QR in child and adult grammar: What children can do and what adults won’t when resolving Antecedent-Contained Deletion

Kristen Syrett\textsuperscript{a} and Jeffrey Lidz\textsuperscript{b}

\textsuperscript{a} Rutgers University and \textsuperscript{b} University of Maryland
Abstract

We show that four-year-olds assign the correct interpretation to ACD sentences because they have the correct representation of these structures, which involves Quantifier Raising (QR). We first show that they treat ACD sentences differently from sentences with conjunction, a plausible alternative if they lacked QR. A baseline of their interpretation of ambiguous VP-ellipsis sentences allows us to investigate sentences involving multiple QR landing sites. When ACD is embedded in a nonfinite clause, children access the embedded and matrix readings. When ACD is embedded in a tensed clause, children and some adults access the supposedly ungrammatical matrix reading. We suggest that this asymmetry reflects the nature of the parser and that the transition to online adult-like sentence comprehension involves developing the ability to implement searches over representations held in memory, but that this can be contrasted with theories predicting chance-like performance with these structures.

keywords: quantifiers, syntax semantics interface, logical form, ellipsis, child language, parsing, language processing, language acquisition, adult language
1. **Introduction**

Much of the literature on children’s acquisition of syntax and semantics can be characterized as showing that children have the same kinds of abstract linguistic representations as adults (Crain (1991), De Villiers and Roeper (1995), Pinker (1989), Poeppel and Wexler (1993); see Goodluck (2007), Lidz (2007) for a review). Such work provides us with strong evidence for the continuity of representation and computation across child and adult grammars. However, if we are to build a theory of language learning with any developmental interest, it is important to identify cases where children and adults differ. Such cases shed light on the ways that a particular language is not uniquely determined by principles inside the child. In addition, they can help us to determine which grammatical properties require development over time and the degree to which this development requires changes in the grammar *per se* or depends on extragrammatical factors such as biological maturation (Wexler (1990)), conceptual development (De Villiers (1998), De Villiers and Pyers (1997)), pragmatic development (Thornton and Wexler (1999)), or development of general computational capacities (Grodzinsky and Reinhart (1993), Reinhart (2004)).

In this paper, we explore a difference between children and adults in the interpretations that these two populations assign to certain ambiguous sentences. We begin by examining children’s knowledge of Antecedent-Contained Deletion (ACD) structures such as

(1) **Lola jumped over every frog that Dora did.**

The meaning assigned to this sentence is that for every frog that Dora jumped over, Lola also jumped over it. We show that children have knowledge of the abstract representations and computational operations required for interpreting such sentences. At the same time, though, we show that for more complex sentences containing ACD, children are more liberal in their
interpretations than adults are.

This discovery leads to an interesting learnability puzzle, since the ways in which children are more liberal than adults fall outside of what could plausibly come to be restricted on the basis of primary linguistic data. Sentences with ACD are far from common in the input, let alone instances attesting to the full range of variability of allowable interpretations. Thus, to the extent that children and adults differ in the range of interpretations that they allow for sentences with ACD, we are forced to look outside of this domain for the evidence that children would need to encounter in order to acquire (or learn to implement) the relevant constraints.

ACD provides an excellent window into the language acquisition process for another reason as well; it provides what is generally considered the strongest piece of evidence for the existence of the operation of Quantifier Raising (QR) (May (1977; 1985)). Other evidence for QR typically comes from ambiguous sentences involving interactions between multiple quantificational NPs or from interactions between a Quantificational NP (QNP) and negation. However, experimental results concerning children’s ability to deal with such ambiguities are somewhat mixed. Thus, definitive evidence that children’s grammars include QR is still lacking.

First, for sentences involving multiple quantificational phrases, such as (2), the issue remains unresolved.

(2) Every boy is riding an elephant.

In some experiments, children have been shown to differ from adults in ways that suggest that their representations of QNPs might be different from those of adults (Drozd (2001), Drozd and van Loosbroek (1999), Geurts (2003), Inhelder and Piaget (1958; 1964), Philip (1995; 1996)). For example, when given a scenario in which there are four elephants, only three of which are being ridden by a boy, children often reject a sentence such as (2), pointing to the elephant that is
not being ridden by a boy as supporting evidence for their rejection. If, for example, these responses lead to the conclusion that the universal quantifier every quantifies over events rather than properties (Philip (1995)), then it suggests that data from these sentences cannot provide evidence for or against the existence of a covert movement operation like QR, which applies only to quantificational NPs. In other experiments investigating these sentences, experimental manipulations have resulted in children patternning more like adults (Brooks and Braine (1996), Crain et al. (1996), Drozd and van Loosbroek (1999), Freeman et al. (1982), Freeman and Stedmon (1986)). This evidence has been taken by some to indicate that children’s errors derive from methodological aspects of the experiments (specifically a failure to control for the effect of pragmatic factors contributing to appropriately restricting the domain of quantification) and not children’s grammatical inabilities. Either way, this piece of the literature on children’s quantificational abilities does not directly speak to the question of QR.

Second, for sentences involving scopal interactions between negation and a QNP in object position, such as (3), the data now clearly indicate that children can compute both readings of the relevantly ambiguous sentences (Gualmini (2003a; b), Lidz and Musolino (2002), Musolino and Lidz (2006); see Musolino (1999a; b) for earlier findings).

(3)   a. Every horse didn’t jump over the fence.
   b. The Smurf didn’t buy an orange.

For a sentence such as (3a) where the universal quantifier every is in subject position, the interpretation corresponding to the surface scope (every > not) is that no horse jumped over the fence. Inverse scope (not > every) allows for some number of horses less than the total to have jumped over the fence. While preschoolers are able to compute both interpretations of these sentences, it is less obvious that doing so requires QR at all. For (3a), it is more likely that the
inverse scope interpretation derives from interpreting the QNP *every horse* in its base position inside the VP (McCloskey (1996)) and thus bears more on the question of children’s knowledge of reconstruction than on their knowledge of QR (cf. Leddon (2006), Leddon and Lidz (2006)). For (3b), the specific reading of the indefinite might not be derived by raising the quantificational phrase covertly but instead by treating this phrase as though it were referential (Fodor and Sag (1982), Kratzer (1998)) or quantificational through existential closure (Abusch (1994), Reinhart (1997), Winter (1997)). Thus, although a sizable literature has developed focusing on how children deal with QNPs, very little of this literature directly addresses the question of whether children’s grammars actually contain the operation of QR.¹

One structure which would provide clear evidence for QR in children is ACD. To our knowledge, only one other published study (Kiguchi and Thornton (2004)) has examined ACD directly, and thus is the only study to directly address the question of children’s QR.² The authors argued that children do indeed interpret sentences containing ACD in an adult-like fashion; however, this study leaves open a number of important questions and, as we will see, may not provide conclusive evidence that children resolve ACD in the same way as adults. Before

¹ Note that for these same reasons, even children’s interpretation of much simpler sentences with a quantified NP in the object position such as (i) cannot be offered as proof of the existence of QR in child grammar, since there is the possibility that for children, there is no type mismatch between the QNP and the transitive verb, or because the quantifier is not treated as a universal quantifier and is interpreted in its base position.

(i) Mary met every student

² See also Guasti and Chierchia (1999/2000), for discussion of reconstruction effects in child language, which also points to the existence of the QR operation in child grammar.
continuing into a discussion of these questions, let us first take a digression into the grammar of ACD.

2. The Grammar of ACD

2.1. Why ACD requires QR

ACD, first discussed by Bouton (1970), is a special case of verb phrase ellipsis (VPE) and is now generally held to provide one of the strongest pieces of evidence for the covert displacement operation of QR (Fiengo and May (1994), Kennedy (1997), Larson and May (1990), May (1985), Sag (1976); but see Baltin (1987), Hornstein (1994)). It is now part of the standard analysis of elided VPs that they are interpreted as identical in reference to another VP in the discourse context (Hankamer and Sag (1976)). For example, in (4), the elided VP (signaled by *did*) is interpreted as identical to the underlined VP (tense aside).

(4)  Lola jumped over every frog and Dora did too.

    = Lola jumped over every frog and Dora did jump over every frog too.

What makes ACD unique, though, is that the elided VP is contained in its antecedent. As we observe in (5), the elided VP is part of the underlined VP.

(5)  Lola jumped over every frog that Dora did.

If we were to replace the elided VP with the matrix VP, there would still be an ellipsis site in the replacement VP:

(5') Lola jumped over every frog that Dora did

[jump over every frog that Dora did ...]

Any attempt to resolve the ellipsis with this antecedent VP results in another elided VP, ad infinitum. As long as the elided VP is contained in its antecedent, the two VPs cannot possibly be
identical and so the ellipsis cannot be properly resolved.\textsuperscript{3} The sentence therefore remains uninterpretable as long as the QNP remains \textit{in situ}. An operation of covert displacement, however, averts both the infinite regress and a lack of parallelism (May (1977)). After movement of the QNP, the elided VP can find a suitable antecedent, as illustrated in (6).

(6)  
\begin{itemize}
  \item a. Lola jumped over [every frog that Dora did]
  \item b. [every frog that Dora did] Lola jumped over t (\textit{after QR})
  \item c. [every frog that Dora did [jump over t]] Lola [jumped over t]
\end{itemize}

\textit{(after QR \& VPE resolution)}

Thus, QR must apply in ACD environments, because if it did not, there would be no way to assign a meaning to the elided VP.\textsuperscript{4}

\footnotesize
\begin{itemize}
  \item \textsuperscript{3} Typically, in a framework assuming PF deletion (Chomsky and Lasnik (1993), Merchant (2000b), Tancredi (1992)), VP identity or parallelism is the problem; in a framework assuming LF copying (Fox (2002), Sag (1976), Williams (1977)), infinite regress is the problem. Either way, QR is the solution. Here we present a bare-bones LF copying approach for concreteness, making no theoretical commitments as to the exact nature of ellipsis resolution. See also Cecchetto and Percus (2006) for discussion of differences between PF-deletion in VPE and VP anaphora.
  
  \item \textsuperscript{4} The problem of antecedent containment is independent of the particular semantic framework adopted and requires special considerations in any theory, even, for example, in approaches whose key feature is direct compositionality (e.g., Jacobson (1992a; b), Steedman (1997)). We use essentially the framework of Heim and Kratzer (1998) so that we can address the acquisition questions that we are interested in with some formal vocabulary. None of the details of the experimental data or the learning problems associated with ACD depend on this framework.
\end{itemize}

\textsuperscript{3} \textsuperscript{4}
2.2. **The Landing Site of QR**

In the preceding discussion, we left the precise mechanics of QR rather unspecified, giving structures in which the landing site of QR is above the surface subject. However, it has been shown that the landing site of QR minimally (and possibly maximally) includes a position lower than the surface position of the subject. Evidence for this conclusion comes from sentences like (7) (Merchant (2000a)).

(7) John didn’t read every book that you did.

This sentence does not mean that John read none of the books that you read, which is the interpretation that corresponds to the universal quantifier *every* outscoping negation. Instead, this sentence means that John read some, but not all, of the books that you read, an interpretation that corresponds to the QNP being inside the scope of negation. If the ACD-resolving movement targeted a position higher than the surface subject, the QNP would then be outside the scope of negation. Since the QNP takes scope inside of negation, there must be a position under negation (e.g., the edge of vP) that serves as the landing site for QR.

A second argument for a clause-internal landing site for QR derives from the contrasts shown in (8a-c) (Fiengo and May (1994), Fox (1995)):

(8) a. You introduced him* to everyone that John, wanted to meet

Germaine to this discussion, we note that Fox (2002)’s account of ACD resolution, which assumes similar semantics, also involves movement of the QNP, although he proposes rightward movement of the QNP, followed by late insertion of the relative clause adjunct. Whether this or the account we assume here is correct, what is important to note is that to the extent that children are able to assign a correct interpretation to ACD sentences, their grammars must have a mechanism for moving the QNP.
b. You introduced him to everyone that John wanted you to

c. He introduced you to everyone that John wanted to

In the absence of ACD, as in (8a), we see a Principle C effect: the direct object pronoun him c-commands the referring expression John inside the indirect object, barring coreference between these two NPs. However, although the pronoun also c-commands into the indirect object at S-structure in (8b), coreference between the pronoun and the name in this example is possible. This can be taken as evidence that QR can bleed application of Principle C (since when the QNP raises, the referring expression antecedent is no longer in the c-command domain of the pronoun). The fact that in (8c) we again see a Principle C effect, suggests that the QNP must target a landing site that is lower than the surface subject. That is, the QNP [everyone that John wanted (you to)] moves covertly in both (b) and (c); however, because in both cases it winds up below the subject (but above the direct object) at LF, in (8b) the name is no longer c-commanded by the pronoun, while in (8c) it is. Thus, QR bleeds Principle C in (8b) but not (8c).5

2.3. Locality of QR

The movement required to resolve ACD is not restricted to a single landing site, however. The QNP may cross nonfinite clause boundaries, as illustrated in (9), generating either an embedded (local) reading (as in (9a)) or a matrix (long-distance) reading (as in (9b)) (cf. Kennedy (1997)).

