In this chapter, we present our experiments testing children’s knowledge of the binding theory in VP ellipsis structures. In sections 4.1–4.4, we discuss experimental methods, introduce the methodology used in our investigation (the truth value judgment task), and present the experimental stimuli and procedures. In sections 4.5–4.8, we report our experimental results, and in section 4.9 we describe a follow-up study.

4.1 Truth Value Judgment Tasks for Children

The underlying assumption from which our methodological considerations follow is that children have an innate module of the brain that is specialized for language, Universal Grammar. A child acquires his or her target grammar, provided some amount of input from the environment is available. Our goal in conducting empirical experiments is to find out what aspects of children’s grammars are given by Universal Grammar, and what aspects develop through experience.

In our study of the properties of VP ellipsis structures, we have identified some that are innate and some that may be governed by pragmatic or real-world knowledge. For example, certain interpretations of VP ellipsis structures are ruled out by principles of the binding theory, purportedly innate principles of Universal Grammar (Chomsky 1981, 1986). Other interpretations are ruled out by the parallelism constraint, which we have suggested may be modularized—one part, structural parallelism, being given by the grammar, and the other part, referential parallelism, being subject to pragmatic considerations. Given the Innateness Hypothesis, we predict that children will demonstrate knowledge of those aspects of the grammar that are innate, but may not show complete mastery of those aspects guided by pragmatic or real-world knowledge.
What methods can be used to determine the representation of a sentence in a child’s grammar? For adult grammars, theoretical linguists use a variety of empirical methods to determine the status of a representation. Most frequently, linguists seek grammaticality judgments from a native speaker. Often, the judgment is made in the absence of context; sometimes, a context is provided. A theory is constructed that explains as much of the linguistic behavior (including judgments) as possible, under the constraint of learnability and the usual scientific constraints of simplicity, elegance, and so on.

In principle, the same requirements hold for studying the status of a representation in a child’s grammar; that is, researchers need to find data (some form of behavior) that allow them to determine the status of a representation (of a sentence, say). However, some of the kinds of evidence that are used to study adult grammars are not available, or not reliable, for studying children’s grammars. In particular, children have difficulty making subtle grammaticality judgments. We assume that cognitive factors are at work here, including, for example, an explicit understanding of the notion “grammatical.”

Empirical work on grammatical development thus concentrates on finding methods that allow researchers to determine the grammatical status of particular sentences in the child’s implicit grammar. Because it is difficult to tap this knowledge directly, researchers must often devise indirect methods. From children’s behavior in these tasks, researchers infer their grammatical knowledge. We focus here on the truth value judgment task (Crain and Thornton 1998; see also Gordon 1996). (For a comprehensive summary of other tasks that have been used with children, see McDaniel, McKee, and Cairns 1996.)

The truth value judgment task takes seriously the idea that language consists of “sentence/meaning pairs” (Wexler and Culicover 1980). A picture, or a story acted out with toys, embodies the context equivalent to the “meaning” part of the sentence/meaning pair. Following the presentation of the picture or the acting out of the story, a puppet usually utters a test sentence. This corresponds to the “sentence” part of the sentence/meaning pair. The child’s task is to judge whether the puppet’s description of the picture or of the events of the story was true or false. From this judgment, an inference is made about the child’s grammar. Tasks based on the truth or falsity of a sentence in context are particularly suitable for testing knowledge of constraints on meaning, that is, constraints that limit the range of meanings that can be assigned to utterances. Among the constraints of this kind are the principle of parallelism constraint.

Truth value judgment tasks vary according to the goal of the experiment, either part of the equation is presented, either prime a value judgment task or as a short story counters and props (the dynamic value task). The task itself allows several variations. The target is not visible by a single picture; in this case, the picture conform to the constraint and the picture constrains it is shown to the child on separate pictures, as used by Chien and Wexler (1990). Alternating two pictures at once and is asked to choose sentence this was the method used by Wexler and colleagues to determine whether children know the c-command property of sentences like Cinderella’s sister is pointing to herself in picture context that depicts the illicit meaning of Cinderella. In the picture context that isulator, Cinderella’s sister points to herself.

In the dynamic truth value judgment task, the context—that is, meanings of a sentence in this version of the task, the two context story, which is acted out for the child’s interpretation under investigation that constraint (e.g., a binding principle) and activities permits for the sentence. The additional sentence is always presented in a felicitous context. We testing the c-command requirement refers to pointing to herself using the dynamic truth value judgment task. The events acted out in the story would be the girl asks her sister to point to her, but her sister instead asks herself instead. Thus, both the illicit meaning points to her (Cinderella), and the grammatically correct meaning of her sister points to herself, are made available within the story.

