REFERENCES
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/1131462?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

First Contact in Verb Acquisition: Defining a Role for Syntax

Letitia G. Naigles and Edward T. Kako

Yale University

NAIGLES, LETITIA G., and KAKO, EDWARD T. *First Contact in Verb Acquisition: Defining a Role for Syntax.* CHILD DEVELOPMENT, 1993, **64**, 1665–1687. The goal of this research was to address 2 questions regarding children's use of syntactic information in acquiring verbs: First, what are children's biases for actions in the absence of syntactic information; and second, how specific is the meaning derived for verbs when syntactic information is present? In 3 experiments we presented nonsense verbs either in syntactic isolation (e.g., "Look! Sebbing!") or embedded within a transitive syntactic frame (e.g., "The frog is sebbing the duck"). These actions were then separated, and the children (mean age = 2 years, 3 months) were asked to select the action which was the referent of the verb. In Experiment 1, Causative actions (in which 1 character forces another to move in some way) were paired with Synchronous actions (in which both characters move simultaneously). In Experiment 2, the same Synchronous actions were now paired with Contact actions (in which 1 character merely touches the other). In Experiment 3, the Contact actions were paired with Causative ones. 2 results emerged: (1) Children have identifiable action biases in the absence of syntactic information and (2) these biases can be shifted by the addition of a transitive syntactic frame. We conclude that the meaning derived from the transitive frame is not specifically Causative or Contact but, more generally, a sense that 1 character is affecting another.

How are verbs acquired by young children? The theory of Syntactic Bootstrapping (Gleitman, 1990; Gleitman, Gleitman, Landau, & Wanner, 1987; Landau & Gleitman, 1985) predicts that children use both syntactic and situational information to learn about verb meanings. The syntactic frames in which a verb is presented allow children to constrain the plethora of meanings available from the situational information by virtue of linkages existing between phrasal syntax and verb semantics. Several recent studies (Fisher, Hall, Rakowitz, & Gleitman, in press; Naigles, 1990; Naigles, Gleitman, & Gleitman, 1993) have provided initial validations of Syntactic Bootstrapping as a procedure for verb learning. Each study has demonstrated that young children will make different conjectures about a verb's meaning based on the syntactic frame in which it is presented (see also Taylor & Gelman, 1988;

Waxman & Gelman, 1986; Waxman & Kosowski, 1990, for research on the use of syntax in noun and adjective acquisition). The present research investigates Syntactic Bootstrapping in more depth, introduces a methodological advance in the form of a separate baseline condition, and asks the question, How specific are the meanings that syntax affords verbs?

Recognition of the need for Syntactic Bootstrapping in children's verb acquisition is based partly on Quine's (1960) statement of the problem of translation between languages, which applies with equal force to children's acquisition of a single language. For any verb-learning situation, there can be any number of possible mappings from the verb (representing the form) to the visuo-spatial scene (encapsulating the meaning). For example, consider the child who hears

(the nonsense word) "blicking" without a syntactic frame, while transporting a toy truck to her parent. She could conjecture that "blick" means *bring* (she is bringing the truck to Mom), *come* or *walk* (she is coming/walking to Mom with the truck), *get* or *take* (Mom is getting/taking the truck from her), *play* (she is playing with the truck), and so on. How is the child to determine which meaning was intended? In contrast, if the verb is presented in a syntactic frame, some of the possible interpretations can be eliminated immediately. For example, if the parent says, "Blick over to me with the truck!" then she probably is referring to *walking* rather than *bringing*; that is, the non-Causative action of the child rather than the Causative one. The opposite would be true if the parent said, "You're blicking the truck!": Because the syntactic frame is transitive, containing both a subject and a direct object, the referent of the verb is more likely to be a Causative action. This is largely due to the semantic implications of the syntactic frames: Causative verb meanings are canonically signaled by transitive frames, while non-Causative meanings are implicated by intransitive frames (i.e., no direct object; see Bowerman, 1974, 1982; Fisher, Gleitman, & Gleitman, 1991; Grimshaw, 1990; Jackendoff, 1983, 1990; Levin, 1985; and Pinker, 1989, for more discussion). Thus, the Syntactic Bootstrapping hypothesis predicts that children who know these correlations between syntax and verb semantics will exploit them while learning verbs.

