 'hiker'), the noun phrases in the relative clauses were reversible (i.e., both were inanimate and non-human), and both were good locatives for the action denoted by the verb in the relative clause.

3.1. Participants

Forty native speakers of English participated in the experiment. Twenty undergraduate and twenty graduate students participated. They were paid $5 for their participation in the experiment, which lasted less than 15 minutes. All of the participants were native speakers of American English.
3.2. Materials and design

A self-paced moving-window reading task was used to explore whether oblique RCs with a short FGD were read faster than oblique RCs with a long FGD when the head noun phrase was an inanimate noun and both types of RCs had a non-canonical word order. There were sixteen sets of experimental sentences and 32 fillers of various types. Sample sentences of each condition are shown in table 1. All the sentences were automatically randomized. After the last word of each sentence, participants used the keyboard to reply to yes-no or wh-comprehension questions.

Table 1. Sample Test Items

<table>
<thead>
<tr>
<th></th>
<th>(12a) Oblique RC A (with a short FGD)</th>
<th>(12b) Oblique RC B (with a long FGD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>the</td>
<td>the</td>
</tr>
<tr>
<td>R2</td>
<td>hilltop</td>
<td>valley</td>
</tr>
<tr>
<td>R3</td>
<td>that</td>
<td>that</td>
</tr>
<tr>
<td>R4</td>
<td>the</td>
<td>the</td>
</tr>
<tr>
<td>R5</td>
<td>hiker</td>
<td>hiker</td>
</tr>
<tr>
<td>R6</td>
<td>wandered</td>
<td>wandered</td>
</tr>
<tr>
<td>R7</td>
<td>to</td>
<td>to</td>
</tr>
<tr>
<td>R8</td>
<td>from</td>
<td>the</td>
</tr>
<tr>
<td>R9</td>
<td>the</td>
<td>hilltop</td>
</tr>
<tr>
<td>R10</td>
<td>valley</td>
<td>from</td>
</tr>
<tr>
<td>R11</td>
<td>was</td>
<td>was</td>
</tr>
<tr>
<td>R12</td>
<td>very</td>
<td>very</td>
</tr>
<tr>
<td>R13</td>
<td>flat</td>
<td>flat</td>
</tr>
</tbody>
</table>

Only the order of the words in the relative clauses is changed for the oblique RC with short FGD as in (12a) or the oblique RC with long FGD as in (12b). The two types of sentences had the same length in terms of number of words across conditions, and each word was considered

2) All test sentences were checked by a group of English native speakers for naturalness.
a region (R). The items were randomized and assigned to one of two lists following a Latin-square design, so that no reader was shown more than one version of each sentence.

3.3. Procedure

The task adopted in the study used non-cumulative, self-paced, word-by-word reading with a moving window display (Just, Carpenter, & Woolley 1982). The E-Prime experimental software package on a PC was used to test participants. Participants sat in front of a computer screen in a booth and they saw one word at a time, beginning with the first word. They pressed a button to call up the next word. Every time they pressed the button, a new word appeared and the preceding word disappeared. After each sentence, participants responded to a yes–no question or a wh-question about its content (e.g., “Where did the hiker walk to?”). Their reading times at all regions in the sentences and their answers to the comprehension questions were recorded for analysis. Before the experiment began, participants were given practice items until they understood how the task worked.

4. Results

Only the reading time data from items with correct answers to the comprehension questions were analyzed.

4.1. Comprehension accuracy

The mean scores on the comprehension questions are shown in figure 1. The overall mean proportion of correct responses to all comprehension questions was 73%. Comprehension accuracy in the oblique RC with a short FGD and oblique RC with a long FGD conditions was 77% and 70%, respectively, and differed significantly across conditions ($F(1, 39) = -2.531; p < .05$). Overall, these results indicate that participants comprehended sentences in condition A better than those in condition B.
4.2. Total reading times (RTs)

For the total reading times, RTs higher than 2.5 SD (standard deviation) of the mean per word position were replaced by a cutoff value within this range. Two different analyses on total RTs were conducted for two sets of regions: (1) R1–R10 and (2) R1–R11.

