Children's difficulty with raising: A performance account

Jinsun Choe & Kamil Deen

Routledge
Taylor & Francis Group

Children's difficulty with raising: A performance account

Jinsun Choe[a] and Kamil Deen[b]

[a]Hankuk University of Foreign Studies; [b]University of Hawaii at Manoa

ABSTRACT

This article explores English-speaking children's acquisition of raising structures with an experiencer (e.g., *John seems to Mary to be happy*). We review and address previously unnoticed issues in the methodologies of existing studies testing the acquisition of raising, thus providing a more reliable picture of children's abilities with respect to raising. We then present three experimental studies, which reveal that children's purported difficulty with raising is significantly reduced when the experiencer is fronted to the beginning of the sentence (e.g., *To Mary, John seems to be happy*) and when the experiencer is pronominal (e.g., *John seems to her to be happy*). While the acquisition of this construction has been argued to be delayed due to grammar-internal reason, we propose an alternative account that attributes the difficulty to a performance limitation, responsible for intervention effects observed in a variety of other structures.

1. Introduction

We investigate the acquisition of a notoriously challenging structure for children to acquire, often referred to as "raising." The NP *the boy* in (1) is semantically linked to the verb phrase *to be happy* in the embedded clause but is syntactically realized as the subject of the matrix clause.

(1) Raising to Subject: [S The boy seems to the girl [S _ to be happy]].

This structure is often referred to as a "raising" pattern, reflecting the fact that some syntactic analyses posit movement of the NP from the subject position in the lower clause to the higher subject position, leaving a gap in the original position (e.g., Postal 1974). For this reason, patterns such as (1) are seen as a classic example of "A(rgument)-movement," since the NP *the boy* appears to move from one argument position (in this case, the subject of the embedded clause) to another argument position (the subject of the matrix clause). Other types of putative A-movement structures include passives (2a) and unaccusatives (2b).

(2) a. Passive: John was kicked _ by Mary.
   b. Unaccusative: John arrived _.

Many studies have investigated children's acquisition of raising patterns. By and large, these studies can be divided into two groups: (i) those that tested simple raising sentences *without* an

CONTACT Jinsun Choe ✉ jinsun@hufs.ac.kr 🖂 Hankuk University of Foreign Studies, Graduate School of Education, Seoul, Korea.

experiencer, as in (3a); and (ii) those that tested raising sentences *with* an experiencer, as in (3b).[1]

(3) a. John seems [_ to be happy].
    b. John seems *to Mary* [_ to be happy].

The first group of studies—on raising without an experiencer—have produced conflicting results in that some studies report difficulty (Hirsch, Orfitelli & Wexler 2007, 2008; Orfitelli 2012), while others do not (Becker 2004, 2006, 2007). Yet there is a general consensus for the second group that children have problems comprehending raising patterns containing an experiencer phrase as in (3b) (Hirsch & Wexler 2007; Hirsch, Orfitelli & Wexler 2007; Hirsch 2011).

The present article aims to provide an explanation for this mixed pattern of findings. We argue that the conflicting results in the first group of studies (testing raising without an experiencer) are due to methodological artifacts, and the results of the second group (testing raising sentences containing an experiencer) can be better explained by reference to an otherwise well-known parsing factor. We therefore propose an explanation that has not been considered before, namely that the source of the difficulty in such raising patterns stems not from a grammatical deficit (e.g., Borer & Wexler 1987; Hyams & Snyder 2005; Orfitelli 2012; Wexler 2004) but from the performance limitations responsible for *intervention effects* in a variety of other constructions.

It is well known that children have difficulty with structures in which an NP intervenes between a filler and the associated gap (e.g., Avrutin 2000; De Vincenzi et al. 1999; Friedmann, Belletti & Rizzi 2009; Friedmann & Lavi 2006). These structures include object relatives (4a), object topicalizations (4b), and object *wh*-questions (4c).

(4) a. Object relative: The boy who the girl kissed _
    b. Object topicalization: The boy, the girl kissed _.
    c. Object *wh*-question: Which boy did the girl kiss _?

The raising pattern in (5) resembles constructions in (4) in that the NP *the girl* intervenes between the dislocated NP *the boy*/*which boy* and its original position.

(5) Raising: The boy seems to the girl [_ to be happy].

Specifically, we propose that the delay in the acquisition of raising (5) and the delays observed in structures such as object relatives (4a), object topicalizations (4b), and object *wh*-questions (4c) should be attributed to a processing effect in which the presence of the NP (*the girl* in (5)) between the displaced NP (*the boy*) and the gap in its original position results in an increased processing cost.

---

[1]Although the majority of raising predicates (e.g., *tend, used to, be about, be going*) do not permit an experiencer argument, a subset of raising predicates, including *seem* and *appear*, do allow such an argument (*Mary* in (ia)) to occur between the raising predicate and the embedded clause ([ia], with [ib] its unraised counterpart).
    (i) a. Raised: John seems to Mary [_ to be happy].
        b. Unraised: It seems to Mary [that John is happy].
While there is no clear syntactic evidence for or against the idea that the experiencer argument is obligatory or optional, Orfitelli (2012:111) posits the existence of an implicit experiencer argument even when it is unpronounced—based on "the native speaker intuition that a proposition cannot 'seem' without there being a 'seem-ee' to experience it." We will proceed on this assumption without making any commitment, as the issue is not directly relevant to the present article. It should be noted, however, that patterns like (ia), where raising takes place across the experiencer, are cross-linguistically rare. In some languages, such as Icelandic, Italian, and Spanish, raising across an experiencer NP is simply forbidden (e.g., Boeckx 1999, 2008).

We refer to this as Performance-based Intervention Effects (PIE). The increased processing cost caused by this intervening NP results in the apparent delay in acquisition of raising, as well as all the structures in (4).

The article is structured as follows. In Section 2, we review previous studies on the acquisition of raising patterns—both without and with an experiencer—and identify several potential concerns about the methodologies used in this work. In Section 3, we consider three grammar-based accounts that have been put forward to account for children's difficulty with raising, and Section 4 proposes a new account of how this difficulty can be attributed to a performance-based limitation. Sections 5, 6, and 7 present three experimental studies that show that English-speaking children's comprehension of raising pattern improves when the experiencer phrase is fronted to the beginning of the sentence and when the intervening experiencer is a pronoun. In Section 8, we interpret these findings to suggest that English-speaking children's difficulty with raising can be accounted for by the Performance-based Intervention Effects that have been observed across multiple constructions.

2. Methodological issues in previous studies

2.1. *Previous studies on raising-to-subject without an experiencer*

Previous research on raising sentences without an experiencer has shown conflicting results: some claiming good comprehension by children (e.g., Becker 2006) and others claiming the opposite (e.g., Hirsch, Orfitelli & Wexler 2008).

Becker (2004) used a sentence judgment task to investigate whether 3- to 5-year-old children can comprehend raising sentences and found that the children had little difficulty with raising sentences. They correctly accepted raising sentences with a compatible lower predicate (e.g., *The hay seems to be on the ground*), and correctly rejected those with an incompatible lower predicate (e.g., #*The hay seems to be excited*). Furthermore, in another sentence judgment task with children ages 3–4 years, Becker (2007) found that both age groups were quite good at accepting raising sentences with a nonreferential subject (e.g., *It seems to be raining*). It should be noted that both studies also tested control verbs (e.g., *want, try*) as well as raising verbs. Becker argues that her participants analyze the control verb as a raising verb by (i) accepting it with a compatible lower predicate (e.g., #*The flower wants to be pink*), (ii) rejecting it with an incompatible lower predicate (e.g., #*The flower wants to fly away*), and (iii) accepting it with an expletive (e.g., #*It wants to be raining*). Therefore, although Becker does not make an explicit claim about children's knowledge of raising verbs, she appears to imply that children have no difficulty with them, since her claim is that children are incorrectly overapplying a raising analysis to the control verbs.

However, Becker's results are compatible with other explanations. One possibility (noted by Becker herself) is that the children may simply be ignoring the main raising verb and interpreting the sentence as a copular structure—hence *John seems to be happy* might be interpreted as "John is happy." Note that this strategy would result in the same response as if they correctly understood raising, since the testing protocol made no distinction between appearance and actual reality (the distinction that raising verbs like *seem* mark). Becker (2006) addresses this in an experiment (Truth-Value Judgment Task [TVJT]) that respected the distinction between appearance and reality. One story described a white dog appearing purple under a purple light. Given the test sentence *The dog seemed to be purple*, a child using a raising analysis should respond *true* since the dog seems to be purple, while a child using a copula analysis should respond *false* since the dog is actually white. She found that most children (3–4 years, $n = 52$) provided appropriate responses, suggesting that they correctly interpreted the main raising verb. Hirsch, Orfitelli & Wexler (2008) point out, however, that copular sentences like *The dog was purple* were not actually tested. Therefore, one is not sure whether the child was really sensitive to the distinction between

appearance and reality. In other words, there is a possibility that children may have responded *true* to the test sentence *The dog seemed to be purple*, by relying on the copular strategy because the dog *was* indeed purple under the light.

To address this point, Hirsch, Orfitelli & Wexler (2008)[2] conducted a modified version of Becker's original task, in which one of the major changes was the addition of the word *really* in each test sentence, serving the role of emphasizing the difference between appearance and reality. A total of 50 children (3.21–7.92 years, mean = 5.51) were tested with the conditions in (6), and the results showed children's poor comprehension only on the raised condition (6d).

(6) a. The dog is really white.   [*Copula*]
      b. The dog really likes to be white.   [*Subject control*]
      c. It really seems that the dog is purple. [*Unraised*]
      d. The dog really seems to be purple.   [*Raised*]

One concern with this study, first raised by Orfitelli (2012), is the use of the word *really*.[3] When it occurs with the copula in (6a), *really* is ambiguous in that it may link the sentence to reality, as in "*In reality, the dog is white*" or function as an intensifier, as in "*The dog is truly white.*" However, in both the unraised (6c) and raised (6d) conditions, *really* most likely functions as an intensifier that highlights the illusion denoted by *seem*, as in "*It truly seems that the dog is purple.*" Orfitelli (2012) dismisses the relevance of this fact on the grounds that children performed poorly only on the raised conditions, even though *really* was also present in the unraised patterns.

In summary, the results from Becker (2004, 2006, 2007) and Hirsch, Orfitelli & Wexler (2008) differ, and these differences may stem from various methodological issues in their respective experiments.

2.2. *Previous studies on raising-to-subject with an experiencer*

Only a few studies have investigated English-speaking children's acquisition of raising patterns containing an experiencer phrase: Hirsch and Wexler (2007),[4] Hirsch (2011), and Hirsch, Orfitelli and Wexler (2007),[5] all of which are comprehension studies.

Hirsch and Wexler (2007) conducted a two-choice sentence-picture matching task with 70 children (3.04–9.96 years, mean = 6.49) to assess comprehension of nonraising (7a) and raising (7b) sentences with an intervening experiencer.

(7) a. It seems to Homer that Marge is pushing a cart.
     b. Homer seems to Maggie to be bowling a ball.

The children were shown two pictures and were asked to choose the picture that best matches the test sentence they heard. The results showed that while the majority of children performed quite well on the unraised condition (93.3%), they had great difficulty with the raised condition (60.6%) until around the age of 7. In particular, most children consistently chose the wrong picture when a raised sentence was presented with a choice between the correct picture and the Double-reversal (DR) foil. For example, the DR foil for the sentence *Bart seems to Lisa to be playing an instrument* had a picture of Bart thinking about Lisa playing an instrument. Based on these results, Hirsch and Wexler (2007)

---

[2]This study is also presented in Hirsch (2011) as Study 5, with further analyses and expanded discussion.
[3]As an aside, it appears from Hirsch's (2011) detailed description of Hirsch, Orfitelli & Wexler (2008) that each child saw the exact same scenario eight times, each time paired with a different test sentence. This unusual experimental design raises concerns that children's performance was affected by repeated exposure to the exact same scenario—concerns that we are unable to address conclusively, as noted by an anonymous reviewer.
[4]This research also appears in Hirsch (2011) as Study 1, with new individual subject analyses and more discussion.
[5]This study is also presented in Hirsch (2011) as Study 4, with further analyses and expanded discussion.

initially proposed the *think*-analysis[6] where children treat the raising verb *seem* as if it were *think*. For example, the sentence *Bart seems to Lisa to be kicking the ball* would be incorrectly interpreted by children as meaning *Bart thinks Lisa is kicking the ball*.