(9) John wanted to read every book that Mary did.

John wanted to read every book that Mary …

5 See Merchant (2000a) for a discussion of cases involving QR and negative polarity items, which must be c-commanded by negation, and Merchant (2000b) for further discussion of the interaction between ACD and Principle C.
a. read
b. wanted to read

If QR were restricted to the first available landing site (the \( \nu P \) of the embedded clause), generating the matrix reading (9b) would be impossible, since the QNP would still be contained within the matrix VP antecedent. Targeting a landing site higher than the matrix VP allows the matrix reading to be generated. However, this movement appears to be bounded by tense. In expressions like (10), the matrix VP is unavailable as an antecedent (10b), and only the embedded reading is permissible (10a).

(10) John said that Mary read every book that Susie did.

        John said that Mary read every book that Susie …

a. read
b. \( \neq \) said that Mary read

Thus, while the covert QR operation can apply across infinitival clauses, tensed clauses create a barrier to this operation.\(^6\)

The locality properties of QR stand as a long-standing puzzle for the theory of movement (cf. Beghelli and Stowell (1997), Cecchetto (2004), Farkas (1981)). If QR is an instance of A-bar movement, as is now standardly assumed (e.g., Kennedy (1997), May (1985)), then we would expect it to show the locality properties of A-bar movement. However, QR seems to show stricter locality restrictions than other A-bar movement processes. Note, for example, that \( wh \)-movement, the canonical form of A-bar movement, is not bounded by tense (cf. (11)).

(11) What did John say that Mary read t ?

These additional locality restrictions raise an interesting acquisition question. Is

\(^6\) See Cecchetto (2004) and Fox (1999) for possible derivations of this clause boundedness.
boundedness by tense a direct consequence of the existence of a movement rule like QR? On the one hand, one might suppose that children would treat QR like any other A-bar movement process, and consequently allow it to apply across tensed-clause boundaries. If this were the case, then we would need to determine what developmental process would lead children to ultimately tighten the locality conditions on this A-bar movement process and not others. On the other hand, since the extra locality conditions associated with QR are unlikely to be extractable from the primary linguistic data that learners are exposed to, one might expect that these conditions are part and parcel of this operation itself. As soon as children are able to understand the relevant structures, they should show evidence of clause boundedness. In essence, the question is whether the locality properties of QR are an essential part of that operation, or whether they derive from an interaction between its A-bar movement properties and something else.

The problem of acquiring the clause-boundedness of scope is complicated by some data that point towards the possibility that QR is not actually clause-bounded as a matter of grammar. Some speakers report that under restricted conditions, for example when the main clause subject must be distributed over ((12), discussed by Farkas and Giannakidou (1996) and (13) M. Steedman (p.c.), or when a *de re* reading requires the phrase to take scope over the matrix verb ((14), discussed by Fiengo and May (1994) and Wilder (1997)), a quantificational expression appears to take scope out of a tensed clause (implicating that QR is potentially not clause-bounded):

(12) A (different) student made sure that every invited speaker had a ride.

= For every invited speaker, there was a student that made sure that speaker had a ride.

(13) A different witness testified that John robbed every bank in town.
John thinks that Mary is taller than Bill does.

= The degree to which John thinks Mary is tall exceeds the degree to which Bill thinks Mary is tall

These data introduce two possibilities with respect to the clause-boundedness of scope. First, these data may indicate that scope is not clause-bounded, and hence that the grammar of QR is no different from the grammar of other forms of A-bar movement. Apparent locality restrictions on scope would then need to be derived from factors external to the grammar (perhaps having to do with parsing routines, discourse processes or a combination of the two). A second alternative is that QR is clause-bounded but that these examples receive the interpretations they do by mechanisms other than QR. For the moment, we will not take a stand on these issues, but we note that children’s behavior may be informative. To the extent that children interpret ACD sentences like adults, we conclude that children have QR in their grammars. If, however, we find that children differ from adults in the interpretations accessed when ACD is embedded in a finite clause, then this may contribute to determining whether the locality properties of QR derive from grammatical or extragrammatical factors.

3. The Acquisition of ACD

We noted in the introduction that very little work on the acquisition of quantification directly addresses the question of whether children can apply QR (or its equivalent) appropriately. Most work on the acquisition of quantification examines children’s abilities in the areas of domain restriction or in the scope of a QNP relative to negation. Neither line of research, however, directly addresses whether children can apply QR. To our knowledge, the only other study that directly addresses this question is Kiguchi and Thornton (2004).

Kiguchi and Thornton (2004) used the interaction between ACD and the binding
principles (Chomsky (1981), Reinhart (1976)), as illustrated in (8), to determine whether children correctly apply QR and target the appropriate landing site for this operation. The authors showed that children, like adults, consistently reject coreference in sentences such as (15).

(15) a. *Darth Vader found her; the same kind of treasure that the Mermaid; did.
   b. *[the same kind of treasure that the Mermaid; did find her; t] Darth Vader found her; t

To identify whether the source of this response pattern was due to a Principle C violation at S-structure or to a Principle B violation at LF, these authors showed that in sentences such as (16), four-year-olds, who typically obey Principle C (Crain and McKee (1985), Crain and Thornton (1998), Kazanina and Phillips (2001), Leddon and Lidz (2006)), allow coreference between a VP-internal pronoun and a name that it c-commands on the surface. Crucially, the lack of Principle C effects occurs only when the name is contained within a constituent that undergoes QR at LF and is consequently not c-commanded by the pronoun at that level of representation.

(16) Dora gave him; the same color paint the Smurf’s father did

Here, the only way to avert the violation of Principle C that would obtain at S-Structure is to QR the QNP the same color paint the Smurf’s father did so that at LF (after QR), the NP the Smurf is no longer in the c-command domain of the pronoun him. Unlike in (15), the name (here, in the possessor position) does not c-command the pronoun at LF. This derivation is illustrated in (17):

(17) a. Dora gave him; [the same color paint the Smurf’s father did]
   b. [the same color paint the Smurf’s father did] Dora gave him; t
   c. [the same color paint the Smurf’s father did [give him; t]] Dora [gave him; t]

The lack of Principle C effects here argues for children’s ability to apply QR. Similarly, these authors argued, children’s responses in (15) must derive from an LF Principle B violation and not from an S-structure Principle C violation, which entails that children are able to apply QR in
order to resolve ACD.

Furthermore, the authors showed that while children’s grammars allow QR to target a VP-external landing site, this movement is restricted to a position lower than the subject. Hence, children allow coreference in (14a) but reject it in (15a), where there is a Principle C violation at LF.

(18) a. *He jumped over every fence that Kermit tried to.

b. *He [every fence that Kermit tried to jump over] jumped over it

Now, while the findings reported by Kiguchi and Thornton (2004) are highly suggestive that children employ QR just like adults do, they do not provide unambiguous evidence that this is the case. Children did interpret these sentences in a way that matches the adult interpretation; however it is possible that this result obtained for a reason other than the children correctly applying QR. Consider the possibility that children lack QR. When faced with a sentence like (14a), children would need to find a way to resolve VPE which did not involve QR, since, by hypothesis, this operation would be unavailable to them. What could they do instead?

One possibility is that children would treat the relative clause as though it were a coordinate clause. Errors of this kind have been reported, for example, when children have been asked to interpret a relative clause in a context that does not meet the felicity conditions on the use of a relative clause (Hamburger and Crain (1982), Tavakolian (1981)). If the problems associated with antecedent containment were not resolvable by QR, children faced with ACD would still need to find some way to resolve VPE. It is possible that here, too, they might turn to

---

7 For example, a relative clause helps to distinguish between multiple members of a contextually relevant set of objects. If there is only one entity in question, the use of a relative clause violates the Gricean Maxim of Quantity (Grice (1989)) and may not seem warranted.
coordination as an alternative structure. Thus, the child might represent (14a) with a structure along the lines of (19):

(19)  \[IP \text{Dora gave him the same color paint]} \text{and} [IP \text{the Smurf’s father did}]

If children adopted this strategy, then the pronoun would also not c-command the name that is it coreferential with and no Principle C violation would result. Thus, a child lacking QR could assign the adult-like interpretation to this sentence without having the adult-like representation.⁸

Now let’s consider children’s interpretations of sentences like (18). Here, children disallowed coreference between the pronoun and the name. This result was argued to derive from children correctly applying QR to resolve the antecedent containment and targeting only a position below the subject as the landing site for QR. But suppose again that children treated relative clauses in ACD constructions as coordinate clauses. This strategy would yield the representation in (20).

(20)  \[He \text{jumped over every fence}] \text{and} [\text{Kermit tried to}]

While there is no Principle C violation in this configuration, the coreference interpretation in which \textit{he} refers to Kermit is highly dispreferred, if not impossible (cf. Evans (1980), Lakoff (1968), Langacker (1969)). Moreover, in the particular contexts presented to children by Kiguchi and Thornton (2004), the children could have rejected the sentence only on the basis of the first clause, provided that they assigned \textit{he} a referent other than Kermit. In this story, only Kermit

⁸ This same line of reasoning would yield the asymmetry between (12) and (14) that Kiguchi and Thornton observed. If children interpreted (12) as a coordinate structure, as in (ii), then upon resolving the VPE, as in (iii), a principle B violation would be incurred:

(ii)  \[\text{Darth Vader found her the same kind of treasure}] \text{and} [\text{the Mermaid did}]

(iii)\[\text{Darth Vader found her the same kind of treasure}] \text{and} [\text{the Mermaid did find her...}]
jumped over each of the three fences, while Cookie Monster, his competitor, stopped after two. Thus, it remains a live possibility that children succeeded on this task for the wrong reason. Because the ACD sentences were not compared with corresponding coordinate structures, we cannot be sure that children were not assigning the coordinate structure interpretation to these sentences. We therefore consider it an open question whether preschoolers’ interpretations of sentences exhibiting ACD are adult-like because they have adult-like representations.

In the remainder of this paper, we report on four experiments concerning the core properties of ACD. In Experiment 1 we show that children have adult-like interpretations of sentences exhibiting ACD, and that these interpretations are a consequence of children’s adult-like representations of ACD. We therefore conclude, in agreement with Kiguchi and Thornton (2004), that children resolve ACD appropriately because they correctly apply the covert QR operation. In Experiment 2, we show that when presented with ambiguous sentences involving VPE without QR, children and adults are able to access either of two antecedent VPs, depending on which one is made true in the experimental scenario. With these results as a backdrop, we turn to two experiments testing participants’ interpretation of ambiguous ACD sentences, now involving both VPE and QR.

In Experiment 3, we show that both children’s and adults’ QR can cross a nonfinite clause boundary. When quantificational expression containing an elided VP is contained within an embedded nonfinite clause, as in (9), repeated here, both children and adults treat both the matrix VP and the embedded VP as available antecedents for the elided VP.

(9) John wanted to read every book that Mary did.

This result enables us to ask whether, by contrast, a tensed clause creates an opaque domain for QR, as has been claimed in the literature. The results of Experiment 4 show that children do not
treat the tensed clause as an opaque domain, but adults do. Given the findings from Experiment 2, we argue that children’s responses to the ambiguous ACD sentences in Experiments 3 and 4 should be attributed to features of the QR operation in their grammar, and not to general VPE resolution strategies. We conclude with a discussion of the possible source of children’s overgeneration of interpretations in the domain of QR and what development to the adult state entails. We suggest that the difference between the age groups may be due to a difference in parsing, rather than the nature of the grammar. Finally, we argue that contra claims by Grodzinsky and Reinhart (1993) and Reinhart (2004), children are able to compare the multiple derivations of ambiguous sentences.

4. Experiments

4.1. Experiment 1

The purpose of this experiment was to determine whether children not only assign the correct interpretation to sentences containing ACD, but also whether they assign the correct representation to such structures.

4.1.1. Method

4.1.1.1. Participants

40 four-year-olds (23 M 17 F, M 4;6, range 4;0-4;11) and 40 adults (undergraduate students at Northwestern University completing an experimental requirement) participated.