Both variations of the truth value judgment task have disadvantages. The experimental results are extremely reliable. Moreover, in this task, syntax and semantics of the sentences being tested are the same, since grammar and pragmatics go hand in hand.

Experimental Investigations
constraints of this kind are the principles of the binding theory and the parallelism constraint.

Truth value judgment tasks vary according to how the context (meaning) part of the equation is presented, either as a picture (the picture truth value judgment task) or as a short story that is acted out with toy characters and props (the dynamic truth value judgment task). The picture task itself allows several variations. The target sentence can be accompanied by a single picture; in this case, the picture depicting the meaning that conforms to the constraint and the picture depicting the meaning that violates it are shown to the child on separate trials. This was the method used by Chien and Wexler (1990). Alternatively, the child is shown both pictures at once and is asked to choose the one that matches the test sentence. This was the method used by Wexler and Chien (1985) in testing whether children know the \( c \)-command property relevant to Principle A in sentences like Cinderella’s sister is pointing to herself, for example. In the picture context that depicts the illicit meaning, Cinderella’s sister points to Cinderella. In the picture context that is compatible with the adult grammar, Cinderella’s sister points to herself.

In the dynamic truth value judgment task, it is always the case that two contexts—that is, meanings of a sentence—are under consideration. In this version of the task, the two contexts are incorporated into a single story, which is acted out for the child. The two contexts represent the interpretation under investigation that is ruled out by a particular constraint (e.g., a binding principle) and a meaning that the adult grammar permits for the sentence. The additional meaning ensures that the target sentence is always presented in a felicitous context. Suppose we were testing the \( c \)-command requirement relevant to Cinderella’s sister is pointing to herself using the dynamic version of the task. The relevant events acted out in the story would be something like these. Cinderella asks her sister to point to her, but her sister won’t do it and points to herself instead.\(^2\) Thus, both the illicit meaning, in which Cinderella’s sister points to her (Cinderella), and the grammatical one, in which Cinderella’s sister points to herself, are made available to the child for consideration within the story.

Both variations of the truth value judgment task have advantages and disadvantages. The experimental results from the dynamic task are extremely reliable. Moreover, in this task, the presuppositions and pragmatics of the sentences being tested are satisfied in the story so that grammar and pragmatics go hand in hand (see Crain and Wexler 1999).
By having the characters in the story discuss the pros and cons of their actions, the task achieves a great deal of precision in matching a target sentence with appropriate contexts that support both its assertion and its denial.

The main disadvantage of the dynamic task is that it is very labor-intensive because each story must be acted out individually for each child. Therefore, it is difficult to test a large number of subjects. To some extent, the fact that the careful experimental conditions of the task elicit precise responses from subjects might overcome this problem. However, in practice, an experimenter often wants to look at subgroups of subjects (e.g., to ask whether those who perform well on X perform well on Y), and in this task, the numbers of subjects in particular subgroups are often small enough that it is difficult to get precise results. It is also difficult to investigate large numbers of items with this task. A maximum of four items testing a particular condition is usual. (Sometimes, only two items are tested, in which case it is not easy to look at properties of individual grammars.) Since roughly 10 stories can be acted out in 30 minutes, it can take three or four sessions to get an in-depth look at any particular phenomenon.

The picture task can be conducted relatively quickly. Large numbers of subjects can be tested on large numbers of items. If only group results are needed (e.g., how 4-year-olds perform compared with 5-year-olds), this task may be sufficient. As illustration of what can be accomplished with the picture task, consider the results that have been standardly taken to demonstrate knowledge of Principle A. To investigate children’s knowledge of this principle, Wexler and Chien (1985) and Chien and Wexler (1990) each tested at least 120 subjects: eight groups of 15 subjects at 6-month age intervals from 2;6 to 6;6. Smaller numbers of subjects might have produced much less stable results, a familiar possibility in experimental psychology.

When the experimental linguist is studying the kind of property whose existence developmental psycholinguists and linguists may not easily believe in, it may be useful to perform experiments with many subjects and with many items per structure, opting for a simpler experimental method, with less rich contextual control, even if this allows a little more random error in the experiment. In this case, the picture task may be the right choice. On the other hand, once a property has been demonstrated (say, that children really do behave in a nonadult way on sentences governed by Principle B), then it may be useful to get a detailed analysis of the phenomenon and a precise picture of the case, the dynamic task is called for.

In the present study, we chose to use the truth value judgment task. Children’s meanings have already been well established, so we look at the phenomenon in a different way. Moreover, the structures are more complex, so the interaction of several events, which may be difficult to recognize that our use of this task measures different patterns of behavior among different.

