LANGUAGE ACQUISITION, *1*(3), 225–295

Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics

Yu-Chin Chien
Department of Psychology
California State University, San Bernardino

Kenneth Wexler
Department of Brain and Cognitive Sciences and
Department of Linguistics and Philosophy
Massachusetts Institute of Technology

We report three experiments concerning English-speaking children's knowledge of locality conditions in the binding of reflexives and pronouns (Principles A and B). The children tested were between the ages of 2;6 and 6;6. By age 6, children know that a reflexive must be locally bound. At the same age, however, they appear to not know that a pronoun may not be locally bound. We suggest that children are missing a pragmatic principle, not the syntactic Principle B. This hypothesis predicts that children will not accept a local antecedent for a pronoun that is a bound variable. Experiment 4 confirms this prediction. We conclude that children know the grammatical principles of binding but do not know a relevant pragmatic principle. We suggest that such dissociation in children might be a useful tool in the study of linguistic theory.

1. INTRODUCTION

In this article, we describe four experimental studies concerning English-speaking children's knowledge of two linguistic principles that concern the

Requests for reprints should be sent to Yu-Chin Chien, Department of Psychology, California State University, San Bernardino, 5500 University Parkway, San Bernardino, CA 92407-2397.

antecedent possibilities for reflexives (e.g., *himself/herself*) and pronouns (e.g., *him/her*). By studying the development of these two principles, we seek to uncover children's knowledge of Binding Theory, which constitutes a major component of adult syntactic knowledge (cf. Chomsky, 1981; Montalbetti & Wexler, 1985). We also hope to elucidate the relationship between the study of linguistic theory and the study of language acquisition. More precisely, we suggest that evidence from language acquisition is directly relevant to the study of linguistic theory, based in part on the result that Principles A and B of the Binding Theory are known to the child, although it takes some nonobvious experimentation to establish the latter.

1.1 Some Basic Issues of Language Acquisition

From our point of view, the major questions of language acquisition concern what children's knowledge of language is at various stages, and how and why development to adult knowledge takes place. To approach these questions we must have: (a) a theory of adult knowledge, (b) a theory of linguistic variation, (c) a theory of learning, and (d) a theory of maturation.

Some of (a–d) might be vacuous in any particular framework. For example, some frameworks might assume that (d) is vacuous, namely that there is no linguistic maturation or it does not apply to a particular case. Other frameworks might assume that (a) is vacuous, or almost so; namely, there is no adult knowledge, only a system of behavior. Here we wish to stress that we are concerned not only with describing children's behavior, but also with seeing how the results fit into theories (a–d). As we have said, from our point of view, this is the only way to answer what the major questions of language acquisition are. Any particular phenomenon concerning child language can be interpreted in various ways, depending (at least) on variations in (a–d). The theories, and interconnections with other phenomena related to the theories, help to narrow down the possible explanations. The point of this article is both to present developmental phenomena concerning binding and to narrow down the class of possible explanations for these phenomena.

The general framework that we adopt (perhaps with some small modifications) is the "principles and parameters" approach. The answer that this framework gives to the question of (a)—the theory of adult linguistic knowledge—is the theory of Universal Grammar (UG). It stipulates that there are innate formal principles of UG known to the child. This framework has received extensive study in the linguistic literature, and we assume particular aspects of it—particularly those concerning Binding Theory—as a given.

The principles and parameters approach also gives a general answer to (b)—the theory of linguistic variation. Namely, in addition to a lexicon that might have language-particular properties, the principles of UG have parameters associated with them, each value of a parameter determining a different language. For each parameter, the children have to learn the value that is correct for their language. They also have to learn the lexicon of their language, to the extent that this is language-particular. This is the general answer that the principles and parameters approach gives to (c)—the theory of learning.

For the phenomena that we are studying in this article, we can actually adopt a somewhat stronger theory than the assumption of linguistic parameters. Namely, we can adopt what Wexler and Chien (1985) called the Lexical Learning Hypothesis, which states that there are no parameters except for varying properties of lexical items. Thus, only lexical items and their properties have to be learned. This is an even more narrow proposal than the parameters proposal because, in any case, a lexicon must be acquired.[1]

The strong result that follows (from either the principles and parameters approach or the Lexical Learning Hypothesis) is that no learning of principles takes place. They are simply innate. The general argument for such a claim is the general basis for linguistic theory, namely the problem of learnability. Simply speaking, no empirically adequate theory of learning can account for the acquisition of the highly specific and structured principles which constitute adult linguistic knowledge. Most contemporary linguistic research (plus the literature on learnability) is a highly detailed argument for this conclusion, and we cannot review this literature here.

Do the answers to (a–c) given so far imply that children will have the same linguistic behavior as adults? Only partially. First, note that there is a learning component, at least of the lexicon, and possibly (if the Lexical Learning Hypothesis is not correct) of principle-related parameters. Thus, some learning might not yet have taken place. (We will consider this question in more detail when we discuss empirical results.) Second, there might be other explanations—not considered in (a–c)—for differing linguistic behavior. An example might be processing limitations or other kinds of cognitive limitations the child might have. The experiments we have

[1]The hypothesis is based on Borer's (1984) hypothesis that language vary only in their lexical and morphological properties. Manzini and Wexler (1987) and Wexler and Manzini (1987) developed the idea, based on their results on variation in binding theory, and called the general statement the "Lexical Parameterization Principle." It is quite likely that properties of lexical *categories* can also vary and must be learned.

conducted on Binding Theory were designed with the goal of overcoming as many of these alternative explanations as possible. Third, we have not yet considered theory (d)—the maturational component of a framework for language development.

Borer and Wexler (1987, 1988a, 1988b) as well as Felix (1984) have argued that some elements of UG might mature. That is, these elements are innate but not available until some time after birth. (This is in addition to whatever nonlinguistic maturation takes place.) The framework we adopt here is that of Borer and Wexler (1988a, 1988b), who developed and provided evidence for the hypothesis of UG-Constrained Maturation. UG-Constrained Maturation states that at every time, the principles of UG constrain the child's grammar. In short, certain capacities might be missing from the child's grammar, capacities that later mature, but no principles are missing. One example that Borer and Wexler discussed is the hypothesis that children do not have knowledge of "argument-chains" until a certain age and thus cannot create verbal passives. No principle is missing. The children do not violate UG in any way. However, they cannot create verbal passives and thus may not have the grammar of the adult language that they are learning. They might interpret sentences concerning verbal passives in other ways—consistent with UG but inconsistent with the adult grammar that they will ultimately develop.

Given UG-Constrained Maturation as the answer to (d), we cannot assume that any principles of UG may be missing from the child's knowledge. Thus, if a child's behavior is different from the adult's in a particular case, a maturational explanation must take a particular form. In addition, the answers that we have given to (a–c) also imply that principles of UG are available to the child at all stages. Any deviance in behavior must follow for some other reason.

The concentration on knowledge of principles also brings with it a particular methodological stance. We are interested in whether children have knowledge of principles of grammar at a certain age. To support a conclusion that children do have such knowledge, we would like children's behavior in a particular experiment to be near perfect (i.e., adult-like), limited only by the extent it seems reasonable that difficulties with an experiment, inattention, and so forth, cause less than perfect behavior. Although this might seem obvious, it is not the way that language acquisition research is always conducted. Rather, experimenters sometimes argue that knowledge is demonstrated whenever children give more "correct" than "incorrect" responses, or, alternatively, whenever children show differing behavior on two kinds of linguistic stimuli (e.g., pronouns and reflexives in our case). Such argumentation, however, does not appear to be related to a particular model of language and its growth. Rather, it appears

related to a hypothesis that "there is a difference in children's behavior on such and such forms." In our case, we are interested in testing children's behavior against predictions of a model. The model is that children have knowledge of certain linguistic principles and that they will behave according to those principles, except for particular difficulties having nothing to do with those principles. Thus, to the extent that we eliminate the experimental difficulties, we expect near perfect performance.

We have devoted some care to these theoretical arguments carefully (although they are still only a sketch) because of the general situation in the study of linguistic-based approaches to language acquisition. Whereas at one time the general assumption was that anything abstract and grammatical had to be learned and therefore would develop late, today it is much more often assumed that children know the abstract linguistic principles innately and that research in language acquisition might be directed toward demonstration of the knowledge of these principles. As already noted, it is important not to confuse knowledge of principles with particular behaviors. One of the major conclusions of the series of experiments reported here (plus others we have reported elsewhere) is that children might have knowledge of linguistic principles even when their behavior in an experiment at first seems to indicate otherwise. To demonstrate such a conclusion requires a careful interplay between precise theorizing and precise experimentation.

1.2 Binding Principles and Their Development

The principles of binding deal with relations between many different kinds of elements in sentences, for example, empty elements such as traces and null pronouns, lexical anaphors such as reflexives, and pronouns. These principles of binding interact with principles from other modules of syntactic knowledge and explain a wide variety of particular syntactic phenomena. From the standpoint of the field of language acquisition, it is essential to understand the development of the knowledge of these binding principles.

In beginning a program to investigate the development of binding phenomena in English, we have chosen to start with the notions of reflexives and pronouns and the corresponding Binding Principles A and B. Binding Principle A as given in (1) may be reformulated as the following: A reflexive must be locally bound. It implies that a reflexive must have an *antecedent,* this antecedent must be *local* (in simple cases, but definitely not in general, this means that the antecedent must be in the same clause as the reflexive), and it must *c-command* the reflexive.

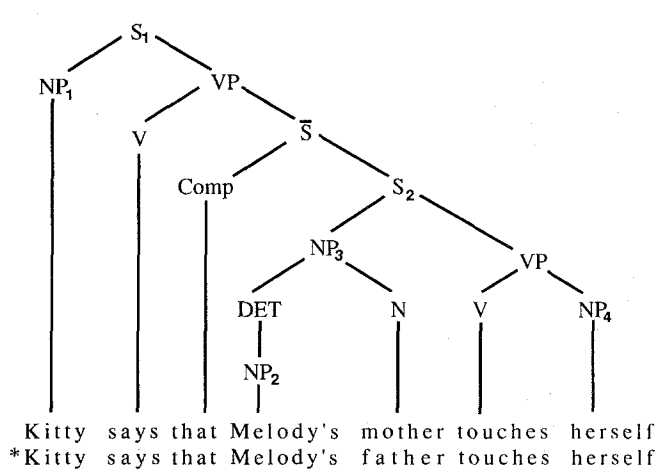
(1) *Binding Principle A:* a reflexive must be bound in its governing category,[2] where "bound" means c-commanded by and coindexed with an antecedent. [*C-command:* in a phrase marker, node A c-commands node B if and only if A and B do not dominate each other, and the first branching node that dominates A also dominates B.]

Sentences (2) and (3) and the corresponding phrase marker (4) illustrate the structural properties exhibited by Principle A.

(2) Kitty$_i$ says [that [Melody's$_j$ mother]$_k$ touches herself$_{*i/*j/k}$].

(3) Kitty$_i$ says [that [Melody's$_j$ father]$_k$ touches herself$_{*i/*j/*k}$].

(4)



In sentence (2) and the ungrammatical sentence (3) (see (4) for the corresponding tree structure), there are three possible antecedents for the reflexive *herself*. They are (NP_1) *Kitty,* (NP_2) *Melody,* and (NP_3) *Melody's mother* (or *father*). Among these three possible antecedents, only *Melody's*

[2]In this article, Wexler and Manzini's (1987, p. 53) parameterized definition of governing category is assumed. The term "governing category" is defined as the following: A is a governing category for B if and only if A is the minimal category which contains B and (a) has a subject or (b) has an INFL or (c) has a TNS or (d) has an indicative TNS or (e) has a root TNS. The value for English reflexive is value (a): The governing category is the minimal category that contains the reflexive and a subject. Although there are different theoretical proposals for the governing category for English, these differences in the definition of "governing category" won't be relevant for the experiments discussed in this article.

mother (or *father*) locally c-commands the reflexive *herself. Kitty* is not a local antecedent even though it c-commands *herself; Melody* is not a c-commanding antecedent even though it is local to *herself.* Sentence (2) is grammatical because the antecedent that locally c-commands the reflexive (i.e., *Melody's mother*) has the same gender as the reflexive. Sentence (3) is ungrammatical because the local antecedent *Melody's father* does not carry the same gender as the reflexive *herself,* so there is no antecedent for *herself.*

In contrast to Principle A, Principle B (as given in (5)) indicates that a pronoun may *not* be locally bound. This implies that, within a sentence, a pronoun may refer only to a *nonlocal* c-commanding antecedent or a *non-c-commanding* antecedent.

(5) *Binding Principle B:* A pronoun must be free in its governing category. [Free = not bound.]

By replacing the reflexive *herself* in (2) with the pronoun *her,* we derive the pronoun sentence (6). In (6), *Kitty* is a possible antecedent for the pronoun *her* because even though *Kitty* c-commands *her, Kitty* does not exist in the local sentence that contains *her. Melody* is also a possible antecedent for the pronoun *her* because even though *Melody* is local to *her; Melody* does not c-command *her.* With the additional possibility that a pronoun may refer to a sentence-external reference in the discourse, sentence (6) is ambiguous in three ways.

(6) Kitty_i says [that [Melody's_j mother]$_k$ touches $\text{her}_{i/j/*k}$].

Aspects of the development of Principles A and B have been studied in different languages, including English, Chinese, Korean, Japanese, Dutch, Icelandic, Italian, and Portuguese (cf. Cairns & McDaniel, 1987; Chien & Wexler, 1987a, 1987b, 1988; Crain & McKee, 1987; Deutsch & Koster, 1982; Deutsch, Koster, & Koster, 1986; Grimshaw, Rosen, & Brown, 1988; Hyams & Sigurjónsdóttir, 1990; Jakubowicz, 1984; Koster, 1984; Lee & Wexler, 1987; Lust, 1986, 1987; McDaniel & Cairns, 1987; McKee, 1988; Otsu, 1981; Sigurjónsdóttir, Hyams, & Chien, 1988; Silva, 1988; Solan, 1983, 1987; Wexler & Chien, 1985, 1988). Wexler and Chien (1985 and later articles) argued that their experimental results demonstrated that (a) children have knowledge of both the c-command and locality aspects of Principle A for reflexives by about age 5;6 (years;months), and (b) children's performance concerning Principle B for pronouns, on the other hand, appears to be greatly delayed compared to Principle A. A reading of

the literature just cited strongly supports this conclusion, as do the first three experiments discussed in this article, which were designed to test not only Principle B but also the locality aspect of Principle A. Unlike adults, in many cases children around age 5;6 still allow a pronoun to be coreferential with a local c-commanding antecedent. (For a discussion of some apparently—but not actually—conflicting data, see Appendix B.)

Recall that the Lexical Learning Hypothesis states that the principles of UG are known to the child; only lexical items and their properties have to be learned. In particular, Principles A and B are unlearned; they are part of the innate endowment that the child brings to the language acquisition task. Lexical items, on the other hand, are learned. This implies that, as long as the child learns that *himself/herself* is a reflexive, and *him/her* is a pronoun, he or she should be able to link these two lexical items to the corresponding principles and to correctly identify their antecedents. (Chomsky, 1981, made this point, deriving it from the theory of UG.) We argue that there is no reason to assume that a pronoun is more difficult for the child to identify as a pronoun than a reflexive is to identify as a reflexive; in fact, the opposite appears to be the case.

Thus, the Lexical Learning Hypothesis (and any theory within the principles and parameters framework that assumes the principles are innate) *appears* to be contradicted by our experimental results (the apparent delay in Principle B). Similarly, we have assumed that maturation is heavily constrained; namely, UG-Constrained Maturation implies that Principle B (as all other principles) must be available to constrain the child from the beginning. Thus, we cannot directly appeal to maturation of Principle B. The answer we give (first suggested, without experimental evidence, in Montalbetti & Wexler, 1985, and Wexler & Chien, 1985) is that the child does know Principle B from the start and that other factors prevent us from seeing this in the earlier experiments. Experiment 4 is the first experimental confirmation of this hypothesis.

In this article, we present experimental data confirming and extending the general findings related to Principles A and B. In particular, we provide evidence concerning children's knowledge of the locality condition for Principle A. In earlier studies (Wexler & Chien, 1985), we demonstrated that children know the c-command aspect of Principle A considerably long before they perform well on Principle B. Given the results of studies in the parameterization of Binding Theory (e.g., Manzini & Wexler, 1987, and Wexler & Manzini, 1987; see footnote 1) it appears that the c-command aspect of Principle A is universal, whereas the locality aspect might be parameterized. In the current studies, we provide evidence that even the locality aspect of Principle A is known before behavior evidenced knowledge of Principle B. If children show good behavior on such a "variable" aspect of linguistic knowledge before the universal aspect of another

principle, it provides even further evidence that the second principle appears to be delayed and that an explanation is necessary.[3]

The remainder of the article is organized as follows. In Section 2, we describe the general methods and procedures for conducting the experiments. In Section 3, we discuss the first three experiments, which focused on children's knowledge of the locality condition in Binding Theory. Section 4 proposes a theoretical explanation for apparent violations of Principle B and discusses an experiment that tests the prediction following the proposed theoretical explanation. Section 5 is the conclusion.

2. GENERAL METHODS AND PROCEDURES

In this section, we discuss the general methodological issues related to the four proposed experiments. The first three experiments concentrate on children's knowledge of the locality condition of Principles A and B. The final experiment tests the hypothesis that children have knowledge of Principle B.

2.1 The Task

In the four experiments, slightly different versions of two experimental techniques were used to test children's comprehension of different sentence constructions containing reflexives or pronouns. These two methods were the Yes/No Judgment task (hereafter, the YNJ task) and the Act-Out task (hereafter, the AO task). In Experiments 1 and 2, a version of the AO task—the Simon-Says game—was applied. In Experiment 3, another version of the AO task—the Party game—was used. In Experiment 4, we employed the YNJ task to elicit subjects' grammaticality judgments of particular sentence constructions. A detailed description of each task is given when individual experiments are discussed.

2.2 The Experimenters

Two experimenters worked as a team in the process of data collection. One administered the task to the subject, the other recorded and observed the subject's behaviors related to the testing situation. The experimenters were native speakers of English carefully selected from a group of undergraduate or graduate students attending the University of California at Irvine. Before

[3]There are a number of other suggestions, besides Manzini and Wexler's model, for accounting for variation in locality conditions on anaphors. These proposals suggest other learning models. For one proposal, see Wexler (1989). However, the general point—that locality varies—remains the same.

actually conducting any experiments, the experimenters were rigorously trained to administer each task with care and accuracy and to present test sentences using natural intonation with a slight stress on the verb preceding the pronoun or the reflexive. By doing this, we made sure that the pronoun or reflexive was not stressed.[4] However, the experimenters were not informed of the underlying hypotheses of each experiment and thus were blind to the predicted results. In this way, we limited the possibility of the experimenters biasing the child's responses.