Using the same methodology and the same pictures, Hirsch (2011) also tested children's comprehension of raising sentences, in which the experiencer phrase was dislocated to the beginning of the sentence (4.25–7.98 years, mean = 6.02, *n* = 60):

(8) To Lisa, Homer seems to be petting a dog.

The results revealed that children generally did well on comprehending raising sentences with a fronted experiencer. However, Hirsch notes that children's performance on one particular foil type did not differ from chance until the age of 7 years: That was the Matrix-reversal (MR) foil that involved switching the character who does the thinking. For example, the MR foil of sentence of (8) would have Homer petting a dog, thinking about Lisa. Given this chance performance on the MR foil, Hirsch argues that the children are ignoring the fronted experiencer in raised sentences and using a copula-analysis in which they interpret the raising verb *seem* as the copula—thereby taking the raised sentence (8) to mean *Homer is petting a dog*. This analysis explains their chance performance with the MR foil, since both pictures (correct and incorrect) presented to a child would then match the test sentence once the experiencer is ignored. Overall, Hirsch (2011) concludes that raising is problematic for young children irrespective of the position of the experiencer phrase.

Careful inspection of the experimental stimuli used in these two studies raises some concerns about the conclusion, however. The first concern involves the use of "thought bubbles" to depict sentences containing the verbs *think* or *seem*, in order to convey the notion of thinking. Consider the picture shown in Figure 1, which is the correct picture for the three sentences in (9).

(9) a. Lisa thinks that Bart is playing an instrument.   [*think*]
    b. It seems to Lisa that Bart is playing an instrument. [*unraised*]
    c. Bart seems to Lisa to be playing an instrument.   [*raised*]

As the term itself suggests, "thought bubbles" are perfectly appropriate for the *think*-condition. However, the children in Hirsch and Wexler's may have incorrectly treated the raising verb *seem* as *think* (or *imagine*) because of the thought bubbles: Since the picture portrays a person thinking about something, it would be natural for children to assume that whatever verb they hear has the same meaning as the verb *think* solely based on the picture, leading them to use either the *seem*-as-*think* or *imagine* analysis.[7] Secondly, the two characters portrayed in the picture seem to differ in their relative salience in that the one who is doing the thinking is given more emphasis than the one who is performing the relevant action. In a sense, this picture is *about* Lisa and the thoughts she has. As for sentence (9c), while truth-conditionally it matches the picture in Figure 1, it is an infelicitous

---

[6]Yet, this literal *seem*-as-*think* analysis is later dismissed in Hirsch (2011), since it entails that children are not paying attention to the nonfiniteness of the embedded clause in the raised sentences, despite previous findings that suggest sensitivity to the finite/nonfinite distinction from around 18 months of age (Pierce 1992; Poeppel & Wexler 1993). Instead, Hirsch (2011) proposes an *imagine*-analysis where children interpret the raising sentences (iia) as (iib), by analyzing the verb *seem* as a raising-to-object verb, like *imagine*. The following examples are taken from Hirsch (2011:110).

    (ii) a. John *seems* to Mary to go to the store every Saturday.
        b. John *imagines* Mary to go to the store every Saturday.

As with the *think*-analysis, the *imagine*-analysis should lead to below-chance performance on the raised condition with DR foils, since the DR foil matches this incorrect interpretation.

[7]This point has been raised by Choe (2012:21) as well as Becker (2014:198). The latter says "the fact that thought-bubbles were used in all of the drawings, and this depiction was explained to children as indicating what a character was thinking about, could easily have biased children to a *seem=think* kind of interpretation of the raised *seem* sentences."

Figure 1. A sample picture used in Hirsch & Wexler (2007) and Hirsch (2011).

description of Figure 1 because the subject of the sentence (*Bart*) is not who the picture is about. Thus, this violates topicality expectations established by the picture itself.

A third concern involves the infelicitous use of the experiencer phrase when there are only two characters in the stimuli, whether it be a picture or a story (Hamburger & Crain 1982). In Hirsch and Wexler (2007) and Hirsch (2011), there are always only two characters depicted in the picture—Bart and Lisa, for example. Notice that the use of the experiencer phrase *to Lisa* in this case is pragmatically infelicitous. There is nobody else to whom Bart seems to be playing an instrument except Lisa, and therefore to explicitly mention the experiencer phrase violates the pragmatic principle of economy. In fact, this may be the very reason why the children in Hirsch's (2011) study seemed to ignore the fronted experiencer.[8] It is quite reasonable for children to ignore the fronted experiencer, since the picture contains no alternative experiencer that contrasts with the "real" experiencer, and thus it maybe infelicitous to topicalize the experiencer phrase to begin with.

There is also a potential concern about Hirsch's (2011) interpretation of the results. Recall that the claim is that children analyze raised sentences such as (10a) as copular sentences such as (10b), thereby ignoring the experiencer phrase and treating the raising verb *seem* as a copula.

(10) a. To Lisa, Bart *seems* to be waving a flag.
     b. Bart *is* waving a flag.

While this analysis successfully captures the pattern of the results, there is an alternative explanation, namely, that children are ignoring the experiencer but correctly interpreting the remainder of the raised sentence. This strategy would also be compatible with Hirsch's experimental findings, since it would lead to chance performance with MR foil type, in which the foil picture for the sentence (10a) would be Bart waving a flag, thinking about Lisa. Crucially, however, this analysis would be incompatible with the claim for the delayed acquisition of raising with a fronted experiencer. In sum, Hirsch's (2011) experimental findings are compatible with both accounts: one in which children ignore the fronted experiencer and *cannot* comprehend raising sentences, thus using the copula-analysis, and one in which children ignore the fronted experiencer but *can* interpret raising sentences.

---

[8]There may be additional reasons for why children were ignoring the fronted experiencer in raised sentences, since those same children had no problems interpreting the fronted experiencer with unraised sentences. Hirsch (2011) does not provide an explanation for this unexpected pattern of results.

Hirsch, Orfitelli & Wexler (2007) used a different methodology from those discussed earlier. They tested children using a Truth-Value Judgment Task (TVJT) on sentences involving raising predicates with an intervening experiencer (4.35–7.95 years, mean = 5.97, *n* = 40). The results show that while children were successful comprehending unraised sentences (e.g., *It seems to Ken that Barbie is wearing a hat*; 97.8%), they showed poor comprehension of raised sentences (e.g., *Barbie seems to Ken to be wearing a hat*; 30.0%). From this finding, Hirsch, Oriftelli & Wexler conclude that children are unable to comprehend raising over an experiencer until around age 7.

While these results are compelling, here too several concerns arise. The first concern has to do with the felicitous use of the verb *seem*. Note that the verb *seem* is frequently used to project uncertainty (*it seems like it might rain today*) or illusion (*it seems to be raining today, but it's just dirty windows*). However, in the contexts used by Hirsch, Orfitelli & Wexler (2007) for the raised condition, it is always the case that what is described in the raising sentence actually matches what happens in the story. For example, in one story, Barbie is wearing a hat, but she does not realize it. Then, Ken sees that Barbie is wearing a hat. The test sentence used in the matching condition is *Barbie seems to Ken to be wearing a hat*. Thus there is neither illusion nor uncertainty for the experiencer (Ken). Rather, the raised subject is under the illusion that she is wearing a hat—a completely infelicitous use of the verb *seem*. Children thus might be puzzled by use of the verb *seem* in a situation where Barbie is actually wearing a hat, so that there is no illusion for Ken, the experiencer.

Secondly, the concern raised regarding the infelicitous use of the experiencer phrase when there are only two characters in the story (Hamburger & Crain 1982) applies here as well. With only two characters (Barbie and Ken), use of the experiencer phrase *to Ken* is unmotivated.

Overall, while the three studies reviewed thus far claim to have found evidence for children's delay in the acquisition of raising with an (intervening and fronted) experiencer until the age of 7, each of the studies suffer from various methodological problems. These problems are addressed in the three experiments presented in this article.

3. Previous accounts

We now turn to review three grammar-based accounts that have been put forward to account for children's difficulty with raising and discuss the predictions they make with respect to the different patterns of raising sentences.

3.1. *Universal phase requirement*

Wexler (2004) proposes the Universal Phase Requirement (UPR), which states that young children take all *v*Ps to be phases. The UPR is couched in the Minimalist framework (N. Chomsky 2000, 2001), which includes the Phase Impenetrability Condition (PIC):[9]

(11) The Phase Impenetrability Condition: An item can be extracted from a phase only if it is in the "edge" position.

However, constructions such as passives, unaccusatives, and raising pose a problem, since they require movement across a phase boundary from a nonedge position, as illustrated in (12).

---

[9]A phase is a subdomain of the syntactic derivations that proceed incrementally. Both *v*P and CP are identified as phases (N. Chomsky 2001).

(12)
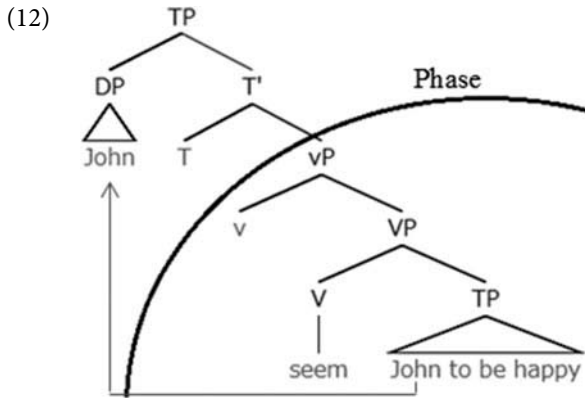
To address this problem, Chomsky proposed that the *v* of those constructions (passive, unaccusative, and raising) is "defective," in that it lacks an external argument and that it therefore does not count as a phase. This in turn allows the embedded subject to move to the matrix subject position without violating the PIC. Drawing on this idea, the Universal Phase Requirement (UPR) claims that phases for children are initially uniformly nondefective, and thus the PIC applies to all phases. So because the phase in a raising construction is "strong," and therefore subject to the PIC, children are unable to raise a subject out of the complement of a raising predicate into the matrix subject position.[10] The presence of the experiencer phrase is irrelevant for UPR, since it is the grammatical operation (i.e., movement, or AGREE, across a phase) itself that is problematic for children. The UPR thus predicts that raising sentences should always be ungrammatical for young children, regardless of whether there is an experiencer phrase.

3.2. *Universal freezing hypothesis*

Hyams and Snyder (2005) seek another explanation to account for delays in raising acquisition by adopting Collins's (2005a) "smuggling" account of raising and propose the Universal Freezing Hypothesis (UFH).

Under Collins's analysis of raising, the experiencer phrase is generated in the specifier position of an applicative phrase, which structurally intervenes between the base position (indicated by _) of the raised argument and its final position in the matrix Spec of the IP:

(13) John seems [$_{vP}$ [$_{ApplP}$ *to Mary* [$_{VP}$ _ to be happy]]].
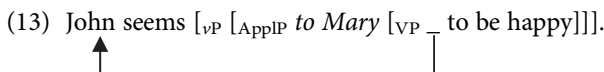
As simple raising of the embedded subject at this point seems to violate locality restrictions on syntactic movement, such as Relativized Minimality[11] (RM, Rizzi 1990), or the Minimal Link Condition[12] (N. Chomsky 1995), Collins's smuggling approach is designed to solve this problem. According to the smuggling analysis, the entire VP containing the subject DP moves to the Spec of the *v*P above the applicative phrase (with the experiencer in its Spec), thereby "smuggling" the subject DP past the experiencer and allowing it to raise to the matrix subject position (see Collins

---

[10]Note that the UPR also predicts the delayed acquisition of verbal passives and unaccusatives, since all three of these structures purportedly involve deficient phases. Yet, as pointed out by an anonymous reviewer, there have been a number of studies that show that children comprehend and produce such structures early and easily (for passives, see Fox & Grodzinsky 1998; O'Brien, Grolla & Lillo-Martin 2006; Messenger et al. 2012; for unaccusatives, see Costa & Friedmann 2012; Friedmann 2007; Snyder, Hyams & Crisma 1994).