4.2 The Truth Value Judgment Task

4.2.1 An Overview

The truth value judgment task requires a description of the task, so far as it narrates the story and manipulates the materials. The second experimenter takes the role of any other character, who interacts with the child. At the end of each story, the puppet asks what the child thought happened. Sometimes the puppet says, “I don’t know, do you think this happened?” (i.e., the “sentence” matches the “meaning”), and the child says, “Yes.” Other times, the puppet says, “This is wrong, don’t you think?” (i.e., the “sentence” does not match the “meaning”). If the child says, “Yes,” the puppet says, “I think you are right. You pay attention, and he describes the scene.” If the child says, “No,” the puppet says, “I think you are wrong. He describes the scene.” The puppet’s description immediately follows the child’s answer, so it is still fresh in the child’s mind. No puppet’s description true or false, the
the phenomenon and a precise picture of how children behave. In this case, the dynamic task is called for.

In the present study, we chose to use the dynamic version of the truth value judgment task. Children's misunderstanding of pronouns had already been well established, so we were in a position to take a closer look at the phenomenon in a different context, namely, VP ellipsis structures. Moreover, the structures are sufficiently complex to necessitate several events, which may be difficult to portray in a single picture. We recognize that our use of this task meant we were able to test fewer children on fewer items. Despite this drawback, we were able to identify different patterns of behavior among different groups of children.

4.2 The Truth Value Judgment Task

4.2.1 An Overview

The truth value judgment task requires two experimenters (for a detailed description of the task, see Crain and Thornton 1998). One experimenter narrates the story and manipulates the toys that participate in the event. The second experimenter takes the role of a puppet (Kermit the Frog or any other favorite character), who watches the stories along with the child. At the end of each story, the puppet's job is to try and say what he thought happened. Sometimes the puppet describes the story correctly (i.e., the "sentence" matches the "meaning"), but sometimes he doesn't pay attention, and he describes the story incorrectly (i.e., the "sentence" does not match the "meaning"). If the puppet says the "right thing," the child feeds him his favorite food. If he says the "wrong thing," the child gives him a bite of something he doesn't like as well, perhaps his second favorite food, or makes him do push-ups or the like. When a child judges the puppet's utterance incorrect, the experimenter playing Kermit the Frog asks, "What really happened?" This question follows up on the child's reason for rejecting the sentence/meaning pair, and it elicits production data that frequently add important insights to children's interpretation of the test sentences.

The experimenter playing the puppet keeps the child involved in the game. The puppet pays close attention to the story, commenting where appropriate and encouraging the child to be observant. The puppet's description immediately follows the end of the story, while the events are still fresh in the child's mind. No matter whether the child judges the puppet's description true or false, the puppet is enthusiastic, so that the
child does not develop a bias toward answering one way or the other. Because it is the puppet’s performance and not the child’s that is being scrutinized, children do not realize that they are being tested. This makes the game much more enjoyable for children.

We have already noted that the story acted out for children incorporates two contexts that correspond to two potential meanings of the test sentence. Which meaning is aligned with the “Yes” response and which meaning is aligned with the “No” response depends on the hypotheses under consideration in the experiment. Following standard terminology, we refer to the experimenter’s own hypothesis as the experimental hypothesis and to the opposing hypothesis as the null hypothesis. For our purposes, the experimental hypothesis will be that children have knowledge of the principles of Universal Grammar, and the null hypothesis will be that children do not necessarily respond in accord with Universal Grammar.

Crain and Thornton (1998) suggest that in experiments testing knowledge of constraints, the “Yes” response should be aligned with the null hypothesis. It is often reported that children tend to respond “Yes” in experimental tasks requiring a “Yes” or “No” judgment. Therefore, if the “Yes” response is equated with the null hypothesis, the experiment is biased toward favoring that hypothesis. This is one way to prevent type I errors (i.e., accepting one’s own experimental hypothesis in circumstances that do not support it). If the “No” response is associated with the experimental hypothesis, children must overcome any bias they may have to say “Yes” in order to judge a sentence/meaning pair as false. Therefore, children’s “No” response to the experimental stimuli can be taken as strong evidence in favor of the experimental hypothesis and against the null hypothesis.

The events within the story are also ordered so as to favor the null hypothesis, and in so doing to avoid type I errors. The part of the story corresponding to the null hypothesis (i.e., the illicit meaning of the test sentence) is staged last in the story. Given that the event that corresponds to the “Yes” response comes last and is presumably fresh in the child’s memory, a “No” judgment leads to the conclusion that the incorrect interpretation is unavailable because it is excluded by the relevant constraint of Universal Grammar.