2.3 The Subjects

For each experiment, we included at least 120 children between the ages of 2;6 and 6;6 (except Experiment 4, in which children between 6;6 and 7;0 were also tested). Only those children who were acquiring English as their first language were included. In Experiments 1, 2, and 3, the children were classified into eight developmental age groups of 6-month periods (with at least 15 subjects in each age group). For example, the first group involved children between 2;6 and 3;0, the final group included children between 6;0 and 6;6. In Experiment 4, the children were classified into four developmental age groups (with at least 40 subjects in each group).[5] The first group involved children under 4, the other three groups included children from 4 to 7 with 1-year intervals. Only those children who had no speech handicap and who were willing to participate in the experiments were included. The child subjects were selected through day care centers, nursery schools, kindergartens, and elementary schools in Orange County, California. In each experiment, we also tested at least 15 adult native speakers using the same methodology. Their results were compared to the results collected from the child subjects. The adult subjects were a random sampling of students attending the University of California at Irvine who participated in the experiment as part of the requirements for an introductory psychology course.

2.4 General Procedures

Each subject was tested individually in an empty classroom by an experimenter. Another experimenter observed and recorded the subject's responses related to the testing situation. Each task was divided into two

[4]We are indebted to Lila Gleitman for this ingenious suggestion as to how to keep our experimenters from stressing the object pronoun or reflexive.

[5]Notice that in each experiment we did not include exactly the same number of subjects in each age group. However, we did set the minimum limit as to include at least 15 subjects in each group for the first three studies and 40 subjects for the final study.

sections: the training section and the testing section. In the former, the subject was first given an opportunity to become familiar with the characters and materials used in each experiment. A set of simple training items was then administered to ensure that the subject understood the testing procedure. The training items did not overlap with the test items in the major structures and the major lexical content. During the training section, the subject's incorrect responses were corrected and the training items were repeated to the subject as many times as needed. In the testing section, a set of several test sentence batteries were given to the subject in a random order according to an established technique tailored for each experiment. The test sentences in each battery were also arranged in a random order. No correction of incorrect responses was allowed during the testing section. The test sentences were not given more than two times. Only natural and positive reinforcement (e.g., "very good," "you have done a good job") was used in the testing section. In each testing period, at most two test batteries were given to each subject. Usually it took about two to four testing periods (on different days) to complete one experiment. The times of visits depended on the number of test items and batteries included in each experiment and the child's willingness to complete the test. In all sessions, a cassette tape recorder was used to record the experimenter's presentation of the sentences and the subject's responses to these sentences.

2.5 Scoring and Analyses

For the first three experiments using the AO task, the subjects' responses were classified into different categories depending on the types of coreference judgments (+CR) they had made for each test construction. For example, +CR with the local c-commanding antecedent, +CR with the nonlocal c-commanding antecedent, +CR with the sentence-external referent. For the final experiment, using the YNJ task, the subjects' responses to different questions were scored as yes or no. The number of yes or no responses or the number of coreference judgments in different categories were then subject to analysis. For simplicity, in the text we discuss only group averages. Appendix A provides several individual subject analyses.

3. EXPERIMENTS 1, 2, AND 3

In this section, we discuss the design and the results of the three experiments formulated to test children's knowledge of the locality condition in Binding Theory.

3.1 Experiment 1

Experiment 1 was designed to test children's knowledge of the locality part of Principle A. The sentences contained two potential antecedents for the reflexive—both antecedents c-commanded the reflexive, but only one was local (i.e., in the governing category of the reflexive). Adults were expected to choose the local antecedent as correct for the reflexive cases, and the point of the experiment was to see if children would do the same. The experiment also tested knowledge of Principle B by using pronouns alternately with the reflexives. Adults were expected to choose the nonlocal antecedent or the sentence-external referent as correct for the pronoun cases, and the point of the study was to see what children would do on these same pronoun cases. Note that the use of both reflexives and pronouns in the same experiment has a particular advantage. As the reflexives and pronouns are in complementary distribution, good performance on both kinds of sentences means that the child cannot be performing according to some nonlinguistic strategy concerning concepts such as "order of mention" or "minimal distance" in terms of word distance.

3.1.1 Method

3.1.1.1 Subjects. In Experiment 1, we tested 156 children and 21 adults. The mean age of the child subjects was 4;6.

3.1.1.2 The design of the test constructions. Three sentence types were included: reflexive sentences (as in (7)), pronoun sentences (as in (8)), and gender control (GC) pronoun sentences (as in (9)).

(7) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should point to $\left\{\begin{matrix}\text{herself}\\ \text{himself}\end{matrix}\right\}$.

(8) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should point to $\left\{\begin{matrix}\text{her}\\ \text{him}\end{matrix}\right\}$.

(9) $\left\{\begin{matrix}\text{Snoopy}\\ \text{Kitty}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should point to $\left\{\begin{matrix}\text{him}\\ \text{her}\end{matrix}\right\}$.

As can be seen from (7–9), each test sentence involved a matrix verb *say* and a tensed complement. The reflexive and pronoun sentences ((7) & (8)) were designed so that the reflexive or pronoun would agree in gender with both the embedded subject and the matrix subject. Thus, on these sentences, knowledge of gender is not sufficient to allow correct performance. That is, gender is not a cue. On the other hand, the gender control pronoun

sentences (9) were designed so that the pronoun would agree in gender only with the correct antecedent (i.e., the matrix subject). These sentences were included to demonstrate that the child *can* refer the pronoun to the nonlocal antecedent. That is, if the child performs appropriately on the gender control sentences, then we cannot conclude that some performance limitation is keeping him or her from performing correctly in the original (nongender control) sentences.

In the test sentences, the child's or adult's name was used as the embedded subject. For example, if the child being tested was Sarah, then sentence (7) would be presented to her like this: Kitty says that Sarah should point to herself. Two sets of matrix subject noun phrases (NPs) were used: Snoopy and Kitty as one set, Garfield and Melody as the other. In the test sentences, there were two potential antecedents for the reflexives or pronouns, of these two potential antecedents, only the child's name (e.g., Sarah or Adam) *locally c-commanded* the reflexive or the pronoun. The matrix subject NP (e.g., Kitty or Snoopy) did not. Five different verbs (*touch, point to, scratch, pinch,* and *tickle*) were used. There were two items for each verb, yielding 10 sentences for each sentence type, and a total of 30 items for each subject.

3.1.1.3 The task. In this experiment, a version of the AO task called the Simon-Says game was used. In the Simon-Says game, the experimenter held two puppets (e.g., Snoopy and Kitty) and read a test sentence, such as (7). The child was asked to perform an action whenever he or she heard "Kitty says" or "Snoopy says." Suppose the name of the chid being tested is Sarah. The sample procedure for the task is as follows:

> Sarah, do you know how to play the game Simon-Says? Each time you hear "Simon-Says" you do what he says, right? Today, we are going to play a game just like that. We will use this girl puppet "Kitty" and this boy puppet "Snoopy" and each time you hear "Kitty-says" or "Snoopy says" you should do what Kitty or Snoopy says, okay? [The experimenter reads a test sentence.] "Kitty says that Sarah should point to herself." [The child performs the action.]

3.1.2 Results and Discussion

The results are illustrated by the lines with black squares in Figures 1, 3, and 4 (the lines with white squares should be ignored for now). Notice that, in all of these graphs, on the abscissa we plot ages in 6-month intervals. Group 1 (G1) consists of children between the ages of 2;6 and 3;0. Group 8 (G8) consists of children from 6;0 to 6;6. On the ordinate, we plot percentage of items.

The major findings of this study are summarized as follows:

Experiments 1, 2, and 3: The Results

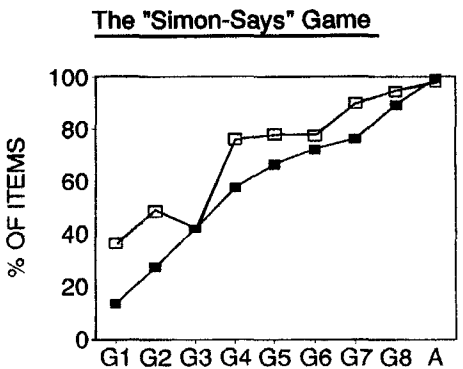


FIGURE 1 Reflexive

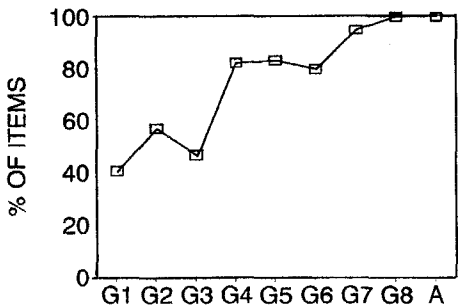


FIGURE 2 Gender Control Reflexive

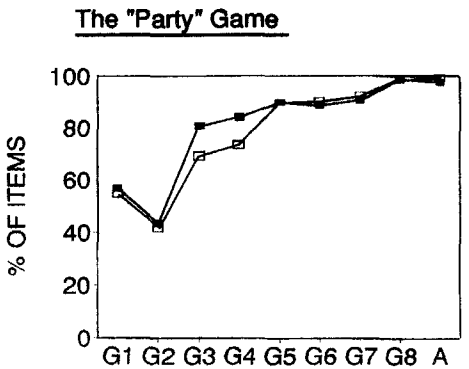


FIGURE 5 Reflexive

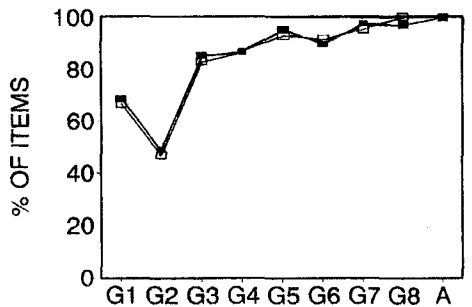


FIGURE 6 Gender Control Reflexive

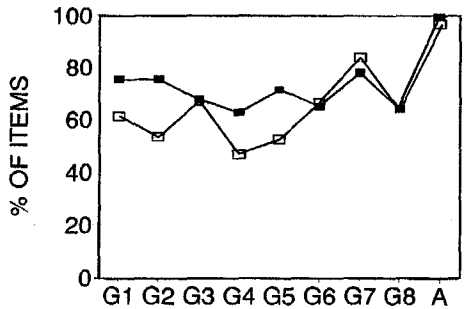


FIGURE 3 Pronoun

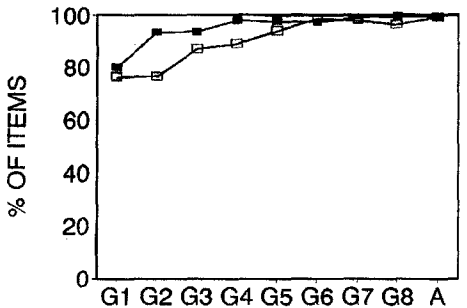
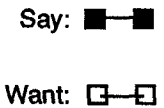
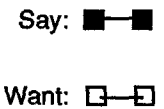


FIGURE 4 Gender Control Pronoun





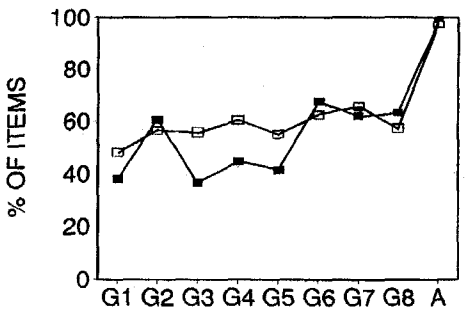


FIGURE 7 Pronoun

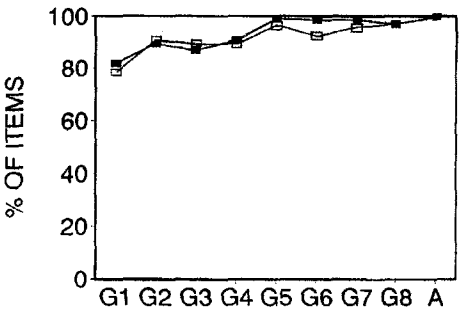


FIGURE 8 Gender Control Pronoun

Note: Correct response: Reflexive = Local C-Commanding Antecedent;
Pronoun = Non-local C-Commanding Antecedent

(A) The line with black squares in Figure 1 illustrates subjects' correct responses to the reflexive sentences (e.g., taking *herself* to be the child Sarah in sentences like "Kitty says that Sarah should point to herself"). As indicated by this line, children older than 6;0 (i.e., the G8 children) knew the major property of reflexives, that the antecedent must be local. Their percentage correct was 89.4%. This line also indicated that children's performance on the locality property of reflexives increased continuously from about the 13% level at age 2;6 to almost perfect performance at 6;0. (Coindexation of the reflexive with the local c-commanding antecedent G1 to G8: 13.30%, 27.30%, 41.90%, 57.50%, 66.60%, 72.50%, 76.50%, 89.40%.)

In many cases (about 39% of the time overall), children made anaphora mistakes by coindexing the reflexive with the nonlocal c-commanding antecedent. The amount of this type of anaphora mistake decreased from 75% to 10% (reflexive = nonlocal c-commanding antecedent: G1 to G8: 74.80%, 66.80%, 49.00%, 41.50%, 33.10%, 27.10%, 23.50%, 10.00%). In a few cases, children took the reflexive to refer to the puppet that was not mentioned in the test sentence (i.e., a sentence-external referent, which disagree in gender with the reflexive). An example would be if Sarah pointed to Snoopy when she heard "Kitty says that Sarah should point to herself." The amount of this type of incorrect responses (i.e., reflexive = sentence-external referent) ranged between 0% to 11.40%. Adult subjects gave perfect responses to the reflexive sentences. About 99.50% of the time, they coindexed the reflexive with the local c-commanding antecedent.

(B) In Figure 3, the line with black squares illustrates subjects' correct responses to the pronoun sentences (e.g., taking the pronoun *her* to be the puppet Kitty in sentences like "Kitty says that Sarah should point to her"). As indicated by this line, children in the age range of 6;0 to 6;6 still did not show the steady behavior predicted by Principle B grammaticality judgments. They did not always obey the Principle B constraint that a pronoun may not have a local c-commanding antecedent. Group 8 children showed only 64% correct. That is, about 36% of the time these G8 children violated Principle B and allowed a pronoun to be coreferential with an antecedent that locally c-commands it. This line also indicates that children's performance on the requirement that pronouns might not have a local c-commanding antecedent stayed roughly flat from 2;6 to 6;6, with only a slight improvement. This flat curve was in direct contrast to the steady increase for the reflexive. (Pronoun = local c-commanding antecedent, i.e., Principle B violation: G1–G8: 10.10%, 20.00%, 24.30%, 36.50%, 27.80%, 34.50%, 21.30%,

36.00%. Pronoun = nonlocal c-commanding antecedent, i.e., performance in accord with Principle B: G1–G8: 75.30%, 75.40%, 67.60%, 78.00%, 62.50%, 71.40%, 64.00%.) In only a few cases, children referred the pronoun to the puppet not mentioned in the test sentence (pronoun = sentence-external referent: ranged between 0% to 13.30%). Adult subjects showed perfect responses to the pronoun sentences (pronoun = nonlocal c-commanding antecedent: 99.50%).

(C) The line with black squares in Figure 4 illustrates subjects' correct responses to the gender control pronoun sentences (e.g., taking the pronoun *him* to be the nonlocal c-commanding antecedent *Snoopy* in sentences like "Snoopy says that Sarah should point to him"). As indicated by this line, when there was a gender match between the pronoun and its correct nonlocal c-commanding antecedent, children paid attention to this gender matching cue and made correct judgments perfectly. Even the youngest children (2;6–3;0) were 80% correct. These results, together with those in Figure 3, show that Principle B does not appear to be constraining the children (as mentioned earlier, we modify this result later), and a behavioral strategy against nonlocal antecedents does not appear to be constraining them. Rather, the pronouns could take either antecedent and the children used other cues to determine their choice. Gender for pronouns was very well established. (Pronoun = nonlocal c-commanding antecedent: G1–G8: 80%, 93.30%, 93.70%, 98.00%, 97.70%, 99.30%, 100%. Adult: 99.50%.) In only a few cases, children mixed up genders (2.5% overall) and referred the pronoun to the different gender puppet that was not mentioned in the sentences. Most of these kinds of mistakes were found in our youngest group. Children younger than 3;0 made this type of mistake about 19% of the time.

To highlight the main findings: Children between the ages of 6;0 and 6;6 demonstrated knowledge of Principle A but appeared to allow violation of Principle B when no extra gender cues were available.

From the standpoint of the Lexical Learning Hypothesis, the developmental results of our reflexive sentences were predicted. That is, a certain period of time was expected and required for a child to complete the learning task that *herself/himself* was a reflexive. This learning, according to our results, was completed around the age of 6;0 (i.e., the age at which almost perfect performance was observed). The question remaining unsolved, however, was why the youngest children consistently chose a nonlocal antecedent for the reflexive. That is, why were the Group 1 children only 13% correct, instead of about 50% correct, which was the chance expectation? Did this represent part of children's grammatical

knowledge at this age, or were other pragmatic factors (e.g., saliency of the puppet in the experiment) determining the results at this age?

On the other hand, the developmental results for our pronoun sentences were not consistent with the prediction generated by the Lexical Learning Hypothesis. Only very little improvement from the youngest to the oldest children was observed. At the age of 6;6, there was still a large number of violations to the requirement that pronouns may not have a local c-commanding antecedent. The question was then: "At what age would this learning be completed and why was there such a time lag between the learning of reflexives as reflexives and that of pronouns as pronouns?"

3.1.3 Rationale for Experiments 2 and 3

There were at least three questions that we sought to study in the two experiments, which are discussed next.

(A) It has been claimed by some researchers (e.g., Solan, 1987) that children show different binding effects between sentences with infinitival complements and sentences with tensed complements, even for pronouns or reflexives in the object position of the complement. Such a result would be at variance with our hypotheses because the governing category for an object in a complement sentence is that sentence, regardless of whether the sentence is tensed or infinitival. Thus, Binding Theory predicts that the behavior of these elements in object position of tensed or infinitival complements is the same.

Because we assume that children know UG (the Lexical Learning Hypothesis) and because the lexical elements are the same, children should either behave appropriately or inappropriately on the sentences regardless of whether the complements are tensed or infinitival. Solan's experiments are not telling on this point because his infinitival complements were control structures, with empty subjects of the complements. Thus, it could simply be that his subjects had not mastered these control structures. (See Wexler, 1989, for the hypothesis that control takes some time to develop, based on a hypothesis compatible with UG-Constrained Maturation.) Thus, it is more relevant to use infinitival complements that contain subjects (though it is conceivable that these develop later than the ages under consideration here; we know of no evidence for this possibility, however).

Our previous experiment used tensed complements. In order to test these differential effects on binding, in the succeeding experi-

ment we employed sentences containing the matrix verb *want* and an infinitival complement using exactly the same methodology, namely the Simon-Says game, and also a different methodology (i.e., the Party game).

(B) We wished to replicate and to investigate in more detail the results from our previous study, especially those involving the younger children's responses to the reflexive sentences and the older children's responses to the pronoun sentences. First, we wanted to find out why young children, when dealing with reflexive sentences, systematically choose the long-distance nonlocal antecedent. Second, the violation of Principle B result seemed so important that we wanted to investigate whether the result would hold up using different experimental methods and with different linguistic materials.