[11]Relativized Minimality (RM): In the configuration … X … Z … Y …, Y cannot be related to X if Z intervenes and Z has certain characteristics in common with X (Rizzi 1990).

[12]Minimal Link Condition (MLC): K attracts *α* only if there is no *β*, *β* closer to K than *α*, such that K attracts *β* (N. Chomsky 1995:311).

2005a for more details). What is relevant to the present discussion is that as the subject DP is raised from the VP that itself has already undergone movement, the smuggling process is considered an exception to the Freezing Principle (Müller 1998), which rules out movement from an already-moved phrase.

Appealing to these processes of smuggling and freezing, Hyams and Snyder's (2005) UFH claims that children overapply the Freezing Principle to all cases of movement:

> (14)　UFH: For the immature child (at least until age 4), the Freezing Principle always applies—No subpart of a moved phrase can ever be extracted.

As a consequence, children have difficulty with structures involving the smuggling process—such as raising sentences and verbal passives (Collins 2005a, 2005b). Note, however, that the UFH predicts that children's comprehension of raising becomes problematic only when an experiencer is present. That is, it predicts no trouble for raising sentences without an experiencer, since it is only when the embedded subject needs to raise across the intervening experiencer that smuggling is required.

3.3.　*Argument intervention hypothesis*

Orfitelli (2012) proposes the Argument Intervention Hypothesis (AIH) to account for children's delay in raising as well as verbal passives:

> (15)　AIH: Children are delayed in acquiring those structures which require A-movement across a structurally intervening argument.

For example, raising sentences with an intervening experiencer (e.g., *John seems to Mary _ to be happy*) are ungrammatical for children, since it involves movement of the embedded subject *John* past the experiencer argument *Mary*.

The AIH is similar to the Universal Freezing Hypothesis (UFH) in that both theories predict difficulty with both raising sentences across an experiencer as well as verbal passives (the latter of which involve a movement of the theme over the agent argument position). However, the AIH makes a prediction different from the UFH with respect to the *seem*-type raising sentences *without* an experiencer. According to the AIH, it is assumed that even when there is no overt experiencer argument in *seem*-type raising sentences, it is syntactically projected (Orfitelli 2012:109–111), leading to the prediction that children's difficulty should persist even in patterns, such as *John seems to be upset*, in which the experiencer phrase is unpronounced.

4.　New account: Performance-based intervention effects

All of the three accounts mentioned—Universal Phase Requirement (UPR), Universal Freezing Hypothesis (UFH), and Argument Intervention Hypothesis (AIH)—have in common that a grammatical deficit underlies children's delayed acquisition of raising. A major virtue of such grammatical deficit approaches is that the proposed deficit in grammar may be generalized beyond just one specific structure. That is, a single (grammatical) deficit might ultimately explain a number of delays in child language. Indeed, this is precisely what each of the aforementioned theories attempts to do: All three theories effectively account for the supposed delay in passives by the same mechanism that purportedly explains the delay in raising across an experiencer. This is because the critical computation involved in raising (different, depending on the theory) is the same as that involved in passives (and unaccusatives).

Such an approach is indeed very appealing. And indeed the approach we advocate here has precisely this property. Anticipating, the approach we outline in the following accounts for the delay in raising through the same mechanism that delays in a variety of other structures have been explained. Crucially, however, our approach does not posit any kind of deficit in grammar. Instead, we account for the difficulty with raising across an experiencer as essentially a performance problem—one that is alleviated as the child's ability to process language matures. Specifically, we believe the difficulty associated with raising patterns in (16) is not unique to children but is merely a reflection of the processing difficulty imposed on both adults and children in a variety of other structures, as in (17).

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

(17) a. Object relative: The boy who the girl kissed _
      b. Object topicalization: The boy, the girl kissed _.
      c. Object *wh*-question: Which boy did the girl kiss _?
      d. Coordination: The boy tickled the girl and _ smiled.
      e. Subject control: The boy promised the girl _ to leave.

Numerous studies have shown that children have difficulty with these structures in which an NP intervenes between a filler and the associated gap (e.g., Avrutin 2000, C. Chomsky 1969; De Vincenzi et al. 1999; Friedmann & Costa 2010; Friedmann, Belletti & Rizzi 2009; Friedmann & Lavi 2006; Friedmann & Novogrodsky 2004; McKee, McDaniel & Snedeker 1998; Tavakolian 1981). For example, Friedmann, Belletti, and Rizzi (2009) show that children (age = 3;07–5;00, mean = 4;06, n = 22) acquiring Hebrew comprehend object relatives (in which the subject NP intervenes between the head noun and the gap from which it was extracted) in a sentence-picture (or -scenario) matching task at rates of 55%, whereas they comprehend subject relatives (in which no NP intervenes between the head noun and its original position) at rates of 90%. This difference has been shown to be quite robust (see aforementioned references). This relative difficulty in object relatives, object *wh*-questions, etc., is caused by what is often referred to as an intervention effect (Gibson 1998, 2000; Warren & Gibson 2002) because the difficulty associated with it is thought to arise from the presence of a noun phrase that interrupts the filler-gap dependency. Note that while this terminology is shared with some grammatical theories, such as AIH and RM (on which the AIH is based), the notion of intervention here is very different—it is not a grammatical intervention but a parsing intervention. We shall return to this subsequently.

A signature property of intervention effects is that their difficulty can be modulated by manipulating the type of the intervener (Arnon 2010; Friedmann, Belletti & Rizzi 2009; Kidd et al. 2007). For example, Arnon (2010) used a picture-selection task with Hebrew-speaking children (age = 3;06–5;03, mean = 4;04, n = 23) and found that they were better at comprehending object relatives with pronominal interveners (*the nurse that **I** am drawing* _; 85%), than ones with lexical NPs (*the nurse that **the girl** is drawing* _; 69%). This advantage will, henceforth, be referred to as the "pronoun advantage" and will be relevant in due course.

Perhaps surprisingly, the same contrast has been found in studies of adult processing: Object relatives have longer reading times and higher error rates in comprehension than do subject relatives (e.g., Caplan & Waters 2002; Carpenter, Miyake & Just 1994; Gibson 1998; Wanner & Maratsos 1978, among others), but the difficulty associated with object relatives disappears when the subject inside the relative clause is pronominal as in (18) (e.g., Gordon, Hendrick & Johnson 2001; Reali & Christiansen 2007).

(18) the barber that you admired _

It seems, therefore, that the presence of an intervening NP between the filler and its gap disrupts the comprehension of that long-distance relationship (see the following for one explanation of why this is). This seems to be true for a number of constructions, and a number of languages, and is therefore something basic about how we comprehend language. On this view, then, it may well be that raising is subject to the same pressures as other constructions that involve long-distance dependencies. If this is on the right track, the delay in the acquisition of raising (19) may be attributed to the increased processing cost associated with the presence of the NP (*the girl* in (20)) between the raised NP (*the boy*) and the gap in its original position.

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

We dub this Performance-based Intervention Effects (PIE). The basis of this PIE hypothesis is Dependency Locality Theory (DLT, Gibson 1998, 2000), which has been put forward to explain the selective intervention effects in adult sentence processing.[13] The DLT states that sentence processing involves integrating a new word into the existing syntactic and discourse structure (which incurs an integration cost) and keeping the structure in memory (memory cost). Importantly, the integration cost is said to increase with the number of new discourse referents that intervene between the elements being integrated within a dependency (Gibson 2000:104–105). Thus, object relatives (20b) are harder to process than subject relatives (20a) because two new discourse elements (*the girl* and *kissed*) intervene between the relative head and the gap in object relatives, while the dependency in subject relatives is resolved immediately without any new discourse element intervening.

(20) a. Subject relative: the boy that _ kissed the girl
       b. Object relative: the boy that the girl kissed _

Furthermore, this theory also explains why the subject-object difference in processing difficulty is reduced when the intervening NP in object relatives is a pronoun (21b), compared to when it is a definite NP (21a) (e.g., Gordon, Hendrick & Johnson 2001): Since the referent of the pronoun is already present in the discourse, it is not considered a new discourse referent and thus incurs less integration cost.

(21) a. The banker that the barber praised _ climbed the mountain.
       b. The banker that you admired _ climbed the mountain.

Additional support for the role of the discourse status of the intervening NP comes from Warren and Gibson (2002), who demonstrate the effects of referential processing on sentence complexity. Using both complexity ratings and self-paced reading tasks, they show that processing difficulty is sensitive to gradations of status between new and old, of the sort proposed in the Givenness Hierarchy (Gundel, Hedberg, & Zacharski 1993) in (22), which correlates NP types with the accessibility of their referents.

(22) The Gundel, Hedberg & Zacharski (1993) Givenness Hierarchy:
      Central   Peripheral
      in focus   < activated   < familiar   < uniquely identifiable   < referential
      I, we, you   *he, she, they*   *Donald Trump*   *the chairman*   *a chairman*

---

[13]It should be noted that there are different accounts of intervention effects, discussion of which can be found later in this article.

For example, in object relatives like (23), the reading time at the main verb was fastest when the subject was a pronoun [*we*] or a famous name [*Donald Trump*], slower for the definite NP [*the chairman*], and the slowest for the indefinite NP [*a chairman*].

(23) The consultant who we/Donald Trump/the chairman/a chairman called advised wealthy companies about tax laws.

Overall, the results suggest that sentences with an NP that is central on the hierarchy and thus highly accessible were associated with lower processing cost than those with an NP that is peripheral on the hierarchy.

As the DLT is a theory that relates sentence processing to the available computational resources, it is also in line with the acquisition theory embodied in the PIE, which attributes children's difficulty with raising to a performance limitation. At this point, it is worth reiterating that while both the AIH and PIE make use of the term "intervention," the underlying notion is fundamentally different. The AIH—a grammatical deficit theory—posits that child grammar does not allow syntactic representations involving intervention (i.e., A-movement over an intervening argument), at least at an early stage. The PIE, on the other hand, makes no such claim. Rather, the assumption that underpins PIE is that of Full Continuity (Pinker 1984; Crain 1991; Clahsen 1996, among many others). Thus, child grammar is identical to that of adults, but intervention (in the PIE sense) occurs because of parsing difficulties, resulting in degraded comprehension. Furthermore, the only requirement for such parsing difficulties is that there be an overt NP within an interpretive dependency, unlike AIH, which allows for intervention even with covert arguments.

Also noteworthy is that of all four accounts considered, only the PIE links the difficulty associated with raising to another well-known difficulty that both adults and children encounter and that is observed in multiple construction types (involving either A- or A-bar-movement). The other three accounts (UPR, UFH, and AIH) limit themselves to the structures involving classic A-movement. In what follows, we present a series of experiments that, when taken together, support the idea of Performance-based Intervention Effects.

5. Experiment 1: Raising across the experiencer

The aim of Experiment 1 is to better assess children's comprehension of raising sentences with an experiencer; it also serves as a baseline for the subsequent experiment. As mentioned earlier, there are several methodological questions about particular aspects of previous studies of children's acquisition of raising:

(i) The felicity conditions for *seem*.
(ii) The felicity conditions for an experiencer.
(iii) Presentation of the characters (their numbers, their prominence, thought bubbles, etc.).

For these reasons, the three experiments presented in this article were designed with the following modifications. First, an illusion is created to properly motivate the raising verb *seem*. For example, in the story illustrated in Figure 2, Mickey thinks Donald is short, while Donald is in fact not short. The reason for Mickey's misconception is that Mickey is looking down at Donald, who is standing inside a hole. From Mickey's relatively high vantage point, Donald looks short—hence the appropriateness of the test sentence *Donald seems to Mickey to be short*.