4.1.1.2. Materials

Participants were presented with sentences similar to either (21a) (ACD) or (21b) (Coordinated Conjunction, CC), depending on the experimental condition to which they were assigned.

(21) a. Miss Red jumped over every frog that Miss Black did.
b. Miss Red jumped over every frog and Miss Black did, too.

Both sentence types have an instance of VPE, signaled by *did*, but the syntactic structure differs; while the elided VP in (21a) is contained within a relative clause, in (21b) the elided VP is inside a conjoined clause. Each participant was presented with four test sentences and three filler sentences, pseudorandomized. Sentence order was counterbalanced across subjects. Filler sentences for the ACD test sentences contained a relative clause without VPE, as in (22a). Fillers for the CC test sentences contained conjunction without VPE, as in (22b).

(22) a. The rhino [made/didn’t make] friends with the hippo that kicked the rock into the water.

b. The rhino picked up the rock and [fell in the water/jumped over the hippos].

The variations in the brackets allowed the experimenter to elicit variable responses based on the participant’s response to the test sentences, so that a bias towards *yes* or *no* responses was avoided. The full set of test and filler sentences is included in Appendix A.

4.1.1.3. Procedure

The procedure used in all four experiments was the Truth Value Judgment Task (TVJT) (Crain and McKee (1985), Crain and Thornton (1998)). In this task, one experimenter tells the child a story using toys and props, while a puppet (played by a second experimenter) watches the story alongside the child. The puppet watches very carefully, and at the end of the story, the puppet says what he thinks happened in the story. The puppet’s statement includes the target construction. The child’s job is to assess the validity of the puppet’s statement with respect to the events in the story. That is, the child must say if the puppet is right or wrong. In our tasks, if the puppet is right, he gets a cupcake, and if he is wrong, he gets a cookie; either way, he receives a sweet, but the child is told that the puppet likes cupcakes more. The child is always encouraged
to tell the puppet why he was right or wrong so that the puppet can learn more.

Our experiment contained three between-subjects factors with two levels each. The first factor was sentence type, comparing coordinate and relative clauses (i.e., CC v. ACD sentences). The second factor was context, comparing contexts that made the coordinate clause structure true and the relative clause structure false against contexts that made the coordinate clause structure false and the relative clause structure true. The third factor was age, comparing children and adults.

Recall from above that if children do not have the QR operation at their disposal, a plausible alternative interpretation for an ACD structure might be Coordinated Conjunction. For this reason, we compared children’s interpretations of sentences such as (18a) and (18b) above. To tease apart these two interpretations, we manipulated the contexts against which the sentences were to be evaluated. This design is illustrated in Figure 1.

Figure 1: Experimental conditions in Experiment 1

```
<table>
<thead>
<tr>
<th>‘One Set’ Condition</th>
<th>‘Two Set’ Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miss Black jumped over</td>
<td></td>
</tr>
<tr>
<td>Miss Red jumped over</td>
<td></td>
</tr>
</tbody>
</table>
```

A description of a typical experimental scenario will make this design clear. In this story, Miss Red and Miss Black are contestants in a frog jumping contest.
**One set condition**

Miss Black and Miss Red are introduced to the four frogs they must jump over in order to win a prize. Miss Black begins to jump first. She jumps over the first frog without any difficulty, and jumps over the second easily but is beginning to get tired. She pauses to wonder whether she will be able to jump over the third frog. She gives it a try and manages to jump over it. She considers jumping over the fourth frog and thinks she might be able to manage it, but reconsiders when she realizes just how tired she is, and stops jumping after the third frog. It is now Miss Red’s turn. She also jumps over the first three frogs, also with increasing difficulty each time. After the third frog, she also comes to the same conclusion that jumping over the fourth frog would be too much, and that three frogs is good enough. Neither girls wins. It is important that in this condition neither contestant jumps over the fourth frog. If Miss Red jumped over all four frogs, while the ACD sentence would still be true, children might focus on the fourth frog, which Miss Black didn’t jump over, when answering. If instead Miss Black jumped over all four frogs, the ACD sentences would be false, as would the CC sentence, so there would be no way to distinguish the two interpretations in children’s rejection of the target sentence.

**Two set condition**

In contrast to the ‘one set’ condition, in the ‘two set’ condition, each contestant is assigned a separate set of frogs to jump over. The rules are therefore slightly different: each contestant must jump over all of her respective frogs in order to win a prize. Miss Black is an experienced frog-jumper, and so is assigned four frogs; Miss Red is only assigned three. Miss Black jumps over her frogs first, and does so very easily. Miss Red then jumps over all of her frogs. In the end, each contestant receives a prize.

There are two important features of the ‘two set’ condition. First, we assigned more frogs
to one contestant (Miss Black) in order to avoid functional readings of the quantifier in the ACD context. That is, if both contestants jumped over each of the three frogs in their set, then it would be appropriate to say “Miss Red jumped over every frog that Miss Black did” on the interpretation that the two participants both jumped over frogs 1, 2 and 3. But, if both contestants jumped over each of their respective frogs, but their sets were of different sizes, then this would make the functional reading infelicitous. This situation is desirable. Second, because the rules specified that each contestant was supposed to jump over her own set of frogs, we expected that rejections of the CC sentences (“Miss Red jumped over every frog and Miss Black did, too”) on the basis that neither jumped over all seven frogs would be minimized. In other words, it was clear from the context that the interpretation of the QNP every frog was restricted to different sets for the different contestants and did not apply to all of the frogs in the experimental workspace.

Subjects were randomly assigned to either the ‘one set’ or the ‘two set’ condition and heard either ACD sentences or CC sentences. At the end of each story, the puppet delivered the test sentence. The truth conditions depend on the set condition, as illustrated in Table 1.

Table 1: Truth conditions in Experiment 1

<table>
<thead>
<tr>
<th>syntactic structure</th>
<th>‘one set’ condition</th>
<th>‘two set’ condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACD</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>Coordinated Conjunction</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Since in the ‘one set’ condition, the two contestants jumped over the same frogs, but did not jump over the fourth frog, the ACD sentence (18a) is true and the CC sentence (18b) false. Since in the ‘two set’ condition, both contestants jumped over all of their respective frogs, but did not jump over the same frogs, the opposite truth conditions hold.
4.1.2. Results

We predicted that if children assign the correct representation to the ACD sentences, their responses to these test sentences should be in inverse relation to the CC sentences, just like adults. In the ‘one set’ condition, participants should accept the ACD sentences and reject the CC sentences; in the ‘two set’ condition, participants should reject the ACD sentences and accept the CC sentences. The percentage of acceptance of the sentences for each of the two age groups and each of the two conditions is shown in Figure 2 (ACD sentences) and Figure 3 (CC sentences). Adult and child responses are in the same direction in both figures. Participants accepted the ACD test sentences in the ‘one set’ condition and rejected them in the ‘two set’ condition. Conversely, they rejected the CC test sentences in the ‘one set’ condition and accepted them in the ‘two set’ condition. Representative examples of children’s justifications for their responses are presented in Appendix B.

Figure 2: Percentage acceptance for ACD sentences in Experiment 1
A 2 x 2 x 2 ANOVA for independent samples (age x sentence type x context) was conducted. We found main effects of age ($F(1, 72) = 5.17, p = 0.026$), sentence type ($F(1, 72) = 8.83, p = 0.004$), and context ($F(1, 72) = 4.17, p = 0.045$) with a highly significant interaction between sentence type x context ($F(1, 72) = 193.83, p < 0.0001$) and an interaction between age x context ($F(1, 72) = 5.17, p = 0.026$). There was no significant interaction between age x sentence type ($F(1, 72) = 1.83, p = 0.180$) and no overall interaction ($F(1, 72) = 0.17, p = 0.681$). Probing these effects further, we conducted 2 x 2 ANOVAs (age x context) for each sentence type. As predicted there was a main effect of context for each sentence type (ACD: $F(1, 19) = 273.03, p < .001$; CC: $F(1, 19) = 49.17, p < .001$) resulting from the fact that participants were likely to accept the ACD sentences, but reject the CC sentences in the ‘one-set’ condition, but reject the ACD sentences, and accept the CC sentences in the ‘two-set’ condition.

While age was not a significant factor for the ACD sentences ($F(1, 19) = 1.07, p = 0.307$), there was a significant interaction between age and context ($F(1, 19) = 5.57, p = 0.024$), deriving from the fact that adults were perfect on this task, giving completely uniform responses...
in each condition. Consequently, while child and adult answers are qualitatively similar, the small amount of variability in the child data makes the children and adults statistically distinct. There was a main effect of age for the CC sentences \((F (1, 19) = 4.66, p = .037)\), driven by adults’ relatively low percentage of acceptance in the ‘two set’ condition (52.5%). Here, it seems that our attempt to restrict the domain of quantification uniquely for each contestant was successful for children, but was less successful for adults. Adults sometimes interpreted every *frog* as referring to every frog in the entire context, not every contextually appropriate frog. This is likely due to some adults over-thinking the task. The interaction was not significant \((F (1, 19) = 1.82, p = 0.186)\).

4.1.3. Discussion

The results of this experiment provide evidence that four-year-olds assign the correct interpretation to ACD sentences and moreover, that they do so because they have the correct representations for these structures. That they treat these sentences differently from those involving coordination indicates that they must be appealing to a covert operation to remove the ellipsis site from its antecedent and resolve the VPE. These results therefore support the conclusion of Kiguchi and Thornton (2004) (and Lidz *et al.* (2004)) that QR is part of children’s grammars.

Before turning to an investigation of children’s interpretation of more complex ACD sentences, we first seek to establish a baseline for children’s resolution of ambiguous VPE sentences. Determining whether children have a bias towards targeting one antecedent over another when resolving VPE allows us to evaluate the locality properties of their QR operation by determining whether the introduction of QR pulls them away from this baseline preference.
4.2. **Experiment 2**

The purpose of Experiment 2 was to provide a baseline in children’s resolution of ambiguous VPE sentences. By presenting participants with sentences that contain an ellipsis site in an embedded clause, and which are therefore ambiguous, we can identify whether children exhibit a bias in the clause they target for VP interpretation. This information will in turn be useful for evaluating the effects of QR in the test sentences used in the experiments that follow this one.

4.2.1. **Method**

4.2.1.1. **Participants**

24 four-year-olds (12 M 12 F, M 4;6, range 3;8-4;10) and 24 Northwestern University undergraduate students participated.

4.2.1.2. **Materials**

Participants were presented with sentences such as (23).

(23) Clifford asked Goofy to read the [big/small] books because Scooby did.

Clifford asked Goofy to read the [big/small] books because Scooby …

a. *read the [big/small] books*

b. *asked Goofy to read the [big/small] books*

In these sentences, there is an ellipsis site within an adjunct clause (signaled by *did*), allowing the VP to be interpreted as identical to either the embedded VP (23)a) or the matrix VP (23)b). (See Section 2.3 for discussion of embedded and matrix readings.) Subjects were randomly assigned to one of two conditions, based on which of these two readings was true in the context. Each participant was presented with four test sentences which made either the embedded or matrix reading true and three fillers in a pseudorandomized order. The filler sentences also contained
VPE, but were unambiguous. Sentence order was counterbalanced across subjects. The full set of sentences is included in Appendix A.\(^9\)

4.2.1.3. Procedure

As in the previous experiment, the method used was the TVJT. We introduced a slight change in the procedure to justify the presence of the *because* clause. After the story, and before the puppet delivered his statement, the experimenter asked the puppet a question, as in (24).

(24) Why did Clifford ask Goofy to read the [big/small] books?

The puppet then answered this question with the test sentence, as in (23).

A typical scenario begins with three dogs, Clifford Goofy, and Scooby, standing in front of two sets of books, big ones and small ones. Scooby asks Goofy to read the big books, reasoning that since they are big, they must contain some very impressive big words. Goofy agrees to read the books, but before he does, Scooby decides to read the small books, which are appropriate for his reading level. When Scooby is done reading, he tells Goofy he is ready for Goofy to read the big books. However, at this point Goofy acknowledges that Clifford has been waiting patiently and asks Clifford what he would like. It is at this point that the two experimental conditions diverge.

In the ‘embedded’ condition, Clifford asks Goofy to read the small books. He explains that if Scooby likes them, they must be good. In the ‘matrix’ condition, Clifford asks Goofy to read the big books.

\(^9\) As we make clear in our description of the experimental scenarios, the choice of the lexical item in the bracketed material in (23) and (24) and all other test sentences in Appendix A depended on the experimental condition. In this example, in the ‘embedded’ condition, Clifford asks Goofy to read the small books, whereas in the ‘matrix’ condition, Clifford asks Goofy to read the big books.
read the *big* books. He notes that Scooby always has good ideas, so the big books must indeed be impressive. If Scooby asked Goofy to read these, then Clifford wants Goofy to read these, too.

