The Origins of Grammar: Evidence from Early Language Comprehension

Kathy Hirsh-Pasek and Roberta M. Golinkoff

The MIT Press

From The MIT Press

MITCogNet

Second printing, 1997

This book was set in Times Roman by Asco Trade Typesetting Ltd., Hong Kong
Printed on recycled paper and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Hirsh–Pasek, Kathy.
The origins of grammar : evidence from early language comprehension / Kathy Hirsh–Pasek and Roberta Michnick Golinkoff.
p. cm.
Includes bibliographical references and index.
ISBN 0-262-08242-X
1. Language acquisition. 2. Grammar, Comparative and general—Syntax. 3. Comprehension in children. I. Golinkoff, Roberta M. II. Title.
P118.H57 1996
401'.93—dc20 95-44507
CIP

MIT Press
0-262-08242-X
Hirsh-Pasek, Kathy and Michnick Golinkoff, Roberta
The Origins of Grammar

Chapter 6

Young Children's Use of Syntactic Frames to Derive Meaning

With Letitia Naigles

In prior chapters we have reported on research that investigated whether infants can detect sentence units (chapter 4) and whether they use the order of these units to derive meaning (chapter 5). In this chapter we focus on both units and relations: on units, because the general question posed is how children learn about verbs; and on relations, because we are concerned with whether children can use a verb's argument structure to derive information about the meaning of that verb.[1] Specifically, then, we discuss how children learn a type of language unit—verbs—that many have considered the cornerstone of the sentence (e.g., see Bloom 1978; Chafe 1971; Chomsky 1981; Bresnan 1978). The larger issue that we address is whether children *can* use syntactic information, combined with observation of the world, to make conjectures about meaning.

Consider the distinction between transitive and intransitive verbs as one example of how grammatical information can be used to predict semantic information. In English a *transitive* verb must be associated with two noun phrases (arguments), one describing the agent of the action and the other describing the patient of the action, as in "Jordan gorps Benjamin." An *intransitive* verb requires no patient: "Allison and Michael blick" is perfectly acceptable. The very fact that there are two arguments in the transitive sentence leads the reader to assume that the first actor must be affecting the second actor. That is, Jordan is either *making* Benjamin do something or *contacting* Benjamin in some way. The reader does not assume such a relation with the intransitive sentence, "Allison and Michael blick," because Allison and Michael make up a conjoined, single argument for the verb "blick." Here, Allison and Michael are presumed to be doing something *together*. No contact or causation is implied. The question that we will address in this chapter is whether learners can capitalize on this relationship between form and meaning to assist them

in the learning of verbs, that is, whether they can capitalize on the subcategorization frames that surround a verb to derive some of the verb's meaning.

How *do* children learn verbs? As Quine (1960) points out, logically a given word might have any of an infinite number of possible meanings. For example, consider a child who hears "Daddy's dancing." She sees her father raising his legs repetitively, smiling, flapping his arms, and moving in a circle. To which aspect of the event—including possibly the event as a whole—does the verb refer? Although the indeterminacy of reference is a general problem for all word learning, many argue that verbs pose a special problem for the learner (Golinkoff et al. 1995; Golinkoff, Jacquet, and Hirsh-Pasek 1994; Tomasello and Merriman 1995). Gentner (1983) entertains a number of hypotheses about why verbs appear to be more difficult to acquire than nouns (but see Bloom, Tinker, and Margulis, in press; Tardif 1994). Although most nouns that appear in early vocabularies label concrete objects that are relatively permanent and tangible, verbs label actions that are often relational by definition, ephemeral in nature, and performed by different actors. Furthermore, languages conflate different aspects of meaning into their verbs. For example, manner is part of the verb in English (as in "hammered on the door" versus "knocked on the door"), whereas in Spanish the manner in which the action is performed is not part of the verb but requires a separate adverb.

One way in which the child might "break into" the verb system is through observation. For example, upon hearing the verb "wash," the child sees a "washer" and a "washee" (perhaps a body part or an inanimate object), but upon hearing the verb "run," the child sees a single individual performing an action. That the first verbs are learned from pairing words with actions is embedded in a position known as the "semantic bootstrapping hypothesis" (Grimshaw 1981; Pinker 1984, 1987). The problem is that children need to learn how the language that they hear maps onto their internal grammar. Under the bootstrapping view, the first basic discoveries about their grammar come when children pair concrete objects with the grammatical category of nouns and actions with the category of verbs. Thus, although semantic bootstrapping is not formally a theory of how children learn verb meanings, the theory posits that they will learn their first verbs by "lifting themselves up by their bootstraps," by pairing actions that they see in situational contexts (meanings) with syntax (forms). It is only after this initial insight that young children

come to use the syntactic frame in concert with observation to inform noun or verb meaning.

An alternative view of initial verb learning posits that even at the outset, observation taken alone offers inconclusive cues to verb identification. Children need to attend to a verb's arguments to deduce its meaning. Gleitman and Gillette (1995) have made the strongest case for this position. They showed adults videotapes of mothers and young children interacting. Every time the mother said the target verb, a beep was substituted for the word. The adults' task was to guess what verb the mother was saying at that moment. Even though adults apparently limited their choice of verbs given the age of the addressee (e.g., they offered "think" rather than "ponder"), they failed dismally in being able to predict the correct verb. In short, although situational context can perhaps reduce the range of what a verb might refer to, observation of the situational context is clearly not sufficient for verb learning. Thus, with only situational observation available, verb learning might be impossible.

Situational observation is surely an important component of verb learning. For example, seeing two individuals involved in a physical altercation can rule out a vast number of candidate meanings for the verbs "fight." Yet Gleitman and Gillette's study indicates that the number of candidate meanings for a potential verb is still too large to be of much use to the child. As Gleitman (1990) writes:

> I think the problem is that words don't describe events *simpliciter*. If that's all words did, we wouldn't have to talk. We could just point to what's happening, grunting all the while. But instead, or in addition, the verbs seem to describe specific perspectives taken on those events by the speaker, perspectives that are not "in the events" in any direct way.... Since verbs represent not only events but the intents, beliefs, and perspective of the speakers on those events, the meanings of the verbs can't be extracted solely by observing the events. (p. 17)

As an example of how the same event can be accurately described using different verbs, consider an event in which a dog is chasing a fox. The fox is *running*, but it is also being *chased*. Though "is running" and "is being chased" are equally plausible descriptors of the event, the sentence "The fox is running from the dog" focuses attention on the fox, whereas the sentence "The dog is chasing the fox" highlights the dog's role in the event. Note that the two verbs participate in distinct syntactic environments. "Run" can take the preposition "from"; "chase" cannot.

The difficulty of learning verbs through observation alone gave rise to a view of verb learning first suggested by Landau and Gleitman (1985) as

the syntactic bootstrapping hypothesis. Under this hypothesis, the child inspects not only the world, but also the syntactic contexts in which a verb is used, to make predictions about its meaning. The syntactic context in which a verb participates acts like a "zoom lens" (Gleitman 1990) that focuses the listener on one aspect or another of the action that is witnessed (in the present case, the chasing or the running).

In contrast to the semantic bootstrapping hypothesis, the syntactic bootstrapping hypothesis *is* a theory of verb learning. It presupposes that children analyze events into predicate-argument structures, and that they link sentences to the event structure that they parse. Children hearing a sentence can map the nouns onto the arguments that they observe. Given this sensitivity to noun-argument linkages, children find the verb (the predicate) through the process of elimination. They then deduce the verb's meaning through the joint contribution of the syntactic context in which the verb is embedded and observation of the extralinguistic scene (Gleitman and Gillette 1995). For example, the child could know that the action of *giving* is being described if she observes an object being transferred and hears the sentence "John is giving a present to Mary." Under this view, the child knows (1) that "John," "present," and "Mary" are nouns in her lexicon and elements in the scene; (2) that John is the agent of the observed action (see Mandler 1992) and that agents typically become sentence subjects (these are "linking rules"; see Pinker 1989; Jackendoff 1972); (3) that the preposition "to" implies directionality. Through the process of elimination the child deduces that the word "giving," the only unspecified part of the sentence, must represent the predicate and hence the verb. Thus, the first part of the syntactic bootstrapping hypothesis allows the child to find the verb. The second part allows the child to deduce the meaning of the verb by noting the arrangement and number of the verb arguments. Upon hearing the verb "give" with its three arguments ("*X* is giving a *Y* to *Z*"), a "to" phrase, and an agent "John" as subject, the child can deduce that the action of transfer is being described from the perspective of the agent. This makes the child likely to interpret the verb as having the meaning associated with "give" rather than the meaning associated with "receive."