4.2.2 An Example

Let us suppose that the aim of the experiment is to test Principle C in matrix sentences like (1).

(1) He dusted Superman.

Principle C rules out the interpretation man and he are coreferential. In other words, a sentence like “He, Superman, dusted himself” cannot be true, because the pronoun “he” must refer to a different individual than the name “Superman.” The sentence “He dusted Superman” is therefore incorrect, and the children are expected to say “No.”

As usual, the meaning associated with the sentence tested is set up as the “Yes” response in the experiment. The sentence “He dusted Superman” is set up as a response to the story; the toy character Superman is the character that corresponds to the “Yes” response in the story. The meaning that corresponds to the story is known as the experimental hypothesis.

A hypothetical story that embodies this meaning goes as follows: Superman has a problem. He has cobwebs to brush them off. Unfortunately, he asks his friend Jimmy the cameraman to help. Jimmy says that although he’d like to help, he has other things to do. Lois Lane also walked into cobwebs than Superman and needs help, Superman does his best to dust the cobwebs.

The experimental design can be seen in the following chart, which shows the two meanings of the sentence “He dusted Superman.” The first meaning (*meaning₁*) is grammatical and true in the story, while the second meaning (*meaning₂*) is grammatical but false in the story. Both meanings are represented by the symbol “false.”

<table>
<thead>
<tr>
<th>Sentence: He dusted Superman</th>
<th>Characters: Superman, Jimmy, Lois Lane</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Meaning₁</em>, true: Superman dusted Superman</td>
<td>Jimmy, Superman</td>
</tr>
<tr>
<td><em>Meaning₂</em>, false: Superman dusted Superman</td>
<td>Jimmy, Superman</td>
</tr>
</tbody>
</table>

The research strategy of making the children think about a sentence and then judge whether it is grammatical or not is called the condition of falsification.
(1) He dusted Superman.

Principle C rules out the interpretation of the sentence in which Superman and he are coreferential. In other words, the sentence cannot mean that he, Superman, dusted Superman, or, put differently, that Superman dusted himself.3 The meaning that violates Principle C ("Superman dusted himself") and the meaning that is compatible with the adult grammar ("He [somebody else] dusted Superman") are both incorporated into the story presented to children. One meaning is made to correspond to a "Yes" judgment of the puppet's description of the story and the other is made to correspond to a "No" judgment. This is so that the child's interpretation of the test sentence can be clearly identified.

As usual, the meaning associated with a violation of the principle being tested is set up as the "Yes" response, and the meaning associated with adult linguistic knowledge is set up as the "No" response.4 The meaning that corresponds to the "Yes" response is made true in the story presented to children; the toy character Superman does in fact dust himself in the story. The meaning that corresponds to the "No" response is made false; it is not the case that he (some male) dusts Superman. Crain and Thornton (1998) call the research strategy of making the sentence false on a particular reading the condition of falsification.

A hypothetical story that embodies the two meanings is as follows. Superman has a problem. He has cobwebs on his shoulders and he needs to brush them off. Unfortunately, he can't reach them very well, so he asks his friend Jimmy the cameraman to brush them off for him. However, Jimmy says that although he'd like to help, he can't. He explains that Lois Lane also walked into cobwebs, and she is covered with more cobwebs than Superman and needs his help more. Since Jimmy won't help him, Superman does his best to dust the cobwebs off himself.

The experimental design can be summarized as in (2). The meaning that is ungrammatical but true in the story (i.e., corresponds to a "Yes" response) is represented by *meaning1, true. The meaning that is grammatical but false in the story (i.e., corresponds to a "No" response) is represented by meaning2, false. This notation will be used throughout the discussion of the experimental stimuli.

(2) Sentence: He dusted Superman
Characters: Jimmy, Superman, Lois Lane
*meaning1, true: Superman dusted himself
meaning2, false: Jimmy dusted Superman
Each story has four components according to Crain and Thornton: background, assertion, possible outcome, and actual outcome. The background can be thought of as a function to which the assertion applies and is represented by an elliptical expression. The assertion fills out the ellipsis. The background can be set up in many different ways, depending on what is taken to be the elliptical expression; in (3), it is the NP in object position and is represented by *so-and-so*. What is stated in the assertion should represent a possible outcome of the story—Jimmy does consider dusting Superman, although this is not what eventually transpires. The actual outcome negates the possible outcome; he (Jimmy) didn’t dust Superman— Jimmy dusted Lois Lane instead and Superman was left to dust himself.