The truth conditions for the two experimental conditions are presented in Table 2.

Table 2: Truth conditions in Experiment 2

<table>
<thead>
<tr>
<th>VP interpretation</th>
<th>embedded condition</th>
<th>matrix condition</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>read the [big/small] books</em></td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td><em>asked Goofy to read the [big/small] books</em></td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

4.2.2. Results

Participants’ responses were analyzed in terms of the percentage of matrix readings (rather than percentage of acceptance), as interpreted from their acceptance or rejection of the puppet’s statement in each condition and supported by their justifications, as captured in Appendix B.\(^\text{10}\) For example, the child in (45) was assigned to the embedded condition and

\(^\text{10}\) Given the ambiguity of these VPE sentences, we are interested in whether participants access the reading that is made true in the experimental condition to which they have been assigned, rather than the other (competing) reading, which is also grammatical. For example, we expect that in the matrix condition, the rate at which participants access the matrix reading should be significantly higher than the rate at which they access the embedded reading, and the rate at which they access the matrix reading in the embedded condition. Capturing the results in terms of the percentage of *matrix readings*, rather than the percentage of *acceptance* in each condition thus allows us to make a clearer comparison between the conditions in Experiments 2-4. The choice of the matrix reading was motivated by the issue of the locality of QR, as discussed in § 2.3. Of course, the relevant comparisons are equivalent if we treat the dependent measure as the
accepts the puppet’s statement *Clifford asked Goofy to read the big books because Scooby did*, offering in the justification the details that Clifford asked Goofy to read the big books because Scooby didn’t know how to read the big books. (This is, in fact, the reason Scooby provides in the story for asking Goofy to read the big books to him, and the reason why the big books are intriguing to Clifford in the scenario.) These percentages are presented in Figure 4.

Figure 4: Percentage matrix readings in Experiment 2

A 2 x 2 ANOVA (age x condition) revealed a main effect of condition (F (1, 23) = 76.04, p < 0.001) and a significant interaction between age and condition (F (1, 23) = 7.26, p < 0.01). Although for both age groups, the percentage of matrix readings is higher in the condition that favors this reading, there is a bit more variability in the children’s data, driving the interaction. We found no reliable effect of age (F (1, 23) = 0.58, p = 0.45). Representative examples of children’s responses are presented in Appendix B.

4.2.3. *Discussion*

The motivation for this experiment was to establish a baseline that would allow us to percentage of embedded readings, or even percentage of acceptance.
determine the effects of QR in the subsequent experiments. Specifically, we sought to determine whether children have a bias towards targeting one VP antecedent over another in ambiguous VPE sentences so that when we present children with ambiguous sentences involving both VPE and QR, we can evaluate what the effects of QR are. We found that when presented with such sentences, both children and adults accessed the interpretation that was made true by the experimental scenario and did not display a strong preference for targeting one VP antecedent or another. If in Experiments 3 and 4 we find a pattern of results that diverges from this one, then we can say with certainty that it is the introduction of QR that is pulling participants away from this baseline preference. For example, if participants do not access the matrix reading as robustly in subsequent test sentences, we have evidence about the locality of QR, since participants were able to access the matrix readings without any difficulty in Experiment 2, and the only difference is that in Experiments 3 and 4, accessing the matrix reading depends on children’s ability to QR the QNP to a higher landing site.

With this in mind, we are now in a position to probe children’s QR operation further. Recall that Kiguchi and Thornton (2004) showed that children target only a vP-level landing site for QR, and not one above the subject. However, in cases where QR is contained inside an embedded nonfinite clause, QR should be able to target the edge of either the embedded or the matrix vP. The authors did not evaluate children’s responses to such ambiguous sentences. The question at hand now is, do children allow the QNP to be QRed out of an embedded clause to the matrix vP, or are they restricted to only the most local landing site?

4.3. Experiment 3

The purpose of this experiment was to determine whether children can access multiple interpretations of ambiguous ACD sentences, thereby indicating that QR in the child grammar
can target multiple landing sites. Recall from Section 2.3 that participants’ ability to access the matrix reading of such sentences depends on their ability to QR to a position outside of the matrix VP, thereby targeting a position higher than the most local one (the vP of the embedded clause).

### 4.3.1. Method

#### 4.3.1.1. Participants

24 four-year-olds (12 M 12 F, M 4;6, range 4;1-4;10) and 30 Northwestern University undergraduate students participated.

#### 4.3.1.2. Materials

Participants were presented with sentences such as (25a) and (25b).

(25)  
   a. Miss Piggy wanted to drive every car that Kermit did.  
   b. Clifford asked Goofy to read every book that Scooby did.

As in the previous experiments, participants were presented with four test sentences and three filler sentences pseudorandomized and counterbalanced across subjects. Two of the test sentences involved subject control, like (22a) with the infinitival subject interpreted as bound by the matrix subject, and two involved object control, like (22b) with the infinitival subject interpreted as bound by the matrix object.\(^{11}\) The filler sentences were the same as in Experiment \(^{11}\) The subject control verbs used were want and need. Because these restructuring verbs have been claimed to be able to unify the domains of the matrix and embedded verb for various word order phenomena in languages such as Spanish, Italian and German (e.g., Aissen and Perlmutter (1983), Rizzi (1978), Wurmbrand (2000)) and affect the range of readings available for sentences such as those in question (e.g., Bruening (2001), Larson and May (1990)), one might expect these contexts to force QR to a higher landing site. The results of this experiment show that for both children and adults, the two verbs are not treated as an inseparable complex.
2. The full set of sentences is given in Appendix A.

4.3.1.3. Procedure

As in the two previous experiments, the procedure was the TVJT. Recall from our example in (9) that sentences in which ACD is embedded in an infinitival clause are ambiguous. Thus, in this experiment, only one of the two interpretations (embedded or matrix) was made true in the experimental scenario. As in Experiment 2, subjects were randomly assigned to one of two conditions, based on which of these two readings was true in the context. The description of a typical experimental scenario follows.

In this story, the two main characters are Kermit and Miss Piggy. Kermit has two sets of cars, a red set and a black set. Miss Piggy is very interested in them and wants to hear more about them. Kermit explains that he has had the red cars for a long time and has driven them so many times that he is now very tired of driving them. The black ones, however, are new; he just received them for his birthday. He is not allowed to drive them yet, but he is very excited to get the chance to drive them, since he thinks they must go very fast. Miss Piggy is very excited about Kermit’s cars and asks to see him drive some. Kermit agrees to do so (albeit reluctantly, since this means he is stuck driving the red ones). After Kermit drives each of the red cars, Miss Piggy is so excited that Kermit decides to let her drive some of his cars and as a bonus allows her to choose which she will drive. Now, it is at this point that the two experimental conditions diverge.

In the ‘embedded’ condition, Miss Piggy says that despite what Kermit thinks about the red cars, she really liked them when he drove them, and she wants to drive those. In the ‘matrix’ condition, Miss Piggy says that she agrees with Kermit that the red cars are not so great and can see why he is so excited about the black cars, so she wants to drive those. Importantly, once she chooses which cars to drive, the story ends. She never drives any cars. By ending the story here,
we avert potential biases towards children responding based on what actually happened (as opposed to what the characters wanted). The difference between these two conditions is illustrated in Figure 5.

Figure 5: Experimental conditions in Experiment 3

This design allows the truth conditions to vary in the two contexts, based on the interpretation of the elided VP, as is captured in Table 3. In the ‘embedded’ condition, Miss Piggy wants to drive every car that Kermit *drives*, not the ones that he wants to drive; in the ‘matrix’ condition, she wants to drive every car that he *wants to drive*, not the ones he drives.

Table 3: Truth conditions in Experiment 3

<table>
<thead>
<tr>
<th>VP interpretation</th>
<th>embedded condition</th>
<th>matrix condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>drove</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>wanted to drive</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

4.3.2. Results

Recall that the possible interpretations of the test sentences covary with the possible
landing sites for QR. If children are able to access both landing sites via application of QR, they should be able to access both interpretations of these sentences.\textsuperscript{12} Because we are most interested in determining if children will be able to access the matrix reading, since this indicates children’s ability to perform long QR, we analyzed the results in terms of the percentage of matrix readings (as in the previous experiment).\textsuperscript{13} These percentages are presented in Figure 6. A 2 x 2 ANOVA

\textsuperscript{12} This is somewhat of an oversimplification. Accessing the highest landing site allows participants to access both interpretations. This is because the highest landing site is contained in neither VP, so either VP is available as the antecedent for the elided VP contained in the QRed expression. If children access both interpretations, it is at least clear that they are not restricted to the lower landing site. T. Roeper (p.c.) has suggested that it may be difficult for children to target the lower site, since it is more deeply embedded in the syntactic structure. Given children’s preference for high adjunct attachment elsewhere (e.g., Trueswell \textit{et al.} (1999)) and a possible processing constraint in which children ‘hold onto’ the first verb while parsing the sentence, a preference for targeting the higher landing site would not be surprising. However, the results demonstrate that children are, in fact, able to access the embedded reading, and so, are able to access the lower VP as an antecedent, either because the lower vP is a target for QR or because this lower antecedent is visible from the higher landing site. Interestingly, the results may suggest that for adults, it is the embedded reading that is the most easily accessed.

\textsuperscript{13} As in Experiment 2, we obtained the percentage of matrix readings from the acceptance or rejection of the puppet’s statement in each condition, and gained additional support for these interpretations from participants’ justifications for their responses, as reflected in Appendix B. It is important to recognize that a child’s yes/no judgment is only an indirect indication of which reading was accessed, whereas the justification for the response provides firmer grounds for
(age x condition) revealed a main effect of age (F (1, 23) = 4.66, p = 0.037) and a significant interaction of age and condition (F (1, 23) = 4.44, p = 0.04). There was no main effect of condition (F (1, 23) = 0.11, p = 0.742).

Figure 6: Percentage matrix readings in Experiment 3

4.3.3. Discussion

The results of Experiment 3 are somewhat complicated and require close inspection. To determining which interpretation the child accessed. For example, one child assigned to the ‘embedded’ condition, whose justification is provided in (49), accepted the puppet’s statement, explaining that Kermit drove the red cars, and that these were the cars that Miss Piggy wanted to drive (thereby indicating that the embedded reading was accessed, and in this case is true). By contrast, another child assigned to the ‘matrix’ condition also accessed the embedded reading (the reading that was false in the experimental condition to which the child was assigned). In this case, the child rejected the puppet’s statement, offering as support for this response the reasoning that Miss Piggy wanted to drive the cars that Kermit did not drive (again, indicating that the embedded reading was accessed, and in this case was false).
begin, neither adults nor children exhibited categorical responses. That is, in neither condition and for neither age group, do we find proportions in the 0% or 100% range. While the TVJ method generally assumes that features of the story contexts can bias participants towards accessing one interpretation over the other on the basis of which interpretation is made true by the context, it appears that other factors may have been at work (to be described momentarily) in determining interpretive preferences in our materials. However, the fact that responses were not categorical does not make them uninterpretable and certainly does not diminish their importance. Recall that the motivation for this experiment was to determine whether children can access multiple interpretations of ambiguous ACD sentences (and specifically, whether they could access a reading generated by long distance QR), and thus determine whether QR in the child grammar can target multiple landing sites. The results answer in the affirmative (in neither condition did the percentage of matrix, or for that matter embedded, readings hover around 0%), and therefore shed new light on the nature of the child grammar.

Take as an example the child whose justification for a response to one of the test sentences is captured in (51) of Appendix B (and repeated below). This child was assigned to the ‘matrix’ condition, the condition in which the scenarios made the ‘matrix’ readings true. In this specific example, the child heard a story about a frog-jumping contest, in which a cowgirl and a cowboy need to jump over frogs. The older and more experienced cowboy demonstrates how to jump over frogs with a set of small frogs, but is ultimately assigned to jump over the bigger ones, given his expertise. After this demonstration, the cowgirl is confident in her skills, and so is also assigned to the bigger frogs. Thus, both participants need to jump over the same set of frogs, although the cowboy has already jumped over the small frogs. At the end of the story, the puppet uttered the sentence in (26).
(26) The cowgirl needed to jump over every frog that the old cowboy did.

This child accepted this statement, offering the justification in (27), appealing to the fact that both participants both needed to jump over the big frogs, a response which has the force of the matrix reading \textit{The cowgirl needed to jump over every frog that the old cowboy needed to jump over}.