(C) Children's almost perfect responses to the gender control pronoun sentences indicated that children did pay attention to this extra cue. This result showed that children were paying attention to the grammatical structure, and also no performance limitation was keeping children from performing correctly in the original (nongender control) sentences. In our succeeding experiments, an additional sentence type, the gender control reflexive sentences, was included. This was designed to investigate whether gender also provides an effective cue to reflexive sentences.

3.2 Experiment 2

Experiment 2 was designed to test infinitival structures and to also use the gender control for reflexives.

3.2.1 Method

3.2.1.1 Subjects. In Experiment 2, we tested 142 children and 20 adults. The mean age of the child subjects was 4;5.

3.2.1.2 The design of the test constructions. Four sentence types were included: reflexive sentences (such as in (10)), pronoun sentences (as in (11)), GC reflexive sentences (as in (12)), and GC pronoun sentences (as in (13)).

(10) $\begin{Bmatrix}\text{Kitty}\\ \text{Snoopy}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to point to $\begin{Bmatrix}\text{herself}\\ \text{himself}\end{Bmatrix}$.

(11) $\begin{Bmatrix}\text{Kitty}\\ \text{Snoopy}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to point to $\begin{Bmatrix}\text{her}\\ \text{him}\end{Bmatrix}$.

(12) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to point to $\begin{Bmatrix}\text{herself}\\ \text{himself}\end{Bmatrix}$.

(13) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to point to $\begin{Bmatrix}\text{him}\\ \text{her}\end{Bmatrix}$.

All sentences included in this experiment contained the matrix verb *want* and an infinitival complement. In the test sentences, there were two potential antecedents (e.g., the matrix subject NP *Kitty* or *Snoopy* and the embedded subject NP—the child's name) for the following reflexives or pronouns. Again, the reflexive and pronoun sentences were designed so that the reflexive or pronoun would agree in gender with both the embedded subject and the matrix subject. The GC pronoun and GC reflexive sentences were designed so that the pronoun or reflexive would agree in gender only with the correct antecedent (i.e., the matrix subject for the pronoun and the embedded subject for the reflexive). In these sentences, only the child's name or the adult's name (e.g., Sarah or Adam) locally c-commanded the reflexive or the pronoun; the matrix subject NP (e.g., Kitty or Snoopy) did not. Two sets of matrix subject NPs were used: Snoopy and Kitty as one set, Garfield and Melody as the other. Five different actional verbs (*touch, point to, scratch, pat,* and *tickle*) were included. There were two items for each verb, yielding 10 sentences for each sentence type and a total of 40 test items for each subject.

3.2.1.3 The task. This experiment employed the same task as Experiment 1 (i.e., the Simon-Says game). The procedures used were exactly the same as those introduced in Experiment 1, except that in the instructions, the word *say* was replaced by the word *want*. For example, the subject was asked to perform an action whenever Kitty or Snoopy wanted him or her to do so.

3.2.2 Results and Discussion

The results of this experiment are illustrated by the lines with white squares in Figures 1 to 4 (the lines with black squares should be ignored for now). In general, these results replicated the results indicated in our previous study. The major findings are summarized as follows:

(A) Considering Figure 1 again, the line with white squares showed that children older than 5;6 (i.e., the G7 & G8 children) knew the major property of reflexives when dealing with the *want*-reflexive sentences

consisting of infinitival complements (e.g., referring *herself* to the child's name Sarah in sentences like "Kitty wants Sarah to point to herself"). We found that the knowledge that the reflexive must have a local antecedent was revealed at a slightly earlier age with these *want*-reflexive sentences than the *say*-reflexive sentences, which involved tensed complements. The line with the white squares in Figure 1 also indicated that children's performance on the locality property of reflexives continuously increased from about 36% at age 2;6 and approximated the adult's level around age 5;6 (reflexive = local c-commanding antecedent: G1–G8: 36.20%, 48.80%, 41.90%, 75.90%, 77.90%, 77.50%, 90.00%, 94.70%). In many cases, the youngest children chose the nonlocal antecedent for the reflexive (G1–G8: 46.80%, 35.30%, 22.30%, 20%, 21.30%, 8.70%, 5.30%).

These results replicated the reflexive results shown in Experiment 1. In only a very few cases (2.90% overall) did children take a reflexive to refer to a sentence-external referent (i.e., the puppet not mentioned). The adult subjects gave almost perfect responses to the *want*-reflexive sentences. They coindexed the reflexive to the local c-commanding antecedent about 98% of the time.

(B) In Figure 3, the line with white squares illustrates subjects' correct responses to the *want*-pronoun sentences (e.g., referring the pronoun *her* to the puppet Kitty in sentences like "Kitty wants Sarah to point to her"). This line indicated that children in the same age range (5;6 to 6;6) still did not show adult-like behavior in accordance with Principle B. Their performance on the requirement that pronouns may not have a local c-commanding antecedent does not change too much in the age range we have studied (2;6 to 6;6). In our oldest age group, it still remains at only about 64% correct. (Pronoun = nonlocal c-commanding antecedent, i.e., performance in accord with Principle B: G1–G8: 61.30%, 53.50%, 66.70%, 46.80%, 52.60%, 66.30%, 83.80%, 64.00%.)

On the average, children appeared to violate Principle B about 33% of the time, allowing coreference between the pronoun *her* and the local c-commanding antecedent (i.e., the child's name; G1–G8: 25.00%, 32.30%, 24.30%, 48.46%, 44.20%, 33.70%, 15.00%, 36.00%). About 4.30% of the time (between 0% and 11.20%), children referred the pronoun to the puppet not mentioned in the sentence. When dealing with the *want*-pronoun sentences, our adult subjects made perfect responses (pronoun = nonlocal c-commanding antecedent: 97%). These pronoun results again replicated our previous pronoun results.

(C) Comparing the line in Figure 2 with the line with white squares in Figure 1, we found that gender match between the reflexive and its

local c-commanding antecedent did not help children very much in making correct coreference judgments (reflexive = local c-commanding antecedent: G1–G8: 40.60%, 57.00%, 46.70%, 82.30%, 83.20%, 80.00%, 95.00%, 100%; adult: 100%).

As indicated in Figure 4, gender match between the pronoun and its nonlocal antecedent, however, had a strong effect on the pronoun responses (pronoun = nonlocal c-commanding antecedent: G1–G8: 76.20%, 76.50%, 87.10%, 89.10%, 94.20%, 98.70%, 98.10%, 96.60%; adult: 99.50%). The result regarding children's responses to the GC pronoun sentences also replicated the result of Experiment 1.

Why doesn't gender information increase performance very much on reflexives, but it increases performance a great deal for pronouns? There are various possibilities. One is that children simply do not know the gender feature of the reflexive, but they do know it for the pronoun. It is quite possible that in languages, there is an implicational relationship such that if gender is distinctive for reflexives it is also distinctive for pronouns, but the converse does not hold. Thus, for some reason, children learn the gender feature more slowly for reflexives than for pronouns. A second possibility is that the errors on reflexives do not represent lack of linguistic knowledge. Rather, children may be making mistakes because they are not paying much attention or their capacities are too strained or for similar reasons. However, the fact that gender information increases performance on pronouns makes this second possibility unlikely.

To emphasize the major findings: This experiment replicates Experiment 1 in showing that children older than 5;6 know the major property of reflexives, Principle A, but they still appear to show violations of Principle B.

3.3 Experiment 3

Recall that the youngest children in the previous experiments tended to select a nonlocal antecedent for both pronouns and—especially strikingly given Principle A—reflexives. On the surface, for reflexives, this seems to contradict the markedness hierarchy for governing categories of reflexives that is induced by the Subset Principle (Manzini & Wexler, 1987; Wexler & Manzini, 1987). This hierarchy predicts that a child will never set the governing category parameter at a value that is higher than the value of the governing category parameter in the language he or she is learning. However, we suspected that at this age the child still had not determined that *himself/herself* was a reflexive. Rather, the child was using a pragmatically based strategy for responding.[6]

[6]Alternatively, the child might know that *himself/herself* were reflexives but might not have set the governing category value. This seems a less preferable hypothesis.

Thus, Experiment 3 was designed to test whether the youngest children believed as a matter of linguistic knowledege that reflexives and pronouns need nonlocal antecedents. We reasoned that if we could manipulate the children's responses by pragmatic cues, their behavior in our earlier experiments was most likely a response bias, rather than a demonstration of linguistic knowledge. For example, perhaps the 2½-year-olds would rather point at a puppet than themselves in our Simon-Says task. Therefore, we created the Party game, a situation in which we expected the response bias to be eliminated or decreased by making the local response (an act of giving to oneself) more attractive to the child. That is, we expected the child to often prefer to give him- or herself a toy than to give the toy to a puppet. If long-distance (i.e., nonlocal) antecedents were truly preferred as a matter of linguistic knowledge, then we expected that the child would not be influenced by this pragmatic bias.

3.3.1 Method

3.3.1.1 Subjects. In this experiment, we tested 174 children and 20 adults. The mean age of the child subjects was 4;5.

3.3.1.2 The design of the test constructions. In this experiment, four sentence types were included: reflexive sentences (such as in (14–17)), pronoun sentences (as in (18–21)), gender control reflexive sentenes (as in (22–25)), and gender control pronoun sentences (as in (26–29)).

(14) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should give $\left\{\begin{matrix}\text{herself}\\ \text{himself}\end{matrix}\right\}$ a car.

(15) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should give a car to $\left\{\begin{matrix}\text{herself}\\ \text{himself}\end{matrix}\right\}$.

(16) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ wants $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ to give $\left\{\begin{matrix}\text{herself}\\ \text{himself}\end{matrix}\right\}$ a ball.

(17) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ wants $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ to give a ball to $\left\{\begin{matrix}\text{herself}\\ \text{himself}\end{matrix}\right\}$.

(18) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should give $\left\{\begin{matrix}\text{her}\\ \text{him}\end{matrix}\right\}$ a popsicle.

(19) $\left\{\begin{matrix}\text{Kitty}\\ \text{Snoopy}\end{matrix}\right\}$ says that $\left\{\begin{matrix}\textit{Sarah}\\ \textit{Adam}\end{matrix}\right\}$ should give a popsicle to $\left\{\begin{matrix}\text{her}\\ \text{him}\end{matrix}\right\}$.

(20) $\begin{Bmatrix}\text{Kitty}\\ \text{Snoopy}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to give $\begin{Bmatrix}\text{her}\\ \text{him}\end{Bmatrix}$ a napkin.

(21) $\begin{Bmatrix}\text{Kitty}\\ \text{Snoopy}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to give a napkin to $\begin{Bmatrix}\text{her}\\ \text{him}\end{Bmatrix}$.

(22) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ says that $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ should give $\begin{Bmatrix}\text{herself}\\ \text{himself}\end{Bmatrix}$ a cup.

(23) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ says that $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ should give a cup to $\begin{Bmatrix}\text{herself}\\ \text{himself}\end{Bmatrix}$.

(24) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to give $\begin{Bmatrix}\text{herself}\\ \text{himself}\end{Bmatrix}$ a crayon.

(25) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to give a crayon to $\begin{Bmatrix}\text{herself}\\ \text{himself}\end{Bmatrix}$.

(26) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ says that $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ should give $\begin{Bmatrix}\text{him}\\ \text{her}\end{Bmatrix}$ a whistle.

(27) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ says that $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ should give a whistle to $\begin{Bmatrix}\text{him}\\ \text{her}\end{Bmatrix}$.

(28) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to give $\begin{Bmatrix}\text{him}\\ \text{her}\end{Bmatrix}$ a spoon.

(29) $\begin{Bmatrix}\text{Snoopy}\\ \text{Kitty}\end{Bmatrix}$ wants $\begin{Bmatrix}\textit{Sarah}\\ \textit{Adam}\end{Bmatrix}$ to give a spoon to $\begin{Bmatrix}\text{him}\\ \text{her}\end{Bmatrix}$.

Half of the test sentences in each type involved the matrix verb *say,* which subcategorizes for a tensed complement; the other half involved the matrix verb *want,* which subcategorizes for an infinitival complement. In addition, we used both types of dative constructions (i.e., Sarah gives him a spoon, and Sarah gives a spoon to him). Again, two sets of matrix subject NPs were included. Set 1 included Kitty and Snoopy. Set 2 included Garfield and Melody. There were 4 items per condition (2 for each of two dative constructions used), yielding a total of 32 sentences for each subject.

3.3.1.3 The task. In this experiment, a version of the AO task called the Party game was used, in which the experimenter set up a situation with a female and a male puppet sitting in front of the child. A big plate containing different small toys or props was put in the middle of a table

between the child and the puppets. In front of the child and each puppet, there was an empty bowl to keep toys. The child was expected to take a small toy from the center plate and put it into either his or her own bowl or one of the puppet's bowls, according to the sentence presented to the child. Suppose the name of the child being tested was Sarah. A sample procedure for the task is as follows:

> Sarah, let us play a game with Kitty and Snoopy, okay? We are going to have some small parties with these two puppets. At the party, Kitty or Snoopy will ask you to help out with something. For example, to get a toy from the big plate and put the toy into your own bowl or one of the puppet's bowls. You should do what the puppet asks, okay? [The experimenter reads a test sentence.] "Kitty says that Sarah should give herself a car." [The child performs the action.]

3.3.2 Results and Discussion

The results regarding the *say*-sentences are illustrated by the lines with black squares in Figures 5 to 8; those regarding the *want*-sentences are illustrated by the lines with white squares in the same figures. The major findings are summarized as follows:

(A) Concerning the reflexive sentences, a very strong within-task (i.e., the Party game) and between-complement-type consistency (*want*-infinitive vs. *say*-tensed) was found in this experiment. As indicated by the two lines in Figure 5, children older than 4;6 (i.e., G5 to G8 children) behaved as if they knew that the antecedent of the reflexive must be local. Only a slightly higher performance was found in G3 and G4 (3;6–4;6) for the *say*-reflexive than the *want*-reflexive sentences. For all the other groups, the response patterns to these two sentence types were almost the same. These two lines also show that children's performance on the locality property of reflexives increases continuously from the chance level (50%) at age 2;6 to almost perfect performance at 6;6. (Reflexive = local c-commanding antecedent: *say* sentences (G1–G8): 56.75%, 43.50%, 80.75%, 84.25%, 90.00%, 88.75%, 90.75%, 98.50%; *want* sentences (G1–G8): 55.00%, 42.00%, 69.25%, 73.75%, 90.00%, 90.00%, 92.00%, 98.50%.)

On the average, about 16.75% of the time, children made anaphora mistakes by referring the reflexive to the nonlocal c-commanding antecedent. This type of anaphora mistake stayed between the levels of 1.50% and 47.50%, and mainly occurred in children younger than 3;6 (i.e., children in G1 and G2). A large decrease in this error type was found at G3. (Reflexive = nonlocal

c-commanding antecedent: *say* sentences (G1–G8): 41.75%, 46.00%, 15.75%, 13.25%, 8.25%, 8.75%, 9.25%, 1.50%; *want* sentences (G1–G8): 38.25%, 47.50%, 25.00%, 22.25%, 7.50%, 10.00%, 6.50%, 1.50%.)

Only in a few cases (both *say* and *want* conditions) did children take the reflexive to have a sentence-external referent. This type of error ranged between the levels of 0% and 8%. Adult subjects gave nearly perfect responses to both types of reflexive sentences (reflexive = local c-commanding antecedent: *say:* 97.50%; *want:* 98.75%).

In general, compared with Experiments 1 and 2, children showed a higher percentage of correct responses to the reflexive sentences with the Party game (*say:* 80.25; *want:* 76.75%) than with the Simon-Says game (*say:* 57.30%; *want:* 57.30%). Comparing Figure 5 with Figure 1, the knowledge that the reflexive must be c-commanded by its local antecedent attained the 90% level in the age range of 4;6–5;0 for the Party game and in the age range of 5;6–6;6 for the Simon-Says game.

It is important to note that when the Party game rather than the Simon-Says game was introduced, the very young children's systematic tendency to coindex the reflexive with the nonlocal referent disappeared. The choice of local or nonlocal antecedents became more random and close to the chance level. (For Group 1, in Figure 1, the Simon-Says game, the local antecedent was selected 13% and 35% for the two kinds of sentences; in Figure 5, the Party game, the local antecedent was selected about 55% for both kinds of sentences.) The results, as we predicted, suggest that the very young children did not think that a reflexive must have a long-distance (nonlocal) antecedent as a matter of linguistic knowledge.

(B) As indicated by the two lines in Figure 7, children at 6;6 still appeared to allow violation of Principle B. In other words, they allowed a pronoun to have a local c-commanding antecedent. This result was consistently found in both the *say*- and the *want*-pronoun constructions. A higher performance was found in G3 to G5 (3;6–5;0) for the *want*-pronoun than the *say*-pronoun sentences. For the other groups, the response patterns are similar for the two types of constructions.

Performance on the requirement that pronouns not have a local c-commanding antecedent again stays relatively flat from 2;6 to 6;6. For both types of constructions, it still remains at only about 60% correct in the oldest age group. (Pronoun = nonlocal c-commanding antecedent: *say* sentences (G1–G8): 38.25%, 60.50%, 36.50%, 44.75%, 41.75%, 67.50%, 61.75%, 63.25%; *want* sentences (G1–G8): 48.25%, 56.50%, 55.75%, 60.50%, 55.00%, 62.50%, 65.75%, 57.25%.)

Concerning the pronoun sentences, an important and consistent finding should be noted, namely that children's Principle B violations adopt a very similar pattern in all three experiments, regardless of the between-task and between-complement-type differences. When the extra gender cue is not available, this type of mistake does not decrease as a function of age. In addition, the youngest children produced more nonlocal responses for the Simon-Says game than for the Party game. (For Group 1, in Figure 3, the Simon-Says game, the nonlocal antecedent was selected 75% and 61% for the two kinds of sentences; in Figure 7, the Party game, the nonlocal antecedent was selected 38% and 48% for the two kinds of sentences.)

(C) The two lines in Figure 6, compared with those in Figure 5, indicate that the gender cue does not help children's coreference judgments very much when reflexive sentences are considered. Children's correct responses to the *say*-GC-reflexive sentences assumed a very similar pattern as their responses to the *say*-reflexive sentences. Their responses to the *say*-GC-reflexive sentences ranged between the level of 48.75% and the level of 97.25% correct. A large increase in correct performance was found at G3 (i.e., children older than 3;6), and a correct response rate over 90% was reached by children older than 4;6 (G5). (GC reflexive = local c-commanding antecedent: *say* sentences (G1–G8): 68.25%, 48.75%, 85.00%, 86.75%, 95.00%, 90.00%, 97.25%, 97.00%.) On the average, about 9.50% of the time (ranged between 0% and 38.25%), children made anaphora errors by coindexing the reflexive to the nonlocal c-commanding antecedent, even though it had a different gender from the reflexive. This type of anaphora error was mainly found in our two youngest groups. In some cases, children referred the reflexive to a sentence-external referent. This type of error ranged between 1.25% and 11.75%.