Recent support for Syntactic Bootstrapping has primarily utilized the transitive-causative correlation. Naigles, Gleitman, and Gleitman (1993) showed that 2–4-year-old children can use syntax to extend the meanings of familiar verbs. They asked children to act out sentences that contained familiar verbs (e.g., *come* and *go*) presented in novel, ungrammatical (to adults) frames (e.g., "The zebra goes the lion"). The children preferred to enact these sentences causatively (e.g., the zebra makes the lion go), altering the meanings of *come* or *go* in accordance with the semantic implications of the transitive frame rather than adhering to their familiar non-Causative meanings.

In a second study, one that is most relevant to the present research, young children (mean age = 2 years, 1 month) were taught the meanings of novel verbs by pairing videotaped multiple novel actions with a single nonsense verb in one of two syntactic frames (Naigles, 1990). Thus, for example, the children would see a duck making a rabbit bend over (a Causative action), while simultaneously both the duck and the rabbit were making arm circles (a non-Causative, Synchronous action). This scene was paired with the nonce verb *krad* in either a transitive frame ("The duck is kradding the bunny") or an intransitive frame ("The duck and the bunny are kradding"). Then, the novel actions were separated, such that the scene of the duck making the rabbit bend over appeared on one video screen, while the scene of the duck and the rabbit making arm circles appeared simultaneously on the other, side-by-side, video screen. The first of these single-action scene trials was presented with a control audio (e.g., "They're different now") to test for stimulus salience. The second and third single-action scene trials were paired with the test audio "Find kradding," and the children's choices during these trials (measured by fixation time to each screen) showed an effect of syntax. Specifically, the children in the transitive audio condition looked longer at the Causative action of the duck making the rabbit bend over, while the children in the intransitive audio condition looked longer at the Synchronous action of the duck and the rabbit making arm circles. Thus, the children used the syntactic frame in which the nonce verb was presented in order to select the action to which it referred.

Previous work, then, has provided an initial validation of the Syntactic Bootstrapping hypothesis, that children use syntax to learn about verb meanings. This paper is concerned with two further questions raised by Naigles's (1990) study. First, does syntax alter children's choice of novel actions when compared with a baseline measure of (syntax-less) action preference? And second, what is the breadth or specificity of meaning that the transitive frame affords the novel verb?

ACTION PREFERENCES AND VERB LEARNING

The usefulness of Syntactic Bootstrapping is best understood in the context of what the child's preference for the mapping between verb form and verb meaning would be, without syntactic information. In the above discussion of the Quinian problem of word acquisition, the child was represented as being ambivalent concerning her choice of form-meaning mappings when no syntactic information is present. Such a depiction of the child's state cannot be entirely correct, as both perceptual and conceptual prefer-

ences are surely operating to facilitate the child's distinguishing actions and/or aspects of actions in a given scene (Gibson & Spelke, 1983). These preferences could be used as "first guesses" concerning the meanings of coincident novel verbs. The existence of such preferences does not necessarily diminish the need for syntax in verb learning, though. In fact, syntactic information now becomes necessary to draw children's attention *away from* their initial preference and *toward* the intended referent of the speaker *when the two are different*. For example, consider the "blick" example presented earlier. Suppose that the child's initial focus in this context was on the agentive. or manipulative action (e.g., Slobin, 1985). In this case, she might initially conjecture that "blick" meant *bring*. If the parent was referring to *walk*, though, this bias would lead to errorful verb learning. Presentation of "blick" in the intransitive frame could redirect the child's focus toward the intended meaning, forestalling such errors (see Fisher et al., in press, for more discussion).

In order to demonstrate more fully the effects of syntactic information on verb learning, then, we need to determine the toddler's initial action preferences. The children in Naigles's (1990) study were not explicitly tested for these action preferences; instead, as described above, the control trial was used to test for stimulus salience. The results from that study showed significantly different looking preferences between the control and test trials, thus suggesting that the test audio did succeed in altering the children's preferences which had been manifested during the control trial. However, it is unlikely that those preferences indicated initial action focus, because they were undoubtedly and unavoidably affected by the just-preceding teaching trials. For example, if the teaching audio is in the transitive frame, the children may exhibit a preference for the Causative action during the control trials, even before being asked to "find kradding" during the test trials. Alternatively, they may prefer the Synchronous action during the control trial simply because it was not singled out by the teaching audio.