(27) The old cowboy and the little cowgirl needed to jump over the big frogs.

Note, though, that accessing the matrix reading is not just a matter of agreeing with the puppet and accessing the reading made true by the experimental scenario. Take as another example a different child, who was assigned to the ‘embedded’ condition, the condition in which the scenarios made the ‘embedded’ readings true. When the puppet uttered the sentence in (28), this child \textit{rejected} the statement, offering the justification in (repeated from (52) of Appendix B).

(28) They needed to jump over different frogs.

This response indicates that the child was accessing the matrix reading \textit{The cowgirl needed to jump over every frog that the old cowboy needed to jump over} and not the embedded reading \textit{The cowgirl needed to jump over every frog that the old cowboy jumped over}. In the ‘embedded’ condition, however, the cowgirl must jump over the small frogs, the ones that the cowboy \textit{did} jump over, and not the big ones, which he needs to jump over (and never does). Thus, while it is widely assumed that children’s responses in a TVJT are determined simply by the truth or falsity of the test sentence given the context, this assumption appears not to be met in this experiment (for similar results, see Musolino and Lidz (2006), Conroy \textit{et al.} (submitted)). What is crucial here, however, is that in their disagreements with the puppet, children are demonstrating their ability to access relevant interpretations in the context.

Likewise, the fact that adults and children differed in their responses does not provide
evidence that their grammars differ. Just the opposite. Both children and adults show evidence of accessing both readings; they differ only in the conditions under which they access them. Consider first the adults. In the ‘matrix’ condition, they provide only 50% matrix interpretations, while in the ‘embedded’ condition, they give even fewer matrix responses (32%, one-tailed t-test: \( t (14) = -1.798, p = 0.047 \)). But as the interaction between age and condition shows, children displayed the opposite pattern. They provided only 38% matrix responses in the ‘matrix’ condition but a numerically higher amount (54%) in the ‘embedded’ condition (one-tailed t-test: \( t (11) = 1.036, p = 0.161 \)). Thus, whereas adults in this experiment were more likely to access a given reading when it was made true, children were more likely to access a given reading when it was made false by the experimental conditions. Such differences are clearly more a matter of extragrammatical factors influencing which interpretation is accessed than they are the grammar’s ability to generate the interpretations in the first place.

Since this pattern has not been reported in previous literature, we gave the data even closer scrutiny. We began with the justifications that each participant provided when accepting or rejecting the puppet’s statement, as captured in Appendix B. In both conditions, children provided the reading aligning with the one made true in the experimental scenario to justify the truth of the puppet’s statement (e.g., (49)-(51)) and the other reading to justify the falsity of the puppet’s statement (e.g., (52)-(56)). For each opportunity a participant had to respond to a test sentence (four in an experimental session for each of the 24 children and 30 adults), we recorded whether a response was given, whether that response was relevant to the story\(^\text{14}\), and finally, whether the response was reliable (i.e., whether one of the two grammatical readings could be inferred from it). The distribution of responses is presented in Table 4.

\(^{14}\) Only five responses were irrelevant: one from a child and four from adults.
Table 4: Distribution of all responses in Experiment 3

<table>
<thead>
<tr>
<th>response type</th>
<th>children</th>
<th>adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>given</td>
<td>79.2% (76/96)</td>
<td>100.0% (120/120)</td>
</tr>
<tr>
<td>relevant</td>
<td>98.7% (75/76)</td>
<td>96.7% (116/120)</td>
</tr>
<tr>
<td>reliable</td>
<td>68.4% (52/76)</td>
<td>80.0% (96/120)</td>
</tr>
</tbody>
</table>

These percentages show that children not only consistently provided justifications for accepting or rejecting the puppet’s statement, but also that the vast majority of these responses (close to 99%) were directly relevant to the plot of the story and most (68%) allowed one of the two readings to be clearly inferred (as illustrated in Appendix B). A further analysis of all of the reliable responses reveals that approximately 54% (28 of 52) of children’s responses and 61% (59 of 96) of adults’ responses corresponded to an embedded reading (16 children and 28 adults), while approximately 46% (24 of 52) of children’s responses and 39% (37 of 96) of adults’ responses corresponded to a matrix reading (13 children and 23 adults). Note that the matrix reading is being accessed in nearly half of the reliable cases and by a little over half of the children, regardless of the condition they were assigned to in the experiment. In addition, seven children provided justifications corresponding to both the embedded and matrix readings within an experimental session. Representative examples of children’s justifications for children’s responses are presented in Appendix B.

Once we filter out the unreliable responses, we are left with a subset of responses that closely resembles the pattern of the original responses presented in Figure 6. Though adult responses remain in the mid-range of the percentage of matrix readings, they are still in the general direction we would expect if participants were accessing the reading that was made true in the experimental scenario (‘embedded’ condition 23%, ‘matrix’ condition 56%). This
asymmetry reveals two things. First, it reveals that adults have a bias, presumably a consequence of the relative weightings assigned to the two structures compatible with the string, in favor of the embedded reading. Second, it reveals that this bias can be overcome as a function of discourse context. When the matrix reading is true, adults are more likely to access that reading.

Children’s responses still display the opposite pattern (‘embedded’ condition 64%, ‘matrix’ condition 31%). They are more likely to access the matrix reading when it is made false in the experimental context. It is as if the children wanted to correct the puppet, rather than giving him the benefit of the doubt. Critically, however, what we find is that both groups access the matrix reading a substantial proportion of the time and that their access to this reading is a function of experimental condition.

Since these percentages still deviate from what normally expects in a TVJT, one might suspect that the individual stories or verbs could be a factor. We conducted a further analysis of the reliable responses, presented in Figure 7.

Figure 7: Analysis of verbs and story types for reliable responses in Experiment 3

Looking at these percentages, there appears to be no pattern attributable to the difference
between subject and object control structures (i.e., want/need vs. ask/invite), or to frequency (want and need are by far more frequent than both ask and invite). If the reason for the observed response pattern cannot be found in the stimuli, we must look to the participants.

One possibility is that the semantics of the matrix verbs and the complexity of the structures in which they appeared presented unique interpretive difficulties to the children, resulting in enough of an increased processing load that children’s normal inclination to agree with the puppet was overridden. Consider the test sentence given in (29):

(29) Miss Piggy wanted to drive every car that Kermit did.

In the story, Kermit expresses his likes and dislikes very clearly and has good reasons for them: he has had the old cars for a long time and is tired of driving them, and he only wants to drive his new cars. When it is Miss Piggy’s turn to choose the cars she wants to drive, her decision is subjectively motivated. While she does provide a reason (namely, in the ‘embedded’ condition, that she liked the way the cars looked when Kermit drove them, or, in the ‘matrix’ condition, that she likes the new cars more), this reason may seem arbitrary to the children. In addition, she never actually drives any cars (by design). The puppet only comments on what Miss Piggy wanted to do. There may therefore be an interaction between Miss Piggy’s desires, the child’s desires with respect to the events of the story, and the child’s willingness to accept Miss Piggy’s motivation. On top of all of this, the child then has to respond to the puppet’s statement (which contains a highly complex syntactic construction).

In essence, what we are suggesting is the following. The matrix verbs in this experiment (want, need, ask, invite) embed propositions about events that never take place (e.g., Miss Piggy

---

15 The frequencies listed in Kučera and Francis (1967) are want 329, need 260, ask 128, and invite 11 (per million).
never drives any cars, Piglet never tastes any treats, and so on) and hence require more mental activity on the part of the child than typical TVJT experiments that deal with concrete actions.\textsuperscript{16} In addition, the real-world correlate of a propositional attitude verb such as\textit{want} (Frege (1892/1997)) is not observable, and therefore also places an increased processing burden on the child participant. Finally, even adults sometimes have difficulty accessing interpretations licensed by ACD structures. This increased cognitive load may have interacted with the typical willingness to agree with the puppet and access the reading made true by the experimental context.

The results of Experiment 3 provide evidence that children can access both the matrix and embedded readings of ambiguous ACD sentences and therefore that children are able to QR out of nonfinite embedded clauses. The QNP is not constrained to one landing site, be it the lower or higher one. An unexpected pattern of results occurred, motivating us to approach the experimental results with slightly more scrutiny and hypothesize about how an increased processing load may have affected participants’ willingness to agree with the puppet. At the end of the day, however, we are able to conclude with confidence that children’s QR does not target fewer landing sites than adult’s QR. Now, whereas Experiment 3 was conducted to test whether there are limits to children’s interpretations of sentences in which an ACD structure is embedded in a\textit{nonfinite} clause (which, we have found, it is not), Experiment 4 was conducted to test whether there are limits to children and adults’ interpretation of sentences in which an ACD

\textsuperscript{16} Intensional transitive verbs, and the subset of propositional attitude verbs such as\textit{want} and\textit{believe}, involve a range of phenomena related to the status of elements in the complement clause that have presented long-standing puzzles in the field of semantics, beginning with Frege (1892/1997).
structure is embedded in a *finite* clause. Recall from § 2.3 that while QR can cross nonfinite clauses, finite (tensed) clauses create a barrier to this operation.

4.4.  **Experiment 4**

In Experiment 4, we seek to obtain empirical support for the claim that adult QR is subject to stricter locality conditions than other forms of A-bar movement, giving rise to clause boundedness. To our knowledge, while the question of whether the matrix reading is a viable one has been raised in the literature (e.g., Cecchetto (2004), Fiengo and May (1994), Wilder (1997)), no study has systematically investigated adults’ ability to access the matrix interpretation of sentences in which ACD is embedded in a tensed clause. In the process of testing this claim, we seek to determine as a comparison whether (in contrast to the results of Experiment 3, where we saw that child QR is not *overly* constrained) the QR in child grammar is *under* constrained.

4.4.1.  **Method**

4.4.1.1.  **Participants**

24 four-year-olds (11 M 13 F, \( M = 4;8 \), range 4;1-4;11) and 24 Northwestern University undergraduate students participated.

4.4.1.2.  **Materials**

Participants were presented with sentences such as (30).

(30)  Clifford said that Goofy read every book that Scooby did.

While the sentences in Experiment 3 were ambiguous, the sentences in this experiment are typically reported to be unambiguous. This lack of ambiguity is classically taken as evidence that QR is clause bounded. Thus, (30) above should only mean that Clifford said that Goofy read every book that Scooby *read*, and not that Clifford said that Goofy read every book that Scooby *said that Goofy read*. The filler sentences were unambiguous VPE sentences with a plot format.
similar in design to the test scenarios. The full set of sentences is given in Appendix A.

### 4.4.1.3. Procedure

As in the previous experiments, the procedure was the TVJT. The stories were similar in design to those in Experiment 3, and, as in Experiment 3, only one of the two interpretations was made true at a time. Subjects were randomly assigned to one of two conditions, based on whether the grammatical embedded or the ungrammatical matrix reading was true in the context. The description of a typical experimental scenario follows. In this story, three dogs, Clifford, Goofy, and Scooby, begin by standing in front of two sets of books, big ones and small ones. Scooby tells the others how much he likes to read, and decides to read each of the small books. After he has finished reading, Goofy proposes a game: Clifford and Scooby will leave the room, and in the meantime, Goofy will read some books. When Clifford and Scooby return, they will guess which books Goofy read. Scooby and Clifford are delighted by this idea and leave.

At this point, we must briefly digress to discuss an important aspect of the experimental design. Because we did not want participants to be influenced by seeing which books Goofy read, we made sure that they never actually saw Goofy read any books. The stories were videotaped, and scenes were then edited and spliced together using Final Cut Pro software. At the point in the story where Scooby and Clifford leave the room, we see Goofy considering both sets of books. Then the scene fades, and the next scene is Goofy telling Scooby and Clifford they can return and make their guesses. Upon their return, Scooby goes first and guesses that Goofy read the big books (the ones Scooby did not read). It is now Clifford’s turn to guess. In the ‘embedded’ condition, Clifford disagrees with Scooby and says Goofy read the small books (the ones Scooby read). In the ‘matrix’ condition, Clifford agrees with Scooby and says Goofy read the big books (the ones Scooby said that Goofy read). This design is illustrated in Figure 8.
Figure 8: Experimental conditions for Experiment 4

As in Experiment 3, this design allows the truth values of the two interpretations to vary across the two conditions, as we see in Table 5.

Table 5: Truth conditions in Experiment 4

<table>
<thead>
<tr>
<th>VP interpretation</th>
<th>embedded condition</th>
<th>matrix condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>read</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>said that Goofy read</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

The format of the filler stories was similar in structure to the test stories. Two characters are engaged in an activity when a third enters and proposes a game in which the two characters continue to participate in their activity while the third leaves the room, and when he returns, he guesses what happened.