(3) Background: He dusted *so-and-so*
Assertion: He (Jimmy) dusted Superman
Possible outcome: Jimmy dusted Superman
Actual outcome: Jimmy dusted Lois Lane
(Superman dusted himself)

One might ask why it is necessary to incorporate the scenario in which Jimmy considers dusting Superman, but decides to dust Lois Lane instead, given that this part of the story does not test whether children allow *He dusted Superman* to mean that Superman dusted himself. The answer is that if a test sentence is presented in a context in which there is no reason to deny it, it is unlikely that children will do so. They might respond according to Principle C and say “No”; but it is more likely that they will either agree with the puppet’s statement because they’re confused or say that they don’t know how to respond. Associated with judging a sentence as false is a pragmatic condition that Crain and Thornton term the *condition of plausible dissent*; it is felicitous to deny a sentence only if the assertion it makes is under consideration. Thus, in the story Jimmy considers dusting Superman but rejects this possibility.

### 4.2.3 Realizing the Actual Outcome

In our example story, the possible outcome is represented by the test sentence itself: *He dusted Superman*. The actual outcome is represented by the negation of the test sentence, which indeed is true of the story; that is, he, Jimmy, *didn’t* dust Superman.

A complication arises for sentences containing quantificational elements such as the universal quantifier *every*. Take a sentence like *Every reindeer brushed him*, one of the stimuli in the knowledge of Principle B. In a story for which the possible outcome is that every reindeer brushed him, the actual outcome entails that every reindeer must be true. The problem is that the two sentences without the ellipsis, depending on whether negation is an NP or not brush him) (see (4b)).

(4) a. NEG S
b. NP NEG VP

The question is, which possibility shows the actual outcome of the story? Crain and Thornton suggest that it is natural to negate the general reading of the verb *brush*; the former takes less cognitive load. This follows from the Principle of Pragmatic Lightness (Bourne and Steedman 1985).

Let us illustrate using the example story acted out for children, there will be some of the interpretation ‘not every reindeer brushed him’ make up the *every reindeer brushes him* true if reindeer A didn’t brush him, if reindeer C didn’t brush him (to name a few, true if none of the reindeer brushed him). In contrast, the reading ‘every reindeer did brush him in the specific circumstance: namely, to brush him. This specific reading takes the reading that is available in the sentence.* Every reindeer brushes him* (some reindeer)

VP ellipsis sentences introduce further conjunctions of events. As illustrated by falsifying the action of the first conjunction only, or by falsifying both conjunctions.

(5) a. NEG S AND S
b. S AND NEG S
c. NEG S AND NEG S
reindeer brushed him, one of the stimuli we used in testing children’s knowledge of Principle B. In a story for which this is the test sentence, the possible outcome is that every reindeer does brush him (some male). However, the actual outcome entails that the negation of the test sentence must be true. The problem is that the test sentence can be falsified in two ways, depending on whether negation is interpreted inside or outside the sentence. That is, it has to be true either that not every reindeer brushed him (see (4a)) or that none of the reindeer brushed him (every reindeer did not brush him) (see (4b)).

(4) a. NEG S
   b. NP NEG VP

The question is, which possibility should be represented in the actual outcome of the story? Crain and Thornton (1998) suggest that it is more natural to negate the general reading of such a sentence than the specific reading; the former takes less cognitive effort to negate than the latter. This follows from the Principle of Parsimony proposed by Crain and Steedman (1985).

Let us illustrate using the example Every reindeer brushed him. In the story acted out for children, there will be several ways to satisfy the truth of the interpretation ‘not every reindeer brushed him’. Suppose three reindeer make up the every reindeer contingent. Then interpretation (4a) is true if reindeer A didn’t brush him, if reindeer B didn’t brush him, or if reindeer C didn’t brush him (to name some of the possibilities). It is even true if none of the reindeer brushed him. This is the general reading. By contrast, the reading ‘every reindeer did not brush him’ is true in only one very specific circumstance: namely, when all of the reindeer failed to brush him. This specific reading takes more cognitive effort to construct. It is suggested, therefore, that the actual outcome of the story accompanying the test sentence Every reindeer brushed him should be that not every reindeer brushed him (some male).5

VP ellipsis sentences introduce further complexities because they involve a conjunction of events. As illustrated in (5), they can be falsified by falsifying the action of the first conjunct only, by falsifying the second conjunct only, or by falsifying both conjuncts.