In sum, the so-called control trials may not adequately reveal the child's initial action preferences in the preferential looking paradigm, and thus those preferences remain unknown. In the current series of studies, we present a second way of determining initial preferences for novel actions. Specifically, we present the same sequence of video trials, but now during the teaching phase, the novel verb is spoken in isolation without a syntactic frame (e.g., "Look! Kradding!"). We then compare children's preferences with this teaching audio with the preferences of their peers from Naigles (1990), when the verb was presented in a syntactic frame (a between-subjects comparison). Changes in looking preference between these audio conditions can then be attributed to the influence of the syntactic frames.

SPECIFICITY OF MEANING AFFORDED BY SYNTAX

The bulk of the research on Syntactic Bootstrapping in children has focused on the acquisition of Causative verbs, primarily because the link between syntactic form and verb semantics for this class of verbs is transparent and has been well studied. Recent research in lexical and conceptual semantics, though, has demonstrated increasingly more systematicity between syntax and semantics in the adult lexicon (see Fisher et al., 1991; Jackendoff, 1983, 1990; and Levin & Rappaport, 1991, for detailed discussions). The present study extends this work into the domain of acquisition to see if children exploit these systematicities in learning verbs. The new class of verbs focused on is Contact verbs (e.g., *rub, stroke, brush, wipe, touch, hit*), which generally depict an agent's action when it comes into contact with a patient. The focus can be on the impact of the action (e.g., *hit, strike*) or on the specific motion itself (e.g., *rub, stroke*). Contact verbs canonically appear in transitive syntactic frames, with the agent in the subject position and the patient/goal in the object position (Jackendoff, 1990; Talmy, 1988).

Interestingly, the surface structures of Contact verbs are very much like those of Causative verbs: Both are canonically transitive, with the agent of the action in subject position. However, the meanings of the two classes of verbs are different: The agents of Contact verbs *act on* their patients, but do not necessarily *cause* their patients to *do* anything. One can *scratch* a door (a Contact verb) without eliciting any corresponding movement, but it would indeed be odd to *close* a door (a Causative verb) without having the door move in some way as a result. This similarity of surface form, accompanied by different meanings, poses something of a conundrum for the theory of Syntactic Bootstrapping. If the form-meaning link that facilitates the acquisition of Causative verbs is a specific one between transitive frames

and causative meanings, then children might be expected to show some delay in acquiring Contact verbs. For example, when they hear "The cat is scratching the door" they may be inclined to search for a Causative action and so be reluctant, at least initially, to assign *scratching* as the referent. Contact verbs might even be particularly troublesome because Contact actions can be followed by patient movement (e.g., *stroking* a responsive as opposed to a passive cat, *scratching* a door which subsequently closes). Data from 2-year-olds indicate, however, that Contact verbs are among the earliest acquired (Goldin-Meadow, Seligman, & Gelman, 1976); therefore, such verbs do not appear to pose any particular problem for acquisition. Two alternatives emerge: Either Contact verbs need not be acquired using syntactic information, or else the verb meaning that is indicated by (and linked to) the transitive frame is less specific than Causation and also includes, at least, Contact. In the following studies, our aim is to distinguish these two alternatives by investigating whether transitive frames can be used by young children to focus attention on Contact actions as well as on Causative actions.

SUMMARY AND PROSPECTUS

In sum, while Naigles (1990) has provided an initial validation of Syntactic Bootstrapping as a procedure for verb learning, questions remain about the children's initial preferences for novel actions and about the specificity of meaning that syntax affords. In this paper we present three experiments that address these questions. We begin with an experiment that extends and validates Naigles's (1990) study: Children view those original videotapes and hear the nonsense verbs presented without a syntactic frame, thus providing a test for initial action focus. We investigate the Meaning Specificity question with two new sets of videotapes: In Experiment 2 we present multiple-action scenes that contain the same Synchronous actions as appeared in Naigles (1990), now paired with Contact actions rather than with Causative ones. These videos test whether the transitive frame can help children to focus on and so acquire agent-patient relations that are not Causative. In Experiment 3 we present both Causative and Contact actions simultaneously, as the multiple-action scene, together with a novel verb in the transitive frame. These videos test whether children have a preference for the transitive-causative linkage between form and meaning, as opposed to the transitive-contact one. We investigate the Initial Focus question in each of these experiments as well, by including a condition in which the novel verb is presented in isolation.

