BRIEF RESEARCH REPORT

Asymmetry in children’s comprehension of raising*

JINSUN CHOE
Hankuk University of Foreign Studies

AND

WILLIAM O’GRADY
University of Hawai’i at Mānoa

(Received 6 June 2015 – Revised 15 November 2015 – Accepted 22 February 2016 –
First published online 11 April 2016)

ABSTRACT

This paper investigates English-speaking children’s acquisition of raising constructions (e.g. John seems to Mary to be happy) and finds an asymmetric effect of NP type on their comprehension: an improvement in performance is observed when a lexical NP is raised across a pronominal experiencer (e.g. John seems to her to be happy) compared to when a pronoun is raised across a lexical NP experiencer (e.g. He seems to Mary to be happy). These results are consistent with a processing-based approach to intervention effects, which reduces children’s difficulty with raising to a performance limitation, rather than a grammatical deficit.

INTRODUCTION

This paper investigates English-speaking children’s acquisition of raising constructions such as Donald seems to Mickey to be short, in which an experiencer argument (Mickey) ‘intervenes’ between seem and the predicate to be short. We argue that children’s poor performance on these patterns,

[*] This work was funded by the NSF Doctoral Dissertation Research Improvement Grant (BCS-1227232). We would like to thank Kamil Deen, Bonnie D. Schwartz, Shinichiro Fukuda, and two anonymous referees for their helpful comments, as well as the audience at the WCCFL (2013) and GALA (2013). All errors remain ours. Address for correspondence: Jinsun Choe, Hankuk University of Foreign Studies – Graduate School of Education, 107 Imun-ro Dongdaemungu, Seoul Seoul 130–791, Republic of Korea. e-mail: jinsun@hufs.ac.kr
which is crucial to a more general understanding of the acquisition of raising patterns in general, is best attributed to a performance limitation commonly referred to as an intervention effect.

**BACKGROUND**

The term ‘raising’ is often used to describe the relationship between the two patterns exemplified in (1).

(1)  
a. It seems [that John is working].  
b. John seems [to be working].

In (1a), the NP *John* is the subject of the verb *work*. In (1b), in contrast, it has ‘raised’ to the subject position in the higher clause, where it triggers agreement on the verb *seem* (see Perlmutter & Soames, 1979, and Postal, 1974, for classic analyses).

We focus here on a subset of raising patterns in which an experiencer argument (e.g. *Mary* in (2)) occurs between the matrix verb and the infinitival clause.

(2)  
a. Unraised: It seems *to Mary* [that John is working].  
b. Raised: John seems *to Mary* [ _ to be working].

Several acquisition studies (Hirsch, 2011; Hirsch, Orfitelli & Wexler, 2007; Hirsch & Wexler, 2007) report that while young children comprehend unraised patterns such as (2a), they have difficulty comprehending their raised counterparts such as (2b). See Choe and Deen (2015) for a comprehensive review of the literature on the acquisition of raising with and without an experiencer.

Children’s difficulty with patterns such as (2b) has often been interpreted as evidence for deficits in child grammar (Borer & Wexler, 1987). Detailed proposals along these lines include the Universal Phase Requirement (UPR; Wexler, 2004), the Universal Freezing Hypothesis (UFH; Hyams & Snyder, 2005), and the Argument Intervention Hypothesis (AIH; Orfitelli, 2012), which offer competing grammar-based explanations for the acquisition path of A-movement structures, such as passives, unaccusatives, and raising. An alternative hypothesis, which we endorse, favors a processing explanation. On this view, the difficulty of raising patterns with an experiencer is attributed to processing-related limitations that are independently manifested in a variety of other constructions, including those involving A’-movement (which underlies the formation of *wh*-questions and relative clauses, among other patterns). Although there are important structural differences between A-movement and A’-movement, the two have in common the existence of a filler–gap dependency, which is a crucial part of the processing-based hypothesis that we pursue.
It is well known that children have difficulty processing structures in which an NP intervenes between a filler and the associated gap in cases of A'-movement (e.g. Avrutin, 2000; de Vincenzi, Arduino, Ciccarelli & Job, 1999; Friedmann, Belletti & Rizzi, 2000; Friedmann & Lavi, 2006). These structures include object relatives, object topicalizations, and object wh-questions.