Children's responses to the *want*-GC-reflexive sentences also assumed a very similar pattern to their non-GC counterparts. Children only showed a slightly higher performance with the presence of the controlled gender cue than without. Their correct responses to the *want*-GC-reflexive sentences ranged between the level of 47.25% and 100% correct. A large increase in correct responses was found at G3 (children older than 3;6), and the greater than 90% rate was reached by children older than 4;6. This response pattern is in close concordance with the response pattern found in the *say*-GC-reflexive sentences. (GC reflexive = local c-commanding antecedent: *want* sentences (G1–G8): 66.75%, 47.25%, 83.50%, 86.75%, 93.25%, 91.25%, 96.00%, 100%.) About 9.75% of the time (ranging between 0% and 38.25%), children made anaphora errors by coindexing the reflexive to the nonlocal c-commanding antecedent, regardless of the gender discrepancy between these two elements. This type of

anaphora error was mainly found in our two youngest groups. In some cases, children referred the reflexive to a sentence-external reference (ranging between 0% and 11.75%).

(D) The two lines in Figure 8 (compared with those in Figure 7) indicate that, with the presence of the controlled gender cue, children make correct judgments almost perfectly for the GC pronoun sentences (i.e., GC pronoun = nonlocal c-commanding antecedent: *say* sentences (G1–G8): 81.75%, 89.50%, 87.00%, 90.75% 99.25%, 98.75%, 98.75%, 97.00%; *want* sentences (G1–G8): 78.25%, 90.75%, 89.25%, 89.50%, 96.75%, 92.50%, 96.00%, 97.00%). Young children sometimes ignored the gender difference between the pronoun and the puppet's name which was not mentioned in the test sentence (ranging between 0% and 6.75%). In a few cases, children violated Principle B by coindexing the pronoun with the local c-commanding antecedent (3.25% overall; ranging between 0% and 7.25%). This type of error mainly occurred in the younger groups. Children's responses to the GC sentences replicate the results found in our previous two experiments.

3.4 A Conclusion Concerning Experiments 1, 2, and 3

A conclusion with regard to the first three experiments is given as follows: When the relative patterns rather than the absolute scores or ages are considered, the results of Experiments 2 and 3 replicated most of the results found in Experiment 1. That is, when the task was to make coreference judgments between the reflexive or the pronoun and the two sentence-internal antecedents, children systematically differentiated reflexive sentences from pronoun sentences in all experiments. They did this regardless of the different complement types used in the test sentences (i.e., tensed complement or infinitival complement) and the different tasks applied in the experiments (i.e., the Simon-Says game or the Party game).

Children at the youngest ages (2;6–3;6) often responded according to a response bias and not according to a locality condition. This was true for both reflexives and pronouns. By age 6 or earlier, they had learned the locality condition for reflexives, whether the complements were tensed or infinitival. Children at this age, however, had still not learned the non-locality condition for pronouns. A developmental delay of Principle B compared to the acquisition of Principle A appears, on the surface anyway, to be consistently revealed. The Lexical Learning Hypothesis (or any hypothesis in which the principles are known to the child from the start), which was confirmed by the data relative to Principle A and apparently

disconfirmed by the data relative to Principle B, needs additional investigation.

4. CHILDREN'S KNOWLEDGE OF PRINCIPLE B

4.1 Explanations and Predictions

There are at least four possible ways out of the dilemma where Principle A data confirmed the Lexical Learning Hypothesis, whereas the Principle B data did not. The first is simply to challenge the Lexical Learning Hypothesis. Instead of stating that only lexical properties are learned, one may argue that principles are also learned, and, for some reason, Principle B is harder to learn than Principle A. However, it seems extremely unlikely that this suggestion could be correct, because correct coindexing between the reflexives or pronouns and their antecedents requires the child to have the ability to converge on knowledge that is not clearly provided by the input data. To hold the argument that principles are learned, one would have to specify what actually constitutes the positive evidence for the child to derive abstract structural notions such as *bound* and *c-command.* In addition, one would still need to explain how these principles are learned and why Principle B is more difficult to learn than Principle A. In particular, notice that if a child does not have Principle A or B, there is no positive evidence in the speech environment to show the child that he or she is wrong.

A second possible explanation is also based on learning, but on the learning of lexical items, not of principles. Perhaps pronouns are simply harder to identify as pronouns than reflexives are to identify as reflexives. Thus, for example, *her* is harder to learn than *herself.* This explanation seems quite unlikely on empirical grounds. As Manzini and Wexler (1987) pointed out, the first use of any of the items we are studying (lexical pronouns and reflexives) is of pronouns, with no grammatical antecedent. Many children produce examples of these, in appropriate contexts, before age 2. For example, suppose a child says "Mary chases him," where *him* receives its reference outside the sentence. On the assumption that the child knows Prinicple A, *him* cannot be a reflexive. Because the child produces enough examples of the pronoun, in diverse enough contexts, it is clear that *him* is not a name. Thus, *him* can only be a pronoun for a child. And this occurs considerably before reflexives are produced. Therefore, we can conclude that pronouns are *not* harder to learn than reflexives.

The third alternative has to do with the theory of syntactic maturation proposed by Borer and Wexler (1987; see also Gleitman, 1981, for a

discussion of the possibility and logic of maturation, and Felix, 1984, for another statement of the maturational idea). Instead of assuming that the formal structures available to children are constant through development, this theory argues that certain structures mature. Like any other instance of biological maturation and the proposed syntactic maturation of "argument chain" (Borer & Wexler, 1987), we might suggest that certain binding principles also mature. The maturational theory indicates a possible way for interpreting the developmental delay of Principle B relative to Principle A.

Following the maturational theory and looking more closely at the three binding principles proposed by Chomsky (1981), we might generate the following working hypothesis. The three binding principles (A, B, and C) all involve the notion of coreference or disjointness between two elements (*X* and *Y*) in a sentence. For example, *X* can be a reflexive, pronoun, or R-expression (i.e., name), and *Y* can be a potential antecedent. A principle involving *disjointness* between *X* and *Y* may mature later than one involving coreference. This will thus predict that Principles B and C will mature later than Principle A. (For reasons of space, we do not review the literature on Principle C in this article. In general, Principle C appears to develop earlier than Principle B; see Crain & McKee, 1987.)

However, we have ignored our assumption that development takes place under the constraints of UG. Borer and Wexler (1988a, 1988b) argued that UG-Constrained Maturation is the most constrained theory that can be empirically held at the moment. All the instances of maturation that they discussed are instances in which *structures* grow, but principles are in place from the beginning. For example, a young child might not have knowledge of argument chains. But the principles of grammar are there from the beginning. It is not possible, for example, that the child might have a non-well-formed argument chain (e.g., one in which the head does not c-command the tail). UG-Constrained Maturation entails a strict *growth* theory. The child adds to his or her stock of grammatical possibilities but never has to replace grammatical possibilities that the child assumes exist but that are not allowed by UG. Thus, language development is not like a caterpillar turning into a butterfly but more like a child turning into an adult (e.g., adding secondary sexual characteristics).

UG-Constrained Maturation, however, will not allow Principle B to mature. Similarly, UG-Constrained Maturation will not allow the more general hypothesis we suggested (that principles involving disjointness mature). This is simply because the principles of UG must hold at all stages. Thus, the maturational suggestions we have made are not consistent with our framework.

The fourth alternative, the one we believe is most promising, suggests that children actually do have Principle B and that the "errors" (non-adult-like behavior) we have uncovered in our experiments do not represent

a violation of Principle B. This proposal was first made by Wexler and Chien (1985; see also Chien & Wexler, 1988; and Montalbetti & Wexler, 1985). How can this be? An answer depends on a closer look into the syntax and semantics of binding, which we can only roughly and informally sketch here.[7]

First, consider Principle A. It says that an anaphor must have the same index as a local c-commanding antecedent. Principle B says the opposite: A pronoun may not have the same index as a local c-commanding antecedent. That is the syntax of Binding Theory (we ignore Principle C).

Now, what is the *semantics* of Binding Theory? That is, what interpretation is to be given to the indices and to the notion of *coindexing?* Considering Principle A and concentrating only on reflexives, it is clear that if two NPs have the same index, then these two NPs refer to the same entity. That is, coindexation implies coreference. (Of course, for reciprocals and other anaphors, we may have to state this differently.) Because anaphors will always be coindexed with their antecedent, we do not have to worry about cases where they have a different index than antecedent.

Consider pronouns and Principle B. Suppose a pronoun has the same index as a local c-commanding NP. This structure is ruled out by Principle B. On the other hand, suppose the pronoun has a different index than a local c-commanding antecedent. Because a sentence such as *Mary likes her* is bad on a reading where *her* is Mary, at first sight we would think that two noncoindexed NPs are disjoint (i.e., noncoreferential).[8] Is this correct?

The crucial question is what does noncoindexing imply? That is, what do we know about two NPs that have different indices? Can we assume that the two NPs are noncoreferential (or that it is *intended* that the two NPs are noncoreferential, as it is well-known that the notion of intended reference is the appropriate notion)?

It is clear from much work in Binding Theory that we cannot simply assume that two noncoindexed NPs are disjoint. For example, Higginbotham (1980) discussed cases involving Principle C, which says that a name may not be coindexed with a c-commanding antecedent. Thus, consider (30).

(30) a. He's wearing John's coat.
b. *He's$_i$ wearing John's$_i$ coat.
c. He's$_i$ wearing John's$_j$ coat.

In (30a), the pronoun *he* cannot be taken to be John. What is the representation of (30a)? Principle C rules out (30b). At first sight, because

[7]We are indebted to Barry Schein for extensive discussion of these issues.

[8]For simplicity, we ignore plurals. Thus, disjointness and noncoreferentiality are equivalent (see Lasnik, 1989, for a discussion).

he cannot be John in (30a), we would think that (30c) could be the representation of (30a), and that noncoindexed NPs must be disjoint, thus yielding the correct reading for (30a).

However, Higginbotham pointed out scenarios similar to (31), which shows that two noncoindexed NPs are not necessarily disjoint.

(31) He's wearing John's coat. Therefore, he must be John.

It is clear that in the first sentence of (31), *he* can be taken to be John. Moreover, the speaker intends that *he* be taken to be John. Because Principle C rules out (30b) as the representation for the first sentence of (31), the representation must be (30c). But then, in this case, noncoindexed NPs cannot be taken to be necessarily disjoint. As Higginbotham argued, Principle C should be taken to mean that a name may not be referentially dependent on a c-commanding NP.

We can create examples comparable to these involving Principle B, though to our knowledge they have not appeared in the literature. Consider (32a), perhaps with the speaker pointing while saying the first sentence.

(32) a. That must be John.
At least he looks like him.
b. That$_i$ must be John$_j$.
*At least he$_i$ looks like him$_i$.
c. That$_i$ must be John$_j$.
At least he$_i$ looks like him$_j$.

In the second sentence of (32a) *he* is taken to be John, and *him* is taken to be John. Thus, *he* and *him* are coreferential. It is clear that the indexing of (32a) should not be as in (32b). This is wrong intuitively, as it is clear that *him* is referentially dependent on *John,* not on *he.* And Principle B would rule out the second sentence in (32b) anyway because *he* c-comands *him.* Rather, it is clear that the indexing of (32a) should be as in (32c), which appropriately maps the referential dependencies. However, once again (32c) makes it clear that noncoindexing does not imply disjointness. In the second sentence of (32a), *he* and *him* are noncoindexed, but they are coreferential. Principle B must state that a pronoun may not be referentially dependent on a local c-commanding NP.

The idea, following much work in Binding Theory, is that noun phrases freely refer, except under coindexing. Principles A and B will specify certin syntactic relations involving coindexing or its lack. Semantically then, two coindexed NPs must corefer, but two noncoindexed NPs are free in reference. They may or may not corefer.

The obvious problem is the following. Why (in a typical simple Principle

B case) is it that the pronoun cannot be coreferential with a local c-commanding antecedent? For example, in the second sentence of (32a), when no context is provided, why do we take the pronoun to be disjoint from the antecedent? This is shown in (33).

(33) a. He looks like him.
b. *He_i looks like him_i.
c. He_i looks like him_j.

Principle B rules out (33b). So the representation for (33a) must be (33c). However, because noncoindexed NPs can refer freely, why, without other context, do we take *him* to be noncoreferential with *he* in (33a)?

Reinhart (1983a, 1983b, 1986) suggested that a pragmatic principle governs the choice of reference in examples like (33). Here we do not discuss the particular pragmatic principle that Reinhart proposed, but we do accept her conclusion that a pragmatic principle governs the choice of reference in (33c) so that *i* and *j* are not coreferential. Whatever pragmatic principle ultimately turns out to be correct, it must be such that coreference is not possible in (33c) but is possible when the same sentential form is used in a particular context (32c). We know of no other proposal that has been made (other than the proposal that pragmatic principles are involved) that would explain the difference between (32) and (33). Basically, it seems to us that the literature on Binding Theory is in agreement on this, although the structure of Binding Theory can differ as can particular binding principles involved.[9]

To reiterate, Principle B is a *syntactic* principle. It governs relationships between syntactic indices. The interpretation of these indices is governed by semantic and pragmatic principles. One of these latter principles (we have been calling it pragmatic, but it might turn out to be more adequately characterized as semantic) is involved in (33c). Let us call the principle that rules out coreference of *i* and *j* in (33c) Principle P (that is, P is a pragmatic, possibly semantic, principle).

What we propose about children then is: They *do* know Principle B. This means that they know the relevant syntactic principle. This hypothesis is in accordance with the framework that we have been assuming, as we have discussed (e.g., the Lexical Learning Hypothesis). What, then, is it that children don't know? Why does it appear that they are violating Principle

[9]Work in progress by Barry Schein and Wexler is attempting to state the pragmatic principle in a somewhat different way than Reinhart. Once again, although the details of the pragmatic theory and the Binding Theory and the notion of referential dependence might differ, Reinhart's important and insightful proposal that some pragmatic principle must be involved in order to explain the relevant distinctions seems correct.

B? With a closer look at Binding Theory, the answer seems obvious. What the children don't know is the pragmatic Principle P. Children know that (33b) is ungrammatical. This follows from Principle B. They know that (33c) is grammatical, because no grammatical principle rules it out. They just don't know that (33c) cannot have an interpretation in which *i* and *j* are interpreted coreferentially. They treat (33c) like the second sentence in (32c).

Our proposal, if true, sheds a very interesting light on children's linguistic knowledge. Modern linguistic theory assumes a modular structure of linguistic knowledge. For example, syntactic and pragmatic principles are in different components. It appears, if our proposal is true, that children may know syntactic principles that are relevant to particular sentences, whereas they may not know a particular pragmatic principle.

Our proposal is quite consistent, then, with Binding Theory. It is consistent with the empirical evidence about Principle B and its acquisition and with the framework of assumptions that underlies our theory, and Principles and Parameters Theory more generally. It turns out, however, that the proposal is capable of an even stronger test. Namely, it makes very concrete, precise, and (on any other grounds) surprising empirical predictions about linguistic development.

To see this, consider that so far (and in the entire preceding literature on the development of Binding Theory), the only interpretation of binding that has been studied is coreference. Thus, even though the child knows Principle B, it has always been possible for the child to appear to violate Principle B by interpreting noncoindexed NPs as coreferential, in violation of Principle P, though not of Principle B. In short, knowledge of Principle B has been confounded with knowledge of Principle P.

However, other semantic operations are possible, in which Principle B and Principle P are not confounded. If we make it clear that the only interpretation of a pronoun in a particular sentence is as a bound variable, then Principle P, which applies to noncoindexed NPs, will be inapplicable. This is because interpreting a pronoun as a bound variable requires that it be coindexed with its antecedent. For example, consider (34).

(34) a. Every man thinks that justice is due him.
b. [Every man]$_i$ thinks that justice is due him$_i$.
c. [Every man]$_i$ thinks that justice is due him$_j$.

Without going through any formal details, the interpretation of (34b) is (35).

(35) For every man *x*, *x* thinks that justice is due *x*.

The other possible syntactic representation of (34a), namely (34c), cannot have (35) as its interpretation. The only interpretation of (34c) is that *him* is a particular individual such that every man thinks that justice is due that individual, in short (36).

(36) There is a man x such that for every man y, y thinks that justice is due x.

In short, coindexation is required for interpretation of a pronoun as a bound variable. Now, consider a Principle B context as in (37).

(37) a. Every woman looks like her.
b. *[Every woman]$_i$ looks like her$_i$.
c. [Every woman]$_i$ looks like her$_j$.

We have seen that (34) is ambiguous between the bound variable and nonbound variable readings. However, (37) is not ambiguous. The bound variable reading can only come from the representation (37b), and this is ruled out by Principle B. The representation (37c) only allows the nonbound variable reading. Thus, the only interpretation of (37) is (38a); (38b) is impossible.

(38) a. There is a woman x such that for every woman y, y looks like x.
b. *For every woman x, x looks like x.

When we were dealing with coreference, there were two routes to such an interpretation: through coindexed NPs and through noncoindexed NPs. Principle P sometimes rules out coreference in the noncoindexed case, but children don't know Principle P, we assume. Therefore, they can take the noncoindexed route to coreference in certain cases that the adult can't (*Mary likes her*, with *her* = *Mary*). But there is no second, noncoindexed route to be a bound variable.

Thus, we make a very clear and precise prediction. Children will interpret a pronoun in sentences like (39a) as being possibly coreferential with the antecedent. That is, they will allow *Mary* to be *her*. However, they will not allow a pronoun to be a bound variable in the same structural context. That is, they will find that (40a) only has the meaning (40c) and not (40b).

(39) a. Mary likes her.
b. Mary$_i$ likes her$_j$.

(40) a. Every woman likes her.
b. *For every woman x, x likes x.

c. There is a woman x such that for every woman y, y likes x.
d. *[Every woman]$_i$ likes her$_i$.
e. [Every woman]$_i$ likes her$_j$.

This is because they will take (39a) to have the syntactic representation (39b), interpreting the pronoun as coreferential (in violation of Principle P, which they don't have). However, the pronoun cannot be a bound variable in (40a) because it cannot be interpreted as a bound variable in (40e), and (40d) is ruled out by Principle B.

In short, children will "appear" to violate Principle B in the coreferential cases like (39) but they will not "appear" to violate Principle B in the bound variable cases (40). In neither case will they actually be violating Principle B. They will simply violate Principle P in (39). A violation of Principle P in (40) is not relevant—an extra interpretation cannot be gained.

Experiment 4 was designed to test this strong prediction.

4.2 Experiment 4

Experiment 4 was designed to test the hypothesis that children have knowledge of Principle B. More specifically, it was designed to test the hypothesis that children will accept a pronoun as being coreferential with a local c-commanding antecedent, but they will not allow a pronoun to be a bound variable when it has a local c-commanding antecedent. We attempted to keep the sentences used for the bound variable and coreferential cases as similar as possible. To make sure that the method used in this experiment did not influence the result, we repeated examples with coreference. That is, we did not depend only on earlier experiments to show children's willingness to accept a pronoun as being coreferential with a local c-commanding antecedent; we actually repeated the relevant examples in this experiment so that we could make the comparison with bound variables directly, given the same method, type of stimuli, and subjects.