(5) a. NEG S AND S
    b. S AND NEG S
    c. NEG S AND NEG S
Consider a key VP ellipsis structure that tests Principle B, *Perry trusted him and Clark Kent did too*. In the actual outcome, the negation of this sentence must be true. This can be implemented in several ways: if Perry did not trust him (some male) but Clark Kent did; or if Perry trusted him but Clark Kent did not; or if neither Perry nor Clark Kent trusted him. The question is, which is the most general reading in this case? It seems that the reading in which neither Perry nor Clark Kent trusted him is more specific and would take more cognitive effort to construct. If so, the ideal actual outcome would be one in which either Perry or Clark Kent trusted him. At the time the experiment was carried out, this was not apparent, and our stimuli for these stories fulfilled the condition of falsification by negating both conjuncts. To our knowledge, this topic has not been investigated, and how crucial this manipulation is remains a matter for future research.

If the VP ellipsis test sentences contain quantifiers, there are even more ways to negate them. Our tests of the binding theory sentences containing the universal quantifier every were negated using the more general not every form of negation.

### 4.3 Materials

This section introduces the stimuli used in the experiment. Three types of stimuli were used, testing (a) children’s knowledge of the binding principles in non-VP-ellipsis structures, (b) children’s knowledge of the parallelism constraint in VP ellipsis structures, and (c) the interaction of VP ellipsis and the binding principles.

In designing the materials, we always adopted the experimental hypothesis that children have knowledge of the principles of Universal Grammar. Our stimuli testing Principle B followed this assumption even though we knew from previous literature that some children allow what look like violations of this principle. We have argued extensively that these apparent violations are in fact acceptances of an alternative (nonadult) local coreference interpretation that does not violate Principle B. This being the case, the sentences testing Principle B and Principle C in VP ellipsis structures also may be potentially ambiguous for some children. In the discussion of the stimuli below, we will make clear which items admit, or potentially admit, the extra local coreference option.

### 4.3.1 Controls for Binding and Local Coreference

We began by evaluating children’s knowledge of matrix sentences. We did so because we knew that children would not treat sentences like (7). That is, we predicted that children would say that the quantificational NP in sentence (7) would violate Principle B because doing so would violate Principle C.

Following results reported by Chierchia, we predicted that children would not treat sentences like (7). That is, we predicted that children would say that the quantificational NP in sentence (7) would violate Principle B because doing so would violate Principle C.

Stories were designed that could in principle be true (7). This step guaranteed that (6) and (8) could be true in context, so that a difference in children’s responses to matrix sentences could not readily be attributed to differences in the items like (8), in which the pronouns were repeated, and the items tested in the same experimental context.

Let us consider the design of the stories. A sample design is illustrated in (9).

### 4.3.1.1 Controls for Binding and Local Coreference

(9) Sentence: Every reindeer brushed him.

Characters: 3 reindeer, 1 child

*Meaning*, true: Every reindeer brushed him.

*Meaning*, false: Every reindeer brushes himself.

The interpretation that is grammatical and consistent with the grammatical one, within the context, is made in the case where the child. The meaning ruled out by the sentence is false, meaning*, is marked by a different grammatical reading in which the children mentioned in the test sentence (here...
4.3.1 Controls for Binding and Local Coreference
We began by evaluating children's knowledge of Principles B and C in matrix sentences. We did so because it was important to identify which of our child subjects would have given the nonadult local coreference response to matrix sentences in previous experiments. These children were tested further on VP ellipsis structures.

Following results reported by Chien and Wexler (1990), we predicted that children would not treat sentences like (6) on a par with sentences like (7). That is, we predicted that children would not treat the pronoun and the quantificational NP in sentences like (6) as anaphorically related, because doing so would violate Principle B. By contrast, we expected that at least some of the children who rejected (6) would accept (7), allowing local coreference between the two NPs. Examples like (8), containing reflexives, were controls to show that children could allow a bound variable interpretation of a pronominal form in a one-clause syntactic configuration.

(6) Every reindeer brushed him.

(7) Bert brushed him.

(8) Every reindeer brushed himself.

Stories were designed that could in principle be used to test both (6) and (7). This step guaranteed that (6) and (7) were tested in the same discourse context, so that a difference in children's responses to the two types of sentences could not readily be attributed to discourse factors. Control items like (8), in which the pronoun was replaced by a reflexive, were tested in the same experimental context.

Let us consider the design of the stories that tested sentences like (6)–(8). A sample design is illustrated in (9).