The second modification involves the presentation of an alternative experiencer needed to create a contextual contrast. For example, if the test sentence is *Donald seems to Mickey to be short*, not only do Donald and Mickey appear in the story, but so does a third character, Daisy, acting as an alternative experiencer, as shown in Figure 3.

This serves two purposes. First, it allows for the felicitous use of the experiencer phrase (e.g., *to Mickey*), because without the alternative experiencer in the context, the experiencer phrase would be superfluous and therefore infelicitous. Second, this third character serves to remind the child to
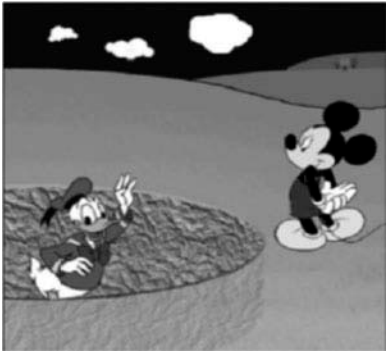
Figure 2. An illusion for the experiencer.

Figure 3. An alternative experiencer.

consider the distinction between what is real and what is illusory. For example, in the story, the third character tries to explain to the character under the illusion why s/he is mistaken (although the character who hears this is not convinced).

The four proposals under discussion make the following predictions. All of the four hypotheses (UPR, UFH, AIH, and PIE) predict that children should be delayed in comprehending raising sentences like *Donald seems to Mickey to be short*—the pattern that is tested in Experiment 1. For the Universal Phase Requirement (UPR), children's difficulties should come from the raising operation itself, which involves movement across a defective phase boundary, while the other three accounts—the Universal Freezing Hypothesis (UFH), Argument Intervention Hypothesis (AIH), and Performance-based Intervention Effects (PIE)—attribute the delay to the presence of an intervening experiencer, albeit for different reasons. For the UFH and AIH, what underlies children's difficulty is a grammatical deficit, whereas for PIE, the difficulty comes from performance limitations. These predictions are summarized in Table 1.

5.1. Participants

The participants in this study included 23 native English-speaking children (age 4;04 to 5;05, mean = 4;08), recruited from the UH Mānoa Children's Center in Honolulu, Hawai'i. There were 14 girls and 9 boys; eleven 4-year-olds and twelve 5-year-olds.

Table 1. Predictions for Experiment 1.

	UPR	UFH	AIH	PIE
Raising across the experiencer	Problematic	Problematic	Problematic	Problematic

5.2. *Procedure*

A Truth-Value Judgment Task (Crain & McKee 1985; Crain &Thornton 1998) was employed. Each participant was presented with animated stories on a laptop computer. At the end of each story, a puppet appeared on the screen, and an experimenter prompted the puppet to make a statement about what happened in the story. The puppet then made a one-sentence statement about what she thought had happened. Participants were asked to determine whether the statement was true or false and to provide justifications for their choice. The whole experiment took no more than 30 minutes for each participant.

5.3. *Materials*

Each child watched 10 stories (2 warm-ups, 2 control items, 4 critical items, and 2 fillers, randomized), after which they judged the truth of a statement made by a puppet. Three types of sentences were used, balanced for match and mismatch:[14] (1) control test items including finite clauses with the verb *think*, (2) unraised sentences with a medial experiencer phrase, and (3) raised sentences with a medial experiencer phrase. These sentences were counterbalanced across six lists to minimize item effects. The complete script of a sample story is shown in (24) (note that these stories were presented with matching visual scenes, making for a fun, cartoon-like experience for the children). A sample set of test items (match and mismatch) is presented in Table 2 (see Table A1 in Appendix A for a full list of test items).

(24) Sample story: This is a story about three friends: Donald, Daisy, and Mickey. One day, Donald and Daisy are playing outside, digging a big hole. Just then, Mickey appears and comes close to Donald and Daisy. Mickey sees Donald, who is inside the hole, and thinks that Donald is very short. Mickey says, "Hey, Donald, I thought you and I are the same height, but I was wrong." Without looking at Mickey, Donald says, "What do you mean? We are the same height." Then, Mickey says, "No, you are so short!" Daisy, who is standing next to Donald, says "No, that's because you are looking at him from up there. Donald is not short." But Mickey says, "What do you mean? Donald is so short." Still without looking at Mickey, Donald says, "Well, if I'm short, then you are short, too." Mickey says "Yeah? Turn around and look then." Donald turns around to look at Mickey, and Donald says, "Uh-oh, you are not short. I was wrong." Mickey says, "See? We are not the same height, after all. You are so short, haha."

The Think condition served as a control, in order to determine whether children have difficulty with the concept of thinking or understanding others' viewpoints, because such knowledge was

Table 2. A Sample Set of Test Items in Each Condition (Exp.1).

Match Items	
Think	At the end of the story, Mickey still thinks Donald is short.
Unraised	At the end of the story, it still seems to Mickey that Donald is short.
Raised	At the end of the story, Donald still seems to Mickey to be short.
Mismatch Items	
Think	At the end of the story, Donald still thinks Mickey is short.
Unraised	At the end of the story, it still seems to Donald that Mickey is short.
Raised	At the end of the story, Mickey still seems to Donald to be short.

[14]Match items are those where the target response is "true"; mismatch items are those where the target response is "false."

required to correctly understand the stories presented. It also made it possible to test the claim of Hirsch and Wexler (2007) that children treat the raising verb *seem* as *think* (i.e., *think*-analysis).

The only verb used in the Unraised and Raised conditions was the verb *seem*, in accordance with Hirsch and Wexler's (2007) observation that this is the most frequently used raising verb in child-directed speech in the CHILDES corpus.[15] The two fillers were unrelated to raising or complementation structures.

5.4. Results

Data from two children (both 5-year-olds) were excluded from the subsequent analysis because they either did not finish the experiment or did not respond correctly to the filler items. The results from the remaining 21 children are presented in Figure 4 in the form of the mean correct percentage for each condition—collapsed across match and mismatch items (i.e., correctly responding "true" to match items and "false" to mismatch items).

As expected, children did well on the Think condition (mean correct: 81.0%) and the Unraised condition (76.2%), scoring significantly above chance level by a one-sample *t*-test (see Table B1 in Appendix B). This shows that they understood the meaning of the raising verb *seem* as well as the meaning of *think*. However, their comprehension of the raised sentences was right at chance (50.0%), suggesting that they had difficulty comprehending the raising sentences.

For further data analysis, we used R (R Foundation 2014) and *lme4* (Bates et al. 2014) to perform a mixed-effects logistic regression analysis. The dependent variable was the accurate response to each item. We entered both participants and items as random effects, and Condition (Think vs. Unraised vs. Raised) as a fixed effect. This model outperformed the null model that only included the random effects, as confirmed by a likelihood ratio test, $\chi^2(2) = 10.13$, $p = .006$, suggesting that Condition was a significant predictor for children's performance. The fitted model revealed that children were significantly less accurate on the Raised condition than the Think condition ($\beta = 1.44$, Z value = 2.84, $p = .004$) or the Unraised condition ($\beta = 1.14$, Z value = 2.37, $p = .018$).

Additional factors that might have affected children's performance include (1) the age of the child, (2) match or mismatch items, and (3) their interaction with Condition. In order to test the possible effects of those factors, we compared the likelihoods of two models: the full model with the effect in question against the model without the effect in question. None of the factors significantly improved data likelihood—Age, $\chi^2(1) = 0.04$, $p = .84$; Age x Condition, $\chi^2(2) = 3.48$, $p = .17$; Match/Mismatch, $\chi^2(1) = 0.98$, $p = .32$; Match/Mismatch x Condition, $\chi^2(2) = 0.37$, $p = .83$—suggesting no significant effect of these factors on the observed results.[16]
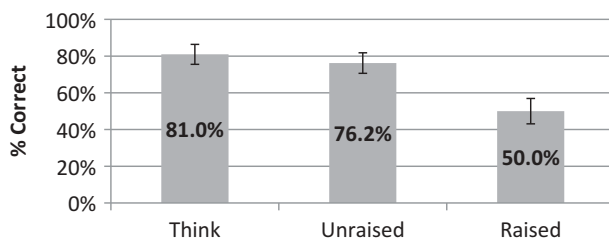
Figure 4. Mean accuracy in each condition (Exp. 1) (error bars represent ±1 SE).

---

[15]Although see Mitchener and Becker (2011), who find the raising predicate *going* (*to*) to be much more frequent than *seem* in the input to children.

[16]As correctly noted by an anonymous reviewer, no significant effect of the factors considered is likely to be due to the small number of items and children per age group. For further data presentation, see Table C1 in Appendix C, which shows the results in terms of the number of children per age group in each condition.

Figure 5. Mean accuracy of children who scored 100% on Unraised (Exp. 1).

Finally, Figure 5 shows the results of only those children who responded to the Unraised condition with perfect accuracy (100%). This requirement resulted in the exclusion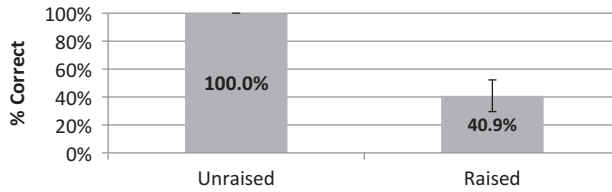 of data from 10 children, leaving data from 11 children, who scored 40.9% on the Raised condition. Overall, the data suggest that even children who comprehended *seem* in the unraised sentences 100% of the time had difficulty with the raised sentences.[17]

5.5. *Discussion*

The primary finding from Experiment 1 is that while children show good comprehension of unraised sentences, they have difficulty with raising sentences that have an experiencer phrase. These results replicate those of the previous studies by Hirsch and Wexler (2007) and Hirsch, Orfitelli & Wexler (2007). Given the methodological improvements we made, this result makes for stronger evidence of this now-well-known fact about difficulty with raising across an experiencer. Thus, our findings are consistent with the UPR, UFH, AIH, and PIE, all of which predict children's delayed comprehension of raised sentences with an intervening experiencer.

6. Experiment 2: Raising with a fronted experiencer

Experiment 2 tests the comprehension of raising sentences in which the experiencer phrase appears at the beginning of the sentence (e.g., *To Mickey, Donald seems to be short*). These sentences are similar to the ones used in Experiment 1 in that they involve raising as well as an experiencer phrase, but they differ in that the experiencer phrase no longer intervenes between the raised NP and the gap. Experiment 2 thus addresses the following question: Is children's difficulty with raising sentences associated with the raising operation *per se* or is it due to an intervening experiencer phrase?

According to the Universal Phase Requirement (UPR), children's delayed comprehension of raised sentences in Experiment 1 follows from their lack of knowledge of the raising operation itself. The presence of the experiencer phrase, as well as its location, should be irrelevant to children's difficulty. If anything, such test sentences might pose an additional difficulty for children, since not only do they involve raising (ruled out by UPR), but they involve the additional displacement of the experiencer to the front of the sentence. Thus, the UPR predicts that the children will fail to comprehend the raising pattern in which the experiencer phrase is fronted, just as they did in Experiment 1. As for the Universal Freezing Hypothesis (UFH), children are predicted to have difficulty with raising to subject even with the fronted experiencer, assuming that fronting of the experiencer takes place after the smuggling process. Finally, under the Argument Intervention Hypothesis (AIH) and the Performance-based Intervention Effects (PIE), the source of difficulty with raising is the presence of an intervening NP (argument for the AIH) within a dependency.

---

[17]We also examined how the excluded 10 children (who scored 50% overall on the Unraised condition) comprehended raising sentences. Their score on the Raised condition was 60%, which is higher than 40.9% from the earlier 11 children, but this difference did not yield significance by a mixed-effects logistic regression analysis ($\beta = 0.35$, Z value = 0.53, $p = .6$). There is thus no reason to think that the excluded children might have acquired the raising pattern before the nonraising pattern.

Table 3. Predictions for Experiment 2.