The syntactic bootstrapping hypothesis rests on the assumption that the presence of certain syntactic frames roughly correlates with meanings. In a series of experiments in English and Italian, Fisher, Gleitman, and Gleitman (1991) showed that there are strong and reliable parallels between the structural and semantic properties of verbs. That is, verbs that

were judged to share certain syntactic properties were also judged to be similar in meaning. For example, verbs that take three arguments, like "give" and "receive," were rated as more similar to each other in meaning (both have to do with transfer of an object) than verbs that do not share this structural property. These results have been replicated in Hebrew (Geyer, Gleitman, and Gleitman 1995) and Mandarin (Li, Gleitman, and Gleitman 1994). With this as evidence, the syntactic bootstrapping hypothesis holds that young children could potentially deduce part of the meaning of the verb by noting the sentence structure—particularly, the number and arrangement of the verb arguments in the sentence.

In sum, different theories have been advanced about how children might come to learn the meaning of verbs. Among these theories, semantic bootstrapping and syntactic bootstrapping are often seen as rival hypotheses. For the semantic bootstrapper, observation of the situational context is the key to entering the grammatical system. The same action observed in multiple contexts provides the basis for the mapping between a verb and its meaning. For the syntactic bootstrapper, although observation is important, verb meaning is not *learned* through observation. Rather, verb meaning is *informed* by observation in conjunction with knowledge of the noun arguments in the sentence frame. The child who can decipher the configuration of noun arguments in the sentence can then figure out not only where the verb must be, but also what the verb is likely to mean. As the child hears a verb used in multiple language frames, the meaning of the verb becomes even more refined.

Although semantic bootstrapping and syntactic bootstrapping have their differences, they also share many similarities. First and foremost, both credit the language learner with a rich, innate linguistic base that must be "linked" to a given language as the learner hears language used in context. Second, both assume that after the initial breakthrough—whether it be through observation as in semantic bootstrapping or syntax as in syntactic bootstrapping—syntax and semantics create a dialectic that allows the child to pair form with meaning and meaning with form. That is, although they have different starting points, both theories hypothesize a learner who is sensitive not only to observational and contextual cues, but also to syntactic information in the input. The question that must be addressed—and the focus of this chapter—is whether young children who have just begun to produce two-word utterances themselves are even capable of attending to argument structure configurations in the speech that they hear.

To explore whether children can exploit the syntactic frames in which verbs are embedded to make inductions about verb meanings, children were presented with contrasting sentences that used either transitive or intransitive sentence frames. Research suggests that there exists a strong, though not perfect, correlation between transitive frames and causal action and between nontransitive frames and noncausal action (at least within the class of motion verbs). If children capitalize on this correlation, then they should, upon hearing a transitive sentence like "Where's Big Bird bending Cookie Monster?", look significantly more at a scene in which Big Bird is causing Cookie Monster to bend than at a scene in which Big Bird and Cookie Monster are both bending. Similarly, children hearing the intransitive sentence "Find Cookie Monster and Big Bird bending" should choose to watch the noncausal alternative in which the two characters are bending together. If children are sensitive to verb subcategorization frames, then they should attribute different meanings to the same verb (here, "bend") when it is used in one sentence frame as opposed to another.

Prior research indicates that young children can attend to verb syntax. For example, children presented with an enactment task were willing to alter the meaning of a familiar verb to conform to the frame in which it was encountered—a "frame-compliant" response. The younger the children, the more frame-compliant (as opposed to "verb-compliant") responses they gave. Using a Noah's ark model and toy wooden animals, Naigles, Gleitman, and Gleitman (1992) presented 2-, 3-, and 4-year-old children with transitive verbs used intransitively (e.g., *"The zebra brings") and intransitive verbs used transitively (e.g., *"The elephant comes the giraffe"). The younger subjects acted out these novel sentences—sentences that really have no right or wrong interpretation—in accord with the demands of the frame, changing the meaning of the verb. So, for example, for the sentence *"The elephant comes the giraffe," a frame-compliant response was scored if the child acted the sentence out by having the elephant make the giraffe move somewhere, as if the verb "come" had become causal. What Naigles, Gleitman, and Gleitman's results show, therefore, is that young children are capable of using verb syntax or, in other words, using the forms in which verbs occur to predict something of their meaning.

In their study Naigles, Gleitman, and Gleitman used children no younger than 30 months because of the demands of enactment tasks. They also used only familiar verbs. The intermodal preferential looking paradigm

enabled us to follow up on their results with even younger children and with both familiar and unfamiliar verbs. Would younger children, who are in the throes of verb learning, give frame-compliant responses with both familiar and unfamiliar verbs? Would even younger children show causal interpretations of verbs that appeared in transitive frames and noncausal interpretations of verbs that appeared in intransitive frames?

In part, then, the studies to be reported in this chapter were undertaken to test the assumption that children are sensitive to verb syntax as would be required by both the semantic (Grimshaw 1981; Pinker 1984, 1994) and syntactic bootstrapping hypotheses (Landau and Gleitman 1985; Gleitman 1990). More specifically, these experiments focus on the role that attention to syntactic frames might play in verb interpretation. To investigate this question in the intermodal preferential looking paradigm, children were shown a causal and a noncausal version of the same action. For example, either Big Bird was seen causing Cookie Monster to bend or Big Bird and Cookie Monster were seen bending together in synchrony. Without accompanying language stimuli, there should be no inherent preference for either of these scenes. Both events have the same characters and both show variants of the same action. If children recognize the link between syntactic forms and meaning, however, the introduction of language stimuli should drive a looking preference for one of these scenes over the other.

To test this hypothesis, four experiments were conducted with children in three different age groups, using known and unknown verbs and several different transitive and intransitive sentence frames. The first (and main) experiment, here called experiment 5, employed a cross-sectional design in which children from 18 to 30 months heard either transitive or intransitive stimulus sentences with known and unknown verbs. Positive results were obtained, but mostly with the older children and only with the transitive sentence frames. Experiment 6 presented four nonsense verbs in transitive sentences to replicate and secure the findings of experiment 5. Finally, by using intransitive frames that gave subjects additional linguistic cues to an intransitive sentence interpretation, experiments 7 and 8 attempted to determine why only the transitive sentences supported the hypothesis in experiment 5. Experiment 7 used intransitive sentence frames in which the oblique marker "with" signaled the intransitive interpretation (e.g., "Big Bird is squatting *with* Cookie Monster!"). Experiment 8 used the auxiliary "are" to buttress the intransitive interpretation of joint, though not causal, action (e.g., "Big Bird and Cookie Monster *are* squatting!").

6.1 Experiment 5: Do Young Children Interpret Verb Meanings According to Their Syntactic Frames? A Cross-Sectional Study of Frame Compliance

Experiment 5 was designed to investigate when children at different ages attend to and use a verb's subcategorization frames to predict meaning. For example, if children are given a transitive sentence frame with an argument on either side of the verb, are they likely to think that the sentence implies a causal relation between the entities referred to by the arguments of the verb? On the other hand, if children hear only a single argument preceding the verb, will they give the verb an intransitive or noncausal interpretation? Children who are aware of the link between syntactic frame structure and meaning should watch a causal scene when they hear a transitive motion verb and a noncausal scene when they hear an intransitive verb.

Another goal of this study was to explore whether children were equally capable of using syntactic frames to guide their semantic interpretations for known ("bend" and "turn") versus unknown verbs ("squat" and "flex"). Both "turn" and "bend" can be used in both transitive and intransitive frames, for example, "She's bending the pipe" and "She's bending (herself)." If children assign different meanings to a known verb (such as "bend") when it appears in a transitive versus an intransitive sentence frame, they may be sensitive to these subcategorization frames and to the meanings with which they correlate. Alternatively, children could have memorized the relation between the frames and the meanings for known verbs. Thus, even if children can look to the correct screen for known verbs, this still leaves open the question of whether they can use sentence frames to predict something of the meaning of what are, for them, novel verbs. If children use syntactic frames to assign different meanings to unknown verbs (such as "squat"), they must be doing so by a general rule that unites syntactic form and meaning. Therefore, a result indicating frame-compliant interpretations with unknown verbs would provide stronger evidence that toddlers can use syntax to predict verb meaning. In addition, the two unknown verbs selected ("flex" and "squat") can actually only be used intransitively. The ability to predict the meaning of these verbs from transitive sentence frames would support the argument that children do not know these verbs and that they are using the subcategorization frames to infer their meaning.

Another possibility, however, and a persistent problem for language studies, is that the meaning of unknown verbs could be assigned by analogy to the meaning of known verbs. That is, children may hear a novel verb (say, "squat") with two arguments and assume, by analogy with the known verb "bend," that it must refer to the event in which one character is making the other character act. This possibility cannot be eliminated, since analogy is one of the mechanisms children use to learn language (MacWhinney 1978). Nonetheless, even if children are using analogy, it could be argued that if they are capable of responding correctly to unknown verbs, they must have made some broader generalizations based on the relations between syntactic surface form and semantic information.

Method

The paradigm used in all experiments was the intermodal preferential looking paradigm as described in chapter 3.