4.4.2. Results

If these sentences are indeed unambiguous (because the tensed clause creates a boundary for QR), we predict that adults will uniformly reject the matrix reading in both conditions. If
children exhibit the same locality conditions on QR as adults are predicted to have, then they, too, should only allow the embedded reading and disallow the matrix reading. If, however, children’s QR is not clause-bounded, then for these participants, the test sentences in this experiment should be ambiguous, with both the embedded and the matrix readings being possible interpretations, as they were in Experiment 3.

As in the two previous experiments, we analyzed the responses in terms of the percentage of matrix readings. These percentages are presented in Figure 9.

Figure 9: Percentage matrix readings in Experiment 4

Two patterns stand out. First, adults, as predicted, appear to strongly disprefer the matrix reading of these sentences, allowing it only around 15% of the time, even in the condition that makes it true. Second, while children are also inclined to reject the matrix reading in the ‘matrix’ condition, they actually allow it more than half of the time in the ‘embedded’ condition (the condition in which the sentence with the matrix reading is false). A 2 x 2 ANOVA (age x condition) revealed a main effect of age (F (1, 23) = 12.28, p = 0.001), a marginally significant effect of condition (F (1, 23) = 3.24, p = 0.079) and a significant interaction between age and
condition \( F (1, 23) = 6.41, p = 0.015 \).

4.4.3. Discussion

The pattern of results observed in Experiment 4 is at the same time strikingly distinct from and similar to the pattern observed in Experiment 3. Looking first at the adults, we notice some key differences. In Experiment 3, adults allowed the matrix reading to hold 40% of the time, accessing this reading more often in the condition that made it true, but in Experiment 4 only allowed it only 16% of the time, a percentage that was consistently low across the two conditions. This tells us that the matrix reading is essentially unavailable to most adults.

Children, on the other hand, exhibit similarities across the two experiments. They allowed the matrix reading to hold 46% of the time across both experiments and in both experiments, they were inclined to access the matrix reading more often in the ‘embedded’ condition – the condition that made this reading false.

As in Experiment 3, let us review an example of a child’s response. Recall that in the scenario in question, the difference between the two conditions is seen in Clifford’s guess: in the ‘embedded’ condition, Clifford disagrees with Scooby and says Goofy read the small books (the ones Scooby read), while in the ‘matrix’ condition, Clifford agrees with Scooby and says Goofy read the big books (the ones Scooby said that Goofy read). However, in both conditions, the puppet says, Clifford said that Goofy read every book that Scooby did. One child, assigned to the ‘embedded’ condition, rejected the puppet’s statement, providing justification (60) in the Appendix, repeated here as (31).

(31) Clifford chose the small ones. Scooby chose the big ones.

This child appears to have accessed the matrix reading (Clifford said that Goofy read every book that Scooby said that Goofy read) and rejected the puppet’s statement, since Clifford and Scooby
made different guesses about which books Goofy read. (Note that Scooby read the small books, so the child saying that Scooby chose the big ones must be interpreted as referencing Scooby’s guess, not his reading.) Additional representative examples of children’s justifications for their responses are presented in Appendix B.

Two conclusions can be drawn from these results. First, adults are strongly disinclined to access the matrix reading. In our data, only three subjects accessed the matrix reading in the ‘matrix’ condition, while the remaining nine did not. We could take this low percentage of matrix readings to simply reflect a certain degree of noise in the data. On this view, the grammar does not generate this interpretation and participants sometimes made mistakes. Alternatively, or we might suppose that the grammar actually generates the structure from which the matrix reading is derived and that adults disinclination to access this reading derives either from the low probability assigned to that structure by the parser, or from some extragrammatical factor contributing to its unavailability. Note that the three subjects who accessed the matrix reading in the ‘matrix’ condition did so for at least two of the four test sentences. Thus it may be that for a small cohort of adults, a tensed clause does not serve as a boundary to QR.

Alternatively, it may be that a tensed clause never serves as a boundary to QR as far as the grammar is concerned, but that other pressures (i.e., pressures from the discourse context or from the parser) make it so difficult to access this reading that it is essentially impossible to generate. That is, the pattern of adults’ responses here – a split between the vast majority who disallow the matrix reading and a small minority who access is regularly without fuss – leaves open the question of whether QR out of a tensed clause is actually made illicit by the grammar, or whether it is so difficult to parse that it might just as well be (but conditions such as those presented in this experiment allow this reading to be accessible where it might otherwise not be).
We return to this issue in the general discussion.

Second, while children pattern with adults by generally rejecting the matrix reading in the ‘matrix’ condition, they are significantly more likely to access it in the ‘embedded’ condition (one-tailed t-test: $t (11) = 2.833, p = 0.008$). The reader may think this finding curious, and we are inclined to agree, given reasons outlined in our discussion of Experiment 3. There we speculated that an increased processing load may have overridden children’s usual willingness to access the interpretation made true by the experimental scenario and to agree with the puppet. Experiment 4 was strikingly similar in its design and the direction of children’s responses. We speculate that allowing certain key events in the stories to be unobserved might have influenced children to be biased against the puppet, and leave it to future research to investigate this possibility further. It is important to realize, however, that although we determined that the matrix reading was available on the basis of rejection of the puppet’s statement in the ‘embedded’ condition, rather than consistent acceptance of it in the condition favoring it, this does not change the observation that this reading was, in fact, available to children. Clearly, four-year-old children can access the matrix reading, and provide clear justifications in their responses, as illustrated in the above example and in Appendix B. We can therefore conclude that children’s QR is not as tightly constrained as adults’.

Let us return now briefly to the results of Experiment 2 for comparison. Recall that in that experiment, we showed that children do not have a bias for the VP they target as the antecedent, and access whichever reading of the ambiguous sentence is made true in the experimental scenario. The results of Experiments 3 and 4 are thus noticeably different than those of Experiment 2. This difference can be made clear by comparing the distribution of individual responses between the three experiments, captured in Table 6. In this table, the
percentage of stories that were assigned a matrix reading is presented in the leftmost column.

The raw numbers of child responses for the test sentences in each of the two conditions in Experiments 2-4 are in the columns to the right. There were 12 children in each condition. The arrows indicate the predicted pattern of responses based on what is allowed by the adult grammar and made true in the experimental scenario.

Table 6: Distribution of children’s responses in Experiments 2, 3, and 4

<table>
<thead>
<tr>
<th>% stories with matrix reading</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
<th>Experiment 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>embedded</td>
<td>matrix</td>
<td>embedded</td>
</tr>
<tr>
<td>0%</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>25%</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>50%</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>75%</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>100%</td>
<td>0</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

As Table 6 indicates, children’s responses in Experiment 2 were in line with the predictions. However, the picture is much different for Experiments 3 and 4. In neither condition of Experiment 3 are children’s responses in the predicted direction. Likewise, in Experiment 4, children’s responses in the ‘embedded’ condition diverge from the direction predicted, and across the two conditions, the responses are skewed in the opposite direction of each other. Thus we cannot say for certain whether the decreased acceptance of the matrix reading in the ‘matrix’ condition of Experiment 4 has a similar motivation as that of adults (i.e., resistance of QR out of a tensed clause), or whether it is the result of children’s general contrariness in these
experimental conditions.

Now, there are two possible reasons why in Experiment 2 children patterned in accord with the assumption that participants in a TVJT will accept a sentence in scenarios that make it true on a reading compatible with the sentence. First, although the same matrix verbs and complement clauses were included in the test sentences, assigning a truth value to the puppet’s statement required deciding whether or not the reason given in the sentence is accurate, based on reasons the characters explicitly outlined in the course of the story. Second, and perhaps crucially, the test sentences in Experiment 2 did not involve QR (only VPE). Thus, it is possible that the sheer choice of landing sites for the QNP gave rise to the divergence from the baseline for VPE interpretation, and invited the semantic factors we have discussed to play a role in children’s response patterns. That is, Experiments 3 and 4 may indicate that when children given ambiguous ACD sentences, we should expect a different pattern of results than those expected by the assumptions of the TVJT, given the complexity of these sentences that invites extragrammatical factors to play a role.

5. General Discussion

The findings of the present experiments add to mounting evidence that the Quantifier Raising operation is part of the grammar 4-year-olds (cf. Guasti and Chierchia (1999/2000), Kiguchi and Thornton (2004), Lidz et al. (2004)). Children not only assign the correct meaning to Antecedent-Contained Deletion (ACD) sentences, which require this covert operation for interpretation, but treat such sentences differently from sentences with Coordinated Conjunction, a structure that children might appeal to, were QR not part of their grammatical repertoire. Furthermore, children are able to access multiple interpretations of ambiguous ACD sentences, attesting to the fact that for children as well as adults, QR is not restricted to one landing site.
When the site of ACD is contained within an infinitival embedded clause, children show no difficulty in accessing either the embedded or the matrix reading, indicating the availability of multiple landing sites for QR. Indeed, children’s QR seems to target even more landing sites than adults’. When the site of ACD is contained within a tensed embedded clause, children robustly access both the matrix and embedded readings, while adults generally access only the embedded reading.

One of the questions we began this paper with concerned the very nature of QR. We noted that the locality properties of QR are odd from the perspective of other forms of A-bar movement (namely wh-movement) since only QR appears to be bounded by tense. This then raised the question of whether children expect QR to be bounded by tense (because this is somehow an inherent characteristic of QR) or whether this restriction needs to be learned as an additional property on top of its A-bar movement properties. Stated more theoretically, we asked whether the locality properties of QR are an integral part of the QR operation itself, or whether these properties derive from other considerations. The data reported here seem to favor the latter option, given that in Experiment 4, a small subset of adults easily and routinely accessed the matrix reading and children robustly accessed the matrix reading in the ‘embedded’ condition. The clause-bounded nature of QR therefore does not fall out directly from the existence of this operation.

This finding, however, opens an interesting learnability puzzle. What developmental process enables children to add constraints to the possible meanings of sentences? How is it that children can subtract an available meaning from their repertoire? The problem is reminiscent of the subset problem discussed by Gold (1967), Berwick (1985), Bowerman (1974), Manzini & Wexler (1987), and Pinker (1989) among others. The set of interpretations allowed by children
includes the set of interpretations allowed by adults; thus, whatever experience children have with the relevant sentences will be consistent with the set of interpretations that they allow. Indeed, as soon as the subset problem was identified as a logical problem, linguists were quick to stipulate that solutions to subset problems are built into the learner such that they should never arise in practice (Berwick (1985), Crain et al. (1994)); wherever a subset problem might arise, the learner defaults to the smaller grammar.

One case where actual subset problems have been discussed extensively is in the domain of verb learning, where it appears that children allow a wider range of argument structure alternations for some verbs than are allowed by the adult grammar (Bowerman (1974), Pinker (1989)). In this domain, two kinds of solutions have been proposed. One solution allows the overgeneration to be restricted via a continual process of subcategorization (Pinker (1989)). As learners discover more detail about the meaning of a verb, that verb gets recategorized into a different subclass and its distributional properties follow from membership in that subclass. An alternative solution (cf. Naigles, Fowler and Helm (1992), Braine and Brooks (1995)) links the whittling away of overly general structures to frequency and indirect negative evidence. Specifically, the claim is that if the verb were actually licensed in a given frame, then the learner would have seen evidence to support this view. Since adults do not use the given verb in the expected context, the expectations about the set of licit frames shrink. It is likely that in the domain of verb learning both solutions are partially at work (Ambridge et al. (2008))

However, neither of these solutions is likely to apply to the case of QR. First, there is no analog in the domain of QR to the lexical subcategories that play a role in determining the set of argument structure alternations allowed by a verb. Second, a frequency-driven account is not possible, because the construction is far too rare. Learning in this fashion would depend on the
learner’s ability to tell that the overly general interpretation that they reached for the sentence was incompatible with the intended interpretation in the context. In other words, learning via indirect negative evidence in this domain would require a very strong error signal from the context telling children that their interpretation did not match what was intended. It seems unlikely that children would be able to first detect this signal and then know that they should change their grammar on the basis of it (as opposed to, say, deciding that their interlocutor was mistaken about the truth of their sentence).  