The experiment involved bound variables. We decided to study bound variables by using explicit quantifiers (e.g., *every*). Because what we were asking the children to understand was so complex, we wanted the materials to be as simple as possible. Thus, we used simple sentences. Also, because the quantification was complex enough, we did not want to confuse the child by presenting more than one situation from which to choose. We decided to present the child with one situation and one sentence at a time and have the child say "yes" or "no" according to whether the sentence matched the situation. In this way, we attempted to reduce the cognitive and behavioral complexity of the task. We found that it was very difficult to set up the pragmatics correctly (i.e., neutrally) in the two-choice situation. (In

some pilot work, we attempted to have the child act out various quantificational sentences, but the behavioral complexity was too difficult for many of the young children. It is worth attempting other methods.)

4.2.1 Method

4.2.1.1 Subjects. The subjects tested in this study were 177 children (between the ages of 2;6 and 7;0) and 20 adults. The child subjects were assigned to four different age groups. Group 1 contained 48 subjects under 4 years old. Group 2 contained 45 subjects between the ages of 4 and 5. Group 3 contained 44 subjects between the ages of 5 and 6. Group 4 contained 40 subjects between the ages of 6 and 7.

4.2.1.2 The task. A Yes/No Judgment (YNJ) task was used in this experiment to invoke children's grammaticality judgments of sentences like those given in (41–48) (see corresponding pictures). In the YNJ task, the child was presented with a cartoon picture, an introductory sentence, and a question related to the picture. After a careful inspection of the picture, the child was expected to answer "yes" or "no" to the question presented. For example, suppose the name of the child being tested was Sarah. The general procedure for the task is as follows: We first presented a picture (e.g., the one given in (41)) to Sarah, then we said to her, "Sarah, look at this picture. This is Goldilocks; this is Mama Bear. Is Mama Bear *touching*[10] herself?" Sarah might answer "yes" or "no" depending on her current stage of language development.

4.2.1.3 The design of the test constructions. Four different types of experimental questions were included in this study (see, e.g., (41–48) and the corresponding pictures). These questions were similar in their syntactic structures but varied with respect to the types of NPs occurring in their subject position or object position. Half of the experimental questions included proper names (e.g., *Mama Bear*) as their subject (e.g., (41) & (42) and (45) & (46)); the other half included quantified NPs (e.g., *every bear*) as their subject (e.g., (43) & (44) and (47) & (48)). The object NPs were varied as to whether they were reflexives (i.e., *himself* or *herself*) or pronouns (i.e., *him* or *her*). According to the NPs occurring in their subject and object positions, we named these four types of experimental questions the *name-reflexive* questions, *name-pronoun* questions, *quantifier-reflexive* questions, and *quantifier-pronoun* questions, respectively. Notice that,

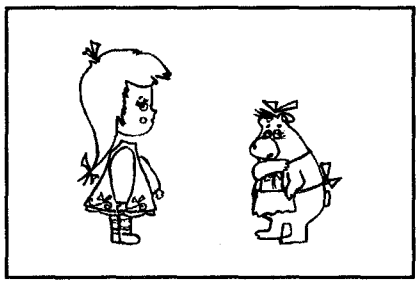
[10]The italicized word indicates stress. When presenting the test sentences, the verb before the pronoun or the reflexive is always slightly stressed. In this way, the reflexive or the pronoun is left unstressed.

Experiment 4: The Experimental Conditions

The "Match" Cases

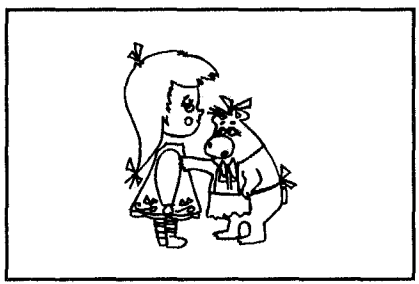
Name-Reflexive

(41) This is Goldilocks; this is Mama Bear.
Is Mama Bear touching herself?



Name-Pronoun

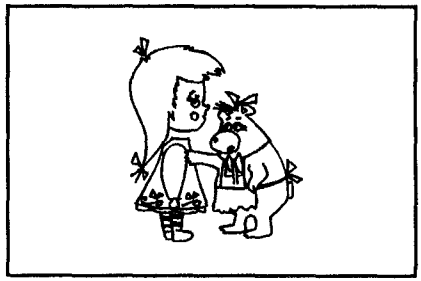
(42) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching her?



The "Mismatch" Cases

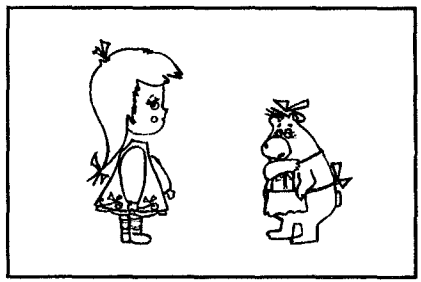
Name-Reflexive

(45) This is Goldilocks; this is Mama Bear.
Is Mama Bear touching herself?



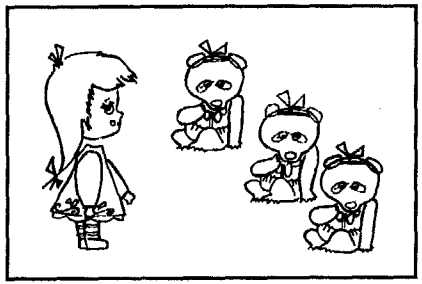
Name-Pronoun

(46) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching her?



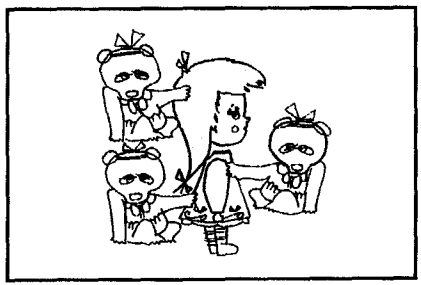
Quantifier-Reflexive

(43) This is Goldilocks; these are the bears.
Is every bear touching herself?



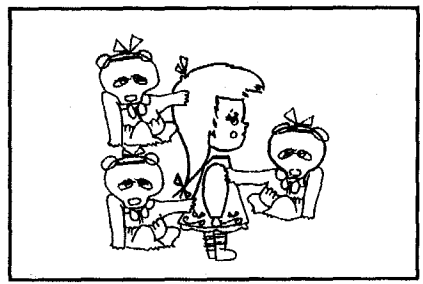
Quantifier-Pronoun

(44) These are the bears; this is Goldilocks.
Is every bear touching her?



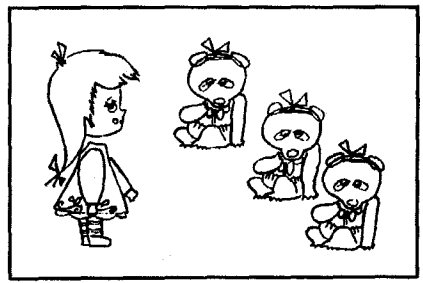
Quantifier-Reflexive

(47) This is Goldilocks; these are the bears.
Is every bear touching herself?



Quantifier-Pronoun

(48) These are the bears; this is Goldilocks.
Is every bear touching her?



paired with each question type, the picture presented to the child matched the question in only half of the cases (e.g., (41–44)); in the other half, there was a mismatch between them (e.g., (45–48)). Adding this match/mismatch factor to each of the four question types, we arrived at eight different experimental conditions.

In addition to the four types of experimental questions, we included three types of control questions. These were designed to see whether children know the concept of quantified NP, such as *every bear* (see (49–54) and the corresponding pictures). The control questions were similar in syntactic structure to the experimental questions. They varied with respect to the types of NPs occurring in the subject position. These subject NPs were proper names, such as *Mama Bear* (e.g., (49) & (52)), quantified NPs, such as *every bear* (e.g., (50) & (53)), or another type of quantified NP—using *all*—such as *all bears* (e.g., (51) & (54)). The NPs occurring in the object position were all of the same type: proper names such as *Goldilocks.* We named these three types of control questions the *name-name* questions, *quantifier every-name* questions, and *quantifier all-name* questions. As with the previous set of experimental sentences, there were also both match (49–51) and mismatch cases (52–54), yielding a total of six control conditions.[11]

Three different verbs (*wash, point to,* and *touch*) were used in each condition, 2 items for each verb, making a total of 6 items for each condition and a total of 84 test items for each child.

4.2.2 Results and Discussion

The results regarding the six control conditions are illustrated by the three lines in Figure 9 and Figure 10, respectively. On the abscissa, we plot ages in 1-year intervals. On the ordinate, we plot percentage of items indicating correct responses. The line with black squares shows the results of the name-name condition; the line with white squares shows the results of the quantifier-name condition involving *every N* as the subject; the line with small crosses shows the results of the quantifier-name condition involving *all Ns* as the subject. The major findings concerning the six control conditions are summarized as follows:

[11]In order to make it easier to compare the similarities and the differences among different test conditions, we have included only one set of cartoon characters, namely, Goldilocks and the bears in the examples. In the real test conditions, there were four different sets of cartoon characters: (a) Goldilocks, Mama Bear, and three female bears; (b) Gargamel, Papa Smurf, and three smurfs; (c) Donald Duck, Mickey Mouse, and three male mice; and (d) Daisy Duck, Minnie Mouse, and three female ducks. The characters "Smurfette" and "Pinocchio" are used in some mismatched *name-name* cases.

(A) As indicated in Figure 9, when the question matched the picture, children's response patterns to these three types of control questions were almost the same. They made correct judgments almost perfectly for all three question types. (Match (G1–G4): name-name: 94.10%, 98.89%, 96.97%, 99.58%; *every*-name: 92.71%, 98.15%, 99.24%, 99.58%; *all*-name: 94.10%, 98.52%, 97.73%, 98.75%.)

(B) As indicated in Figure 10, when there was a mismatch between the question and the picture, this mismatch, in general, depressed children's correct responses. This indicated that children might have a tendency to say "yes" more frequently than "no." The depression was larger for the two conditions involving quantified NPs (i.e., the *every*-name and the *all*-name conditions) than the name-name condition. This depression was also more significant for children younger than 5 than for children older than 5. These results indicated that around age 5, children knew the concepts of quantified NPs such as *every N* and *all Ns*. Their percentage correct was beyond the level of 90%. (Notice, this meant that the 5-year-olds were saying "no" 90% of the time to the mismatch cases. Mismatch (G1–G4): name-name: 60.42%, 83.33%, 95.45%, 95.00%; *every*-name: 32.99%, 71.48%, 93.94%, 97.29%; *all*-name: 34.03%, 72.22%, 96.21%, 99.17%.) As shown in Figures 9 and 10, our adult subjects gave perfect responses to all of the control items.

The results concerning the eight experimental conditions are illustrated in Figures 11–14. In order to make the results easier to understand, we repeat the related test conditions with the figures, on top of and beneath each figure. Again, on the abscissa we plot ages in 1-year intervals; on the ordinate, we plot percentage of correct responses.

The major findings concerning the eight experimental conditions are summarized as follows:

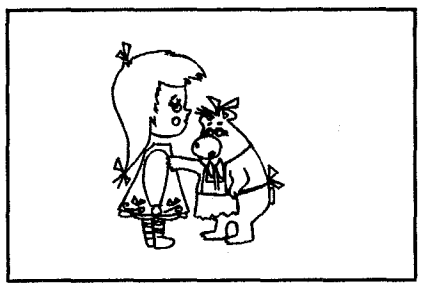
(A) Figure 11 concerns the results of the name-reflexive questions. The line with black squares illustrates the results of the match cases, the line with white squares the results of the mismatch cases. As can be seen from these two lines, children at age 5 and older clearly demonstrated knowledge of Principle A. They know that a reflexive must be locally bound. In both the match and the mismatch conditions, their percentage correct was higher than 90%. In other words, children at age 5 and older not only correctly identified local binding for reflexives but also correctly rejected nonlocal binding for reflexives (correct responses: match (G3–G4): 96.97%, 96.67%; mismatch (G3–G4): 92.80%, 99.17).

Experiment 4: The Control Conditions

The "Match" Cases

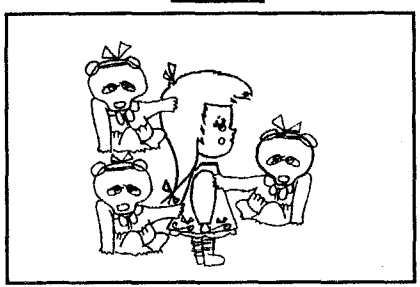
Name-Name

(49) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching Goldilocks?



Quantifier "every" - Name

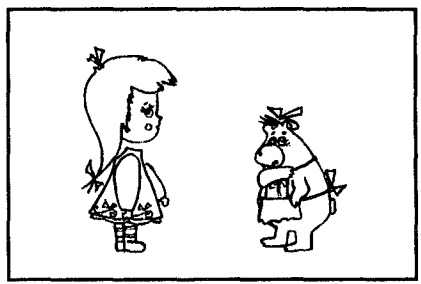
(50) These are the bears; this is Goldilocks.
Is every bear touching Goldilocks?



The "Mismatch" Cases

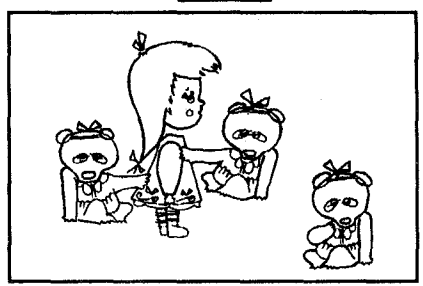
Name-Name

(52) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching Goldilocks?



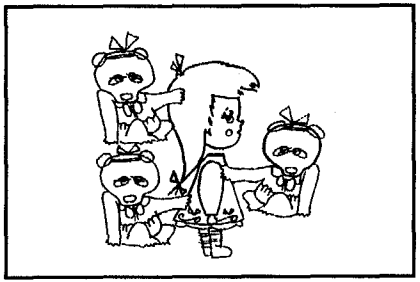
Quantifier "every" - Name

(53) These are the bears; this is Goldilocks.
Is every bear touching Goldilocks?



Quantifier "all" - Name

(51) These are the bears; this is Goldilocks.
Are all the bears touching Goldilocks?



The Results of the 3 Match Conditions

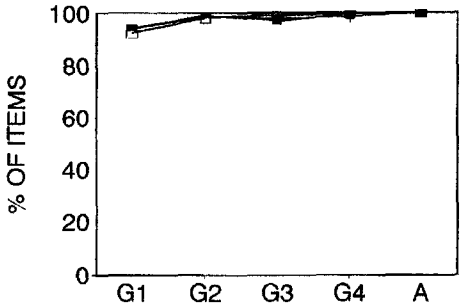


FIGURE 9

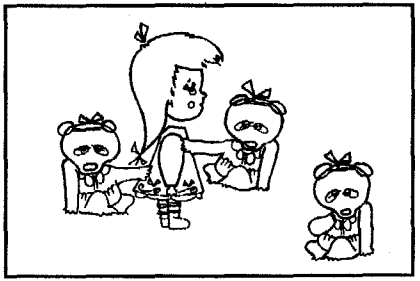
Name-Name:

Q "every" - Name:

Q "all" - Name:

Quantifier "all" - Name

(54) These are the bears; this is Goldilocks.
Are all the bears touching Goldilocks?



The Results of the 3 Mismatch Conditions

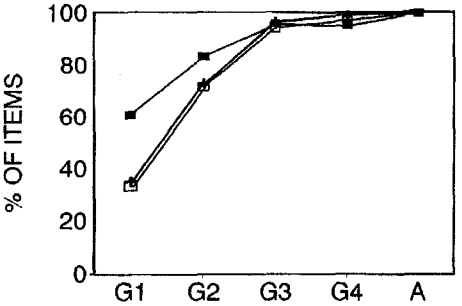


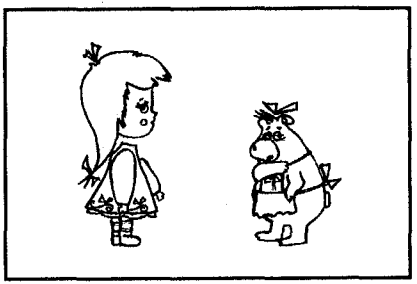
FIGURE 10

Experiment 4: The Results of the Experimental Conditions

[The Name-Reflexive and the Quantifier-Reflexive Cases]

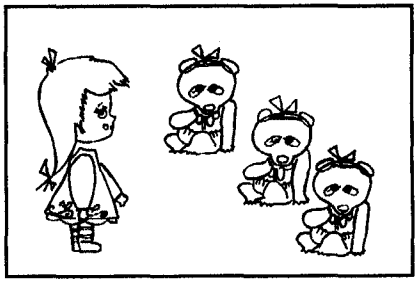
Name-Reflexive (Match)

(55) This is Goldilocks; this is Mama Bear.
Is Mama Bear touching herself?



Quantifier-Reflexive (Match)

(57) This is Goldilocks; these are the bears
Is every bear touching herself?



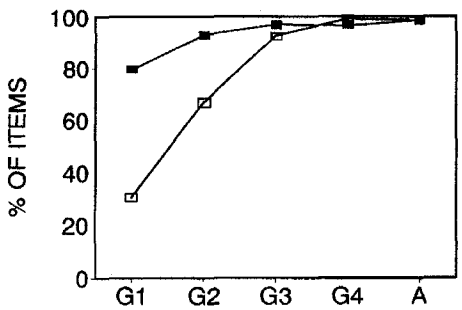


FIGURE 11

Match:

Mismatch:

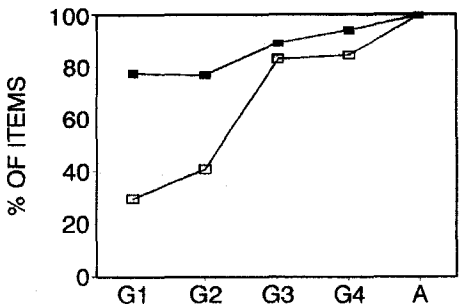
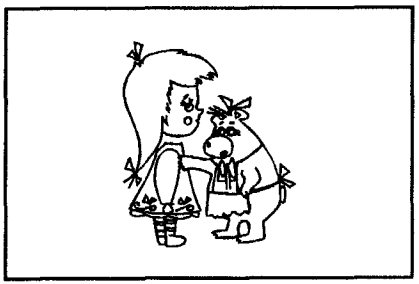


FIGURE 12

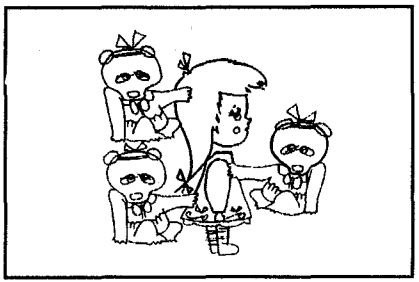
Name-Reflexive (Mismatch)

(56) This is Goldilocks; this is Mama Bear.
Is Mama Bear touching herself?