Design of the Video Events The same video events were used in all four experiments; only the linguistic stimuli paired with the events differed. For this reason, the video events are described only for experiment 5; the description is not repeated for experiments 6–8.

Table 6.1 shows the layout of the tapes for the verb "squat" and the linguistic stimuli for experiment 5. Figure 6.1 shows the visual events for the verb "squat," with the causal event on the left. As table 6.2 indicates, the other three verbs used in experiments 5, 7, and 8 were "turn," "bend," and "flex." These were presented in exactly the same format as displayed in table 6.1. The stimuli were chosen as examples of known ("turn," "bend") and unknown verbs ("flex," "squat")—a categorization reaffirmed by all parents when they were questioned about their child's comprehension vocabulary in the laboratory. The verbs were also chosen so that they could be equally well portrayed as transitive or intransitive events. For example, in the transitive version of "squats," Big Bird caused Cookie Monster to squat by pushing him down. In the intransitive version, Cookie Monster and Big Bird each squatted independently in synchrony.

The transitive and intransitive stimulus tapes were identical in structure down to the number of frames per episode and the 3-second intertrial intervals. Further, the two tapes portrayed an equal number of "squat" actions. Because of this level of precision, the two tapes could begin

Table 6.1
Experiment 5: Video events and accompanying transitive linguistic stimuli for the verb "squat"

Tape 1	Linguistic stimuli	Tape 2
	Sequential trials	
Big Bird (BB) pushes Cookie Monster (CM) into a squatting position.	"See, squatting!"	Blank screen
Blank screen	{Center light} "See, squatting!"	Blank screen
Blank screen	"Look, squatting!"	BB and CM squat side by side.
	Simultaneous trial	
Blank screen	{Center light} "See, squatting!"	Blank screen
BB pushes CM into a squatting position.	"See, squatting!"	BB and CM squat side by side.
	Test trials	
Blank screen	{Center light} "Find BB squatting CM!"	Blank screen
BB pushes CM into a squatting position.	"Find BB squatting CM!"	BB and CM squat side by side.
Blank screen	{Center light} "Look at BB squatting CM!"	Blank screen
BB pushes CM into a squatting position.	"Look at BB squatting CM!"	BB and CM squat side by side.

at exactly the same frame and could remain in synchrony trial for trial; moreover, at any given time the same amount of action appeared on both screens.

As is evident in table 6.1, and as was the case in the other studies reported here, children were given an opportunity to visually examine the events on the tapes before the test trials began. The first sequential trial, for example, allowed children 6 seconds to look at the causal event that would later be accompanied by the transitive audio. This exploration trial was accompanied by a nondescript audio that was neither transitive nor intransitive: "Squatting! See, squatting!" The second sequential trial, accompanied by the same nondescript audio, showed the intransitive or

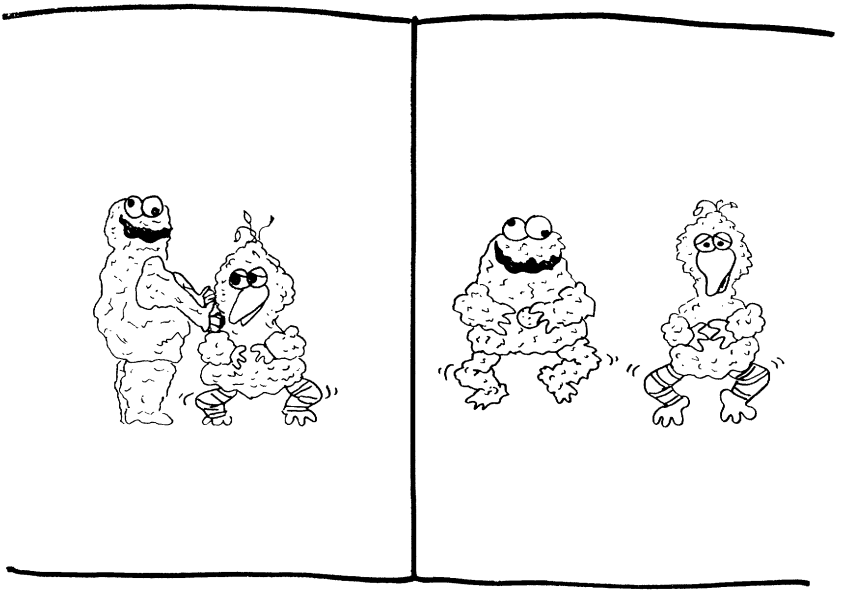

Figure 6.1
A sample stimulus set showing Cookie Monster causing Big Bird to squat on the left panel and Cookie Monster and Big Bird squatting together on the right panel

Table 6.2
Experiment 5: The video events and linguistic stimuli for the verbs "turn," "bend," and "flex"

Causal events	Linguistic stimuli*	Noncausal events
1. Cookie Monster (CM) stands still and turns Big Bird (BB) around in tight circles.	TR: "Look at CM turning BB!" IN: "Look at CM and BB turning!"	CM and BB turn side by side in tight circles.
*3. BB holds CM's shoulders and makes CM bend at the waist.	TR: "See BB bending CM?" IN: "See BB and CM bending?"	BB and CM bend side by side from the waist.
4. CM makes BB do a deep knee-bend as BB's arms come forward.	TR: "Watch CM flexing BB!" IN: "Watch CM and BB flex!"	CM and BB do deep knee-bends side by side.

*TR = transitive; IN = intransitive
**The second event, as seen in table 6.1, was squatting. That table also shows the actual structure of the videotapes for the verb "squat," and by analogy, for each of the other verbs.

noncausal event on the opposite screen. During the simultaneous trial, children saw both the causal and the noncausal scenes again, this time side by side and with the neutral audio. Finally, children saw the two test trials. Here, the video scenes were displayed simultaneously and the accompanying linguistic stimulus was either transitive (for half of the children) or intransitive (for the other half). The transitive audio was "See Big Bird bending Cookie Monster!"; the parallel intransitive audio was "See Big Bird and Cookie Monster bending!" Thus, this experiment had a between-subjects design.

The independent variables were (1) the form of the linguistic stimuli (transitive vs. intransitive), (2) verb type (known vs. unknown), (3) age (18–22 months, 23–25 months, and 27–30 months), and (4) sex (male vs. female). The dependent variable was the total amount of visual fixation time allocated to the matching and nonmatching screens across each pair of test trials. Four variables were counterbalanced. First, the order of mention of the two characters on the sequential trials was counterbalanced. On half of the trials, Cookie Monster was mentioned first; on the other half, Big Bird was mentioned first. Second, the screen side of the match was counterbalanced by showing half of the children tape 1 in deck 1 and half of the children tape 1 in deck 2. Third, the match occurred an equal number of times on the left and right screens. Fourth, the side of the appearance of the first sequential trial was counterbalanced.

Subjects Ninety-six toddlers, half boys and half girls, participated in experiment 5, which was conducted exclusively in the Temple laboratory. The children were equally distributed among three age groups: 18–22 months (mean age = 19.28 months), 23–25 months (mean age = 24.15 months), and 27–30 months (mean age = 28.3 months). Their average productive vocabularies were 102 words, 230 words, and 315 words, respectively, as assessed on the Rescorla Language Inventory (1991). According to parental report, the average number of words per sentence was 1.6 for the youngest group, 3.0 for the middle group, and 3.8 for the oldest group.

Subjects were contacted as described in chapter 3. To obtain the final sample of 96 subjects ($n = 32$ at each of 3 ages), 90 additional children were tested. Subject loss distributed itself approximately equally in each age group (youngest, 30 subjects dropped; middle, 31 dropped; oldest, 29 dropped) and condition (intransitive group, 44 dropped; transitive group, 46 dropped). Of those whose data were discarded, 25 had shown a side

bias (75% or greater looking preference to either the right or left screen during the test); 32 failed to return to the center light before four or more of the intertrial intervals; 15 were too fussy and did not complete the experiment; 14 were discarded because of technical error; and 4 showed language delay (that is, they were tested by mistake, having been identified as delayed for the purposes of another study).

Results and Discussion

Calculation of Stimulus Salience from the Simultaneous Trials Recall that the audio during the simultaneous trials did not direct the children to watch either the causal or the noncausal events. When inspection of the means indicated the possibility of a "match" level by age interaction, three separate analyses of variance, one on the data from the simultaneous trials for each age group, were performed. Each analysis of variance contained the two between-subjects factors of sex (male vs. female) and type of linguistic stimulus (transitive vs. intransitive), and one within-subjects factor—"match" level (match vs. nonmatch). Thus, for half of the children in each age group, the "matching" event during a simultaneous trial (which actually contained neutral audio) was the *causal* event; children were to hear transitive sentences describing this event during the test trials but had not yet been exposed to the transitive audio. For the other half of the children, the "match" was the *noncausal* event, accompanied by neutral audio during the simultaneous trials and by the intransitive audio during the test trials. Children's mean visual fixation times to the four simultaneous trials that preceded each of the four sets of test trials were entered into the three-way analyses of variance. It was predicted that since pains were taken to balance these events for stimulus salience, neither the causal nor the noncausal screen would receive more attention from either linguistic stimulus group.