(3) a. Object relative: The boy who the girl kissed _

b. Object topicalization: The boy, the girl kissed _

c. Object wh-questions: Which boy did the girl kiss _?

Raising patterns resemble these constructions in that the experiencer argument intervenes between the raised NP and its original position, as illustrated in (4).

(4) Raising: The boy seems to the girl _ to be happy.

A signature property of intervention effects is that their difficulty can be modulated by manipulating the features of the intervener (Arnon, 2010; Friedmann et al., 2009; Kidd, Brandt, Lieven & Tomasello, 2007). For example, Kidd et al. (2007) report that English-speaking children are better at repeating object relatives with pronominal interveners, as in (5a), than ones with lexical NP interveners, as in (5b). Friedmann et al. (2009) and Arnon (2010) report a parallel result for Hebrew-speaking children on a comprehension task.

(5) a. Pronoun intervener: 
   the nurse [that I am drawing _]
   ↑ ___________________ ↑

b. Lexical NP intervener: 
   the nurse [that the girl is drawing _]
   ↑ ___________________ ↑

Crucially, Choe and Deen (2015) found that children’s comprehension of raising sentences significantly improved with a pronominal experiencer, as in (6a), compared to a lexical NP experiencer, as in (6b).

(6) a. Pronoun intervener (81.5% accuracy): 
   Bart seems to her [_ to be studying].
   ↑ ___________________ ↑

b. Lexical NP intervener (40.9% accuracy): 
   Bart seems to Lisa [_ to be studying].
   ↑ ___________________ ↑
This finding suggests that children’s difficulty with raising patterns is sensitive to the same sort of influences that contribute to intervention effects in other structure types.

Different theories have been put forward as to the exact motivation for intervention effects in language acquisition and processing, including a grammatical account built on Relativized Minimality (Belletti, Friedmann, Brunato & Rizzi, 2012; Friedmann et al., 2009; Rizzi, 2013) and a processing-based account that draws on Dependency Locality Theory (DLT; Gibson, 2000).¹

Relativized Minimality (RM) invokes a syntactic locality principle, proposed by Rizzi (1990, 2004), which rules out a dependency involving X and Y in the configuration ... X ... Z ... Y ..., if Z is too ‘similar’ to X. Friedmann et al. (2009) argue that children apply a stricter version of RM than adults, requiring the features of X and Z to be entirely distinct. This idea predicts that young learners should have trouble with object relatives such as (5b), repeated as (7) below, in which the relativized head the nurse (X) and the intervener the girl (Z) share the NP feature.

(7) X Z Y
the nurse [that the girl is drawing _]
[+NP] [+NP]
↑________________________↑
('+NP' designates a full lexical noun phrase.)

As we have seen, children do in fact have trouble comprehending object relative clauses of this type. And, predictably, they have been reported to do better on patterns such as (5a), repeated as (8) below, in which the intervener is a pronoun rather than a lexical NP.

(8) X Z Y
the nurse [that I am drawing _]
[+NP] [+pronoun]
↑________________________↑

While the RM account focuses on the similarity of the intervener to other NPs in the construction, the Dependency Locality Theory (Gibson, 1998, 2000) focuses on its referential accessibility. For example, Warren and Gibson (2002) show that processing difficulty is sensitive to the information-related gradations that make up the Givenness Hierarchy (Gundel, Hedberg & Zacharski, 1993) in (9), which correlates NP types with the accessibility of their referents.

¹ See Snyder and Hyams (2015) for a similar grammatical approach using RM, and see Gordon et al. (2001, 2004) and Van Dyke (2007) for different versions of the processing-based account relating to the memory-based interference.
The Gundel et al. (1993) Givenness Hierarchy:

Central

Peripheral

in focus < activated < familiar < uniquely identifiable < referential

I, we, you he, she, they Donald Trump the chairman a chairman

For example, in object relatives like (10), the reading time at the main verb was fastest when the subject was a pronoun (we), slower for a definite NP (the chairman), and slowest of all for an indefinite NP (a chairman).

(10) The consultant who [we/the chairman/a chairman] called advised wealthy companies about tax laws.

Warren and Gibson (2002) interpret these results as evidence that an intervener whose referent is highly accessible has a less disruptive effect on dependencies than one whose referent requires more effort to recover.