	UPR	UFH	AIH	PIE
Raising with a fronted experiencer	Problematic	Problematic	Unproblematic	Unproblematic

Therefore, both theories predict better comprehension of raising sentences with a fronted experiencer because there is no intervening element in these sentences. Table 3 summarizes the predictions made by the four hypotheses.

6.1. Participants

A total of 28 native English-speaking children (age 3;03 to 5;08, mean = 4;07) who had not participated in the previous experiment were recruited for this study from the UH Mānoa Children's Center in Honolulu, Hawaiʻi. There were 10 girls and 18 boys: six 3-year-olds, ten 4-year-olds, and twelve 5-year-olds.

6.2. Procedure and materials

The procedure and materials used in Experiment 2 were identical to Experiment 1, except that the experiencer phrase occurs at the beginning of the test sentences in the Unraised and Raised conditions.[18] In addition, the test sentences in the current experiment no longer begin with the phrase *at the end of the story* because it would not be as natural to have two presubject prepositional phrases in these sentences. Instead, this phrase was incorporated into the experimenter's prompt to the puppet, so that "Hey puppet, can you tell us something about the story?" in Experiment 1 was changed to "Hey puppet, can you tell us what happens at the end of the story?" in Experiment 2. A sample set of test items (match and mismatch) is presented in Table 4 (see Table A2 in Appendix A for a full list of test items).

6.3. Results

Due to poor performance with filler items, the data from three children (one 4-year-old and two 5-year-olds) were removed from the subsequent analysis, leaving the data from 25 children. As illustrated in Figure 6, the children demonstrated fairly good comprehension of all of the test sentences, scoring 80% in each of the three conditions. An important point is that the children's comprehension of raised sentences with a fronted experiencer was on par with their comprehension of unraised sentences. Furthermore, the results of a one-sample *t*-test (see Table B2 in Appendix B) found their performance in all conditions to be significantly above chance.

Table 4. A Sample Set of Test Items in Each Condition (Exp. 2).

	Match Items
Think	Mickey thinks Donald is short.
Unraised	To Mickey, it seems that Donald is short.
Raised	To Mickey, Donald seems to be short.

	Mismatch Items
Think	Donald thinks Mickey is short.
Unraised	To Donald, it seems that Mickey is short.
Raised	To Donald, Mickey seems to be short.

---

[18]Fronting the experiencer phrase (thereby topicalizing it) was motivated and natural, as each story contained another potential experiencer that contrasted with the real experiencer in the test sentence.
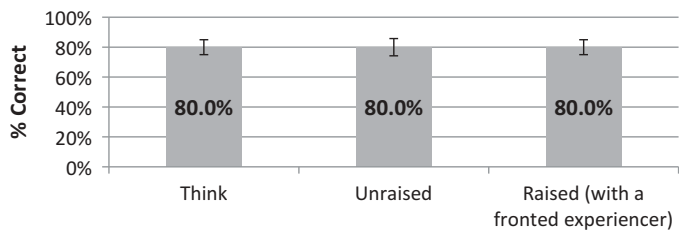
Figure 6. Mean accuracy in each condition (Exp. 2).

The mixed-effects logistic regression analyses were performed similarly to those of Experiment 1. As evident in Figure 6, Condition did not turn out to be a significant predictor for children's responses, $\chi^2(2) = 0$, $p = 1$), indicating that their performance did not differ among the three conditions. Further analyses showed that none of the potential factors significantly improved the fit over the null model: Age, $\chi^2(1) = 0.002$, $p = .97$; Age x Condition, $\chi^2(2) = 1.94$, $p = .38$; Match/Mismatch, $\chi^2(1) = 1.59$, $p = .21$; Match/Mismatch x Condition, $\chi^2(2) = 3.26$, $p = .20$.[19]

To further verify these findings, the results of only those children who responded to the Unraised condition with perfect accuracy (100%) were considered separately. This procedure resulted in excluding data from nine children.[20] The data for the remaining 16 children is presented in Figure 7, showing that their score on the Raised condition is quite high, 87.5%.

Summarizing, the overall results of Experiment 2 are presented together with those of Experiment 1 in Figure 8.

We fitted a mixed logistic model in which Condition and Experiment (and their interaction) were included as fixed effects and participants and items as random effects. The analysis revealed that there was no significant difference between experiments for the Think ($\beta = -0.06$, Z value $= -0.11$,
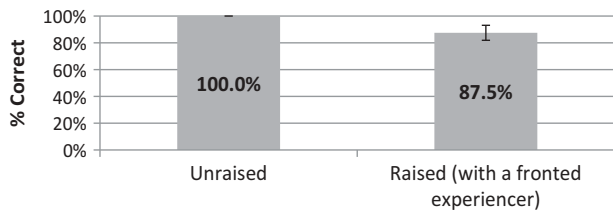
Figure 7. Mean accuracy of children who scored 100% on Unraised (Exp. 2).

	Think	Unraised	Raised
Experiment 1	81.0%	76.2%	50.0%
Experiment 2	80.0%	80.0%	80.0%
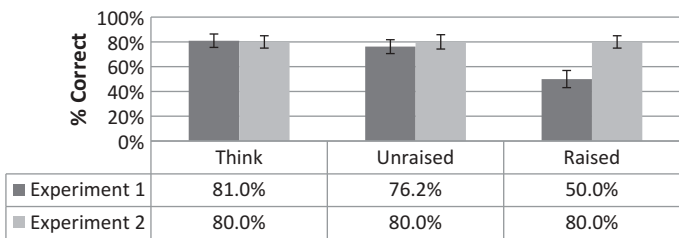
Figure 8. Comparison of Experiments 1 and 2.

[19]For further data presentation, see Table C2 in Appendix C, which shows the results in terms of the number of children per age group in each condition.
[20]The overall scores of the nine excluded children were 44.4% on the Unraised condition and 66.7% on the Raised condition.

$p = .92$) and Unraised conditions ($\beta = 0.25$, Z value $= 0.49$, $p = .63$). However, crucially, in the Raised condition, their comprehension was significantly better in Experiment 2 (with a fronted experiencer, 80.0%) than in Experiment 1 (with an intervening experiencer, 50.0%; $\beta = 1.38$, Z value $= 2.92$, $p = .003$), indicating that their difficulty with the raising pattern is significantly reduced when the relationship between the raised NP and the infinitival clause is not interrupted by an intervening NP.

6.4. *Discussion*

In Experiment 2, children showed good comprehension of raising sentences with a fronted experiencer, as well as of unraised sentences. These results, taken together with those of Experiment 1, indicate that while children are delayed in comprehending raising sentences with an intervening experiencer, this difficulty is significantly reduced with a fronted experiencer, suggesting that the presence of an intervening experiencer phrase in raising patterns, rather than the movement of the subject NP *per se*, is the source of children's difficulty with raising sentences. Furthermore, the current study provides evidence that an intervening experiencer is indeed an additional crucial factor for testing children's comprehension of raising sentences and that the failure to take it into account has obscured the real picture of the acquisition of raising.

Children's good comprehension with raising sentences in Experiment 2 is therefore not consistent with the UPR, which predicts children's delayed comprehension of raising sentences, irrespective of the location of the experiencer. It is also inconsistent with the UFH, assuming that smuggling takes place even in experiencer-fronted raising sentences. Rather, these results are in line with the predictions of the AIH and PIE. However, the two proposals differ in a crucial way: Whereas the AIH associates the difficulty of the intervening argument with a grammatical deficit, the PIE attributes it to a performance limitation. The difference between the two proposals will be investigated more closely in the following section.

7. Experiment 3: Raising across a pronominal experiencer

The findings of the previous experiment indicate that what underlies children's difficulty with raising patterns is the presence of an intervening experiencer phrase, rather than the raising operation itself. However, it still remains unclear whether this difficulty stems from a grammatical deficit (Argument Intervention Hypothesis, AIH) or a parsing difficulty (Performance-based Intervention Effects, PIE). As noted earlier, the signature property of the kind of parsing effect invoked by PIE is the pronoun advantage, namely that pronominal interveners cause less difficulty than lexical NP interveners. Thus, if children's difficulty with raising is due to the very same limitation, manipulation of the type of intervening experiencer should have a similar effect on children's comprehension. That is, children's comprehension should improve with a pronominal experiencer. Experiment 3 investigates this possibility by changing the type of intervening noun phrase to a pronoun, while the raised NP remains a lexical NP, as in *Bart seems to **her** to be studying*. Note that it is only the PIE hypothesis that predicts the pronoun advantage. In contrast, grammatical-deficit accounts such as the Universal Phase Requirement (UPR), Universal Freezing Hypothesis (UFH), and AIH predict that the type of intervener should not be crucial for children's comprehension. For the UPR, it is the raising operation itself that should be problematic to children, and for the UFH and AIH, it is the mere presence of an intervening argument of any type that poses difficulties for children. Table 5 summarizes the predictions made by the four competing hypotheses.

Table 5. Predictions for Experiment 3.

	UPR	UFH	AIH	PIE
Raising across a pronominal experiencer	Problematic	Problematic	Problematic	Improved performance

7.1. Participants

Nineteen native English-speaking children (age 3;06 to 6;08, mean = 4;07) who had not participated in the previous experiments were recruited for Experiment 3 from the UH Mānoa Children's Center in Honolulu, Hawai'i. There were 6 girls and 13 boys: eight 3-year-olds, five 4-year-olds, two 5-year-olds, and four 6-year-olds.

7.2. Procedure and materials

The procedure and materials were identical to those in Experiment 1, except for two changes: there is no Think condition,[21] and the experiencer (the intervening NP) is a pronoun rather than a lexical NP for the unraised and raised sentences. A lead-in sentence was used by the puppet right before the test sentence to establish the antecedent of the pronoun and to increase the naturalness of the pronoun (e.g., *This is a story about Mickey, and* ... [test sentence]). Each critical condition (Unraised, Raised) had three items, and four lists were created in a Latin square design. A sample set of test items follows (see Table A3 in Appendix A for a full list of test items): In half of the six critical items the referents of the raised NP and the experiencer NP had the same gender (Table 6), while in the other half they had different genders (Table 7).

7.3. Results

The data from one 3-year-old were excluded, as he did not finish the experiment. We also excluded the data from four children (one 3-year-old, two 4-year-olds, and one 6-year-old) who failed to respond correctly to the filler items. The results from the remaining 14 children are presented in Figure 9.

Table 6. A Sample Set of Test Items in Each Condition: Same Gender (Exp. 3).

	Match Items
Unraised	This is a story about Mickey.
	And at the end of the story, it seems to him that Donald is short.
Raised	This is a story about Mickey.
	And at the end of the story, Donald seems to him to be short.
	Mismatch Items
Unraised	This is a story about Donald.
	And at the end of the story, it seems to him that Mickey is short.
Raised	This is a story about Donald.
	And at the end of the story, Mickey seems to him to be short.

Table 7. A Sample Set of Test Items in Each Condition: Different Gender (Exp. 3).

	Match Items
Unraised	This is a story about Lisa.
	And at the end of the story, it seems to her that Bart is studying.
Raised	This is a story about Lisa.
	And at the end of the story, Bart seems to her to be studying.
	Mismatch Items
Unraised	This is a story about Bart.
	And at the end of the story, it seems to him that Lisa is studying.
Raised	This is a story about Bart.
	And at the end of the story, Lisa seems to him to be studying.

---

[21]As for cognitive control, the Unraised condition was deemed to be sufficient.
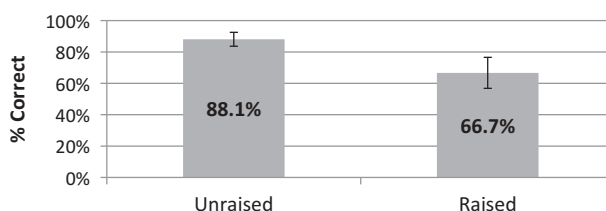
Figure 9. Mean accuracy in each condition (Exp. 3).