The first analysis looked at the RTs for R1–R10; (13a) and (13b) are examples showing the words included in this region set. The results are shown in figure 2.

(13) a. OBL(A): The hilltop [that the hiker wandered to _ from the valley]
    
    b. OBL(B): The valley [that the hiker wandered to the hilltop from _ ]

The grammatical function of the distance (short FGD vs. long FGD) was used as an independent variable, and total RTs were used as the dependent variable. Data were submitted to a one way ANOVA to compare differences in total RTs for each condition. The total RT for the oblique RC (A) condition was numerically shorter than the total RT for the oblique RC (B) condition (5422 ms vs. 5712 ms). However, there was no significant difference between oblique (A) and (B) conditions.

---

3) As one of the reviewers pointed out, region-by-region analyses would be a more appropriate way to investigate processing difficulty in RCs with filler-gap dependency. Such analyses, which would require not raw reading times but residual reading times (RTs) because the words in each region differ in length, must be left for future study.
The second analysis looked at total RTs in R1–R11; region 11 was the main verb region, as shown in (14a) and (14b).

(14) a. OBL(A): The hilltop [that the hiker wandered to _ from the valley] was  
    b. OBL(B): The valley [that the hiker wandered to the hilltop from _] was

The total RT results in these regions are shown in figure 3. There was a significant effect of length ($F(1, 39) = -3.6; p = .001$).

The RT data analyses show that the participants actually read sentences in the oblique RC (A) condition, with short FGDs, quicker than they read sentences in the oblique RC (B) condition, with long FGDs. In sum, based on reading time measurements, the oblique relatives with long FGDs were more difficult than the oblique relatives with short FGDs.
5. Discussion

Both comprehension scores and total RTs show a difference between oblique relative clauses with different lengths between the filler and the gap. The research question of this study asked whether the type of oblique relative clause (i.e., short FGD vs. long FGD) has an effect on sentence processing. As reported above, the results of the comprehension scores indicate that English oblique RCs with short FGDs are easier for English-speaking adults to comprehend and judge than oblique RCs with long FGDs (77% vs. 70%).

The analyses of the total reading times (RTs) also support the claim that participants had more difficulty processing oblique RCs with long FGDs (B) than oblique RCs with short FGDs (A), showing a significant difference in the RTs between the two types of oblique relatives. This difference was caused by elevated RTs in both region sets tested: (1) relative clause with a head noun; (2) relative clause with a head noun and main verb. The slowdown that appeared in each analysis for the oblique relative (B) condition suggests that it is the length of the filler-gap that affects syntactic processing.

The first region set consisted of ten words (the hilltop that the hiker wandered to from the valley in the oblique (A) condition vs. the valley that the hiker wandered to the hilltop from in the oblique (B) condition), which was critical to the experimental research question. A distance effect was detected in this set, which included only the relative clause with the head noun, with participants in the study spending more time on the oblique (B) than the oblique (A) condition (5712 ms vs. 5422 ms). Moreover, if readers employ an active filler-gap strategy (Frazier & d’Arcais 1989), they will consider the head noun the subject of the relative clause and tend to put the gap in the subject position after the relative pronoun that in both types of oblique relative clauses. As the readers would first interpret the head noun as subject and then, at R4, realize that their initial syntactic interpretation is wrong, they would need to engage in additional retrieval in the relative clause region with the head noun (R1–R10). The second region set additionally included the main verb (R11).
At R11, the matrix verb was read much faster in the oblique (A) condition (1627 ms vs. 2139 ms), and the differences between the two conditions were larger for this region set. Again, the difference in RTs leads to the conclusion that the readers had greater difficulty in the oblique (B) condition than in the oblique (A) condition. As already pointed out, it is the distance of the filler-gap dependency that causes more difficulty in the processing of the former condition.