For the children in the youngest group—the 19-month-olds—none of the main effects or interactions were significant. The mean visual fixation time to the match and nonmatch in the causal scenes during the simultaneous trials was 2.38 seconds; to the match and nonmatch in the noncausal scenes, the mean was 2.83 seconds ($F > .15$). For the 24-month-olds, there was a significant main effect of "match" ($F(1, 28) = 3.90$, $p = .05$) and no other significant effects. This means that the middle group preferred to watch the noncausal event during the simultaneous trials. The mean visual fixation time to the causal event was 2.34 seconds; to

the noncausal event, 2.83 seconds. Any significant findings in the main analysis with this age group will need to be interpreted in light of this unexplained preference for the noncausal event. The analysis of the 28-month-old children's visual fixation times indicated that there were no significant effects or interactions.

In summary, a preference for the noncausal event was found for the 24-month-olds during the simultaneous trials. Neither of the other age groups showed any preference for either type of event.

Test Trials Two separate four-way mixed analyses of variance were conducted on the data from each linguistic condition (transitive vs. intransitive). The two between-subjects factors were age (19-, 24-, and 28-month-olds) and sex. The two within-subjects factors were verb type (known vs. unknown) and match level (match vs. nonmatch).

The Transitive Sentence Group For children hearing the transitive audio, the only main effect was that of the match variable ($F(1, 42) = 7.95$, $p < .007$). Children in this group preferred to watch the match (the causal scene) significantly more than the nonmatch (the noncausal scene). The mean visual fixation time to the match was 2.92 seconds; to the nonmatch, 2.42 seconds. Two significant interactions, however, indicated that these results were not carried by all age groups. A match by age interaction ($F(2, 42) = 6.01, p < .006$) that collapsed across the two types of verb (known and unknown) was superseded by an age by verb type by match interaction ($F(2, 42) = 5.78, p < .007$). The interactions were tested with a priori one-tailed t tests since in all cases we predicted that the match would be watched more than the nonmatch. Further, not all means were tested against all other means; only those that made sense theoretically were compared. Table 6.3 gives the mean visual fixation times to the match and the nonmatch by age and type of verb.

The t tests revealed that the youngest group showed no preference for the match for either the known or the unknown verbs. The middle group watched the match (mean = 3.48 seconds) more than the nonmatch (mean = 2.05 seconds) for the known verbs ($t(42) = 3.20, p < .05$). The difference between match and nonmatch for the unknown verbs was not significant and in the wrong direction (nonmatch mean = 2.91 seconds; match mean = 2.06 seconds). Thus, the 24-month-olds were able to find the match only for the verbs they were acquainted with ("turn" and "bend") and could not use the syntactic frame in which the verb appeared to help them solve the task for novel verbs. This pattern must also be

Table 6.3
Experiment 5: Mean visual fixation time (in seconds) for each age and sex group to the matching and nonmatching screen for known and unknown verbs during the test trials in the transitive sentence condition

		Verb type			
		Known*		Unknown	
Age (in months)	Sex	Match	Nonmatch	Match	Nonmatch
19	Male	2.73	2.74	2.44	2.54
	Female	2.50	2.28	2.16	2.97
	Mean	2.62	2.51	2.30	2.76
24	Male	3.16	2.70	2.17	2.87
	Female	3.80	1.39	1.94	2.64
	Mean	3.48	2.05	2.06	2.91
28	Male	3.51	2.15	3.80	3.01
	Female	3.14	2.77	3.80	2.03
	Mean	3.32	2.50	3.80	2.02
Overall mean		3.14	2.35	2.72	2.56

*The known verbs were "turn" and "bend"; the unknown verbs were "squat" and "flex."

interpreted in light of the 24-month-olds' preference for the noncausal event during the simultaneous trials. The 24-month-olds maintained the pattern of looking more toward the noncausal event when presented with unknown verbs during the test trials. They were, however, able to overcome this preference for the noncausal event in the case of the known verbs. This suggests that for known verbs, the structure of the linguistic stimulus (a transitive sentence frame) had a powerful effect. The linguistic stimulus could have influenced children to move only to a "neutral" position, with attention to the causal and noncausal stimuli about equal. Instead, it influenced children not only to abandon their preference for the noncausal event, but in fact to prefer the causal event (the match) significantly more than the noncausal event (the nonmatch).

By the time children reached 28 months of age, just about 4 months later, they had no difficulty using the linguistic stimulus to guide their watching even for the unknown verbs. Further, unlike the 24-month-olds, this group did not have to overcome any preference for either event during the simultaneous trials. These children watched the match more than the nonmatch for both known and unknown verbs. This group's

mean visual fixation times for the known verbs were 3.32 seconds for the match and 2.46 seconds for the nonmatch ($t(42) = 1.90, p < .05$). For the unknown verbs, this group's means were 3.80 seconds for the match and 2.02 seconds for the nonmatch ($t(42) = 3.96, p < .05$).

To see if individual subjects' results would match the overall results from the analysis of variance, we next counted how many children in each age group had mean visual fixation times greater to the match than to the nonmatch. In line with the analysis of variance results, 13 of the 16 children in the oldest group preferred to look at the matching screen rather than the nonmatching screen. In the middle group, 11 of 16 children preferred the matching screen, and in the youngest group, only 8 of 16 children, or half, preferred to look at the matching screen.

Analyses by Sex Although the present analysis failed to turn up any significant main effects of or interactions with the sex factor (all p's $> .20$), inspection of the means in table 6.3 indicated that the boys and the girls in some age groups responded differently to the transitive test sentences. In order to evaluate the effects of the sex factor more closely, two two-way analyses of variance were conducted within each age group, one for boys and one for girls. The factors in these analyses were verb type (known vs. unknown) and match level (match vs. nonmatch).

In the youngest group, as the prior analysis indicated, neither boys nor girls watched the match more than the nonmatch, regardless of whether the verbs were known or unknown. In the middle group, the boys responded like the children in the youngest group, failing to watch the match more than the nonmatch regardless of the verb type. The girls, however, showed a marginally significant effect of match ($F(1, 7) = 4.37$, $p < .08$) and a marginally significant verb type by match interaction ($F(1, 7) = 5.28, p < .054$), which indicated that the match was watched more than the nonmatch *only* for the known verbs (mean to match = 3.80 seconds; to nonmatch = 1.39 seconds) (see table 6.3). When the two key theoretical comparisons (match vs. nonmatch for known and unknown verbs, respectively) were analyzed with an a priori one-tailed t test, only the difference between match and nonmatch for the known verbs was significant ($t(7) = 2.51, p < .05$). Indeed, comparing the mean visual fixation time to the match and nonmatch overall, 7 of the 8 24-month-old girls demonstrated preference for the matching screen in the transitive condition with known verbs. The boys' responses were considerably weaker: only 4 out of 8 showed overall means to the match greater than

ing. Conjoined subjects (like the ones in the stimuli for experiment 5) appear to be a source of difficulty for children: once conjoined subjects were removed in experiment 7 and "with" was added, some children did use the intransitive frame. Another earlier-mentioned interpretation of the failure to achieve expected results in experiment 5 is a bit more complicated. Although the conjoined subjects may indeed have caused difficulty, the problem may have been compounded by the paucity of other grammatical cues to intransitivity. For example, the form of the verb used (a gerund such as "squatting") was unmarked for number and certainly did not provide the child with any cue to the fact that the sentence subject was in the plural. Further, the sentence frame did not include an auxiliary verb (such as "is" or "are") marked for number.

Experiment 8 assessed the hypothesis that if an auxiliary verb marked for number (e.g., "are") were included in the test sentences, children would be able to find the scene that matched the intransitive sentence. Whereas in experiment 5 the audio was "See Cookie Monster and Big Bird turning," in experiment 8 it was "See, Cookie Monster and Big Bird *are* turning." If children do *not* attend to the plural auxiliary, then their responses should look like those of the group tested in the intransitive condition of experiment 5; that is, there should be no difference in visual fixation time between the matching and nonmatching screens. If, in contrast, the children are assisted by the addition of the plural auxiliary, then they should prefer to look at the noncausal scene.

Method

Subjects The subjects were 47 children tested in the Delaware and Temple laboratories. Sixteen children (8 boys and 8 girls) were from 23 to 25 months of age, with a mean age of 24; 12 and a mean productive vocabulary of 276 words. Thirty-one children (17 boys and 14 girls) were from 28 to 30 months of age, with a mean age of 29; 13 and a mean productive vocabulary of 305.33 words.