(9) Sentence: Every reindeer brushed him
Characters: 3 reindeer, Bert
*Meaning\textsubscript{1}, true: Every reindeer brushed himself
Meaning\textsubscript{2}, false: Every reindeer brushed Bert

The interpretation that is ungrammatical in the adult grammar is lined up with the grammatical one, within the context of the story presented to the child. The meaning ruled out by Principle B, in which every reindeer brushed himself, meaning\textsubscript{1}, is made available alongside meaning\textsubscript{2}, a grammatical reading in which the pronoun refers to a character not mentioned in the test sentence (here, Bert). The illicit meaning\textsubscript{1} is true,
and is designed to evoke a “Yes” response in order to bias the experiment against our hypothesis that children will obey Principle B. The grammatical meaning is false, and is designed to evoke a “No” response from children who have adult grammatical knowledge of Principle B.

The story line that embodies the design features in (9) is roughly as follows. Bert and three reindeer friends have a snowball fight, and they all get covered in snow. When they go inside, Bert is shivering, so he asks the reindeer to brush the snow off him. Two of the reindeer (separately) refuse, saying they have too much snow to deal with, and they brush themselves. The third reindeer helps Bert a little bit, but then brushes the snow off himself. Bert thanks the helpful reindeer for starting to brush him. He says he’s sorry he can’t reciprocate by helping brush the reindeer; he needs to finish brushing all the snow off himself because he’s still very cold. In the story, then, it is true that every reindeer brushed himself. Therefore, a child who allows meaning1 should respond “Yes” to the puppet’s description Every reindeer brushed him. On the other hand, a child to whom meaning2 is unavailable because it is ruled out by Principle B should take the pronoun to refer to Bert, as in meaning1. This meaning is false, because it isn’t the case that every reindeer brushed Bert; only one of the reindeer did. Therefore, a child who interprets the sentence as having meaning2 should say “No.”

It is worth checking to see how the story satisfies the truth value judgment task “recipe.” The four factors are listed in (10).

(10) Background: Every reindeer brushed so-and-so
    Assertion: Every reindeer brushed Bert
    Possible outcome: Every reindeer brushed Bert
    Actual outcome: Only one reindeer brushed Bert

The background behind the story is that Bert has a problem: he wants the reindeer to brush the snow off him. The assertion that applies to the background is that every reindeer brushed him. In the story, it is possible that all of the reindeer could have brushed Bert. They all consider it, but, as it turns out, two of them say they have too much snow to deal with, and they brush themselves; so only one reindeer brushes Bert. The actual outcome represents the negation of the possible outcome: not every reindeer brushed him. The truth of meaning1 is fulfilled in the final part of the story. The third reindeer brushes himself after helping Bert, so at this point every reindeer has brushed snow off himself.

Experimental Investigations

Now consider how items in which the antecedent, like (7), were tested. Suppose to test this sentence, Bert brushed him, the pronoun referred to as brushed. (Recall that the same story was ever presented to a child more than once, and sentences like (6) and sentences like (7), were ever presented to a child more than once, and sentences like (7), were given to (11).

(11) Sentence: Bert brushed hi
    Characters: Bert, helpful
    *Meaning1, true: Bert brushed himself
    Meaning2, false: Bert brushed himself

The reading corresponding to meaning2 (alternatively, the reading that violates B), is available, meaning2, the pronoun referred to as brushed, is mentioned in the test sentence itself (here) who understood Bert brushed him as having meaning2, however. This sentence to be false and should respond to the story, because Bert did not brush the brushed.

It was necessary to verify that trials like (8) and B (see Chien and Wexler 1990). This was to whether the stories did, in fact, allow the reflexive when it was legitimate. Since reflexives were variable reading, a reflexive was substituted (8). The features of this design are shown in (12).

(12) Sentence: Every reindeer brushed himself.
    Characters: 3 reindeer, Bert
    Meaning1, true: Every reindeer brushed himself.
    *Meaning2, false: Every reindeer brushed himself.

Here, meaning1 corresponded to a grammatical reading, as in (6). In the reflected himself, so children were expected to read. Every reindeer brushed himself. Since this
a grammatical reading, we did not expect
Now consider how items in which the pronoun has a referential NP antecedent, like (7), were tested. Suppose the same reindeer story was used to test this sentence, *Bert brushed him*, that was used to test *Every reindeer brushed him*. (Recall that the same story pattern could be used to test both sentences like (6) and sentences like (7), though in fact no particular story was ever presented to a child more than once.) The features of the design are given in (11).

(11) Sentence: Bert brushed him
Characters: Bert, helpful reindeer
*Meaning₁, true: Bert brushed himself
Meaning₂, false: Bert brushed the helpful reindeer

The reading corresponding to meaning₁ is the local coreference reading (alternatively, the reading that violates Principle B). On the meaning that is available, meaning₂, the pronoun refers to some other character, not mentioned in the test sentence itself (here, the helpful reindeer). Children who understood *Bert brushed him* as having meaning₁ would say “Yes” because Bert did in fact brush himself after the helpful reindeer stopped brushing him. Children with the adult grammar would understand the sentence to have meaning₂, however. These children should judge the sentence to be false and should respond “No” to it as a description of the story, because Bert did not brush the reindeer who helped him. Bert explicitly says that he can’t brush the helpful reindeer, because he’s still cold.