Quantifier-Reflexive (Mismatch)

(58) This is Goldilocks; these are the bears.
Is every bear touching herself?



Children younger than 5 did not clearly show behavior predicted by Principle A. Although they know that it was correct to coindex a reflexive to the local antecedent, in many cases, they allowed the reflexive to refer to the external referent that was mentioned in the introductory sentence preceding the question (correct responses: match (G1–G2): 79.51%, 92.96%; mismatch (G1–G2): 30.56%, 67.04%). For example, when we presented the picture in (56) to the children, read the introductory sentence "This is Goldilocks; this is Mama Bear," and asked the question "Is Mama Bear *touching* herself," children younger than 4 answered "no" only 30% of the time, and children between the ages of 4 and 5 answered "no" about 67% of the time. These results replicated the results shown in our previous studies, namely that children demonstrated their knowledge of Principle A between the ages of 5 and 6. Children younger than 5 were not firm in their belief that a reflexive must have a local c-commanding antecedent. In this study, the tendency to answer "yes" often overrode the knowledge for these young children.

(B) Figure 12 concerns the results of the quantifier-reflexive questions. The line with black squares shows the results of the match cases, and the line with white squares the results of the mismatch cases. These two lines also indicate that children around age 5 clearly demonstrate knowledge of Principle A. When the question matched the picture, they made correct judgments about 90% of the time (correct responses: match (G3–G4): 89.39%, 94.17%). When there was a mismatch between the question and the picture, children between the ages of 5 and 6 correctly said "no" about 83% of the time; children older than 6 correctly said "no" about 85% of the time (correct responses: mismatch (G3–G4): 82.95%, 84.58%).

Children younger than 5 did not clearly show behavior predicted by Principle A when dealing with questions involving a reflexive and a quantified antecedent. Similar to their responses to the name-reflexive questions—although they knew that it was correct to coindex a reflexive to the local antecedent—in many cases they allowed the reflexive to be coreferential with an external referent (correct responses: match (G1–G2): 77.43%, 76.67%; mismatch (G1–G2): 29.51%, 40.74%).

(C) Comparing the results in Figure 12 to those in Figure 11, we found that children's correct responses to the quantifier-reflexive questions were systematically lower than their responses to the name-reflexive questions. This result was not unpredicted, however, because, as the results in our control conditions showed, children did not demonstrate clear knowledge of the concepts of quantified NPs until the age

of 5. By the time they knew what *every* meant and what *himself/herself* meant, they demonstrated knowledge of Principle A. They knew that a reflexive must be bound by its local antecedent, whether this antecedent was a definite NP or a quantified NP. As indicated in Figures 11 and 12, adults gave almost perfect responses to the reflexive sentences.

(D) Figure 13 summarizes the results of the name-pronoun questions. The line with black squares indicates that when the question matched the picture, children made correct judgments almost perfectly. Even our 2- and 3-year-olds seemed to know that a nonreflexive pronoun could refer to a definite NP outside the sentence. Their percentage correct was around the 90% level (match (G1–G4): 91.67%, 88.52%, 90.15%, 94.58%).

The line with white squares indicates that when there was a mismatch between the question and the picture, this mismatch significantly lowered children's correct responses regarding the name-pronoun questions. In many of the name-pronoun cases, they allowed a nonreflexive pronoun to be interpreted as coreferential with its local c-commanding antecedent. For example, when we presented the picture in (60) to the children, read the introductory sentence "This is Mama Bear; this is Goldilocks," and asked the question "Is Mama Bear *touching* her?" children younger than 4 answered "yes" about 70% of the time; children between 4 and 5 answered "yes" about 60% of the time; children between 5 and 6 answered "yes" about 50% of the time; children between 6 and 7 answered "yes" about 24% of the time. The results indicated that in the age range of 5 to 6, children still did not know that a pronoun must not refer to the local c-commanding antecedent (correct "no" responses: mismatch (G1–G4): 30.90%, 39.26%, 49.24%, 76.67%). Although children in the same age range already demonstrated their knowledge of Principle A, they apparently still did not show their behavior in accordance with Principle B (i.e., they violated Principle P, according to our hypothesis, not B).

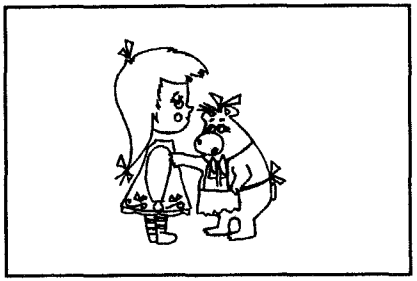
To summarize, the results presented so far replicated our previous findings concerning children's acquisition of Principle A and Principle B (i.e., the results shown in Experiments 1 to 3). At the time children first clearly demonstrated their knowledge of Principle A, they did not clearly demonstrate their knowledge of Principle B for pronouns that were not bound variables (i.e., those cases in which accidental coreference was possible). (Remember, we interpret this as meaning that Principle B was not violated.) The appearance of a difference between children's development of Principle A and Principle B was replicated by using a completely

Experiment 4: The Results of the Experimental Conditions

[The Name-Pronoun and the Quantifier-Pronoun Cases]

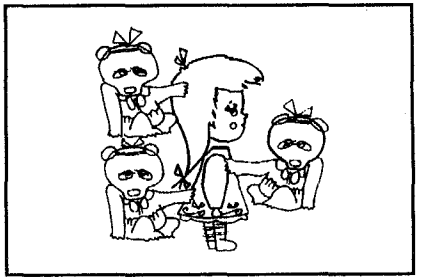
Name-Pronoun (Match)

(59) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching her?



Quantifier-Pronoun (Match)

(61) These are the bears; this is Goldilocks.
Is every bear touching her?



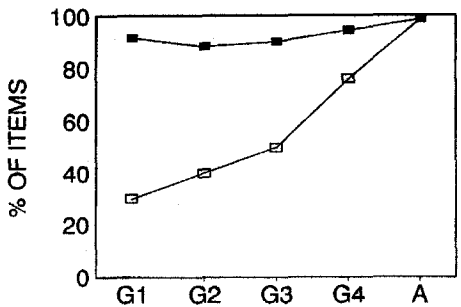


FIGURE 13

Match:

Mismatch:

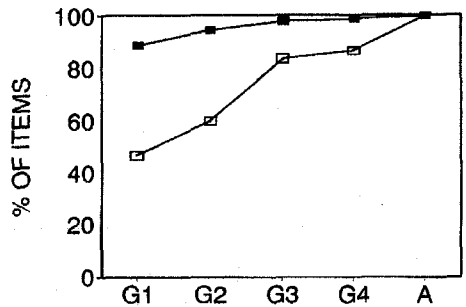
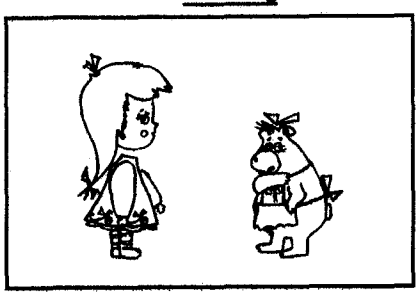


FIGURE 14

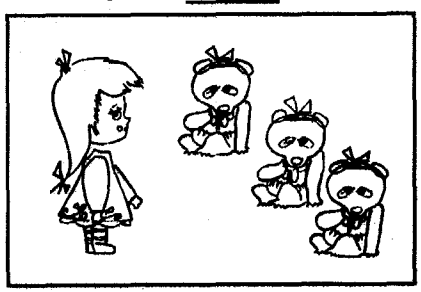
Name-Pronoun (Mismatch)

(60) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching her?



Quantifier-Pronoun (Mismatch)

(62) These are the bears; this is Goldilocks.
Is every bear touching her?



different research methodology and different linguistic materials. Further discussion, however, shows why this is not the correct conclusion. Rather, it is the pragmatic Principle P that is missing in young children.

(E) Figure 14 summarizes the results related to children's knowledge of Principle B as applied to the bound variable cases, that is, children's responses to the quantifier-pronoun questions. The line with black squares shows the results of the match cases, whereas the line with white squares shows the results of the mismatch cases. These two lines indicate that children, in the age range of 5 to 6, know Principle B. When the question matched the picture, children in G3 and G4 gave correct responses more than 90% of the time (correct responses: match (G3–G4): 97.93%, 98.75%). When there was a mismatch between the question and the picture, children between the ages of 5 and 6 gave correct "no" responses about 84% of the time; children older than 6 correctly answered "no" about 87% of the time (correct responses: mismatch (G3–G4): 83.71%, 86.67%).

Children younger than 5 did not behave according to Principle B when dealing with questions involving a pronoun and a quantified antecedent. Similar to their responses to the name-pronoun questions—although they knew that a pronoun can refer to an external referent—in many cases they allowed the pronoun to be coindexed with a local quantified antecedent. However, recall that children at this age also showed poor knowledge of the meaning of the quantified NPs. Thus, they might very well have knowledge of Principle B but performed poorly on these true binding cases simply because they did not understand the quantifier. The poor performance on the quantifier-name cases at these ages means that the poor performance on the quantifier-pronoun cases cannot be taken as evidence against knowledge of Principle B for these ages (correct responses: match (G1–G2): 88.54%, 94.44%, mismatch (G1–G2): 46.88%, 60.00%).

(F) Comparing the results in Figure 14 to those in Figure 13, we found that children's correct responses to the quantifier-pronoun questions were much higher than their responses to the name-pronoun questions. This result is important, especially when we look at the mismatch cases and consider the responses given by our G3 children (i.e., children between the ages of 5 and 6). About 50% of the time, the G3 children allowed a pronoun to be coreferential with a local definite NP (about the chance level). These children, however, did not often allow the pronoun to be locally bound by a quantified NP. They coindexed the pronoun with its local quantified antecedent only 16% of the time. This indicated that, for children, pronouns as

bound variables must be free in their governing category. Our adult subjects again showed almost perfect performance responding to the pronoun sentences.

To summarize, children between the ages of 5 and 6 demonstrated their steady knowledge of the concepts of quantified NPs such as *every bear*. Children in the same age range also clearly demonstrated their knowledge of Principle A. They knew that a reflexive must be interpreted as coreferential with its local antecedent, no matter whether this antecedent was a definite NP or a quantified one. In addition, these children also demonstrated their knowledge of Principle B in cases where accidental coreference was not allowed (i.e., coindexed cases). They knew that a nonreflexive pronoun must be interpreted as not bound by any local quantified NP. In cases where accidental coreference was allowed, the same children in the same age range allowed optional coreference. That is, they allowed a nonreflexive pronoun to be interpreted as coreferential with a proper name that locally c-commanded it. Thus, it seems quite clear that the best interpretation is that children know Principle B and that in the coreference cases they are lacking the pragmatic principle. See Appendixes A and B for additional discussion.

5. CONCLUSION

The results of Experiment 4 replicated and extended the results of our previous experiments. These data are crucial on two accounts. First, to a large degree the results confirm the Lexical Learning Hypothesis. Namely by the time the child shows control of the concept of *every N* and learns that *himself* is a reflexive and *him* is a pronoun, the child is able to link these lexical items to the corresponding Principle A and Principle B and to correctly identify their appropriate antecedents and correctly rule out the inappropriate ones. Principle A specifies the conditions for an obligatory binding interpretation; Principle B determines the conditions for an obligatory nonbinding interpretation. These two principles are part of the innate endowment that the child brings to the language acquisition task. They are unlearned. For those cases where accidental coreference is possible (e.g., our name-pronoun condition), children allow optional coreference. At what age will children demonstrate adult-like knowledge and know that if a name locally c-commands a coreferential pronoun, the sentence is therefore bad? And how will they acquire this knowledge? These are still open questions.

The second important conclusion to be drawn from our data is that it is important to be careful not to immediately deduce that children do not have

principles of Universal Grammar whenever a discrepancy between child and adult language is discovered. Closer experimental investigation may reveal that, in fact, the child has a deep universal principle. For example, our experiments show that children know that a pronoun must not be locally bound, although initial experiments on the referential cases suggested otherwise. Our conclusion is that children do know Principle B. What they are showing us is the true nature of the principle, uncolored by other factors. To what degree this conclusion is correct, of course, must be determined by further theoretical and empirical investigation. A series of experiments adopting this approach is now being conducted in our language acquisition laboratory. In these experiments, we use different research methods and different linguistic materials. We hope to attain more converging data to support and extend our current findings.

One intriguing result of our studies concerns the question we raised earlier: Is it possible that results of studies in child language may be useful in the study of linguistic theory itself? It seems to us that the behavior patterns exhibited by children and the theoretical conclusions drawn from these patterns provide significant support for and direction to thinking about binding in linguistic theory. The fact that the principles governing syntactic coindexation are autonomous (i.e., modular) with respect to the principles governing semantic interpretation of indices is somewhat difficult to see in adult linguistic studies. It is veiled by many correlations and uncovered only by highly special cases. However, young children appear to lack the pragmatic principle that provides so much overlap in the behaviors (judgments) of adults. Therefore, their system demonstrates the modularity much more directly. Children, in lacking certain pragmatic principles, appear to demonstrate the true nature of the syntactic principles of UG (which they do have) in a much clearer way than adults. The study of language acquisition, if it shows more of these cases, might ultimately play a significant role in the study of linguistic theory itself.[12]

[12]An anonymous reviewer suggested that the conclusion regarding modularity might not be as strong as we have indicated. The reviewer points out that Borer and Wexler (1987) argued that move-NP is not present in young children and states that similar arguments to the ones made in the text would "lead to the suggestion that the two parts of move-alpha are in different modules since they are apparently acquired at different times. We do not make this assumption since of insights given by linguistic theory."

However, we think that this quotation does not represent the actual situation. The idea of the reviewer is that if children have move-*wh* but do not have move-NP, then these two subparts of the rule are actually in different modules, not simply different manifestations of one general move-alpha rule. However, move-NP and move-*wh* are not two different rules. There is only one rule, move-alpha. Whether the result is good or bad depends on the independent principles that apply. For example, super-raising will be bad for autonomous reasons. Raising *wh* to C position rather than Specifier C position will be bad. However, these results are not stated in move-alpha.

ACKNOWLEDGMENTS

The studies presented in this article were supported by the Alfred P. Sloan Foundation Grant 78-2-10 and by the NSF Grants BNS-84-19475, BNS-88-20439, and BNS-88-20585.

We thank Noam Chomsky, Lila Gleitman, Irene Heim, Dana McDaniel, Tom Roeper, Barry Schein, and a very thorough anonymous referee for comments on an earlier version of this article. We thank the School of Social Sciences, University of California at Irvine (UCI) for providing an excellent research environment. We also thank graduate and undergraduate research assistants at UCI for their assistance in data collection. We are grateful to Wendy Snyder for editing an earlier version of this article, Joyce Mochizuki for editing and proofreading the article, and Xuan-Huong T. Hoang for data coding and scoring. We also thank those children who took part in the experiments and the school directors and principals for their cooperation in research.

A short version of this article, which included the first three experiments, was presented at the 18th Annual Child Language Research Forum at Stanford University in 1987. A short version of the final experiment was presented at the 13th Annual Boston University Conference on Language Development in 1988. A preliminary version of this article was presented at the 11th meeting of Generative Linguists of the Old World (GLOW), Budapest in 1988.

REFERENCES

Borer, H. (1984). *Parametric syntax.* Dordrecht, The Netherlands: Foris.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 123-172). Dordrecht, The Netherlands: Reidel.

The actual assumption of Borer and Wexler is that children cannot assign thematic roles noncanonically. Because the moved NP will not have a theta role, the argument chains would result in a violation of the theta criterion. This statement is not a part of move-alpha. The structures that result from move-alpha are actually determined by principles of different modules. Therefore, in saying that children cannot form argument chains, Borer and Wexler *are* assuming (following the standard view) that argument chains are the result of interacting modules, only one of which is move-alpha. Children *do* have move-alpha in its entirety (it does not *have* subparts). However, children *do not* have the capacity in some other module (having to do with theta role assignment) that allows the creation of argument chains. Similarly, we have argued in this article that children have Principle B but not Principle P. Thus, the situation with respect to the lack of argument chains is exactly the same as the situation with respect to binding. Adult competence is the result of a complex network of autonomous modules. Children may be missing some of these modules; at least the ones whose absence does not result in structures that are not compatible with UG.

Borer, H., & Wexler, K. (1988a, March). *On the acquisition of agreements.* Paper presented at the 11th Generative Linguists of the Old World (GLOW) Colloquium, Budapest, Hungary.

Borer, H., & Wexler, K. (1988b). *The maturation of syntactic principles.* Unpublished manuscript, University of Massachusetts, Amherst and Massachusetts Institute of Technology.

Cairns, H., & McDaniel, D. (1987, October). *The status of the binding principles in the grammars of young children.* Paper presented at the 12th Annual Boston University Conference on Language Development, Boston.

Chien, Y.-C., & Wexler, K. (1987a). Children's acquisition of reflexives and pronouns. *Papers and Reports on Child Language Development, 26,* 30–39.

Chien, Y.-C., & Wexler, K. (1987b, October). *A comparison between Chinese-speaking and English-speaking children's acquisition of reflexives and pronouns.* Paper presented at the 12th Annual Boston University Conference on Language Development, Boston.

Chien, Y.-C., & Wexler, K. (1988, October). *Children's acquisition of binding principles.* Paper presented at the 13th Annual Boston University Conference on Language Development, Boston.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht, The Netherlands: Foris.

Crain, S., & McKee, C. (1985). Acquisition of structural restrictions on anaphors. In S. Berman, J. W. Choe, & J. McDonough (Eds.), *Proceedings of the Sixteenth Annual Meeting of the North Eastern Linguistics Society (NELS)* (pp. 94–110). Montreal: McGill University.

Crain, S., & McKee, C. (1987, October). *Cross-linguistic analysis of the acquisition of coreference relations.* Paper presented at the 12th Annual Boston University Conference on Language Development, Boston.

Deutsch, W., & Koster, J. (1982). Children's interpretation of sentence-internal anaphora. *Papers and Reports on Child Language Development, 21,* 39–45.

Deutsch, W., Koster, C., & Koster, J. (1986). Chidlren's errors in understanding anaphora. *Linguistics, 24,* 203–225.

Felix, S. (1984). Maturational aspects of universal grammar. In C. Cripper, A. Davies, & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 133–161). Edinburgh, Scotland: Edinburgh University Press.

Gleitman, L. (1981). Maturational determinants of language growth. *Cognition, 10,* 103–114.

Grimshaw, J., Rosen, S., & Brown, C. (1988, May). *Antecedents for pronouns and the development of Principle B.* Paper presented at the Workshop on Language Processing and Acquisition, University of Massachusetts, Amherst.

Higginbotham, J. (1980). Anaphor and GB: Some preliminary remarks. In J. Jensen (Ed.), *Proceedings of the 10th Annual Meeting of the North Eastern Linguistics Society (NELS)* (pp. 223–236). Ontario: University of Ottawa.