Given that no known solution to the subset problem can be applied in this domain, we would like to consider an alternative, namely that there is no subset problem. It is, at least in principle, possible that QR is not bounded by tense from the perspective of the grammar. Rather, the grammar allows QR to move freely out of tensed clauses just like other forms of A-bar movement. Instead, QR only appears to be bounded by tense because of pressures from the parser. Children, not yet being efficient parsers, do not yield to this pressure. Consequently, children’s behavior reveals the true nature of QR, namely that it is free to cross tensed clause boundaries. Let us call this hypothesis the “extragrammatical source of QR locality hypothesis”, or the extragrammatical hypothesis for short. This hypothesis can be contrasted with the standard account, which we will call the “grammatical source of QR locality hypothesis, or the grammatical hypothesis for short.

It is important to recognize at this point that the fact that adults consistently judged the sentences with ACD embedded in a tensed clause as lacking the matrix interpretation is not necessarily evidence in favor of the grammatical hypothesis. Rather it is simply evidence that in the contexts we provided, participants failed to access a particular interpretation. Of course, this

\footnote{See Goro (2007) for related discussion.}
Evidence is consistent with the grammatical hypothesis. But it is equally consistent with the extragrammatical hypothesis, i.e., the view that the grammar does generate this structure but that certain factors outside of the grammar make this structure exceedingly difficult to access. This account would treat the case of QR out of a tensed clause as analogous to the case of center embedded relative clauses, which are generally uninterpretable despite being generable by the grammar (Miller and Chomsky (1963)).

In order to contrast these two hypotheses, it is important to identify precisely what parsing steps are involved in resolving ACD sentences and how the efficient implementation of these steps could lead to the appearance that a structure generated by the grammar is not possible. By hypothesis, ACD sentences involve grammatical operations that target the left edge of VP: Quantifier Raising and anaphora resolution of the elided VP. Assume that the interpretation of the sentence in (32a) requires moving the QNP into a position at the left edge of VP, as seen in the corresponding LF representation in (32b) (Heim and Kratzer (1998)). Assume also that part of this interpretation involves creating an S-structure representation, as seen in (32c).

   b. \[[IP John [VP every boy] [VP saw t,]]\] = LF
   c. \[[IP John [VP saw every boy]]\] = S-structure

Given that the parser cannot know that QR is required for a given quantificational expression until that expression is encountered (here, upon reaching the QNP in object position), it follows that the relevant portions of the S-structure representation are constructed prior to the completion of an LF representation containing the moved quantificational expression. Thus, given a QNP in object position, the parser must retrospectively search the already constructed S-
structure representation to identify a VP to serve as a landing site, and then must move the QNP there.\(^\text{18}\) Similarly, when the parser gets evidence that a VP is missing from a structure, as in (33a), it must retrospectively search the already constructed representation in (33b) in order to identify an antecedent VP to fill in as the content of the elided VP (33c).

(33) a. Lola jumped over the frog and Dora did too.
    b. [IP Lola [VP jumped over the frog]] and [IP Dora [I did [VP _] too]]
    c. [IP Lola [VP jumped over the frog]] and [IP Dora [I did [VP jump over the frog] too]]

Putting these two ideas together leads to the conclusion that in parsing a sentence containing ACD, the parser must implement two searches for VPs, one to resolve QR and one to fill in the ellipsis.

In Experiment 3 we saw that when interpreting sentences that provided multiple grammatical landing sites for QR, adults have a (defeasible) preference for the embedded reading (the reading that involves QR to the closest landing site). This behavior contrasted with the results of Experiment 2, in which the sentences involved only VPE but no QR, where adults showed no such preference. Combined, these results indicate that from the perspective of the parser, QR and VPE resolution have different locality properties with respect to the target of multiple VPs.

In Experiment 4 we found that adults showed a nearly categorical preference embedded reading when interpreting sentences in which ACD was contained inside a finite embedded clause. This pattern contrasted with the data from children, who were able to access both the

---

matrix and embedded readings. Under the extragrammatical hypothesis, adults’ preference for
the embedded reading in these sentences reflects a parsing bias that could not be overcome by
the experimental contexts provided, rather than constraints imposed by the grammar. The fact
that with the Experiment 3 sentences, where both the matrix and embedded landing sites are
made possible by the grammar, adults still showed a preference for the embedded VP as a
landing site for QR, makes it straightforward to suppose that this preference is amplified in the
Experiment 4 sentences, perhaps because of the extra processing load introduced by the
interpretation of Tense and a new discourse referent in the subject of the embedded clause.
Indeed, J. Merchant (p.c) notes that making the embedded subject a bound pronoun also seems to
weaken the boundary introduced by Tense, allowing for sloppy identity in (34):

(34) Clifford said that he read every book that Scooby did.

   Clifford said that he read every book that Scooby…
   a. read
   b.? said that he read.

Examples such as (12)-(14), repeated here, provided further support for the possibility
that the grammar does generate structures requiring QR out of an embedded clause:

(12) A (different) student made sure that every invited speaker had a ride.

   = For every invited speaker, there was a student that made sure that speaker had a ride.

(13) A different witness testified that John robbed every bank in town.

(14) John thinks that Mary is taller than Bill does.

Now, despite adults’ preference for the embedded VP as a landing site for QR, adults
displayed no preference for one or the other VP in Experiment 2, where the sentences involved
VPE resolution and no QR. This contrast may be illuminating about children’s performance in
Experiment 4. It is possible that for children, activating the matrix VP for the purposes of ellipsis resolution somehow makes it easier to activate that VP for the purposes of targeting a landing site for QR. That is, it is possible that children are less able than adults to keep the two VP searches – one for a QR landing site and the other for an ellipsis antecedent – separate and exhibit a kind of facilitation effect on the matrix VP, making that VP more available as a QR landing site than it would be otherwise. Since under the extragrammatical hypothesis, the grammar places no restrictions on the matrix VP as a landing site, children are free to choose it.

The transition to adulthood, then, would involve becoming more efficient at keeping the search for a VP as a QR landing site separate from the search for a VP as an antecedent for ellipsis. As this ability develops, presumably as a consequence of increased efficiency in performing such searches during online sentence comprehension, children’s behavior should also become more adult-like, and no particular triggering experience should be required to initiate this development. If this account is on the right track, children should be less likely to QR out of a tensed clause when no VPE is involved, since in such an example there is no interfering activation of the matrix VP as a potential landing site. We leave testing this prediction for further research.

Now let us consider the grammatical hypothesis. On this view, when the parser conducts a search for VPs to serve as landing sites for QR, it finds only one – the embedded VP – because QR is prohibited from crossing a finite clause boundary. Although the search for potential antecedents for ellipsis yields two VPs (the embedded and matrix VPs), only the embedded VP is available, since the choice of the matrix VP for ellipsis resolution and the embedded VP for QR would mean that the QNP was still contained in its antecedent (a problem we discussed in § 2.1). It is possible that – in a situation analogous to the one described for the extragrammatical
hypothesis – a coordinated search for VPs would lead to the parser erroneously considering the
matrix VP as a potential landing site for QR. As a result, the parser would temporarily consider
as acceptable an interpretation that is not actually licensed by the grammar. This type of situation
has gained considerable attention in the adult sentence processing literature in recent years (e.g.,
Arregui et al. (2006), Gordon, Hendrick and Johnson (2001), Van Dyke and Lewis (2003), Van
Dyke and McElree (2006), Lewis and Vasishth (2005), McElree, Foraker and Dyer (2003)). It is
possible that with the ACD sentences we are discussing, while adults generally recover from this
interference, children are more susceptible to it. On this view, the children have acquired the
correct locality conditions for QR, but parsing factors make it appear that they have not (see
Conroy et al. (submitted) for related discussion). Development, on this view, would involve
children’s parsers improving the ability to overcome interference effects during online sentence
processing.

Both the extragrammatical hypothesis and the grammatical hypothesis explain the
transition to adulthood in terms of developmental changes in parsing strategies by saying that
children become more efficient at conducting and maintaining as separate concurrent searches
for VPs. Both approaches therefore also treat development in the domain of sentence processing
to be driven, at least in part, by developmental changes to the efficient deployment of the parser.
We have outlined here two plausible scenarios that account for how children can become more
adult-like without requiring any changes to their grammar per se. Teasing apart the differences in
the approaches will require significant subsequent research.

Before concluding, we’d like to turn our attention towards an alternative hypothesis
concerning the role of memory in children’s abilities to deal with certain types of ambiguity
offered by Grodzinsky and Reinhart (1993) and Reinhart (2004). Their hypothesis, which we
will refer to as the *processing overload hypothesis*, holds that the maintenance of multiple derivations in working memory for the purpose of comparing them is beyond the scope of children’s memory capacities. Thus, when faced with grammatical constructions requiring a comparison of multiple items, the child’s memory system is overloaded and so the child is unable to find an interpretation. In Reinhart’s (2004) words, “The processing poses too big a problem on their working memory, … [which] is not developed enough to hold the materials needed to complete the execution of the task – namely constructing an alternative derivation while holding the previous one and then comparing the two derivations. Failing the execution, they resort to some strategy” (122). This account, like the two accounts put forth above, treats children’s non-adult behavior as a consequence of developmental differences in the parser. But unlike the accounts described above, this account assumes that difficulties accessing working memory can derail the sentence processing mechanism altogether. The data presented here provide strong evidence against the processing overload hypothesis, on the assumption that computing the meaning of sentences with QR involves a comparison of multiple derivations (Fox (1995; 1999)). In order to see how this works, we must first consider an account of scope economy.

The essence of scope economy is that QR is licit only when applying the operation yields an interpretation that is distinct from a derivation lacking the operation (Chomsky (1993), Fox (1995; 1999)). In the case of ACD sentences, such as those as in our Experiment 1, repeated here, QR is licit because failing to apply QR leads to certain problems (e.g., infinite regress, lack of parallelism between the elided VP and its antecedent, or a type mismatch between the QNP in the object position).

(1) Lola jumped over every frog that Dora did.
In the case of ambiguous ACD sentences where ACD is in a nonfinite clause, such as (25) from Experiment 3, repeated here, the matrix reading is licensed because QRing into the matrix clause yields an interpretation distinct from QRing only as high as the embedded vP.

(25) Miss Piggy wanted to drive every car that Kermit did

In the case of sentences where ACD is embedded in a finite clause, such as (30) from Experiment 4, repeated here, QRing into the matrix clause is argued to be blocked because there is a required step along the way that does not alter the interpretation.

(30) Goofy said that Scooby read every book that Clifford did

Because that move is not motivated, the move to the higher landing site, and consequently the matrix reading, should not even be considered (cf. Cecchetto (2004)). Consider the derivation outlined in (35) and the corresponding tree illustrated in (36).

(35) Goofy said that Scooby read [every book that Clifford did]

a. QNP is QRed to the lowest [spec, vP]

   = Goofy said that for every book that Clifford read, Scooby also read it

b. QNP is QRed to embedded [spec, CP] (*because meaning is indistinct from result of QR in (a))

   = Goofy said that for every book that Clifford read, Scooby also read it

c. QNP is QRed to matrix [spec, vP] (would be licensed by a distinct meaning)

   = For every book that Clifford said that Scooby read, Goofy also said that Scooby read it

Step (b) of this derivation is supposedly illicit because it does not generate a meaning that is distinct from the one that results from the earlier step. Consequently, as the story goes, the QP is trapped inside the tensed clause. We will not discuss the merits of this proposal, but rather
assume that it is valid, so that its consequences for the processing overload hypothesis can be examined.\textsuperscript{19}

Now, the hypothesis states that a comparison of derivations and the interpretations that result from them is beyond the memory capacities of children. Consequently, in tasks requiring children to understand sentences that require such comparisons, they should show chance performance, picking an interpretation at random. In all three of our experiments involving ACD, then, children should be at chance in accessing a given reading. The facts are clearly otherwise.

\textsuperscript{19}The processing overload hypothesis presupposes that some variant of the grammatical hypothesis is correct. If we adopt some version of the extragrammatical hypothesis, whereby the matrix reading is grammatical but highly dispreferred by the parsing mechanism, then the scope economy hypothesis would need to be adjusted (or rejected altogether) in order to account for why movement to the higher landing site is licensed.
In Experiment 1, children showed nearly perfect performance and demonstrated nothing even approaching chance behavior. When the ACD sentence was true in the context, children recognized it as such, and when the ACD sentence was false in the context, they did the same. As we have noted, however, failure to apply QR in an ACD sentence would give rise to an uninterpretable structure. It could be that faced with this possibility, children somehow manage to overcome their memory limitations.