While the children correctly responded to the Unraised condition 88.1% of the time, their accuracy rate fell to 66.7% for the Raised condition. This difference was significant by a mixed-effects logistic regression analysis (β = −1.72, Z value = −2.42, p = .015) with Condition as the fixed effect and participants and items as random effects, showing that children were better at comprehending unraised sentences than raised sentences. A one-sample t-test (see Table B3 in Appendix B) also showed that the children's comprehension of raised sentences was no different from chance level (50%), whereas their score on the Unraised condition was at a rate greater than chance. Further analyses showed no significant effect of the following factors: Age, $\chi^2(1)$ = 0.03, p = .86; Age x Condition, $\chi^2(1)$ = 0.04, p = .84; Match/Mismatch, $\chi^2(1)$ = 0.99, p = .32; Match/Mismatch x Condition, $\chi^2(1)$ = 1.04, p = .31.[22] The critical items were also divided into Gender match and Gender mismatch and were compared for each condition.[23] The analysis revealed no difference between gender match and gender mismatch items on any condition (Unraised: β = −2.19, Z value = −1.60, p = .11; Raised: β = 0.30, Z value = 0.33, p = .74).

Turning back to the children's overall results, it is clear that the children's comprehension was better on the Unraised condition than the Raised condition. But does this mean that children have the same level of difficulty with raising sentences containing the pronominal experiencer as they do with a lexical NP experiencer? Before drawing this conclusion, let us first examine the results of those children who scored 100% on the Unraised condition. As illustrated in Figure 10, once we exclude the data from the five children who failed to comprehend the verb *seem*,[24] the remaining nine children comprehended the raised sentences with pronominal experiencers at far higher rates (81.5%) than our initial results suggest.

Let us compare the results of Experiment 3 to the previous experiments. First, the overall results of both Unraised and Raised conditions in Experiments 1, 2, and 3 are presented in Figure 11.

We fitted a mixed-effects logistic regression model in which Condition and Experiment (and their interaction) were included as fixed effects and participants and items as random effects. Condition was a significant predictor, $\chi^2(1)$ = 8.28, p = .004, showing that children were significantly better at

---

[22]For further data presentation, see Table C3 in Appendix C, which shows the results in terms of the number of children per age group in each condition.

[23]The purpose of this analysis was to see if the children had difficulty finding the correct antecedent of the pronoun when the gender of the pronoun was matched with the gender of the referent of the matrix subject (gender match) as in (iii).
   (iii) Donald seems to him to be short.
          Coreference between *him* and *Donald* in this case is ruled out by Principle B of the Binding Theory (N. Chomsky, 1981).
   (iv) Principle B: A pronoun may not be bound within its governing category.
As the governing category for the pronoun *him* in sentence (iii) is the matrix clause, Principle B rules out a referential dependency between *him* and *Donald*. However, prior acquisition studies show that children around age 5 often violate Principle B, allowing the ungrammatical interpretation (e.g., Chien & Wexler 1990; Grodzinsky & Reinhart 1993). Because the children in the current study are in a similar age range, it is crucial to establish that their responses on the gender-match items did not result from a difficulty with Principle B. Thus, the children's scores on gender-match items were compared with their scores on gender-mismatch items, where the gender cue prevents the pronoun from being coreferential with the local subject (e.g., *Bart seems to her to be studying*). No difference was observed between gender-match and gender-mismatch items on any condition, indicating that the children's responses were not confounded with any effect related to Principle B.

[24]The overall scores of the five excluded children were 73.3% on the Unraised condition and 60.0% on the Raised condition.
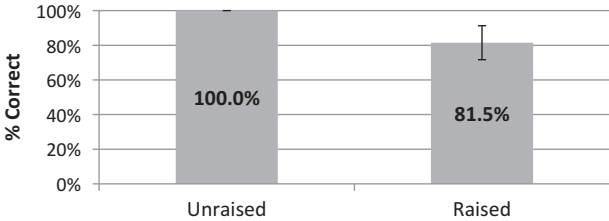
Figure 10. Mean accuracy of children who scored 100% on Unraised (Exp. 3).
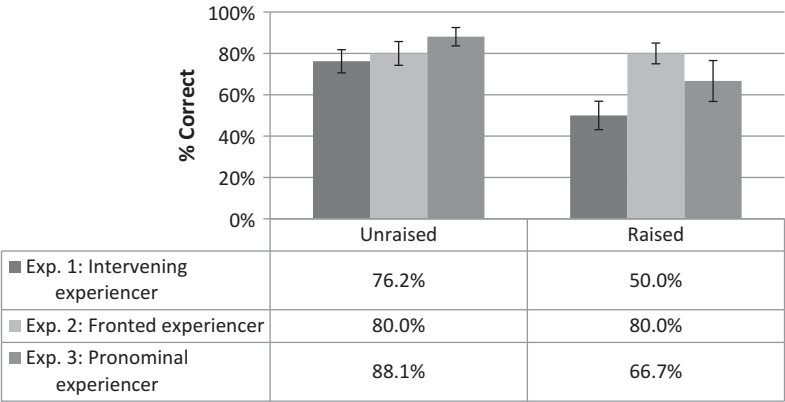
Figure 11. Comparison of Experiments 1, 2, and 3 (Unraised and Raised).

comprehending unraised sentences than raised sentences. Experiment also turned out to be a significant predictor, $\chi^2(2) = 7.52$, $p = .02$, suggesting that children's scores significantly differed by experiments. But there was no significant interaction between Condition and Experiment, $\chi^2(2) = 4.03$, $p = .13$. A mixed logistic model only for the Unraised condition showed that children's comprehension of unraised sentences in Experiment 3 did not differ from those in Experiment 1 ($\beta = -0.84$, Z value $= -1.40$, $p = .16$) or Experiment 2 ($\beta = -0.62$, Z value $= -1.04$, $p = .30$). However, a model only for the Raised condition revealed that children's comprehension of raising sentences was significantly better in Experiment 3 than in Experiment 1 ($\beta = -1.08$, Z value $= -2.05$, $p = .04$), and no significant difference was observed between Experiments 2 and 3 ($\beta = 0.43$, Z value $= 0.82$, $p = .41$).

Finally, Figure 12 presents a comparison of the results of the Raised condition from children who scored 100% on the Unraised condition in Experiments 1, 2, and 3.

A mixed logistic model was fitted with Experiment as the fixed effect and participants and items as random effects. The analysis showed that Experiment (1 vs. 2 vs. 3) was a significant predictor of
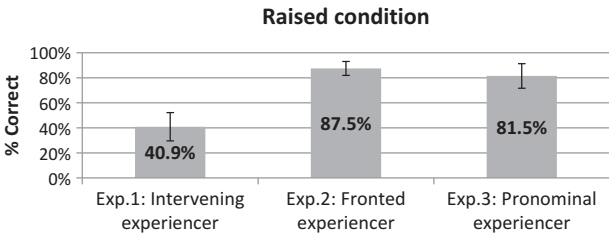
Figure 12. Comparison of Experiments 1, 2, and 3 (Raised).

children's performance on the Raised condition, $\chi^2(2) = 12.67$, $p = .0017$, and it revealed that the children's score on the Raised condition in Experiment 3 (intervening pronominal experiencer) was significantly higher than in Experiment 1 (intervening lexical NP experiencer; $\beta = -2.80$, Z value = -2.09, $p = .036$), while there was no significant difference, when compared to Experiment 2 (fronted experiencer; $\beta = 0.15$, Z value = 0.15, $p = .88$). This suggests that children's difficulty with raising is significantly reduced when the intervening experiencer is a pronoun, to the extent that the results are almost on par with those of Experiment 2, where the experiencer was fronted.

7.4. *Discussion*

Experiment 3 was conducted to explore the effect of the type of the intervening NP on children's comprehension of raising sentences by changing it to a pronoun. When compared with the results of Experiment 1, in which the raising sentences such as (25a) contain a lexical NP experiencer, the results of Experiment 3 demonstrate that having a pronominal experiencer, as in sentence (25b), facilitates children's comprehension.

  (25) a. Experiment 1: Donald seems to Mickey to be short.
       b. Experiment 3: Donald seems to him to be short.

Children's improved comprehension with a pronominal intervening experiencer demonstrates that their difficulty with raising patterns is sensitive to the same factors as other structures for which intervention effects have been posited and not to any grammatical effects *per se*. Thus, these findings are inconsistent with grammatical-deficit hypotheses of the kind considered in the field to date, such as the AIH, because it predicts that children's difficulty should persist as long as there is an intervener, regardless of its NP type. In contrast, this pattern of results is consistent with the idea of Performance-based Intervention Effects, which places children's difficulty with raising at the performance level and not in their competence.

8. General discussion

In this article, we addressed several methodological shortcomings of previous studies on the acquisition of raising. Moreover, we sought to tease apart the reasons for the previously claimed difficulties with raising. To this end, we conducted three experiments to examine children's comprehension of raising sentences across the experiencer (Experiment 1), with a fronted experiencer (Experiment 2), and across a pronominal experiencer (Experiment 3).

Experiment 1 showed that, despite the methodological improvements, children still had difficulty comprehending raising sentences across the experiencer (e.g., *Donald seems to Mickey to be short*)—a finding consistent with the results of Hirsch and Wexler's (2007) and Hirsch, Orfitelli & Wexler's (2007) studies. Experiment 2 was conducted to investigate whether the source of the difficulty comes from the operation of raising or the presence of an intervening experiencer phrase. Experiment 2 therefore tested raising sentences where the experiencer phrase is moved to the beginning of the sentence (e.g., *To Mickey, Donald seems to be short*). We found that children's difficulty with raising sentences is significantly reduced with a fronted experiencer. That is, when the experiencer no longer intervenes between the raised NP and the gap, children's difficulty with raising disappears, to the extent that their comprehension of raising sentences becomes similar to their comprehension of unraised sentences. Thus, the results of Experiments 1 and 2 are incompatible with the Universal Phase Requirement (UPR; Wexler 2004) and Universal Freezing Hypothesis (UFH; Hyams & Snyder 2005), which predict that children have trouble comprehending raising sentences irrespective of the location of the experiencer. Rather, these results support accounts that attribute the difficulty to the intervening experiencer phrase: the

Argument Intervention Hypothesis (AIH; Orfitelli 2012) and Performance-based Intervention Effects (PIE), advocated in this article.

Finally, Experiment 3 tested raising patterns in which a lexical NP is raised across a pronominal experiencer (e.g., *Donald seems to him to be short*). The results revealed that children were able to comprehend the raising sentences with a pronominal experiencer significantly better than those with a lexical NP experiencer (Experiment 1: *Donald seems to Mickey to be short*) and that their comprehension was as accurate as when there was no intervener (Experiment 2: *To Mickey, Donald seems to be short*). That is, having a pronominal experiencer facilitated the comprehension of raising sentences, even though the pronoun interrupts the filler-gap dependency. This pronoun advantage argues against grammar-based approaches to the acquisition of raising, such as the UFH and AIH, which claim that a grammatical deficit makes raising impossible for children.[25] However, this lexical NP/pronoun difference is in fact expected under the PIE hypothesis, as this kind of intervention effect is known to be sensitive to the NP type. Furthermore, according to the Dependency Locality Theory (DLT), as the intervener is a pronoun, a more referentially accessible noun phrase than a lexical NP, it incurs less processing cost. Thus, the results of Experiment 3 provide a key piece of evidence that children's difficulty with raising is due to the same limitations that affect other structures for which intervention effects have been found. [26]

Overall, the results from the three experiments suggest that the difficulty associated with raising patterns cannot be attributed to children's underdeveloped syntactic abilities or to grammatical deficits, as has previously been suggested (e.g., Borer & Wexler 1987; Hyams & Snyder 2005; Orfitelli 2012; Wexler 2004). Rather, we are led to the conclusion that the observed difficulty with raising is due to grammar-external factors (Arnon 2005; Bloom 1990; Valian 1991) and that the operation of raising *per se* is essentially intact in child grammar.