To obtain 47 children, 70 were tested. Children's data were eliminated for the following reasons: 1 for side bias; 7 for failing to center on four or more trials; 11 for experimenter error or technical problems; and 4 for fussiness.

Stimuli Again, the video stimuli were as presented in tables 6.1 and 6.2. The linguistic stimuli, however, were constructed to include the auxiliary

Table 6.4
Experiment 5: Mean visual fixation time (in seconds) for each age and sex group to the matching and nonmatching screen for known and unknown verbs during the test trials in the intransitive sentence condition

		Verb type			
		Known*		Unknown	
Age (in months)	Sex	Match	Nonmatch	Match	Nonmatch
19	Male	2.58	2.62	2.74	2.29
	Female	3.28	2.52	2.94	1.94
	Mean	2.93	2.57	2.84	2.12
24	Male	2.59	2.62	2.09	2.32
	Female	2.30	2.96	3.32	2.44
	Mean	2.45	2.79	2.71	2.38
28	Male	2.69	3.41	3.12	1.88
	Female	3.23	2.78	3.12	2.79
	Mean	2.97	3.09	3.12	2.33
Overall mean		2.78	2.82	2.89	2.28

*The known verbs were "turn" and "bend"; the unknown verbs were "squat" and "flex."

The Intransitive Sentence Group The results for children who heard the intransitive audio are very different from those for children who heard the transitive audio (see table 6.4). There were no results on any analyses. Across ages, only 24 of the 48 children looked longer at the noncausal scene when hearing the intransitive sentences. Even children in the oldest group responded randomly in this condition, with only 7 of 16 children looking more to the matching than the nonmatching screen. Although the trends were often in the right direction, they did not reach significance.

Although the children did not use the argument structure of the intransitive verbs presented to glean something of their meaning, their uniform failure on these sentence types does illustrate that they could discriminate one- from two-argument verbs. This in itself is an interesting finding since it shows that the stimulus sentences in the transitive condition were clearly driving children's visual fixation. This finding alone makes the results with the transitive verbs seem more valid, since if an artifact was operating, it should have influenced these results as well.

General Summary

The results of experiment 5 suggest that children have some knowledge of the significance of the transitive verb form and its implications within the class of motion verbs. By $2\frac{1}{2}$ years of age, children seem to comprehend that verbs in transitive sentences often refer to causal relations between the referents. Children do not seem to conclude the reverse—that intransitive frames are related to noncausal scenes. The distinction found between known and unknown verbs in the study also points to a developmental path in the young child's knowledge of the relationship between syntactic structure and meaning. It appears that 24-month-olds (at least, girls) enter a verb in the lexicon with its surrounding syntax. When they have memorized enough verbs in context, children can then generalize across similar frames, eventually using the frame itself to derive something about an unknown verb's meaning. That is, children may use analogy to begin solving the verb-learning problem.

Although these results are interesting, they leave much unexplored. Were the older children in the transitive condition overly influenced by the presence of the known verbs? For example, did they use their answers with the known verbs to influence their responses with the unknown verbs? Did the children in the intransitive condition really fail to understand the implications of the intransitive sentence frame or was there something about the particular intransitive frame used in the experiment that proved problematic? Are 24-month-old boys truly less sensitive to syntactic frames (like the 19-month-olds) or was this just an unusual sample? These questions provided the grist for three follow-up experiments.

6.2 Experiment 6: Do Young Children Exhibit Frame Compliance with Unknown Transitive Verbs?

As noted above, the older children may have responded correctly to the unknown verbs in experiment 5 because they used their knowledge of form-meaning pairings with known verbs to predict where to look when they heard unknown verbs. That is, the solution to the "unknown" verb problem was found by *analogy* to known verbs.

One could argue that exposure to the known verbs caused the children to always look toward the causal scenes and to ignore the noncausal scenes. In addition, perhaps children thought the known verbs were always

causal. The data, however, do not bear this out. Children distributed their attention randomly to the causal and noncausal scenes when there was no directive audio during the simultaneous trial. Further, the 24-month-old group responded correctly to the known verbs, but not to the unknown verbs.

Even though the interpretation that invokes analogy is unlikely, it must be conclusively ruled out. For this reason, experiment 6 was conducted, using four unknown verbs and a new pool of subjects from the oldest age group.

Method

Subjects Sixteen children were tested in the Yale University laboratory of Letitia Naigles, and 20 were tested in the Delaware laboratory. In each laboratory, half the subjects were boys and half were girls. The children were 27 to 30 months of age (mean age = 29;0) and had an average vocabulary of 312 words. To obtain results from 36 children, 61 children were tested. Potential subjects were dropped for the following reasons: 5 because they failed to return to the center light during intertrial intervals on more than four test trials; 4 because of fussiness; 8 because of equipment failure or experimenter error; 4 because of side bias; 3 because the parent watched the videotapes; and 1 because the child did not know the characters' names.

Stimuli The video events were the same as those used in experiment 5 (see tables 6.1 and 6.2). All of the test sentences were transitive and followed the syntax shown in table 6.2. The only difference was that instead of actual English words, four nonsense words—"glorp," "blick," "dax," and "krad"—served as the verbs.

Results and Discussion

Calculation of Stimulus Salience from the Simultaneous Trials As in experiment 5, the simultaneous trial data showed that without any directive linguistic stimuli, children look randomly toward the causal and noncausal scenes. The mean visual fixation time to the causal scenes (averaged over the four verbs) was 2.40 seconds; to the noncausal scenes, 2.60 seconds.

Table 6.5
Experiment 6: Mean visual fixation time (in seconds) to the matching and nonmatching screen for nonsense verbs in transitive sentences ($n = 20$; mean age = 29;0)

	Mean visual fixation time		
Stimulus*	Match	Nonmatch	Mean**
1	3.07	2.04	2.56
2	2.64	2.64	2.64
3	2.79	2.31	2.55
4	2.65	2.00	2.33
Mean	2.79	2.25	

*A description of the stimuli and the order in which they appeared is found in table 6.2.
**A main effect of stimulus was due to the difference between stimulus 2 and stimulus 4.

Test Trials The data were tested in a three-way mixed analysis of variance with the between-subjects factor of sex (male vs. female) and the within-subjects factors of match level (match vs. nonmatch) and stimulus (the four verbs). The analysis revealed two main effects—match level ($F(1, 34) = 6.64, p < .02$) and stimulus ($F(3, 102) = 3.34, p < .03$)—and no significant interactions. The match level main effect occurred because children watched the match significantly more than the nonmatch (mean to the match = 2.79 seconds; to the nonmatch = 2.25 seconds) (see table 6.5). Thus, the results of this experiment parallel and secure the results of the transitive condition in experiment 5 (compare the means in tables 6.3 and 6.4). By 29 months children are able to use a verb's arguments to derive a causal interpretation of that verb—whether it is known or completely novel. That is, they are able to make frame-compliant responses to the transitive sentences that they hear.

The main effect of stimulus was due (according to Tukey tests) to the difference between the second stimulus and the fourth. The mean visual fixation times to these stimuli were 2.64 seconds and 2.33 seconds, respectively. This fall-off of approximately one-third of a second is probably attributable to the fact that children became a bit bored with the videotapes by the fourth block of trials. Although no stimulus by match level interaction emerged, the means in table 6.5 show that for the second

stimulus event, the match was watched for exactly the same amount of time as the nonmatch.

The important finding is that children watched the match more than the nonmatch for three completely novel verbs that they could not have heard prior to this experiment. In short, then, these results replicate and secure the findings of experiment 5: children were capable of using a novel verb's argument structure to predict that it had a causal component conflated in its meaning. They watched the causal event significantly longer than they watched the noncausal event in the presence of a linguistic stimulus that used the nonsense verb in a transitive sentence frame. Furthermore, this result is not attributable to an overall preference for the causal event; children did not watch the causal event more when it was accompanied by a neutral audio during the simultaneous trials.

The next experiment reexamines a previously nonsignificant finding, namely, why children hearing intransitive sentence frames in experiment 5 failed to watch the noncausal event significantly more than the causal event.

6.3 Experiment 7: Do Children Know the Noncausal Implications of Intransitive Sentence Frames Containing the Grammatical Marker "With"?

There are several reasons why the children who participated in experiment 5 might have failed to watch the matching screen when they heard the intransitive stimuli. First, children of this age may in general be incapable of responding to intransitive frames. That is, they may master the meaning implications of transitive sentence frames before they master the meaning implications of intransitive sentence frames.

Second, the children may not have understood the *particular* frame chosen in experiment 5 to test for comprehension of intransitivity. The frame selected was quite bare, containing only a conjoined subject whose two parts were linked by an unstressed "and" (e.g., "See Big Bird and Cookie Monster glorping"). Children did not even hear a plural auxiliary verb (as in "See, Big Bird and Cookie Monster *are* glorping!") to buttress the interpretation that the sentence contained a conjoined subject. If the children did not process or comprehend the significance of the "and" in creating a complex noun phrase, they might well fail to interpret the sentence as an intransitive. Thus, the conjoined subject might have con-

fused them. Unsure of what it signaled, children might have watched the causal and noncausal events approximately the same amount since both named characters were shown on both screens.