Choosing between the RM and DLT accounts of intervention would obviously shed important light on children’s difficulty with raising patterns and possibly even on the nature of intervention effects in general. We report next on an experiment on raising that helps tease apart the claims made by the two approaches.

EXPERIMENT

METHOD

As mentioned in ‘Background’, Choe and Deen (2015) have found that children’s comprehension is better for patterns such as (11a) than (11b).

(11) a. Bart seems to her to be studying.
    b. Bart seems to Lisa to be studying.

Both the RM and DLT accounts predict this result – the RM because the pronoun is sufficiently different from the lexical NP that raises over it, and the DLT because the pronoun has a highly accessible referent. However, the predictions of the two accounts diverge in the following pattern, in which the matrix subject is a pronoun and the intervening experiencer is a lexical NP:

(12) He seems to Lisa to be studying.

On the one hand, the RM account predicts no increase in difficulty, compared to the pronominal experiencer case in (11a), since the raised NP and the intervening NP are still of different types – one is a pronoun and the other a lexical NP. On the other hand, the DLT-based account predicts an increase in difficulty, compared to the pronominal experiencer pattern, since the intervener is a lexical NP whose referent is lower on the
accessibility hierarchy. In order to tease apart the two accounts, we conducted a comprehension experiment.

**Participants**
We initially tested thirty-nine native English-speaking children, all of whom were recruited from the University of Hawai‘i Mānoa Children’s Center, Montessori Community School, and Kaimuki Christian School in Honolulu, Hawai‘i. Eleven children were subsequently excluded on the basis of their failure to complete the experiment or to properly interpret the filler items.

**Design**
The experiment employed a mixed design with one between-participants factor and one within-participants factor. Participants were assigned to one of the two conditions: (1) the LEXICAL NP–PRONOUN condition (14 children, aged 3;6–6;8, mean = 4;6), which tested patterns in which a lexical NP is raised across a pronominal experiencer, and (2) the PRONOUN–LEXICAL NP condition (14 children, aged 3;0–5;11, mean = 4;3), which examined raising patterns in which a pronoun is raised across a lexical NP experiencer.

Participants were also tested on the unraised counterpart of the raising pattern (within-participants with two levels: unraised vs. raised). Table 1 illustrates the design of the experiment, along with the test sentences for each condition.

**Procedure**
Following the format commonly used in a Truth-Value Judgment Task (Crain & McKee, 1985; Crain & Thornton, 1998), we showed children illustrated stories on a laptop computer. At the end of each story, a puppet appeared on the screen and made a one-sentence statement about what she thought had happened in the story. Children were then asked to determine whether the puppet’s statement was a true or false description of what

<table>
<thead>
<tr>
<th></th>
<th>Lexical NP–Pronoun</th>
<th>Pronoun–Lexical NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unraised</td>
<td>It seems to him that Donald is short.</td>
<td>It seems to Mickey that he is short.</td>
</tr>
<tr>
<td>Raised</td>
<td>Donald seems to him to be short.</td>
<td>He seems to Mickey to be short.</td>
</tr>
</tbody>
</table>

This condition originally appeared in Choe and Deen (2015), with a different set of comparisons and discussion.
happened in the story and to provide a justification for their choice. The whole experiment took no more than thirty minutes for each child.

**Materials**

For each main condition (Lexical NP–Pronoun or Pronoun–Lexical NP), a child watched ten short stories in total, of which two were warm-up items, two were fillers, and the rest were critical items, divided into Unraised and Raised patterns. For the Raised pattern in the Lexical NP–Pronoun condition, a lexical NP was raised across a pronominal experiencer, as in *Donald seems to him to be short*. In the Pronoun–Lexical NP condition, a pronoun was raised across a lexical NP experiencer, as in *He seems to Mickey to be short*. The items were balanced for match and mismatch, and the critical test sentences were counterbalanced across four different lists in order to minimize any item effects. The complete script of a sample story and the corresponding pictures shown to children are presented in Figure 1. A sample set of test sentences for each condition is also presented in Tables 2 to 5. In half of the six critical items, the referents of the raised NP and the experiencer NP had the same gender, as in Tables 2 and 3, while in the other half they had different genders, as in Tables 4 and 5.4

In order to increase the naturalness of the pronoun and establish its antecedent, the experimenter’s prompt and the puppet’s lead-in sentence were presented immediately before the test sentence. For example, in the match item in (13) for the Pronoun–Lexical NP condition, only the intended referent *Donald* is explicitly mentioned in the first sentence of the experimenter’s prompt. Then, in the second sentence, a pronoun is used to refer to the same referent while the experimenter points to the picture of *Donald*. Then the puppet also mentions the intended referent, making *Donald* the undisputed topic, which leads children to easily choose *Donald* as the antecedent of the pronoun in the test sentence.