Overall, this study shows that both comprehension accuracy and total RTs support a distance effect on sentence processing; the study’s detailed analyses demonstrate that the difficulty takes place in both the relative clause region and the main verb region.

6. Conclusion

This study investigated how the effects of linear distance between the head and the gap in RCs arise when parsers read two types of oblique RCs. Previous works comparing either subject with indirect object relatives or direct object with oblique relatives have supported the argument that the distance between the filler and the gap position in RCs causes processing difficulty. However, they left open the possibility that direct object relatives are easier to process than oblique relatives because of frequency: direct object relatives are more common than oblique relatives. For this reason, this study of the distance effect was conducted by comparing both comprehension accuracy and total reading times for two types of oblique RCs.

The advantage of this comparison is that both conditions follow non-canonical patterns in word order and they are infrequently used in actual speech. In addition, both oblique NPs are typically inanimate, allowing this comparison to avoid issues related to animacy.

The results from this experiment confirm that the distance factor between the filler and the gap can predict processing difficulty. O’Grady (2011) and Hawkins (2004) affirmed that the memory difficulties of distance increase when there is more intervening material. Although oblique relatives with a short FGD have two intervening lexical items (e.g., the NP the hiker and the verb wandered), oblique relatives with a long FGD
have one more (e.g., the NP *the hiker*, the verb *wandered*, and the NP *the hilltop*) and thus when the head noun is integrated at the gap position, parsers have extra processing difficulty in oblique relatives with a long FGD.

The comprehension score and total RT data analyses showed that participants comprehended oblique RCs with a short FGD better and quicker than oblique RCs with a long FGD, with a significant difference between the two types. Reading time slowdown in the oblique RC with long FGD condition fully supports this idea. Therefore, the effect of the factor of distance in processing difficulty is supported by the present study. The fact that identical words were used in the two conditions, but comprehension accuracy and total reading times nonetheless differed significantly, indicates that differing levels of syntactic complexity give rise to asymmetrical patterns; in addition, different syntactic patterns are associated with greater or lesser working memory loads in relative clause processing.

The results of this study show similar patterns to those found by Kim (2013), who observed that oblique RCs with a short FGD were easier to produce than oblique RCs with a long FGD. The current findings suggest that the native speakers of English in this study tended to differentiate between the two types of oblique RCs, while the findings of previous works on L1 acquisition have revealed that English- and Korean-speaking children are better at producing oblique relatives with a short FGD than oblique relatives with a long FGD (Kim 2013). These findings for both young children and adults are compatible with predictions based on the distance factor (Gibson 1998, 2000; O'Grady 2011; Kim & O'Grady 2015).

This study used a new comparison (two oblique relatives) and controlled the factors of animacy, word-order canonicity, and frequency to investigate the effect of the length of the filler-gap dependencies. However, there is still a possibility that oblique relatives with a short FGD are easier to process than oblique relatives with a long FGD because of PP order. English allows two different patterns: (i) S-V-OBL (source)-OBL (goal); (ii) S-V-OBL (goal)-OBL (source). These elements can be freely ordered with respect to each other because of the presence of the contrasting prepo-
sitions from and to. Thus, additional research is needed to examine the effect of PP order by including stimuli with the S-V-OBL (source)-OBL (goal) order, as in (15).

(15) a. The hilltop [that the hiker wandered from_ to the valley]
   b. The valley [that the hiker wandered from the hilltop to _]

For this reason, further studies of the distance effect are needed. Also, given the possibility that second language learners of English are influenced by the same pressures, further research needs to be conducted to see how L2 learners process parallel contrasts (e.g., oblique RCs with short or long FGDs).

References


The Distance Effect in the Processing of English Oblique RCs

Language Acquisition 25.3, 433-448.

Chae-Eun Kim
Department of English Language and Literature
Korea University
145 Anam-ro, SeongBuk-gu
Seoul 136-701, Korea
Email: alohachaeun@gmail.com

Received: November 2, 2015
Revised version received: December 19, 2015
Accepted: December 23, 2015