To see whether children could show understanding of the meaning implications of intransitive frames, two more experiments were conducted that gave additional grammatical cues for an intransitive interpretation. Experiment 7 signaled the intransitive frame with a prepositional phrase (e.g., "Where's Cookie Monster turning with Big Bird?"); experiment 8 introduced a plural auxiliary verb to underscore the conjoined subject.

For experiment 7, the inclusion of "with" should block the interpretation of "Big Bird" as a direct object of the verb if children appreciate the function of the preposition. That is, children should not construe the sentence "Where's Cookie Monster turning with Big Bird?" as though it was "Cookie Monster is turning Big Bird!" If children do not attend to the function of the prepositional phrase "... with Big Bird," they will derive the wrong interpretation: under that scenario, children should interpret the sentence as a transitive and watch the causal scene more than the noncausal scene.

Method

Subjects The subjects were 32 children (half boys and half girls), all tested in the Temple laboratory. Subjects were very close in age to the two older age groups tested in experiment 5. There were 16 children from 22 to 25 months of age (mean = 23; 20 months), with a mean vocabulary (as indicated on the Rescorla Language Inventory (1991)) of 235 words. There were also 16 children from 26 to 30 months (mean = 28.5 months), with an average productive vocabulary of 300 words. To obtain a final sample of 32 subjects, 41 children were tested. Children's data were eliminated for the following reasons: 2 for side bias; 3 for failing to center on four or more trials; 3 for fussiness; and 1 for technical problems.

Stimuli The stimulus videotapes were the same ones used in experiment 5 (see tables 6.1 and 6.2). Here, however, the linguistic stimuli for the intransitive test sentences were on the order of "Find Cookie Monster turning with Big Bird!", for all four of the original verbs ("turn," "bend," "squat," and "flex").

Results

Calculation of Stimulus Salience from the Simultaneous Trials A two-way analysis of variance with the between-subjects factor of age (2 levels) and the within-subjects factor of "match" (noncausal vs. causal) was first conducted. No main effects or interactions resulted. Thus, there was no preference for the causal over the noncausal scene during the simultaneous trials for either age group. The mean visual fixation time to the causal scene (averaged across all four verbs) was 2.69 seconds; to the noncausal scene, the mean was 2.94 seconds.

Test Trials A four-way mixed analysis of variance with age (23 vs. 28 months), sex (male vs. female), verb type (known vs. unknown), and match level (match vs. nonmatch) was conducted. Table 6.6 presents the results. The only significant effect was an interaction of age by sex by match ($F(1, 28) = 3.96, p < .03$).

A priori one-tailed t tests performed on the means to the match and nonmatch within each age and sex group revealed the following pattern of results. In the older group only the boys watched the match significantly more than the nonmatch (mean to match = 3.01 seconds; to nonmatch = 2.24 seconds) ($t(28) = 1.86, p < .05$). Although the girls' means were in the right direction, they failed to reach significance (mean to match = 3.07 seconds; to nonmatch = 2.45 seconds). However, despite the failure of the girls' mean difference to reach significance, overall, 15 of the 16 older

Table 6.6
Experiment 7: Mean visual fixation time (in seconds) for each age and sex group to the matching and nonmatching screen during the test trials in the intransitive "with" condition

		Mean visual fixation time	
Age (in months)	Sex	Match	Nonmatch
23	Male	2.26	3.21
	Female	3.07	2.45
	Mean	2.67	2.83
28	Male	3.01	2.24
	Female	3.22	2.37
	Mean	3.12	2.31
Overall mean		2.89	2.57

children had higher visual fixation times to the matching than to the nonmatching screen.

The picture was different with the 23-month-olds. Here, the girls outperformed the boys. The girls' means were 3.22 seconds to the match and 2.37 seconds to the nonmatch ($t(28) = 2.05, p < .05$). Further, of the 8 girls tested, 6 had overall means in favor of the match. The boys' means indicated that they preferred to watch the nonmatch over the match (mean to match = 2.26 seconds; to nonmatch = 3.21 seconds). In fact, the overall means for 7 of the 8 boys indicated that they watched the nonmatch more than the match! Ironically, had we performed one-tailed t tests in the *opposite* direction (i.e., predicting that the mean to the nonmatch would be greater than the mean to the match), the difference between the boys' means would have been significant ($t(28) = -2.29$, $p < .05$). This finding means that the 23-month-old boys treated the intransitive "with" sentences as though they were active sentences, either ignoring "with" or failing to understand its function in the sentence. Thus, an intransitive sentence signaled by the use of a prepositional phrase (e.g., "See? Big Bird is bending *with* Cookie Monster!) was interpreted by the boys as though it were a transitive sentence ("See? Big Bird is bending Cookie Monster!"). These boys watched the causal event far more than the noncausal event.

Discussion

By contrast to the result for the intransitive condition in experiment 5, here the older group and the girls in the 23-month-old group showed a significant preference for the matching (noncausal) scene when the intransitive frame was created with the use of a prepositional phrase. This experiment revealed that children were able to comprehend the noncausal implications of *some* intransitive sentence frames. When intransitivity was signaled by the preposition "with," at least the 28-month-olds and the 23-month-old girls were able to find the matching screen.

The outcome of experiment 7 suggests the following conclusions about, and possible explanations for, children's failure in experiment 5. First, because of the 23- and 28-month-olds' success in experiment 7 we can rule out the possibility that children are simply unable to comprehend intransitive sentences. Second, children may understand intransitive sentences, but the presence of a conjoined subject in experiment 5 may have confused them. That is, in the intransitive frame used in experiment 5 ("See

Cookie Monster and Big Bird turning"), Cookie Monster and Big Bird "shared" the subject slot. Yet children may not have understood the function of the conjunction "and." Thus, upon hearing the conjoined subject, they may have thought that since both screens depicted both of the mentioned characters, either screen could be a match for the sentence. Third, as will be discussed further under experiment 8, the intransitive frame used in experiment 5 was bare of any additional cues to intransitivity.

In experiment 7 the conjoined subject was not a problem, since it was omitted. Instead, only the name of the first-mentioned character was the subject of the sentence; the name of the second-mentioned character became the object of a prepositional phrase signaled by "with" (e.g., "See Cookie Monster turning with Big Bird"). Thus, children had to discriminate between Cookie Monster turning *with* Big Bird on one screen and Cookie Monster turning by himself on the other screen, where the causal action was taking place (i.e., on the screen where Big Bird was seen making Cookie Monster turn).

Impressively, the 28-month-old boys and the 23-month-old girls appear to be sensitive to the preposition "with" used in this experiment. If they had attended only to the predicate, the arguments, and word order, then they would have given the opposite, incorrect response, treating the sentence as an active. In fact, this may be what the 23-month-old boys did, since their mean visual fixation times were clearly in favor of the nonmatch. These younger boys may not yet have gone beyond the simpler "semantic" strategy of noting the order of actor, agent, and actor. This result is interesting in light of the findings from experiment 5. There, even for known verbs, the boys in the transitive condition looked more to the causal scene but not significantly so.

The 23-month-old girls and the 28-month-old boys appear to have gone on to a more "syntactic" strategy, focusing on unstressed oblique markers in the sentence frame. (The older girls' means favored the match, although the difference in visual fixation times did not reach significance.) This developing syntactic sophistication permits them to notice even subtle differences in sentence frames and to give frame-compliant responses in this task.

6.4 Experiment 8: Do Children Know the Noncausal Implications of Intransitive Sentence Frames Containing the Auxiliary "Are"?

The results of experiment 7 suggest that children as young as $2\frac{1}{2}$ years are capable of using intransitive argument structure to derive sentence mean-

ing. Conjoined subjects (like the ones in the stimuli for experiment 5) appear to be a source of difficulty for children: once conjoined subjects were removed in experiment 7 and "with" was added, some children did use the intransitive frame. Another earlier-mentioned interpretation of the failure to achieve expected results in experiment 5 is a bit more complicated. Although the conjoined subjects may indeed have caused difficulty, the problem may have been compounded by the paucity of other grammatical cues to intransitivity. For example, the form of the verb used (a gerund such as "squatting") was unmarked for number and certainly did not provide the child with any cue to the fact that the sentence subject was in the plural. Further, the sentence frame did not include an auxiliary verb (such as "is" or "are") marked for number.