(13) <Match item>

Experimenter’s prompt: That was a fun story about Donald and his friends. He (*pointing to Donald*) is playing with Daisy, and then Mickey comes along. Hey, puppet, can you tell us what happens next?

Puppet: Donald is in a hole. He’s playing down there, and . . .

(Unraised) it seems to Mickey that he is short.

(Raised) he seems to Mickey to be short.

---

3 Match items are those where the target response is ‘true’; mismatch items are those where the target response is ‘false’. As there were six critical items per child, there were three match items and three mismatch items.

4 As correctly noted by an anonymous reviewer, the number of gender match/mismatch items was not completely balanced, as we did not control for this variable during the initial design stage of the experiment.
In order to assure children’s understanding, the stories were presented to children with matching visual scenes and moving characters. In a previous study (Choe & Deen, 2015), the same stories were paired up with a slightly different test sentence (To Mickey, Donald seems to be short), in
which the experiencer does not intervene between the two subject positions. On this condition, children (age 3;3 to 5;8, mean = 4;7, N = 28) had a success rate of over 80% accuracy, demonstrating that they had no trouble understanding the stories that preceded the test items.

**RESULTS**

**Figure 2** shows the mean correct percentage for each pattern (Raised and Unraised), across the two main conditions, Lexical NP–Pronoun and Pronoun–Lexical NP. For data analysis, we performed a mixed-effects logistic regression analysis, using R (R Core Team, 2014) and lme4 (Bates,
Maechler, Bolker, & Walker, 2014). The dependent variable was the correct response to each item. Pattern (Raised vs. Unraised) and Condition (Lexical NP–Pronoun vs. Pronoun–Lexical NP) were entered as fixed effects, and participants and items were included as random effects. With the unraised pattern, the accuracy rate was similar between the two conditions (Lexical NP–Pronoun: 88.1% vs. Pronoun–Lexical NP: 78.6%), showing no statistically significant difference ($\beta = -0.84$, $Z$ value $= -1.14$, $p = .26$). Crucially, children’s comprehension of the raised pattern was significantly better in the Lexical NP–Pronoun condition (66.7%) than in the Pronoun–Lexical NP condition (38.1%; $\beta = -1.40$, $Z$ value $= -2.27$, $p = .023$).

Additional analyses with the maximum likelihood ratio test indicated that none of the potential factors (the match/mismatch items and the age of the child) provided a significant improvement in fit over the null model (all $p$s > .1). In other words, there was no difference in children’s rate of success on match items and mismatch items in any pattern, and no effect of age (as well as its interaction with Condition) within patterns (all $p$s > .1).

---

5 As we were concerned about the small number of participants for each counterbalanced list, we also conducted a Mann–Whitney $U$ test, a non-parametric test, which revealed a significant difference on the raised pattern between two conditions ($Z = -2.01$, $p = .04$).

6 Half the critical items had a female experiencer (e.g., Donald seems to her to be short), creating a gender mismatch with the raised subject, and half had a male experiencer (him), creating a gender match that might have misled the children into thinking that two were co-referential—especially if they had not fully mastered Principle B. Crucially, however, children’s performance on the two patterns did not differ, allaying any concerns of a confound involving co-reference.
Finally, the results were further analyzed by looking at the mean accuracy of the sixteen children who scored 100% on the unraised pattern, thus filtering out those children who had trouble comprehending the verb seem itself. As can be seen in Figure 3, the difference in performance on the two conditions becomes sharper, with a success rate of 81.5% on the raised pattern in the Lexical NP–Pronoun condition, compared to just 38.1% in the Pronoun–Lexical NP condition. This difference was significant by a mixed-effects logistic regression analysis ($\beta = -2.60$, $Z$ value $= -2.16$, $p = .031$), with Condition as the fixed effect and participants and items as random effects.