It was necessary to verify that trials like (6) were a valid test of Principle B (see Chien and Waxler 1990). This was achieved by checking to see whether the stories did, in fact, allow a bound variable interpretation when it was legitimate. Since reflexives provide a grammatical bound variable reading, a reflexive was substituted for the pronoun in (6), as in (8). The features of this design are shown in (12).

(12) Sentence: Every reindeer brushed himself
Characters: 3 reindeer, Bert
Meaning₁, true: Every reindeer brushed himself
Meaning₂, false: Every reindeer brushed Bert

Here, meaning₁ corresponded to a grammatical reading, not to a reading ruled out by the grammar, as in (6). In the story, every reindeer did brush himself, so children were expected to respond “Yes” to the description *Every reindeer brushed himself*. Since the “Yes” answer corresponded to a grammatical reading, we did not expect that children would choose
meaning. To interpret the sentence in this way, children would have to allow the reflexive to refer outside the sentence to Bert, which is unlikely, given that this interpretation is not a possible grammatical interpretation. Unlike the target sentences that contain pronouns, this control sentence is unambiguous.

Four trials of each of the sentence types in (6) and (7) were presented to children. The verbs used in these trials were brush, wipe, blow on, and fan. We purposely chose verbs that are not inherently reflexive, to safeguard against the possibility that children would interpret the test sentence without interpreting the pronoun in object position (on this issue, see Wexler and Chien 1985; Chien and Wexler 1990; Grimshaw and Rosen 1990). Two trials of sentences like (8) were presented to children.

We also wanted to demonstrate that children who reject bound variable pronouns that are locally bound as in (6) do not do so because they cannot access bound variable interpretations of pronouns, as Grimshaw and Rosen (1990) suggest. To put it the other way around, we wanted to show that children can indeed access bound variable interpretations of pronouns (in addition to bound variable interpretations of reflexives), thus replicating the results of Chien and Wexler’s (1991) study. Therefore, we included another control, using test sentences of the form Every Cabbage Patch boy said Superman likes him. In such sentences, a bound variable interpretation of the pronoun is grammatical. By demonstrating children’s ability to assign bound variable interpretations to pronouns with quantificational antecedents when it is possible, we aimed to lend credence to our test of Principle B in (6), where the bound variable interpretation is ruled out. The test situation is summarized in (13).

(13) Sentence: Every Cabbage Patch boy said Superman likes him
Characters: 3 Cabbage Patch boys, Superman, Pokey the Horse
Meaning1, true: Every Cabbage Patch boy said Superman likes him (Cabbage Patch boy)
Meaning2, false: Every Cabbage Patch boy said Superman likes Pokey

Because the purpose of the control was to demonstrate that children accept bound variable interpretations of pronouns, this interpretation was set up as the “Yes” response. Given that children have a tendency to say “Yes” if their grammar allows the interpretation in question, the availability of the bound variable really boosted. Thus, meaning1 corresponded to every Cabbage Patch boy claimed that Cabbage Patch boy). However, the alternative, meaning2, was also true as pronoun refers to an individual not mentioned in the Horse. Because both “Yes” and “No” responses to a sentence like (13) is difficult to assign a bound variable interpretation on a preference for the deictic interpretation clearly that the pronoun can be a bound variable sentences like (13) rejects ones like (6), which are due to Principle B and not to any bound variables as suggested by Grimshaw.

Insofar as possible, the plots of the story kept parallel. There were three characters: Superman, (here, the Cabbage Patch boys), and Pokey (the Horse), and the character named in the story line is as follows. Each Cabbage Patch boy that Superman likes him, giving a reason thinks Superman likes Pokey, but the boy would like to go for a ride on Pokey.9 If him would therefore cause children to consider every Cabbage Patch boy said Superman likes him as a possible outcome. The requirements for the story can be summarized as follows:

(14) Background: Every Cabbage Patch boy likes so-and-so
Assertion: Every Cabbage Patch boy likes Pokey
Possible outcome: Every Cabbage Patch boy likes Pokey
Actual outcome: Only one Cabbage Patch boy likes Pokey (Every Cabbage Patch boy likes him)

There were two trials of this type.

The final control testing children’s knowledge in matrix sentences investigated their ability to