Regarding intervention effects in general, the findings of the article add to the growing body of literature demonstrating the presence of intervention effects in a number of sentence types (e.g., object relatives, object topicalizations, object *wh*-questions), suggesting that this phenomenon is not limited to a single structure. As mentioned earlier, intervention effects have been observed typically in structures that involve *wh*-movement. However, following Friedmann and

---

[25]An anonymous reviewer (correctly) points out that Relativized Minimality or other similar frameworks allow for the possibility of pronouns and nouns having different feature sets and thus not counting as like elements. As currently stated, the AIH does not invoke such a possibility and simply states that movement may not occur over an intervening argument (nouns and pronouns alike). Were such a system to be adopted by the AIH, it might prove fruitful to explain some of our results.

[26]It should be noted that different theories have been put forward to explain the exact nature of intervention effects in language acquisition and processing and that not all accounts are processing-based explanations like the Dependency Locality Theory (DLT). For example, some (Belletti & Rizzi 2010; Friedmann, Belletti & Rizzi 2009) suggest that the observed intervention effects can be explained in terms of Relativized Minimality (RM; Rizzi 1990). This locality constraint states that in the configuration, … X … Z … Y …, Y cannot be related to X if Z intervenes and Z bears a certain similarity to X. Extending this idea to language acquisition, Friedmann, Belletti & Rizzi (2009) argue that children require the relevant features of X (the target) and Z (the intervener) to be entirely distinct in order to move an element over an intervener. In other words, the child's principle allows only patterns in which there is no overlapping feature between the moved element and the intervener. Thus, extraction in object relatives like (v) is disallowed in the child system, as the relative head *the boy* and the intervener *the girl* share the same "+NP" feature, since both have a lexical head noun.

  (v) Object relative: The boy that   the girl kissed _
                    [+R, +NP]   [+NP]   [+R, +NP]
                    (where "+R" designates a relative pronoun feature, "+NP"
                    designates a full lexical noun phrase.)

In turn, the RM account predicts that if the relative head and the intervener are made sufficiently different (for example, if one is a lexical NP and the other is a pronoun), then extraction should be possible even in the child system. As such, we do not necessarily deny the possibility that the findings of the current article, namely the pronoun advantage in raising acquisition, may also be explained by a grammar-based approach such as the RM account. However, further evidence for a performance effect comes from Choe, Deen & O'Grady (2014), who found an asymmetry in the effect of NP type on children's comprehension of raising sentences: An improvement in performance was observed when a lexical NP was raised across a pronominal experiencer (e.g., *Bart seems to her to be studying*) but not when a pronoun was raised across a lexical NP experiencer (e.g., *He seems to Lisa to be studying*). Such results are not expected under the RM account but are predicted by the DLT-based account of intervention effects.

Costa ([2010](#)), who report intervention effects in coordinated structures (e.g., *Grandma tickled the girl and _ smiled*), the current findings on raising provide additional evidence that intervention effects may occur in structures involving a broader range of dependencies, including A-movement dependencies.

While the current findings have shed some light on how children comprehend raising sentences with an experiencer, further investigation is necessary to examine children's comprehension of raising sentences without an experiencer phrase (e.g., *John seems to be short*). It was noted earlier that the results of the previous studies are controversial in that some studies claim delayed acquisition of such sentences (e.g., Hirsch, Orfitelli & Wexler [2008](#)), while others claim no delay (e.g., Becker [2006](#)). However, each of the previous studies raises methodological concerns that need to be addressed before any definitive conclusion can be drawn. As for the predictions of the PIE, it predicts that such structures should not be difficult for children, since there is no intervening nominal within the dependency of the raised NP and the gap in its original position. In fact, this prediction was borne out, at least indirectly, in Experiment 2, where we tested raising sentences with a fronted experiencer. The results (showing children's good comprehension) show that when the experiencer is fronted, children's difficulty with raising disappears, suggesting that children should have no difficulty with experiencer-less raising sentences either.

We are thus left with the conclusion that the raising facts in child language are best understood in terms of the performance factors that we have been considering rather than any fundamental grammatical deficits in the realm of A-movement structures.

References

Arnon, Inbal. 2005. Relative clause acquisition in Hebrew: Towards a processing-oriented account. In Alejna Brugos, Manuella R. Clark-Cotton & Seungwan Ha (eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development (BUCLD)*, 37–48. Somerville, MA: Cascadilla Press.

Arnon, Inbal. 2010. Rethinking child difficulty: The effect of NP types on children's processing of relative clauses in Hebrew. *Journal of Child Language* 37(1). 27–57.

Avrutin, Sergey. 2000. Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. In Yosef Grodzinsky, Lewis P. Shapiro & David Swinney (eds.), *Language and the brain: Representation and processing*, 295–312. San Diego, CA: Academic Press.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. http://CRAN.R-project.org/package=lme4.

Becker, Misha. 2004. Learning verbs that lack argument structure: The case of raising verbs. In Jacqueline van Kampen & Sergio Baauw (eds.), *Proceedings of GALA 2003 (Generative Approaches to Language Acquisition)*, 115–125. Utrecht: LOT.

Becker, Misha. 2006. There began to be a learnability puzzle. *Linguistic Inquiry* 37(3). 441–456.

Becker, Misha. 2007. Animacy, expletives, and the learning of the raising-control distinction. In Alyona Belikova, Luisa Meroni & Mari Umeda (eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA 2)*, 12–20. Somerville, MA: Cascadilla Proceedings Project.

Becker, Misha. 2014. *The acquisition of syntactic structure: Animacy and thematic alignment.* Cambridge: Cambridge University Press.

Bloom, Paul. 1990. Subjectless sentences in child language. *Linguistic Inquiry* 21. 491–504.

Boeckx, Cedric. 1999. Conflicting c-command requirements. *Studia Linguistica* 53. 227–250.

Boeckx, Cedric. 2008. Raising across experiencers cross-linguistically. In *Aspects of the syntax of agreement*, 150–161. London: Routledge.

Borer, Hagit & Kenneth Wexler. 1987. The maturation of syntax. In Thomas Roeper & Edwin Williams (eds.), *Parameter setting*, 123–172. Dordrecht: Reidel.

Caplan, David & Gloria Waters. 2002. Working memory and connectionist models of parsing: A reply to MacDonald and Christiansen (2002). *Psychological Review* 109. 66–74.

Carpenter, Patricia, Akira Miyake & Marcel Just. 1994. Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. In M. Gernsbacher (ed.), *Handbook of psycholinguistics*, 1075–1122. San Diego, CA: Academic Press.

Chien, Yu-Chin & Kenneth Wexler. 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1(3). 225–295.

Choe, Jinsun. 2012. *Children seem to know raising: Raising and intervention in child language*. Honolulu, HI: University of Hawai'i at Mānoa dissertation.

Choe, Jinsun, Kamil Ud Deen & William O'Grady. 2014. Children seem to know raising: Intervention effects in child language. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL)*, 107–114. Somerville, MA: Cascadilla Proceedings Project.

Chomsky, Carol. 1969. The acquisition of syntax in children from 5 to 10. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael J. Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Clahsen, Harald (ed.). 1996. *Generative perspectives on language acquisition. Empirical findings, theoretical considerations & cross-linguistic comparisons*. Amsterdam: John Benjamins.

Collins, Chris. 2005a. A smuggling approach to raising in English. *Linguistic Inquiry* 36(2). 289–298.

Collins, Chris. 2005b. A smuggling approach to the passive in English. *Syntax* 8. 81–120.

Costa, João & Namma Friedmann. 2012. Children acquire unaccusatives and A-movement very early. In Martin Everaert, Marijana Marelj, & Tal Siloni (eds.), *The theta system: Argument structure at the interface. Oxford Studies in Theoretical Linguistics 37*, 354–378. Oxford: Oxford University Press.

Crain, Stephen. 1991. Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 4. 597–650.

Crain, Stephen & Cecile McKee. 1985. The acquisition of structural restrictions on anaphora. In Stephen Berman, Jaewoong Choe & Joyce McDonough (eds.), *Proceedings of the 16th Annual Meeting of the North East Linguistic Society (NELS 16)*, 94–110. Amherst, MA: GLSA Publications.

Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar: A guide to experiments in the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

De Vincenzi, Marica, Lisa Saskia Arduino, Laura Ciccarelli & Remo Job. 1999. Parsing strategies in children comprehension of interrogative sentences. In Sebastiano Bagnara (ed.), *Proceedings of European Conference on Cognitive Science*, 301–308. Rome: Istituto di Psicologia del CNR.

Fox, Danny & Yosef Grodzinsky. 1998. Children's passive: A view from the by-phrase. *Linguistic Inquiry* 29. 311–332.

Friedmann, Naama. 2007. Young children and A-chains: The acquisition of Hebrew unaccusatives. *Language Acquisition* 14. 377–422.

Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119. 67–88.

Friedmann, Naama & João Costa. 2010. The child heard a coordinated sentence and wondered: On children's difficulty in understanding coordination and relative clauses with crossing dependencies. *Lingua* 120(6). 1502–1515.

Friedmann, Naama & Hedva Lavi. 2006. On the order of acquisition of A-movement, Wh-movement and V-C movement. In Adriana Belletti, Elisa Bennati, Cristiano Chesi, Elisa Di Domenico & Ida Ferrari (eds.), *Language acquisition and development*, 211–217. Cambridge: Cambridge Scholars Press/CSP.

Friedmann, Naama & Rama Novogrodsky. 2004. The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language* 31. 661–681.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita & Wayne O'Neil (eds.), *Image, language, brain*, 95–126. Cambridge, MA: MIT Press.

Gordon, Peter C., Randall Hendrick & Marcus Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27. 1411–1423.

Grodzinsky, Yosef & Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24(1). 69–101.

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Referring expressions in discourse. *Language* 69. 274–307.

Hamburger, Henry & Stephen Crain. 1982. Relative acquisition. In Stan Kuczaj (ed.), *Language development, Vol. 1: Syntax and semantics*, 245–274. Hillsdale, NJ: Erlbaum.

Hirsch, Christopher. 2011. *The acquisition of raising*. Cambridge, MA: MIT dissertation.

Hirsch, Christopher, Robyn Orfitelli & Kenneth Wexler. 2007. When seem means think: The role of the experiencer-phrase in children's comprehension of raising. In Alyona Belikova, Luisa Meroni & Mari Umeda (eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA 2)*, 135–146. Somerville, MA: Cascadilla Proceedings Project.

Hirsch, Christopher, Robyn Orfitelli & Kenneth Wexler. 2008. The acquisition of raising reconsidered. In Anna Gavarro & M. João Freitas (eds.), *Proceedings of Generative Approaches to Language Acquisition (GALA 2007)*, 253–262. Cambridge: Cambridge Scholars Press.

Hirsch, Christopher & Kenneth Wexler. 2007. The late acquisition of raising: What children seem to think about seem. In William D. Davies & Stanley Dubinsky (eds.), *New horizons in the analysis of control and raising*, 35–70. New York: Springer.

Hyams, Nina & William Snyder. 2005. Young children never smuggle: Reflexive clitics and the universal freezing hypothesis. Paper presented at the 30th Annual Boston University Conference on Language Development [BUCLD 30], Boston University, Boston, MA.

Kidd, Evan, Silke Brandt, Elena Lieven & Michael Tomasello. 2007. Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes* 22. 860–897.

McKee, Cecile, Dana McDaniel & Jesse Snedeker. 1998. Relatives children say. *Journal of Psycholinguistic Research* 27. 573–596.

Messenger, Katherine, Holly Branigan, Janet McLean & Antonella Sorace. 2012. Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language* 66(4). 568–587.

Mitchener, William Garrett & Misha Becker. 2011. A computational model of learning the raising-control distinction. *Research on Language and Computation* 8. 169–207.

Müller, Gereon. 1998. *Incomplete category fronting.* Dordrecht: Kluwer.

O'Brien, Karen, Elaine Grolla & Diane Lillo-Martin. 2006. Long passives are understood by young children. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development (BUCLD 30)*, 441–451. Somerville, MA: Cascadilla Press.