Hyams, N., & Sigurjónsdóttir, S. (1990). The development of "long-distance anaphora": A cross-linguistic comparison with special reference to Icelandic. *Language Acquisition, 1,* 57–93.

Jacubowicz, C. (1984). On markedness and binding principles. *Proceedings of the Fourteenth Annual Meeting of the North Eastern Linguistics Society (NELS)* (pp. 154–182). Cambridge, MA: Massachusetts Institute of Technology.

Kaufman, D. K. (1988). *Grammatical and cognitive interactions in the study of children's knowledge of Binding Theory and reference relations.* Unpublished doctoral dissertation, Temple University, Philadelphia, PA.

Koster, J. (1984). On binding and control. *Linguistic Inquiry, 15,* 417–460.

Lasnik, H. (1989). *Essays on anaphora.* Dordrecht, The Netherlands: Kluwer Academic.

Lee, H., & Wexler, K. (1987, October). The acquisition of reflexives and pronouns in Korean: From the cross-linguistic perspective. Paper presented in the 12th Annual Boston University Conference on Language Development, Boston.

Lust, B. (Ed.). (1986). *Studies in the acquisition of anaphora: Volume I. Defining the constraints.* Dordrecht, The Netherlands: Reidel.

Lust, B. (Ed.). (1987). *Studies in the acquisition of anaphora: Volume II. Applying the constraints.* Dordrecht, The Netherlands: Reidel.

Manzini, R., & Wexler, K. (1987). Parameters, Binding Theory, and learnability. *Linguistic Inquiry, 18,* 413–444.

McDaniel, D., & Cairns, H. S. (1987). *The child as informant: Eliciting linguistic intuitions from young children.* Unpublished manuscript, City University of New York.

McDaniel, D., Cairns, H. S., & Hsu, J. R. (1990). Binding principles in the grammars of young children. *Language Acquisition, 1,* 121–139.

McKee, C. (1988). *Italian children's mastery of binding.* Unpublished doctoral dissertation, the University of Connecticut, Storrs.

Montalbetti, M., & Wexler, K. (1985). Binding is linking. *Proceedings of the West Coast Conference on Formal Linguistics IV* (pp. 228–245). Stanford, CA: Stanford University Press.

Otsu, Y. (1981). *Towards a theory of syntactic development.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Reinhart, T. (1983a). *Anaphora and semantic interpretation.* London: Croom Helm.

Reinhart, T. (1983b). Conference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy, 6,* 47–88.

Reinhart, T. (1986). Center and periphery in the grammar of anaphora. In B. Lust (Ed.), *Studies in the acquisition of anaphora: Vol. I. Defining the constraints* (pp. 123–150). Dordrecht, The Netherlands: Reidel.

Sigurjónsóttir, S., Hyams, N., & Chien, Y.-C. (1988). The acquisition of reflexive and pronouns by Icelandic children. *Paper and Reports on Child Language Development, 27,* 97–106.

Silva, C. (1988). *Language acquisition of pronouns: Some aspects of Binding Theory.* Unpublished doctoral dissertation. University of California, Irvine.

Solan, L. (1983). *On the acquisition of pronominal reference.* Dordrecht, The Netherlands: Reidel.

Solan, L. (1987). Parameter setting and the development of pronouns and reflexives. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 189–210). Dordrecht, The Netherlands: Reidel.

Thornton, R. (1990). *Adventures in long-distance moving.* Unpublished doctoral dissertation, the University of Connecticut, Storrs.

Wexler, K. (1989, March). *Some issues in the growth of control.* Paper presented at the conference on control, Massachusetts Institute of Technology, Cambridge. (To appear in J. Higginbotham, S. Iatroadov, V. Lahiri, & R. Larson [Eds.], *Papers from the workshop on control.* Norwell, MA: Kluwer Academic.)

Wexler, K., & Chien, Y.-C. (1985). The development of lexical anaphors and pronouns. *Papers and Reports on Child Language Development, 24,* 138–149.

Wexler, K., & Chien, Y.-C. (1988, March). *The acquisition of locality principles in Binding Theory.* Paper presented at the 11th Generative Linguists of the Old World (GLOW) Colloquium, Budapest, Hungary.

Wexler, K., & Manzini, R. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds), *Parameter setting* (pp. 41–76). Dordrecht, The Netherlands: Reidel.

APPENDIX A. INDIVIDUAL SUBJECT ANALYSES AND MODELS OF KNOWLEDGE

A.1 Knowledge Models and Strength Models

As is often the case in studies of linguistic development, the observed behavior of children in many of the experiments reported in this article approaches the behavior of adults over time. For example, the mean probability that the group of subjects will take a local antecedent for a reflexive increases almost monotonically over a variety of experiments (see Figures 1, 2, 5, and 6). As we suggested, this increase means that at any stage some children do not know that the reflexive (*himself/herself*) is a reflexive. However, as soon as they know that *himself/herself* is a reflexive, children will assign a local antecedent to it. This interpretation assumes that there is such a thing as linguistic knowledge, along with the possibility that a child might or might not have this knowledge. The models on which these interpretations are based will be called knowledge models.

There is another possible interpretation that is more consistent with traditional nonlinguistic approaches to language acquisition. This interpretation suggests that it is not a case of whether or not a particular child knows the principle (or parameter value) of grammar. Rather, the increasing proportion of "correct" (adult-like) behavior shown by children simply reflects an increasing "strength" of the correct response behavior. This implies that children simply increase in their "tendency" of giving a correct response because of the increasing strength of this behavior. The models on which these interpretations are based will be called strength models.

Although strength models could possibly play a role in associative learning (this is disputable), we do not view this as having any plausibility with respect to linguistic development. For example, what would it mean to have a strength associated with the claim that a reflexive must be c-commanded by an antecedent? How would this strength be assigned? Given positive evidence, how would this strength grow?

Nevertheless, it is worth looking at the differing experimental predictions made by the strength and knowledge models. If the predictions from the strength model were correct, we would have to rethink the knowledge model. Although what constitutes the logical basis for the strength model is not clear, at least its statement is observationally adequate for the mean developmental curves reported in this and many other articles on linguistic development. This is because the knowledge and strength models do not differ on mean behaviors.

However, as is well known, these models do differ in predictions with

regard to distributions of individual subject patterns. The strength model says that at any given strength (e.g., measured by a probability of correct performance) subjects will show performance reflecting that strength. If we assume that the subjects at a given age all have the same strength, then their probability correct on any given item should be the same.[13] This does not mean that we should expect each observed child to show the same number correct. Rather, because each stimulus sentence is an independent trial, on which the child is correct with probability equal to the mean strength (i.e., these are Bernoulli trials), we would expect the number correct for a given age to be binomially distributed around the mean strength. More crucially, given enough trials and subjects, this distribution would have only one mode (with a bell shape like a normal distribution). (Informally speaking, if the mean proportion correct for a given age group were 3 out of 10, i.e., .3, we would expect the histogram of number correct to look like the histogram of number of heads, for a weighted coin in which the probability of a head turning up on a given trial were .3.)

Let us now consider the prediction about the histogram of number correct over individual subjects for the knowledge model. The essential idea of the knowledge model as developed here is that subjects either *know* or *do not know* the relevant linguistic properties. If they know the relevant properties, then, in an ideal situation, they should respond correctly with probability approaching 1 (i.e., they should respond correctly at each trial like adults). What should we assume about the child's behavior if he or she does not know the relevant linguistic properties? (For example, if the child does not know that *herself/himself* is a reflexive and only treats it as some kind of anaphoric element, which must get its reference in some way, then what kinds of response pattern should we expect?) Grammatical theory does not tell us the answer to this question. It is not clear what theory does tell us the answer. Presumably, the child responds according to a constellation of factors when more than one response is possible and grammar does not indicate what the appropriate analysis of a sentence is. The child might have pragmatic biases. There might be salience effects. There might be preferences for certain kinds of actions. A child might respond to ambiguity in all sorts of ways. This situation is quite typical in studies of cognitive development (in fact, in experimental psychological studies of all kinds, from signal detection theory to cognitive psychology).

Therefore, the knowledge model cannot provide an exact prodiction about how subjects will respond in detail when they do not know the

[13]This is an idealization. If we made a more finely correct assumption, for example, the observed subjects at a given age were drawn from a population that had strengths normally distributed around a mean strength, we would not change the essence of our conclusions. Therefore, we simplify here for this informal presentation.

grammatical properties. However, this knowledge model *does* make one crucial prediction. It predicts that the subjects will be partitioned into at least two different categories, namely a class of subjects who have the linguistic properties and a class of subjects who do not have the linguistic properties. Ideally, the first group of subjects will behave in adult-like fashion (in this case 100%), and the second group of subjects will behave otherwise. As long as the second class of subjects (those who do not have the linguistic properties) do not show a mode of response at 100% correct, then the knowledge model will predict a histogram of correct responses that is not unimodal. The histogram will be at least bimodal, with one mode at 100% and the other someplace else. If we assume that the distribution of correct responses for the subjects who do not have the linguistic knowledge is binomial around a mode that does not indicate 100% correct, then we would generally expect a bimodal histogram, with one mode at 100% and another somewhere else. However, if the children use a mixture of strategies we might obtain a multimodal, even a uniform distribution.

The amount of data to be presented quickly becomes large, so for ease of presentation we only present some of the data from some of the experiments. However, the patterns we present are in general typical of all the data.

A.2 Reflexives in Experiment 1

Consider Experiment 1, which investigated sentences involving reflexives and two potential antecedents such as sentence (7) given in the main body of the article. One instance of (7) is repeated here as (63).

(63) *Kitty* says that *Sarah* should point to *herself*.

The "correct" response is to refer *herself* to the local antecedent, by Principle A, so that Sarah should point to herself, not to Kitty. The distribution of subjects across correct responses for each age group is reported in Table 1. The results confirm the predictions of the knowledge model and disconfirm the predictions of the strength model.

The strength model predicts that for any age group the responses are binomially distributed around a single mode. Consider G3 through G6. For all of these, the two values of number correct scored by the largest numbers of subjects are 0 correct and 10 correct—at the opposite ends of the scale. For example, in G3, about 42.86% of the subjects have none correct, whereas 19.05% of the subjects have 10 correct. None of the other numbers correct have as large values as these. Consider G4, in which 25% of the subjects have none correct and 25% have 10 correct; none of the other numbers correct between 0 and 10 have as many subjects. These results are *extremely* nonunimodal. In fact, they tend to look bimodal, with 0 and 10 as the two modes. The percentages slide down as we go up from 0 and down from 10,

TABLE 1
Subject Analysis for Experiment 1: Numbers and Percentages of Subjects in Each Group by Number of Correct Responses to the Reflexive Sentences[a]

	Reflexive Sentences Numbers and Percentages of Correct Responses										
Group	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Group 1											
n = 15	11	1	0	1	0	0	0	1	0	1	0
2;08 (28)	73.33	6.67	00.00	6.67	00.00	00.00	00.00	6.67	00.00	6.67	00.00
Group 2											
n = 15	8	2	0	0	1	0	0	0	2	1	1
3;03 (00)	53.33	13.33	00.00	00.00	6.67	00.00	00.00	00.00	13.33	6.67	6.67
Group 3											
n = 21	9	2	0	0	0	1	0	2	0	3	4
3;09 (08)	42.86	9.52	00.00	00.00	00.00	4.76	00.00	9.52	00.00	14.29	19.05
Group 4											
n = 20	5	0	0	2	2	0	0	1	1	4	5
4;03 (10)	25.00	00.00	00.00	10.00	10.00	00.00	00.00	5.00	5.00	20.00	25.00
Group 5											
n = 28	5	1	1	1	1	0	0	3	1	3	12
4;09 (08)	17.86	3.57	3.57	3.57	3.57	00.00	00.00	10.71	3.57	10.71	42.86
Group 6											
n = 27	6	0	0	0	0	1	0	2	0	3	15
5;02 (19)	22.22	00.00	00.00	00.00	00.00	3.70	00.00	7.41	00.00	11.11	55.56
Group 7											
n = 15	1	2	0	0	0	0	1	0	1	1	9
5;08 (28)	6.67	13.33	00.00	00.00	00.00	00.00	6.67	00.00	6.67	6.67	60.00
Group 8											
n = 15	1	0	0	0	0	0	0	0	1	4	9
6;03 (02)	6.67	00.00	00.00	00.00	00.00	00.00	00.00	00.00	6.67	26.67	60.00
Total											
n = 156	46	8	1	4	4	2	1	9	6	20	55
4;06 (11)	29.49	5.13	0.64	2.56	2.56	1.28	0.64	5.77	3.85	12.82	32.56

[a]Correct response: reflexive = local c-commanding antecedent. Example: *Kitty$_i$ says that Sarah$_j$ should point to herself$_j$.*

with almost no subjects in the middle of the scale. This result is consistent with the knowledge model and is inconsistent with the strength model.[14]

Of course, if we look at the youngest and oldest subjects, the distributions tend more toward the unimodal, in a way quite understandable by the

[14]As we indicated earlier, the knowledge model does not in itself predict where the second (non-100%) mode will occur. That the second mode occurs at the zero is a consequence of the bias that children seem to have when faced with an anaphoric form with properties (e.g., anaphor or pronoun) that are not clear to them. This bias is for the child (e.g., Sarah) to point to the doll held up by the experimenter when she heard the sentence "Kitty says that Sarah should point to herself." We can only speculate on what the source of this bias is, but there is no reason to think that it is linguistic.

knowledge model. Consider the youngest children (i.e., children in G1). Out of 10 attempts, about 73% of the youngest children gave no correct responses. No children gave 10 correct responses, and only one child (6.67%) gave 9 correct responses. Basically, almost all of these G1 children appear to be in the state of no knowledge of the relevant linguistic properties. A few more children might show knowledge in G2, and by G3, more than 33% of the children have 9 or 10 correct.

On the other hand, among the oldest children (i.e., G8 children), just one child (6.67%) seems to be in the state of no linguistic knowledge (i.e., that child has 0 correct). Thirteen out of the 15 children (86.67%) have either 9 or 10 correct and the remaining child has 8 correct. In G7, it appears as if three subjects (1 or 2 correct) have no linguistic knowledge.

The youngest and oldest children show unimodal distributions, with the mode on opposite ends of the scale for these children. This result is in accordance with the knowledge model, which partitions the children into two categories: those who know and those who do not know the corresponding linguistic principle. It is possible that in the youngest age group almost none of the children know the principle. It is also possible that in the oldest age group almost none of the children do not know the principle. In the ages at the middle we see a strong bimodal distribution.

In summary, it is quite clear that the distribution of individual response patterns for the reflexive sentences (in Experiment 1) is inconsistent with the strength model. This model is ruled out by the data given, as we would have expected on theoretical grounds. On the other hand, the knowledge model seems to be completely confirmed.

A.3 Pronouns in Experiment 1

Let us discuss the distribution of individual responses for the pronoun sentences of Experiment 1. The related test sentences are like the one given in (8) in the main body of the article, which is repeated in (64).

(64) *Kitty* says that *Sarah* should point to *her*.

According to Principle B, Sarah should point to Kitty and not to herself (Sarah). As we have extensively discussed in this article, children do not improve much in this kind of response over four years. Unlike the case for the reflexive, the data for the pronoun do not appear in the mean as gradual increase toward perfect (adult) performance, at least for the ages that we have studied.

However, the issue of strength versus knowledge models still exists. Does a mean response rate of about 70% for the pronoun indicate that the strength model is correct or that the knowledge model is correct? According

to the strength model, children's responses will be binomially distributed around a single mode (i.e., about .7, or 7 correct responses out of the 10 trials). According to the knowledge model, children's responses will be distributed around a mode of 100% correct (i.e., 10 correct responses out of 10 trials) and at least one additional mode, or a somewhat uniform distribution of the other responses.

The results are summarized in Table 2. Again, the experimental results (as can be seen from Table 2) are quite consistent with the knowledge rather than the strength model.

For every one of the eight groups, the modal response is 10 out of 10

TABLE 2
Subject Analysis for Experiment 1: Numbers and Percentages of Subjects in Each Group by Number of Correct Responses to the Pronoun Sentences[a]

	Pronoun Sentences Numbers and Percentages of Correct Responses										
Group	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Group 1											
n = 15	0	0	1	1	1	1	0	2	1	3	5
2;08 (28)	00.00	00.00	6.67	6.67	6.67	6.67	00.00	13.33	6.67	20.00	33.33
Group 2											
n = 15	1	1	0	0	0	1	0	1	4	2	5
3;03 (00)	6.67	6.67	00.00	00.00	00.00	6.67	00.00	6.67	26.67	13.33	33.33
Group 3											
n = 21	0	2	1	1	2	3	0	2	1	0	9
3;09 (08)	00.00	9.52	4.76	4.76	9.52	14.29	00.00	9.52	4.76	00.00	42.86
Group 4											
n = 20	0	2	4	0	0	2	1	1	3	2	5
4;03 (10)	00.00	10.00	20.00	00.00	00.00	10.00	5.00	5.00	15.00	10.00	25.00
Group 5											
n = 28	2	1	0	3	0	0	5	2	2	0	13
4;09 (08)	7.14	3.57	00.00	10.71	00.00	00.00	17.86	7.14	7.14	00.00	46.43
Group 6											
n = 27	1	3	1	2	3	1	0	2	1	4	9
5;02 (19)	3.70	11.11	3.70	7.,41	11.11	3.70	00.00	7.41	3.70	14.82	33.33
Group 7											
n = 15	0	0	1	1	1	1	0	2	0	1	8
5;08 (28)	00.00	00.00	6.67	6.67	6.67	6.67	00.00	13.33	00.00	6.67	53.33
Group 8											
n = 15	0	2	2	0	1	1	2	0	0	1	6
6;03 (02)	00.00	13.33	13.33	00.00	6.67	6.67	13.33	00.00	00.00	6.67	40.00
Total											
n = 156	4	11	10	8	8	10	8	12	12	13	60
4;06 (11)	2.56	7.05	6.41	5.13	5.13	6.41	5.13	7.69	7.69	8.33	38.46

[a]Correct response: Pronoun = nonlocal c-commanding antecedent. Example: *Kitty$_i$ says that Sarah$_j$ should point to her$_i$.*

correct. The numbers at this perfect mode range from 25% for G4 to 53.33% for G7. There is no hint of a mode around the mean correct response rate of .7. In general, the subjects who are not at (or near) perfection do not seem to show a particularly strong tendency; the rest of the distribution is somewhat flat, or jagged, but with no particular direction. There is not much of a tendency for the subjects to be wrong almost all the time, as there was with the reflexives. In other words, no strong cluster of subjects who show only 0 or 1 correct was found.

To take one example, consider G7. Out of the 15 subjects in this group, 9 (60%) gave 9 or 10 correct responses. If we assume that the remaining 6 subjects were in the no knowledge stage, we would expect 1.33 subjects to show 0 or 1 correct responses. The reason is that 6 out of the 15 subjects are spread out over 9 possible response patterns (i.e., 0 through 8 correct), therefore only 1.33 subjects should indicate either 1 correct or 0 correct [6 (subjects) divided by 9 (possible patterns) times 2 (i.e., 0 correct or 1 correct) equals 1.33]. However, as the data given in Table 2 show, no subjects in G7 gave 0 or 1 correct responses. In some of the other groups we do find a few more subjects having 0 or 1 correct, perhaps equal to what would be expected on a uniform assumption (e.g., G2). It seems that there is a definite cluster of subjects at every age who behave like adults and that there is no particular pattern, or perhaps there are a number of varying patterns that are not discernible for the other subjects.