Experiment 8 assessed the hypothesis that if an auxiliary verb marked for number (e.g., "are") were included in the test sentences, children would be able to find the scene that matched the intransitive sentence. Whereas in experiment 5 the audio was "See Cookie Monster and Big Bird turning," in experiment 8 it was "See, Cookie Monster and Big Bird *are* turning." If children do *not* attend to the plural auxiliary, then their responses should look like those of the group tested in the intransitive condition of experiment 5; that is, there should be no difference in visual fixation time between the matching and nonmatching screens. If, in contrast, the children are assisted by the addition of the plural auxiliary, then they should prefer to look at the noncausal scene.

Method

Subjects The subjects were 47 children tested in the Delaware and Temple laboratories. Sixteen children (8 boys and 8 girls) were from 23 to 25 months of age, with a mean age of 24; 12 and a mean productive vocabulary of 276 words. Thirty-one children (17 boys and 14 girls) were from 28 to 30 months of age, with a mean age of 29; 13 and a mean productive vocabulary of 305.33 words.

To obtain 47 children, 70 were tested. Children's data were eliminated for the following reasons: 1 for side bias; 7 for failing to center on four or more trials; 11 for experimenter error or technical problems; and 4 for fussiness.

Stimuli Again, the video stimuli were as presented in tables 6.1 and 6.2. The linguistic stimuli, however, were constructed to include the auxiliary

"are." Thus, for all four verbs, the test sentences were of the type, "See? Cookie Monster and Big Bird are turning!"

Results and Discussion

Calculation of Stimulus Salience from the Simultaneous Trials Two separate one-way analyses of variance, one for each age group, were conducted on the amount of time children watched the causal (taken to be the nonmatch) and noncausal (taken to be the match) scenes during the simultaneous trials that preceded the test events. The results indicated that there was no preference for the causal or the noncausal scene in the simultaneous trials. That is, the factor of match was not significant in either analysis ($F(1, 250) = 1.93, p > .16$ for the younger children; $F(1, 250) = 1.61, p > .21$ for the older children). Across ages, the mean visual fixation time to the causal scene was 2.61 seconds; to the noncausal scene it was 2.32 seconds. Since there was no preference for either type of event during the simultaneous trials, any difference between visual fixation times on the test trials can only be a result of the linguistic stimulus drawing children's attention to one or the other of these types of events.

Test Trials The results for the test trials are presented in table 6.7. A four-way mixed analysis of variance was conducted with the between-subjects factors of age (24 vs. 29 months) and sex (male vs. female) and the within-subjects factors of verb (known vs. unknown) and match level

Table 6.7
Experiment 8: Mean visual fixation time (in seconds) for each age and sex group to the matching and nonmatching screen during the test trials in the intransitive "and-are" condition

		Mean visual fixation time	
Age (in months)	Sex	Match	Nonmatch
24	Male	3.46	2.82
	Female	2.83	1.55
	Mean	3.15	2.19
29	Male	2.62	2.58
	Female	2.84	2.47
	Mean	2.73	2.53
Overall mean		2.94	2.36

(match vs. nonmatch). The only two significant effects were a main effect of match and an age by match interaction ($F(1, 43) = 13.88, p < .001$; $F(1, 43) = 5.51, p < .03$, respectively). Each age group's mean visual fixation times to the match versus the nonmatch were tested with a priori t tests, given our hypotheses. These tests yielded a surprising result: the match was preferred to the nonmatch only by the 24-month-old subjects. Their mean visual fixation time to the match was 3.07 seconds; to the nonmatch, 1.82 seconds ($t(43) = 3.76, p < .05$). For the 29-month-old group, the mean visual fixation time to the match was 2.77 seconds; to the nonmatch, 2.49 seconds. This result does not approach significance.

The fact that 24-month-olds can comprehend the semantic implications of intransitive sentences when the sentences include an auxiliary verb replicates the results of Naigles (1990).[2] At 25 months of age, Naigles's subjects, hearing intransitive sentences with the same structure as the sentences heard by the children in this experiment, watched the matching screen significantly more than the nonmatching screen. In addition, Naigles's subjects heard only nonsense verbs—more difficult linguistic stimuli than were used in the present experiment. Thus, there are two studies securing the result that by 24 months of age, children can use the meaning implications of some kinds of intransitive sentence frames (in particular, a frame that includes an auxiliary verb) to watch noncausal events more than causal events.

Why should the younger children appear to do better than the older children? By our rudimentary assessment of their linguistic capabilities (i.e., size of productive vocabulary and length of word combinations), the older children who participated in experiment 8 do not appear to differ from the older children tested in experiments 5, 6, and 7. Nor do the 24-month-olds appear to be particularly accelerated. It may be that the 29-month-old boys are the source of this noneffect, although a sex by age by match interaction does not show up in the main analysis.

What do the present findings say about children's inability to comprehend the semantic implications of the intransitive sentences in experiment 5? The bare intransitive sentence frame used in experiment 5, which was not marked for number and was marked for intransitivity only by the lack of a direct object, may have been misunderstood by the children. They may have been uncertain about how to interpret the conjunction ("and") in the absence of additional information about subject number. They appear to have watched both the causal and noncausal scenes to the same degree, perhaps guided by the fact that each scene showed both named

characters. With the addition of the plural auxiliary in the linguistic stimuli of experiment 8, children were able to overcome the difficulty and to render a noncausal interpretation of the intransitive. Thus, these children seem to be looking for syntactic cues in the input—and even for subtle, unstressed syntactic cues—to make responses that comply with the syntactic frame.

6.5 General Discussion

In this chapter we have presented four experiments designed to test how children learn about verb meaning. The arguments presented here turned on the position that one significant route to the learning of verb meanings is through attending to sentential relations. That is, the present studies tested the hypothesis that by noting a verb's arguments, in addition to the situational cues, the learner can glean important information about verb meaning.

The present studies specifically examined the correlations between semantics and syntax for transitive and intransitive motion verbs. Transitive sentence frames, which include direct objects, are canonically (although not invariably) associated with expressing a causal relationship between the entities to which the arguments refer. Intransitive sentence frames, which omit direct objects, do not usually imply causation. The question addressed by these experiments was whether young children, at an early point in language acquisition, are capable of exploiting such regularities to watch causal events when they hear transitive verbs and noncausal events when they hear intransitive verbs. Children in three different age ranges (discussed here in terms of each group's mean age) were selected for study. Eighteen-month-olds, just beginning to combine words and shown in chapter 5 to be sensitive to at least the order of constituents in sentences, made up the youngest group. Twenty-four-month-olds, typically speaking in two- and three-word sentences and utilizing the rudiments of grammar, constituted the middle group. Twenty-nine-month-olds, typically using much longer sentences and capable of at least comprehending sentences that contain hierarchical structure (see chapter 7), made up the oldest group. Since both syntactic and semantic bootstrapping rest on children's ability to parse sentences into different units, and on their having had at least some experience with the semantics-syntax correlations implied by the bootstrapping construct,

it was important to investigate at what point in their linguistic development children are capable of attending to verb syntax in the language that they hear.

Summary of the Results of Experiments 5–8

Experiment 5 asked whether children could, hearing two known and two unknown verbs, use transitive and intransitive sentence frames to show a looking preference for causal and noncausal events, respectively. Except for the 18-month-olds, children watched causal events when they heard transitive sentences. All age groups failed to watch noncausal events when they heard intransitive sentences.

Experiment 6, using transitive sentences only, explored whether children could watch the causal event if they heard sentences containing nonsense verbs (i.e., verbs they had never heard before). A positive result would be evidence for the claim that children can use syntactic frames to learn something of the meaning of new verbs. The data from 29-month-olds indicated that novel verbs posed no problem; children watched the causal more than the noncausal event. This finding replicates the work of Naigles (1990), who used nonsense verbs with 25-month-olds.

Experiments 7 and 8 probed the null finding with intransitive sentences from experiment 5. The hypothesis underlying both studies was that the syntactic frame signaling the intransitive in experiment 5 may have been too bare for young children. Experiment 7, using a different intransitive construction (where the intransitive was accompanied by a prepositional phrase), found that the children in the two older groups (except for the 24-month-old boys) were now able to use the sentence frame to watch the noncausal more than the causal event. Thus, without the conjoined subject used in the bare sentence frame of experiment 5, these children could even override a straight word order strategy to interpret the intransitive. However, the 24-month-old boys did interpret these intransitive sentences as though they were transitives, watching the causal event significantly longer than the noncausal event. Finally, in experiment 8, where the only change in the linguistic stimulus relative to experiment 5 was the inclusion of a plural auxiliary verb ("are"), 24-month-old children showed their ability to use the new sentence to watch the noncausal event. This finding also replicates the work of Naigles (1990), who used this sentence frame with novel verbs. The 29-month-old group's visual fixation means surprisingly did not reach significance.

How Did the Children Represent the Stimulus Sentences?