To investigate the potential effects of the children’s age ($N = 9$ [age 3;8 to 6;8, mean = 4;6] for Lexical NP–Pronoun; $N = 7$ [age 3;0 to 5;10, mean = 4;3] for Pronoun–Lexical NP), we compared the likelihoods of two models: the full model with the effect in question against the model without the effect in question. The results revealed that whereas Condition turned out to be a significant predictor ($\chi^2(1) = 6.50$, $p = .01$), neither age nor its interaction with Condition significantly improved the fit over the null model: Age ($\chi^2(1) = 0.10$, $p = .75$); Age × Condition ($\chi^2(1) = 2.04$, $p = .15$).

CONCLUSION
The main finding of our experiment is that whereas children’s comprehension of raising sentences improves when the experiencer is a

---

7 Although the mean age of the children in the Pronoun–Lexical NP condition was numerically lower than that of the children in the Lexical NP–Pronoun condition, there was no statistically significant difference in age between the two groups by an independent $t$-test ($t(14) = 0.57$, $p = .58$).
pronoun, as in (14a), they still have difficulty with raising sentences in which a pronoun is raised across a lexical NP experiencer, as in (14b).

(14)  a. Bart seems to her _ to be studying.
     b. He seems to Lisa _ to be studying.

Taken together, these results demonstrate an asymmetric effect of NP type on children’s comprehension of raising. Whereas a pronominal facilitated comprehension when it occurred in the experiencer position, as in (14a), no such effect was observed when it occurred in the matrix subject position, as in (14b).

This finding is not expected under the grammar-oriented RM account of intervention (Friedmann et al., 2009; Rizzi, 2013), which focuses on the (dis)similarity of the two NPs. This theory predicts no increase in difficulty for the raising sentence in (14b), compared to the pronominal experiencer case in (14a), since the raised NP and the intervening NP are sufficiently different in terms of their feature specifications—one is a pronoun and the other is a proper name. In contrast, the results are consistent with the Dependency Locality Theory-based approach, whose appeal to the referential accessibility of the intervening NP predicts improved comprehension only when it is a pronoun.8

Given our view that an intervention effect is responsible for children’s non-target performance on both raising (an A-movement structure) and object relative clauses (an A’-movement structure), a question arises as to why children are able to do well on object relative clauses one to two years before they master comprehension of raising (Hirsch, 2011). The likely answer is that there is no reason to expect intervention effects to be overcome at the same time in relative clauses and raising patterns, given the many differences between the two patterns. These differences include the type of movement (from a lower subject position to a higher subject position in raising patterns but not in a relative clause), the patterns’ internal structure (raising patterns involving an infinitival clause), and frequency of occurrence in the input (object relatives are far more frequent than raising patterns with an experiencer argument9).

---

8 Previous studies on experiencer-less raising sentences (e.g., *The dog seems to be white*) have yielded conflicting results, with some reporting difficulty for children (Orfitelli, 2012) and others noting no such problem (Becker, 2004, 2006, 2007). Although our experiment consisted only of test items that contain an experiencer, we predict that patterns without an intervening experiencer should be easier than those with such a constituent. Interestingly, Choe and Deen (2015) report a higher level of success by children when the experiencer appears at the beginning of the sentence, as in “To Mickey, Donald seems to be short”, rather than between the two subject positions.

9 See Choe (2012), who conducted a corpus study to find that English-speaking children very rarely hear or produce raising patterns with an experiencer. In particular, out of the entire CHILDES corpus (both in the child-directed and in the child-produced speech), there were
In sum, the results reported here strongly support an account of comprehension asymmetries in raising that draws on a familiar theory of processing difficulty. It would of course be unwise to suggest that this is a complete account of children’s problems with raising patterns. We do not rule out the possibility that grammatical factors of a type yet to be determined may also be in play, nor do we rule out the possibility that future work will alter the line between grammar and processing. For now, it suffices to note that the results for the contrast that our study was designed to investigate are highly compatible with what is traditionally thought of as a processing account, but not with any currently known grammatical account.

REFERENCES


no instances of *seem* utterances in which an experiencer intervenes between the verb *seem* and an infinitival complement.