Let us turn now to ambiguous sentences in Experiment 3. Recall that there are two possible landing sites for QR, the higher site being licensed because it yields an interpretation that is distinct from the one resulting from targeting the lower landing site. This is therefore an issue of scope economy, so the processing overload hypothesis predicts that children should be at chance in choosing an interpretation. Now, Reinhart (2004)’s predictions are actually a bit more subtle than pure chance performance across the board. She offers two possibilities when semantic disambiguation is involved. The first is individual guessing, which should basically be a flip of the coin across trials for individual subjects. The second, which she says should manifest itself in “dual-choice experiments” involving semantic disambiguation, is the adoption of a consistent default interpretation. Thus, the overall group performance should still hover around 50%, because it is the result of a bimodal distribution. The prediction, then, is that across our conditions, we should see a mean of 50%, resulting from either guessing or a random choice of a default interpretation.

Reinhart (2004) discusses at some length Crain et al.’s (1994) claim that the parser, in seeking to reduce the risk of making commitments that will later need to be altered, chooses by default an interpretation that allows for a minimal commitment. When the question is one of entailment, the default interpretation will be the one that entails the others. Reinhart offers other
motivations and says that the more general characteristic of a default interpretation is that it gives the child a greater chance of meeting a relevant goal, be that curiosity, success, minimal mistakes and revisions, etc. Since in our experiment, one reading does not entail the other, any choice of a default interpretation would have to be motivated by one of these other factors.

Let us see if this prediction is borne out in the data. Table 6 (presented in the discussion of Experiment 4) captures the distribution of the responses across Experiments 2-4. Since the test sentences in Experiments 2 do not involve QR, and therefore do not involve an issue of scope economy, let us leave them aside. For each condition, there were 4 trials, so each of the 12 children in a given condition had the possibility of accessing the matrix reading 0%, 25%, 50%, 75%, or 100% of the time. Now, according to the processing overload hypothesis, we should see one of two patterns hold across the board—a binomial distribution clustered at 50% or a bimodal distribution with peaks at 0% and 100%.

Let us begin by entertaining the first possibility. A binomial distribution with a probability of 0.5 would mean that four to five children accepted a given reading half of the time, approximately three children accepted it once or three out of four times, and either one or no children accepted it for none or all four of the trials. An examination of the responses in Experiments 3 and 4 reveals that this prediction is not borne out. A Chi-square analysis comparing the expected and observed frequencies in each of these conditions revealed a significant or marginally significant difference in each of these conditions, except for the ‘matrix’ condition of Experiment 3 (Expt. 3, embedded: $\chi^2 = 24.22, p < 0.0001$; Expt. 3, matrix: $\chi^2 = 5.56, p = 0.2345$; Expt. 4, embedded: $\chi^2 = 9.22, p = 0.0558$; Expt. 4 matrix: $\chi^2 = 39.89, p < 0.0001$). In the ‘matrix’ condition of Experiment 3, where the distribution is not statistically significant from a binomial distribution with a probability of 0.5, we do see the most children of
all four conditions (five) at 50%. In the ‘embedded’ condition of Experiment 3, only one child was at 50%, and in both of the two conditions of Experiment 4, only two children were at 50%. However, even in this condition, we do not observe the tails we would expect from a strictly normal distribution: six children accessed the matrix reading 25% of the time or less, while only one accessed it 75% of the time or more. Interestingly, adults’ percentage of matrix readings in this same condition were clustered around 50%, a pattern which presumably does not derive from their limited working memory capacities.

Now, the second possibility suggested by Reinhart (2004) is a bimodal distribution resulting from children choosing a default interpretation. Of the four conditions, only the ‘embedded’ condition of Experiment 3 appears to have such a distribution, with five children accessing the matrix reading 25% of the time or less and six children accessing it 75% of the time or more. However, in this condition, we still see five of the twelve children accessing both readings within the experimental session, only one of which is at chance level. Thus, while the two conditions in Experiment 3 appear at first glance to have the characteristics that were expected according to the processing overload hypothesis, closer inspection shows that they do not fit the picture entirely. Moreover, they each show different distributions, something that is unexpected from the hypothesis.

The distributions of the two conditions in Experiment 4, which do not resemble either type of predicted distribution, also differ dramatically from each other. In the ‘embedded’ condition, seven of the twelve children accessed the matrix reading 75% of the time or more, and three 25% or less, while in the ‘matrix’ condition, where there is a clear positive skew, only two of the twelve children accessed this reading 75% of the time or more, while eight accessed it 25% of the time or less. This pattern is striking, since with these test sentences, scope economy is
given as the reason for the lack of a matrix reading. Grodzinsky and Reinhart (1993) would predict that children should be at chance, because the computation requires the comparison of multiple derivations. Instead, in the ‘embedded’ condition, children were more likely to access the matrix reading, and in the ‘matrix’ condition, they were more likely to access the embedded reading. Finally, it is unclear how one would reconcile a guessing strategy with the justifications offered by children and presented in Appendix B, which were both coherent and apropos.

In sum, there is no evidence from any of our ACD experiments in support of the processing overload hypothesis. Whatever processing difficulties children may have with these sentences, their behavior does not reflect an inability to compare competing derivations. Instead, we have suggested that children’s behavior reflects a grammar unimpaired by an efficient parsing mechanism, allowing them to access interpretations that most adults consistently disallow. Where children differ from adults with respect to the sentences in the current study – that is, why adults find it so hard to access interpretations requiring QR out of a tensed clause, but children do not – appears to lie outside of the grammar. We have suggested that this difference very likely has something to do with how the parser implements a search of representations held in memory during online sentence comprehension. While we have only outlined the spirit of such an account, we take the construction of a theory of LF parsing and its relation to development to be a critical goal both for the theory of sentence understanding and for the theory of linguistic development.
References


Brain Sciences 14, 597-612.


Larson, R. and R. May (1990) Antecedent Containment or Vacuous Movement: Reply to Baltin,


Comprehension of Verb-Phrase Ellipsis,” *Journal of Memory and Language* 58, 879-906.


Musolino, J. (1999b) *Universal Grammar and the Acquisition of Semantic Knowledge: An Experimental Investigation into the Acquisition of Quantifier Negation-Interaction in English*, Doctoral dissertation, University of Maryland, College Park, Maryland.

Quantification,” *Linguistics* 44, 817-852.


Sag, I. (1976) *Deletion and Logical Form*, Doctoral dissertation, Massachusetts Institute of


Massachusetts, Amherst.


Appendix A

Experiment 1

Test Sentences

ACD

(1) Miss Red jumped over every frog that Miss Black did.
(2) Miss Red fed every snake that Miss Black did.
(3) Lady Bug drove every race car that Mister Bug did.
(4) The black butterfly landed on every butterfly that the green butterfly did.

Coordinated Conjunction

(5) Miss Red jumped over every frog, and Miss Black did, too.
(6) Miss Red fed every snake, and Miss Black did, too.
(7) Lady Bug drove every race car, and Mister Bug did, too.
(8) The black butterfly landed on every butterfly, and the green butterfly did, too.

Filler Sentences

ACD

(9) The rhino [made/didn’t make] friends with the hippo that kicked the rock into the water.
(10) The cow [played/didn’t play] with the horses that jumped over the barn.
(11) [The lizards that ate the lobsters/the lizard that didn’t eat the lobster] got sick.

Coordinated Conjunction

(12) The rhino picked up the rock and [fell in the water/jumped over the hippos].
(13) The [cow/black horse] jumped over the [log/barn] and kicked the [black horse/cow].
(14) Some lizards ate lobsters and [got/didn’t get] sick.

*Experiment 2*

Test Sentences

(15) Miss Piggy wanted to drive the [red/black] cars because Kermit did.

(16) The cowgirl needed to jump over the [small/big] frogs because the old cowboy did.

(17) Clifford asked Goofy to read the [big/small] books because Scooby did.

(18) Winnie the Pooh invited Piglet to taste the [cookies/cupcakes] because Tigger did.

Filler Sentences

(19) The [monkey/elephant] pushed the rock further than the [elephant/monkey] did.

(20) The [shark/lizard] ate a lobster before the [lizard/shark] did.

(21) The rhino rode the motorcycle, [and the hippo did, too/but the hippo didn’t].

*Experiment 3*

Test Sentences

(22) Miss Piggy wanted to drive every car that Kermit did.

(23) The cowgirl needed to jump over every frog that the old cowboy did.

(24) Clifford asked Goofy to read every book that Scooby did.

(25) Winnie the Pooh invited Piglet to taste every treat that Tigger did.

Filler Sentences

The filler sentences were the same as in Experiment 2.
Experiment 4

Test Sentences

(26) Miss Piggy said that Fozzy drove every car that Kermit did.

(27) The genie said that the cowgirl jumped over every frog that the old cowboy did.

(28) Clifford said that Goofy read every book that Scooby did.

(29) Winnie the Pooh said that Piglet tasted every treat that Tigger did.

Filler Sentences

(30) The lion said that the [monkey/elephant] pushed the rock further than the [elephant/monkey] did.

(31) The dinosaur said that the [shark/lizard] ate a lobster before the [lizard/shark] did.

(32) Ernie said that the rhino rode the motorcycle, [and the hippo did, too/but he said that the hippo didn’t].

Appendix B

Experiment 1

Justification of the truth of the puppet’s statement

ACD

(33) Those two fed these three snakes. (‘one set’ condition)

(34) Both didn’t jump over all of them. (‘one set’ condition)

CC

(35) Ladybug went on her cars. Mr. Bug went on his cars. (‘two set’ condition)

(36) They jumped over all their frogs. (‘two set’ condition)

(37) Miss Black jumped over HER frogs, and Miss Red did, too. (‘two set’
condition)

Justification of the falsity of the puppet’s statement

ACD

(38) The black butterfly landed on HIS dinosaur friends and the green butterfly landed on HIS dinosaur friends. (‘two set’ condition)

(39) The big bug drove all THESE cars, and the little bug drove all THESE cars. (‘two set’ condition)

(40) The red one fed THESE snakes and the black one fed THESE snakes. (‘two set’ condition)

CC

(41) They jumped that one and that one and that one but NOT that one. (‘one set’ condition)

(42) They fed all the snakes except the last one. (‘one set’ condition)

Experiment 2

Justification of the truth of the puppet’s statement (acceptance of statement)

(43) Because Kermit drove the red cars and Piggy thinks … and she wanted to drive them. (‘embedded’ condition)

(44) Because the cowboy teaches the little girl to jump over the frogs (pointing to small frogs). (‘embedded’ condition)

(45) Clifford asked Goofy to read the big books ’cause Scooby didn’t know how to read the big books. (‘matrix’ condition)

(46) ’Cause Mr. Piggy and Mr. Froggy wanted to drive the black cars, but he not allowed to. (‘matrix’ condition)
Justification of the falsity of the puppet’s statement (rejection of statement)

(47) The girl had to jump over the little frogs, and the cowboy had to jump over the big frogs. ('embedded' condition)

(48) Kermit did the red, but Miss Piggy wanted to do the black. ('matrix' condition)

Experiment 3

Justification of the truth of the puppet’s statement (acceptance of statement)

(49) Kermit drove the red ones. Miss Piggy wanted to drive the red ones. ('embedded' condition)

(50) The cowboy showed the cowgirl with the little frogs, and the cowgirl needed to jump over the small ones, and the cowboy needed to jump over the big ones. ('embedded' condition)

(51) The old cowboy and the little cowgirl needed to jump over the big frogs. ('matrix' condition)

Justification of the falsity of the puppet’s statement (rejection of statement)

(52) They needed to jump over different frogs. ('embedded' condition)

(53) Pooh invited Piglet to taste the cookies, and that’s what Tigger was nibbling on. Tigger wanted Piglet to eat the other treats. ('embedded' condition)

(54) Miss Piggy wanted to drive the new cars. Kermit drove the old cars. ('matrix' condition)

(55) Miss Piggy wanted to drive the cars Kermit didn't drive. ('matrix' condition)

(56) Clifford wanted Goofy to read THESE books, but Scooby read THOSE. ('matrix' condition)
Experiment 4

Justification of the truth of the puppet’s statement (acceptance of statement)

(57) Genie said what the … the same … so he went on the cowboy's side. ('matrix' condition)

(58) Experimenter: What did MP say?
Child: She go with Kermit. ('matrix' condition)

Justification of the falsity of the puppet’s statement (rejection of statement)

(59) Winnie the Pooh said Piglet tasted the cookies and Tigger – he said Piglet tasted the icing. ('embedded' condition)

(60) Clifford chose the small ones. Scooby chose the big ones. ('embedded' condition)

(61) The Genie said the little frogs. The Old Cowboy said she jumped over the big frogs. ('embedded' condition)

(62) First Scooby read the little ones, then they [Clifford and Scooby] hid, and then he [Goofy] read a book. They guessed the big ones. ('matrix' condition)

(63) Winnie the Pooh agreed with Tigger. ('matrix' condition)