The ability of children in the two older groups to watch the causal or noncausal events in response to transitive or intransitive sentence frames suggests that, at a minimum, these children were able to identify which character was which and to find the agents and patients in the sentences they heard. However, these data suggest that in addition to parsing and performing a semantic analysis, children were performing syntactic analyses. This is most forcefully suggested by experiment 7, where children heard sentences of the type "Big Bird is squatting with Cookie Monster," with the subject (the name of one intransitive actor) at the beginning and the object of the preposition (the name of the other intransitive actor) at the end. The simplest (but incorrect) sentence analysis would have been to ignore "with" and treat the sentence as though it were transitive (as in "Big Bird is squatting Cookie Monster"—that is, causing Cookie Monster to squat). In fact, this appears to have been just the analysis performed by the 23-month-old boys. That the majority of the children did not make this mistake suggests that in this experiment, and possibly in the others as well, children may have been computing syntactic relations such as subject and predicate of the sentence. Regardless of the particular nature of children's representation of the sentences used in these experiments, however, the fact that children consistently watch the screen that matches the sentence frame they hear indicates that children can use these frames to partially derive verb meaning.

What exactly is the meaning children derive from these sentences? This, too, is somewhat indeterminate despite the fact that children were able to watch the matching screen. As Naigles (1990) points out, children could have assumed that the transitive frames they heard were implicating not causal events but "unergative" transitives (Perlmutter 1978; Levin 1985). In an unergative transitive the entity whose name is in subject position acts as an experiencer and the entity whose name is in direct object position acts as the patient. For example, in the transitive sentence "Adam eats the fish" or "Mary sees John," no causation is implied. If children in the present studies interpreted the transitive sentences as unergatives, they could have thought that "Cookie Monster is daxing Big Bird" meant something like 'Cookie Monster is holding Big Bird while Big Bird waves his arm'. That is, the explicit causal component would be absent, and Cookie Monster's role would be reduced to that of helper or accompanist rather than the cause of Big Bird's action.

A similar argument can be made about the intransitive sentences used here (see Naigles 1990 for a lengthier discussion). Not all intransitives express noncausal meanings. It is therefore possible that "Big Bird and Cookie Monster are daxing" could be interpreted as 'Big Bird is performing a symmetric action with someone using his arm', an interpretation that is not specifically noncausal. To determine which interpretation children in our experiments made requires further research. In particular, it requires presenting children with scenes that pit these interpretations against each other and seeing which they choose to watch in the presence of certain sentence frames. As an example of this kind of work, Naigles and Kako (1993) conducted a set of experiments to determine whether children (1) can interpret transitive frames as referring to contact events (e.g., "pat"), and (2) prefer to assign causal or contact interpretations to transitive sentences. To answer these questions, they presented videotapes to 27-month-olds in the intermodal preferential looking paradigm. Their second experiment indicated that toddlers were indeed capable of learning a new contact verb when they heard a transitive sentence frame. More important for the present point is the result of their third experiment. When they pitted a contact event against a causal event, children preferred to watch the causal event without a linguistic stimulus, a preference that continued once an audio using a transitive sentence frame was introduced. Naigles and Kako concluded that children interpret transitive frames as mapping to both causation and contact—in other words, as a more general notion of object affectedness (see Gropen et al. 1991). Clearly, then, it would appear that children in the present experiments who heard transitive frames were at minimum interpreting them in terms of an affected object instead of merely in terms of one character accompanying another while the first carried out an action (e.g., Big Bird simply being there while Cookie Monster performed an action).

Another reason to attribute causal and noncausal interpretations to children for the sentence frames used here stems from the results of Fisher et al. (1994) with 3- and 4-year-old children. When children were shown a videotaped scene that could be interpreted either causally or noncausally, they reliably used the sentence context that accompanied the scene to give it either a causal or a noncausal interpretation. For example, these children were shown a rabbit feeding an elephant, who ate. They then heard either "The bunny is nading the elephant" or "The elephant is nading." The former sentence led children to interpret the verb causally; the latter

sentence led them interpret it noncausally. The nature of the dependent variables used in Fisher et al.'s study increases our confidence in the inference that children in the experiments reported here made causal and noncausal interpretations, even though Fisher et al.'s subjects were older than ours. In that study, children were asked to paraphrase the test sentences (a task not possible for the younger children in our experiments to perform): their paraphrases were something like "The bunny is feeding the elephant" and "The elephant is eating."

In sum, Naigles and Kako's (1993) results in the intermodal preferential looking paradigm and Fisher et al.'s (1994) results, using a different dependent variable with older children, suggest that subjects in experiments 7 and 8 may well have made causal and noncausal interpretations of the stimulus sentences.

The Developmental Picture That Emerges from the Experimental Data

The data from experiments 5–8 paint a picture consistent with that sketched above of young children exploiting sentence frames and the known nouns contained therein, as well as their observation of real-world scenes, to interpret the meaning of novel verbs and of familiar verbs with more than one possible reading (e.g., "turn" can be both causal and noncausal). However, children may not start the process of language learning using formal structural information—even when they are familiar with the meaning of the noun arguments that verbs take. The 18-month-old children in experiment 5 showed little ability to use either the transitive or the intransitive frames they were offered despite recognizing Big Bird and Cookie Monster and despite observing videotaped scenes. By 24 months of age, and certainly by 28 months, many children have become sensitive to the phrase-structural implications of transitive sentence frames for describing causal events (experiments 5 and 6), although there is a suggestion that they are more so with known than with unknown verbs. Children were able to use the syntax as a guide to meaning even when the scenes were accompanied by sentences that did not contain an auxiliary verb but only the main verb surrounded by the two noun arguments (e.g., "Find Big Bird bending Cookie Monster!").

For intransitive sentences, bare sentence frames apparently do not work quite as well. For intransitive sentence frames, experiments 5, 7, and 8 clearly suggest a developmental progression for exploiting ever more subtle nuances of syntactic frames. None of the children in these studies

could use intransitive frames with a conjoined subject and a verb that failed to mark subject number (experiment 5) (e.g., "Find Big Bird and Cookie Monster bending!"). Once structural information was added to the intransitive frame, however, both older age groups began to show their ability to use syntax to predict meaning. For example, experiment 7 separated the noun arguments, thereby eliminating the conjoined subject, and added the preposition "with." The boys in the 24-month-old group continued to interpret these sentences as though they were active declarative sentences. Whether their failure was due to unfamiliarity with the preposition "with" or an inability to exploit an intransitive sentence frame for meaning is unclear. However, the 24-month-old girls were able to use these frames to look for noncausal events. Furthermore, some children at this age are capable of understanding the function of a conjoined subject when they hear a sentence in which an auxiliary verb marked for number is also provided (experiment 8) (e.g., "See, Cookie Monster and Big Bird are bending!"). Surprisingly, the oldest group did not show a clear effect in experiment 8.

Thus, the intermodal preferential looking paradigm has revealed that children—at least by 23 months of age—are sensitive to phrase-structural information and can use it to predict verb meaning. Transitive constructions seem easier than some intransitive constructions. When additional syntactic information is added to intransitive frames, however, children use them to look for noncausal events.

There are several caveats we should raise before attempting to extrapolate these findings to the way in which children approach verb learning in the real world. First, in the real world, children are not given two scenes from which to choose. Instead, they are usually observing a single scene whose beginning and end may not be clearly demarcated and that may contain overlapping elements from other events. In other words, the paradigm used in these experiments may make it easier for children to decide which aspects of the scene are important for word learning. Second, the studies reported here use absolutely minimal vocabulary, requiring only that children know the videotaped characters' names. In the real world, children probably hear sentences all the time in which both the nouns and the verb are unfamiliar to them. As Gleitman and Gillette (1995) have shown, it is much easier to predict noun meaning from scene observation than it is to predict verb meaning. Still, we have made it easier for children to reveal their ability to use phrase-structural information by so severely

limiting the stimulus nouns. Third, not all subjects who are tested in this paradigm produce usable data. This paradigm may work best with children who have longer attention spans relative to their peers.

What these three factors add up to is the possibility that it may take children a bit longer to manifest the skills they show here in real-world language learning.[3] However, whether children use syntactic frames in real life a month or two after they show their ability to do so in the intermodal preferential looking paradigm is not as important as the fact that they can exploit form to learn something of meaning. Further, by the age of 30 months, the earlier-mentioned work of Naigles, Gleitman, and Gleitman (1992) bolsters the present findings on children's sensitivity to syntactic information for determining verb meaning. Their research used a different paradigm—an enactment task with standard verbs like "come" and "bring" embedded in noncanonical constructions.

Additional research must continue to disentangle exactly what aspects of the structural information available to young children in the sentence frames they hear are being exploited by them. Also, input studies must be conducted to examine the range of frames available to the language-learning child (for one such study, see Naigles and Hoff-Ginsberg 1992). In theory, syntactic bootstrapping is most effective as a means of verb learning when several different frames, each giving unique syntactic information about the permissible range of uses of a verb, are available to the child.