One might argue that the mode at 10 of 10 correct (like the approximately 70% overall correct response rate) does not necessarily indicate that the children at this mode know the properties indicating that the local antecedent *Sarah* is not a possible antecedent for the pronoun in (64). In other words, one might argue that some of the children, especially the younger ones, were in the no knowledge state; they might just follow the bias toward pointing to the doll (Kitty) in (64) rather than the required linguistic principles. However, this argument seems to be quite unlikely for the pronoun condition, because there is strong reason to believe that children know a pronoun is a pronoun at an early age, as Wexler and Manzini (1987) pointed out. For example, it is well-known that the earliest productively produced anaphoric forms are pure pronouns, used without an intrasentential antecedent, as in (65).

(65) *Mary* saw *him*.

In (65), *him* does not have an antecedent; therefore, it must be a pronoun, not an anaphor, according to the Binding Theory. The fact that children use such forms productively indicates that they know they are pronouns. Moreover, the older children are not often influenced by this bias toward the doll, which was indicated in the reflexive case. As a result, it

seems most likely that a large number of the children (especially the older ones) who get 10 out of 10 correct are in fact aware of the linguistic properties (Principles B and P together, as discussed in this article) that make the local antecedent inappropriate in (64).

In summary, the knowledge model is strongly supported for both the reflexive and the pronomial data in Experiment 1. Children partition into two categories: those who show linguistic knowledge and thereby respond almost perfectly, and those who do not show such knowledge and thereby respond according to a variety of factors. Such strong support from individual subject analyses lends further credence to one major aspect of the underlying theory as presented here—that the final state is to be characterized as a knowledge state and that children may move into this knowledge state. Alternative theories, which do not characterize humans as having knowledge but rather characterize probabilities of responding correctly, are quite inconsistent with the data.

A.4 Experiment 4: Pronouns With Quantified and Name Antecedents

The general results of individual subject analyses for Experiments 2 and 3 replicate those for Experiment 1, especially with regard to the choice between the knowledge and the strength models. To save space we do not present these data. However, Experiment 4 is a crucial experiment in the demonstration that children know Principle B. Therefore, we report and discuss an individual subject analysis for the pronoun sentences from this experiment.

The numbers of subjects who had particular numbers correct are summarized in Tables 3 and 4. Table 3 shows the results for the name-pronoun mismatch condition, in which sentences like (66) together with a picture of Mama Bear touching herself were presented.

(66) This is Mama Bear; this is Goldilocks. Is Mama Bear touching her?

For concreteness, focus on G3. (This is the group that is central to the comparison between the name-pronoun and the quantifier-pronoun conditions because G1 and G2 children do not know concepts involving the quantifier.) Does it appear as if the results of G3 are generated from a strength model? No. The strength model predicts that the responses are binomially distributed around the mean value. It is clear on inspection that there are far too many values at the extremes (0 and 6 correct) for such a binomial model to be correct. We know that the mean correct value for this

TABLE 3
Subject Analysis for Experiment 4: Numbers and Percentages of Subjects in Each Group by Number of Correct Responses to the Name-Pronoun Mismatch Sentences[a]

	Name-Pronoun Mismatch Sentences Numbers and Percentages of Correct Responses						
Group	*0*	*1*	*2*	*3*	*4*	*5*	*6*
Group 1							
n = 48	13	13	7	6	4	2	3
3;05 (04)	27.08	27.08	14.58	12.50	8.33	4.17	6.25
Group 2							
n = 45	11	8	7	7	3	3	6
4;06 (05)	24.44	17.78	15.56	15.56	6.67	6.67	13.33
Group 3							
n = 44	7	5	5	9	7	6	5
5;05 (00)	15.91	11.36	11.36	20.45	15.91	13.64	11.36
Group 4							
n = 40	2	2	3	3	3	7	20
6;04 (04)	5.00	5.00	7.50	7.50	7.50	17.50	50.00
Total							
n = 177	33	28	22	25	17	18	34
4;10 (09)	18.64	15.82	12.43	14.12	9.60	10.17	19.21

[a]Question: This is Mama Bear; this is Goldilocks. Is Mama Bear touching her? Picture: Mama Bear is touching (Mama Bear) herself (i.e., pronoun = local c-commanding antecedent). Correct response: "No."

group is about .5. Therefore, the probability of a child's getting zero correct under the binomial model is .5 to the sixth power (= .0156). That is, we expect less than 2% of the subjects to have 0 correct. However, the results are that 15.91% (7 of 44) have 0 correct. Similarly, we expect that less than 2% will have 6 correct. However, 11.36% (5 of the 44) of the subjects have 6 correct.

It is clear that the binomial model is not correct for G3. The same is true for G1 and G2, although for those groups there is a stronger tendency to have 0 correct. Presumably this has to do with a tendency to say yes when the child's analysis is that the sentence is ambiguous and that either yes or no could be correct. However, this is only a tendency, and some complex biasing process determines whether the child will say yes or no to any given sentence.

By G4, 50% of the subjects have all 6 correct. There is a major jump in performance from G3 to G4. Twenty-five percent of the G3 subjects have 5 or 6 correct, whereas 67.5% of the G4 subjects have 5 or 6 correct. The average age for G4 is 6 years and 4 months. It appears that around this age a good number of the subjects have acquired Principle P. However, 25% of

TABLE 4
Subject Analysis for Experiment 4: Numbers and Percentages of Subjects in Each Group by Number of Correct Responses to the Quantifier-Pronoun Mismatch Sentences[a]

	Quantifier-Pronoun Mismatch Sentences Numbers and Percentages of Correct Responses						
Group	*0*	*1*	*2*	*3*	*4*	*5*	*6*
Group 1							
n = 48	9	4	11	7	4	6	7
3;05 (04)	18.75	8.33	22.92	14.58	8.33	12.50	14.58
Group 2							
n = 45	5	3	4	9	6	8	10
4;06 (05)	11.11	6.67	8.89	20.00	13.33	17.78	22.22
Group 3							
n = 44	0	1	2	3	7	7	24
5;05 (00)	00.00	2.27	4.55	6.82	15.91	15.91	54.55
Group 4							
n = 40	0	1	3	0	4	7	25
6;04 (04)	00.00	2.50	7.50	00.00	10.00	17.50	62.50
Total							
n = 177	14	9	20	19	21	28	66
4;10 (09)	7.91	5.08	11.30	10.73	11.86	15.82	37.29

[a]Question: These are the bears; this is Goldilocks. Is every bear touching her? Picture: Three bears (B1, B2, & B3); B1 is touching (B1) herself; B2 is touching (B2) herself; B3 is touching (B3) herself (i.e., pronoun = local c-commanding antecedent). Correct response: "No."

the subjects get 3 or fewer correct.[15] Again, the knowledge model seems to be correct. A certain group of subjects knows the relevant linguistic property, getting all or almost all the sentences correct. The other subjects do not know the relevant property and either respond randomly or according to some response bias.

Table 4 shows the response distribution for mismatch sentences with quantified antecedents. An example is given in (67). The corresponding picture depicts that each of three bears is touching herself.

(67) These are the bears; this is Goldilocks. Is every bear touching her?

Principle B dictates a no answer to (67). We only discuss G3 and G4 because, as discussed earlier, the evidence from control sentences shows

[15]The results are not different than the results for the same age subjects (G8) in Experiment 1 (see Table 2). It is difficult to directly compare the numbers because of different experimental methods and stimuli.

that the G1 and G2 children have not yet learned relevant properties of the quantifier.

About 55% of the G3 children have all 6 correct. More than 70% have 5 or 6 correct (31 of 44 subjects). None of the children have 0 correct; only one child has one correct and only three children have 2 or fewer correct. It is quite clear that there is a very small tendency for the children to respond in an ungrammatical fashion. There appears to be no bias or pattern to select the ungrammatical response, because none of the children show 0 correct responses. This analysis together with the set of data for the same G3 children indicated in Table 3 make it clear that most of the children know Principle B, whereas many still do not know Principle P.

Consider G4. Even more of the children show knowledge of Principle B. Thirty-two out of 40 children (80%) get 5 or 6 correct. And 90% get 4 or more of the 6 correct. The best summary for G3 and G4 seems to be that all (with perhaps a very small number of exceptions) of the children know Principle B and that the few incorrect responses represent experimental noise, inattention, processing difficulties, and so on.

Data from G1 and G2 are somewhat difficult to interpret because many of the children do not understand the true meaning of a quantifier. However, if we compare Table 4 with Table 3, it is clear that more of the subjects in the quantifier-pronoun condition get 5 or 6 of the sentences correct (Table 4) than in the name-pronoun condition (Table 3). The numbers are summarized as follows: in G1, 13 subjects for the quantifier-pronoun condition and 5 for the name-pronoun condition. In G2, 18 for the quantifier-pronoun condition versus 9 for the name-pronoun condition. In short, even at these younger ages, the differential effects of Principle B and Principle P show up. We have to assume that the responses in these groups that are inconsistent with Principle B are the result of lack of knowledge of the quantifier. The control sentences independently demonstrated this lack of knowledge. Given that a child does not know the interpretation of the quantifier, or does not even know that it is a quantifier, he or she must resort to some strategy of responding.

In summary, the distribution of individual subject responses concerning the pronoun cases tested in Experiment 4 confirms the knowledge model. The set of data is in strong contradiction to the strength model. There is no indication that children learn by increasing the probability of a particular behavior. Rather, there is evidence indicating that even very young children know Principle B although the knowledge of Principle P may develop later.

APPENDIX B. SOME APPARENTLY CONFLICTING DATA

Kaufman (1988) claims that she has obtained a different set of results than those we have reported. She claims that (a) children make many fewer

mistakes in which a pronoun is made to be coreferential with a local c-commanding antecedent, and (b) there is no difference between children's responses to the pronoun sentences with a quantified antecedent and their responses to those with a referential antecedent. There are a number of problems with Kaufman's methods and analyses. To save space, we do not provide a detailed critique here but restrict ourselves to a discussion of the "truth value judgment task" adopted by Kaufman and some related data.

The method used by Kaufman was adopted from Crain and McKee (1985). Children were made to observe a situation acted out by an experimenter, hear a description of the situation given by a puppet "Kermit," and then asked to feed Kermit a cookie if the description was correct or a rag if it was incorrect. It is a method of obtaining yes/no judgments, with the logic being similar to the methods used in our Experiment 4. An example of Kaufman's task is as follows: The child hears the sentence "The boy sees the dog" and sees a dog licking itself, then the child is given a test sentence such as (68).

(68) *The dog* licked *him.*

Given the situation described, (68) is a mismatch, and the child should feed Kermit a rag (i.e., respond "no", which is the description we use from now on).

What are the results? (To save space we will concentrate on correct rejections and not on acceptances. Essentially children did very well on acceptances.) For simple sentences like (68), both groups of children tested by Kaufman correctly rejected the situation, indicating a coreferential relation between the pronoun (*him*) and the sentence internal antecedent (i.e., *the dog*) about 88% of the time. However, for complex sentences that included sentences like (69) and some other types, the youngest children only correctly rejected the local antecedent for the pronoun about 56% of the time; the older children did this about 64% of the time.

(69) *The dog* wanted *the boy* to wash *him.*

Kaufman argues that the results on (68) indicate that the children know everything that adults do about disjoint reference, and that the poor results on (69) are due to the difficulty of processing or understanding complex sentences. However, we do not recommend this proposal because of the following reasons: The older group of subjects tested by Kaufman were between the ages of 5;0 and 6;5. It is quite well known that children understand and process complex sentences considerably before this age, and the results of many Binding Theory studies using complex sentences confirm this view. For example, as evidenced in Experiments 1 through 3 reported in this article, children do quite well with complex sentences when

knowledge of Principle A is required. Therefore, the small number of correct responses on complex sentences are a real problem for the hypothesis proposed by Kaufman, which assumes that there is no difference at all in children's and adult's competence concerning pronouns. (Recall that the theory proposed in our article is that children have Principle B but not P. Kaufman does not have such a possibility open as she assumes that children are lacking nothing relevant to either binding or coreference.)

Consider the results concerning the simple sentences again. Why did Kaufman's children do so well on simple sentences like (68)? The 88% correct response rate was considerably high compared to the approximately 50% correct for the 5–6-year-olds and the 76% correct for the 6–7-year-olds in our Experiment 4 (as indicated in Figure 13). An example of our test sentences and the corresponding picture is given in (60) in the main body of the article, which is repeated in (70). (In the corresponding picture, Goldilocks is facing Mama Bear; Mama Bear is touching herself.)

(70) This is Mama Bear; this is Goldilocks. Is *Mama Bear* touching *her?*

In our study, when children are presented with a picture indicating that Mama Bear is touching herself and a pronoun sentence like (70), their correct performance is low. However, when they are presented with the same picture (or a different picture in which Mama Bear is touching Goldilocks) along with a sentence such as (70) substituting the reflexive *herself* for the pronoun *her,* their correct performance approached the adults' level. Kaufman did not include reflexive sentences in her study; there is no way of seeing the relevant comparison.

Moreover, and even more problematically, the lack of reflexives created a large bias in Kaufman's stimuli. Looking at her sentences, it is clear that a *self-action* is *never* the correct response. In other words, a reflexive response is never the correct response. Therefore, if the children have any tendency at all to prefer a nonreflexive action, this is a bias that can be extended throughout the experiment. In our Experiment 4 (and in our other experiments, as a matter of methodological principle), both pronouns and reflexives were included in equal amounts and randomly interspersed.

Consider (68) again. When a child hears "The boy sees the dog," sees the dog licking itself, then hears the test sentence "The dog licked him," what is the child likely to think? A Gricean principle of relevance would suggest that the boy is more relevant to the second sentence than the dog. This would further imply that there should be a tendency to make *him* have *the boy* as its antecedent. This is not necessarily true for all sentences, but it seems to be often (or even always) true in Kaufman's study. Given the lack

of clear correct responses with a reflexive action, it is no wonder that children may develop a tendency to perform nonreflexive actions.

Consider (69) from Kaufman. The experimenter creates a situation in which he or she makes the dog say, "I'm clean. You're dirty, go wash yourself boy." At the same time, the experimenter shows the child a scenario in which a boy is washing himself. Then the child hears the mismatched test sentence "*The dog* wanted *the boy* to wash *him*." In this case, the nonlocal antecedent *the dog* already has a relevance established to the embedded clause, namely the dog *wants* the proposition implied in the embedded clause to occur. It is not necessary for the child to maintain relevance between *the dog* and the pronoun *him* by assigning *the dog* as the antecedent for the pronoun. As a result, the child can sometimes go along with the tendency to give an incorrect yes response to the test sentence (69). (Everything else being equal, there is much evidence for this kind of response shown in our studies as well as Kaufman's and other related studies.) In summary, the lack of incorrect coreference in (68) is due to an artifact. Local binding of the pronoun in (69) confirms our results.

Let us know consider Kaufman's second claim, which stated that children do not perform differently in reaction to pronoun sentences involving a quantified antecedent than to those pronoun sentences involving a referential antecedent. The sentence given in (71) is an example of a test construction with a quantified antecedent.

(71) *Every dog* licked *him*.

We would like to argue that it is quite difficult for children's responses to be better on the quantified antecedent cases (e.g., (71)) than the simple referential antecedent cases (e.g., (68)), because in Kaufman's study children's correct response rate to the referential cases (e.g., (68)) is approaching adults' level (i.e., 88% correct) (due to some response biases and pragmatic effects, as discussed earlier in this appendix).

However, consider a case of a complex sentence with quantified antecedent, as in (72).

(72) Every dog wanted the boy to wash him.

On complex sentences such as (72), children only showed 64% correct rejections. In other words, when a situation in which a boy is washing himself was presented to the child and the child heard (72), he or she only responded "no" 64% of the time. Children's correct response rate regarding this type of quantifier sentences (e.g., (72)) is very close to their correct response rate regarding complex pronoun sentences with referential antecedents. Based on this set of data, Kaufman argued that children do not

perform differently in reaction to sentences involving a quantified antecedent than to those involving a referential antecedent. However, looking closely to the stimulus sentences used by Kaufman, it is clear that the appearance of the quantifier in sentences like (72) is irrelevant to our claim that children do not make mistakes on pronouns locally bound by quantifiers. This is because the quantifier *every dog* in (72) does *not* locally bind the pronoun *him*. A sentence like (73) would be more relevant, however, Kaufman did not test children on this type of construction.

(73) *The dog* wanted *every boy* to wash *him*.

In short, the quantifier sentences used by Kaufman were irrelevant to the claim about children's knowledge of Principle B versus Principle P. Kaufman also tested children on some other quantifier sentences with adjuncts. Two examples are given in (74) and (75).

(74) Every monkey scratched him, while the elephant was watching.

(75) Every girl splashed her, when the lady was at the pool.

The results indicated that children did not make too many errors in (75), only 13 out of 120, approximately 11% errors, on a par with our results for Principle B in Experiment 4. (Kaufman did not separate analyses of sentences including referential antecedents from those including quantifier antecedents.) However, there were many more errors in (74), 43 out of 120. The question is why doesn't Principle B prevent the errors here? Kaufman assumes children cannot "process" these sentences, but we have already pointed out that this is an untenable assumption.

We propose one possible alternative here, which has to do with the verbs used in Kaufman's study. For example, notice how *scratched him* can be easily replaced by *scratched*. The word *scratch* is a kind of inherent reflexive, and it is very likely that children may have read the sentences in that way, by ignoring *him*, as the correct thematic roles are obtained without it. The only other two examples of this kind of sentence are (76) and (77).

(76) Every baby fed her, while the chicken was watching.

(77) Every horse hurt him, when the pig came to visit.

The verbs used in both (76) and (77) seem to share the inherent reflexive properties of *scratched*. The children made numerous errors on both (45 and 26, respectively). Only one of the verbs clearly does not seem to have

this inherent reflexive meaning: *splash,* as in (75). *Mary splashed* does *not* mean *Mary splashed herself.* This is the one verb that produces very few errors when the pronoun has a quantified antecedent. In short, Kaufman seems to have produced Principle B errors by choosing verbs that caused the child to ignore the pronoun.

Much more could be discussed; however, it is clear that our approach is quite consistent with the data. We assumed in our approach that pronoun sentences are ambiguous for children because accidental coreference is possible to more than one NP and this NP can be pragmatically determined. In Kaufman's data, the violations for complex sentences of the constraint against local c-commanding antecedents for a pronoun are too many and too strong. These should not allow her to claim that children behave in the same way as adults when pronominal anaphora is considered. For the most part, Kaufman's data fall within the scope of the generalizations made in our article and confirmed by results in Cairns and McDaniel (1987); McDaniel, Cairns, and Hsu (1990); and Thornton (1990).