Orfitelli, Robyn. 2012. *Argument intervention in the acquisition of A-movement.* Los Angeles, CA: University of California, Los Angeles dissertation.

Pierce, Amy E. 1992. *Language acquisition and syntactic theory: A comparative analysis of French and English child grammars.* Dordrecht: Kluwer.

Pinker, Steven. 1984. *Language learnability and language development.* Cambridge, MA: MIT Press.

Poeppel, David & Kenneth Wexler. 1993. The full competence hypothesis of clause structure in early German. *Language* 69. 1–33.

Postal, Paul M. 1974. *On raising.* Cambridge, MA: MIT Press.

The R Foundation. 2014. The R Project for statistical computing. http://www.R-project.org/.

Reali, Florencia & Morten H. Christiansen. 2007. Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language* 57. 1–23.

Rizzi, Luigi. 1990. *Relativized minimality.* Cambridge, MA: MIT Press.

Snyder, William, Nina Hyams & Paola Crisma. 1994. Romance auxiliary selection with reflexive clitics: Evidence for early knowledge of unaccusativity. In *Proceedings of the 26th Annual Child Language Research Forum*, 127–136. Stanford, CA: CSLI.

Tavakolian, Susan L. 1981. The conjoined-clause analysis of relative clauses. In Susan L. Tavakolian (ed.), *Language acquisition and linguistic theory*, 167–187. Cambridge, MA: MIT Press.

Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40. 21–81.

Wanner, Eric & Michael Maratsos. 1978. An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. Miller (eds.), *Linguistic theory and psychological reality*, 119–161. Cambridge, MA: MIT Press.

Warren, Tessa & Edward Gibson. 2002. The influence of referential processing on sentence complexity. *Cognition* 85. 79–112.

Wexler, Kenneth. 2004. Theory of phasal development: Perfection in child grammar. *MIT Working Papers in Linguistics* 48. 159–209.

Appendix A

Table A1. Test Sentences for Experiment 1.

Item	Condition	M/MM	Test Sentence
1	Think	Match	At the end of the story, Mickey still thinks Donald is short.
1	Unraised	Match	At the end of the story, it still seems to Mickey that Donald is short.
1	Raised	Match	At the end of the story, Donald still seems to Mickey to be short.
1	Think	Mismatch	At the end of the story, Donald still thinks Mickey is short.
1	Unraised	Mismatch	At the end of the story, it still seems to Donald that Mickey is short.
1	Raised	Mismatch	At the end of the story, Mickey still seems to Donald to be short.
2	Think	Match	At the end of the story, Piglet still thinks Pooh is wearing a blue shirt.
2	Unraised	Match	At the end of the story, it still seems to Piglet that Pooh is wearing a blue shirt.
2	Raised	Match	At the end of the story, Pooh still seems to Piglet to be wearing a blue shirt.
2	Think	Mismatch	At the end of the story, Pooh still thinks Piglet is wearing a blue shirt.
2	Unraised	Mismatch	At the end of the story, it still seems to Pooh that Piglet is wearing a blue shirt.
2	Raised	Mismatch	At the end of the story, Piglet still seems to Pooh to be wearing a blue shirt.
3	Think	Match	At the end of the story, Lisa still thinks Bart is studying.
3	Unraised	Match	At the end of the story, it still seems to Lisa that Bart is studying.
3	Raised	Match	At the end of the story, Bart still seems to Lisa to be studying.
3	Think	Mismatch	At the end of the story, Bart still thinks Lisa is studying.
3	Unraised	Mismatch	At the end of the story, it still seems to Bart that Lisa is studying.
3	Raised	Mismatch	At the end of the story, Lisa still seems to Bart to be studying.
4	Think	Match	At the end of the story, Lisa still thinks Bart plays the saxophone well
4	Unraised	Match	At the end of the story, it still seems to Lisa that Bart plays the saxophone well.
4	Raised	Match	At the end of the story, Bart still seems to Lisa to play the saxophone well.
4	Think	Mismatch	At the end of the story, Bart still thinks Lisa plays the saxophone well.
4	Unraised	Mismatch	At the end of the story, it still seems to Bart that Lisa plays the saxophone well.
4	Raised	Mismatch	At the end of the story, Lisa still seems to Bart to play the saxophone well.
5	Think	Match	At the end of the story, Homer still thinks Lisa is crying.
5	Unraised	Match	At the end of the story, it still seems to Homer that Lisa is crying.
5	Raised	Match	At the end of the story, Lisa still seems to Homer to be crying.
5	Think	Mismatch	At the end of the story, Lisa still thinks Homer is crying.
5	Unraised	Mismatch	At the end of the story, it still seems to Lisa that Homer is crying.
5	Raised	Mismatch	At the end of the story, Homer still seems to Lisa to be crying.
6	Think	Match	At the end of the story, Goofy still thinks Mickey is purple.
6	Unraised	Match	At the end of the story, it still seems to Goofy that Mickey is purple.
6	Raised	Match	At the end of the story, Mickey still seems to Goofy to be purple.
6	Think	Mismatch	At the end of the story, Mickey still thinks Goofy is purple.
6	Unraised	Mismatch	At the end of the story, it still seems to Mickey that Goofy is purple.
6	Raised	Mismatch	At the end of the story, Goofy still seems to Mickey to be purple.

Table A2. Test Sentences for Experiment 2.

Item	Condition	M/MM	Test Sentence
1	Think	Match	Mickey thinks Donald is short.
1	Unraised	Match	To Mickey, it seems that Donald is short.
1	Raised	Match	To Mickey, Donald seems to be short.
1	Think	Mismatch	Donald thinks Mickey is short.
1	Unraised	Mismatch	To Donald, it seems that Mickey is short.
1	Raised	Mismatch	To Donald, Mickey seems to be short.
2	Think	Match	Piglet thinks Pooh is wearing a blue shirt.
2	Unraised	Match	To Piglet, it seems that Pooh is wearing a blue shirt.
2	Raised	Match	To Piglet, Pooh seems to be wearing a blue shirt.
2	Think	Mismatch	Pooh thinks Piglet is wearing a blue shirt.
2	Unraised	Mismatch	To Pooh, it seems that Piglet is wearing a blue shirt.
2	Raised	Mismatch	To Pooh, Piglet seems to be wearing a blue shirt.
3	Think	Match	Lisa thinks Bart is studying.
3	Unraised	Match	To Lisa, it seems that Bart is studying.
3	Raised	Match	To Lisa, Bart seems to be studying.
3	Think	Mismatch	Bart thinks Lisa is studying.
3	Unraised	Mismatch	To Bart, it seems that Lisa is studying.
3	Raised	Mismatch	To Bart, Lisa seems to be studying.
4	Think	Match	Lisa thinks Bart plays the saxophone well
4	Unraised	Match	To Lisa, it seems that Bart plays the saxophone well.
4	Raised	Match	To Lisa, Bart seems to play the saxophone well.
4	Think	Mismatch	Bart thinks Lisa plays the saxophone well.
4	Unraised	Mismatch	To Bart, it seems that Lisa plays the saxophone well.
4	Raised	Mismatch	To Bart, Lisa seems to play the saxophone well.
5	Think	Match	Homer thinks Lisa is crying.
5	Unraised	Match	To Homer, it seems that Lisa is crying.
5	Raised	Match	To Homer, Lisa seems to be crying.
5	Think	Mismatch	Lisa thinks Homer is crying.
5	Unraised	Mismatch	To Lisa, it seems that Homer is crying.
5	Raised	Mismatch	To Lisa, Homer seems to be crying.
6	Think	Match	Goofy thinks Mickey is purple.
6	Unraised	Match	To Goofy, it seems that Mickey is purple.
6	Raised	Match	To Goofy, Mickey seems to be purple.
6	Think	Mismatch	Mickey thinks Goofy is purple.
6	Unraised	Mismatch	To Mickey, it seems that Goofy is purple.
6	Raised	Mismatch	To Mickey, Goofy seems to be purple.

Table A3. Test Sentences for Experiment 3.

Item	Condition	M/MM	Test Sentence
1	Unraised	Match	This is a story about Mickey, and it seems to him that Donald is short.
1	Raised	Match	This is a story about Mickey, and Donald seems to him to be short.
1	Unraised	Mismatch	This is a story about Donald, and it seems to him that Mickey is short.
1	Raised	Mismatch	This is a story about Donald, and Mickey seems to him to be short.
2	Unraised	Match	This is a story about Piglet, and it seems to him that Pooh is wearing a blue shirt.
2	Raised	Match	This is a story about Piglet, and Pooh seems to him to be wearing a blue shirt.
2	Unraised	Mismatch	This is a story about Pooh, and it seems to him that Piglet is wearing a blue shirt.
2	Raised	Mismatch	This is a story about Pooh, and Piglet seems to him to be wearing a blue shirt.
3	Unraised	Match	This is a story about Lisa, and it seems to her that Bart is studying.
3	Raised	Match	This is a story about Lisa, and Bart seems to her to be studying.
3	Unraised	Mismatch	This is a story about Bart, and it seems to him that Lisa is studying.
3	Raised	Mismatch	This is a story about Bart, and Lisa seems to him to be studying.
4	Unraised	Match	This is a story about Lisa, and it seems to her that Bart plays the saxophone well.
4	Raised	Match	This is a story about Lisa, and Bart seems to her to play the saxophone well.
4	Unraised	Mismatch	This is a story about Bart, and it seems to him that Lisa plays the saxophone well.
4	Raised	Mismatch	This is a story about Bart, and Lisa seems to him to play the saxophone well.
5	Unraised	Match	This is a story about Homer, and it seems to him that Lisa is crying.
5	Raised	Match	This is a story about Homer, and Lisa seems to him to be crying.
5	Unraised	Mismatch	This is a story about Lisa, and it seems to her that Homer is crying.
5	Raised	Mismatch	This is a story about Lisa, and Homer seems to her to be crying.
6	Unraised	Match	This is a story about Goofy, and it seems to him that Mickey is purple.
6	Raised	Match	This is a story about Goofy, and Mickey seems to him to be purple.
6	Unraised	Mismatch	This is a story about Mickey, and it seems to him that Goofy is purple.
6	Raised	Mismatch	This is a story about Mickey, and Goofy seems to him to be purple.

Appendix B

Table B1. One Sample *t*-test for Each Condition (Exp. 1).

Condition	Percent Correct	t	df	p
Think	81.0%	5.70	20	0.00
Unraised	76.2%	4.69	20	0.00
Raised	50.0%	0.00	20	1.00

Table B2. One Sample *t*-test for Each Condition (Exp. 2).

Condition	Percent Correct	t	df	p
Think	80.0%	6.00	24	0.00
Unraised	80.0%	5.20	24	0.00
Raised	80.0%	6.00	24	0.00

Table B3. One Sample *t*-test for Each Condition (Exp. 3).

Condition	Percent correct	t	df	p
Unraised	88.1%	8.60	13	0.00
Raised	66.7%	1.69	13	0.12

Appendix C

Table C1. Number of Children per Age Group in Each Condition (Exp. 1).

	Age 4 (n = 11)			Age 5 (n = 10)		
Correct/Total Tokens	Think	Unraised	Raised	Think	Unraised	Raised
0/2	0	0	0	0	0	4
1/2	4	7	9	4	3	4
2/2	7	4	2	6	7	2

Table C2. Number of Children per Age Group in Each Condition (Exp. 2).

	Age 3 (n = 6)			Age 4 (n = 9)			Age 5 (n = 10)		
Correct/Total Tokens	Think	Unraised	Raised	Think	Unraised	Raised	Think	Unraised	Raised
0/2	0	0	0	0	1	0	0	0	0
1/2	3	2	1	3	4	4	4	2	5
2/2	3	4	5	6	4	5	6	8	5

Table C3. Number of Children per Age Group in Each Condition (Exp. 3).

	Age 3–4 (n = 9)		Age 5–6 (n = 5)	
Correct/Total Tokens	Unraised	Raised	Unraised	Raised
0/3	0	1	0	0
1/3	0	3	0	2
2/3	3	1	2	0
3/3	6	4	3	3