All three experiments use the preferential-looking paradigm recently developed by Golinkoff and Hirsh-Pasek (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1991; see also Naigles, 1990; Reznick, 1990). This is a more sensitive language comprehension paradigm than the more common "acting out" method because it simply requires the child to *look* at one of two simultaneously presented video events. If the utterance the child hears is understood correctly, then she should focus on the one scene that is consistent with that utterance. Otherwise, the child should look randomly at either scene.

Our instantiation of the preferential-looking paradigm involved two changes from earlier studies: First, in the studies of Golinkoff et al. (1987) and Naigles (1990), the children's screen preferences were recorded on-line, that is, while the children were watching the videos. In the present three sets of experiments, the children were videotaped as they watched the TV screens, and their screen preferences were subsequently recorded off-line, which allowed us to assess the reliability of the coding. Second, the only dependent variable analyzed by Naigles (1990) was that of duration of visual fixation to the matching versus nonmatching video screens. In the present experiments, following Golinkoff et al. (1987), we also compared the children's direction of first look to the matching versus nonmatching screens. This second measure provided a potentially corroborating indication of the children's screen preferences, for if they understand the linguistic audio, they should also look first at the screen that matches what they hear.

Experiment 1

Experiment 1 was designed to extend and validate Naigles's (1990) findings by addressing the question of initial action focus when Causative and Synchronous actions are presented together in one scene.

METHOD

Subjects

Sixteen children completed the experiment. Twenty-six children were tested, but 10 children were eliminated from the final

TABLE 1

DETAILED SUBJECT INFORMATION

Experiment and Condition	N	Words[a]	Verbs[b]	Age (Range)[c]	Multiword Utterances[d]	Parental Education[e]
Experiment 1:						
Frameless	16	272 (99.6)	31 (12.8)	28 (26–29)	.875	3.0
Transitive (Naigles, 1990)	12	224	28	25 (23–27)	.67	4.0
Intransitive (Naigles, 1990)	12	253	32	25 (23–27)	.75	4.5
Experiment 2:						
Frameless	16	260 (103.9)	31 (13.4)	27 (26–28)	.875	2.5
Transitive	20	291 (74.6)	34 (9.8)	28 (26–29)	.950	3.0
Intransitive	16	290 (89.2)	33 (11.8)	26 (25–27)	.938	3.0
Experiment 3:						
Frameless	16	279 (97.6)	33 (37.3)	27 (26–28)	1.000	2.5
Transitive	20	308 (60.9)	36 (7.8)	28 (26–29)	1.000	4.0

[a] Mean number of words (SD) used spontaneously by parental report, out of 354 on Rescorla's (1991) checklist.
[b] Mean number of verbs (SD) used spontaneously by parental report, out of 40 on Rescorla's (1991) checklist.
[c] Mean age (range) in months.
[d] Proportion of children who produced three-word or longer utterances by parental report.
[e] Mean level of parental education, in years past high school.

sample for the following reasons: Fixation bias of more than 75% to a single side, averaged across trials (four), parent watching video along with child (two), refusal of the child to watch the videos for more than half of the trials (one), inattention to the centering light (one), and inability of two out of five coders to agree on child's fixations (two). The final sample included eight males and eight females with a mean age of 2 years, 3 months (SD = 28 days). The detailed subject information for each audio condition is presented in Table 1. The 10 children who were eliminated did not differ significantly from the final subject pool in any of the listed variables. In this and all the other experiments reported here, all of the children were being raised in monolingual English-speaking homes; nearly all were of European heritage. Names of potential subjects were obtained from files of newspaper birth announcements. Parents were contacted by telephone.

Apparatus

The basic set-up is shown in Figure 1. The child was seated on the parent's (usu-

ally mother's) lap and observed two different, simultaneously presented video events on two side-by-side video monitors. Between the two monitors, which were 12 inches apart, a hidden auditory speaker played a message that matched only one of the video scenes. Mounted above the speaker was an arrangement of lights, consisting of a 15-watt light bulb under three concentric circles of "chase lights," both of which were illuminated between trials to attract the child's attention. Children were placed 2 feet away from the center of the two video displays. The children's faces were filmed by a camcorder hidden behind the arrangement of lights. In order that the mother be blind to the experimental condition, she was explicitly instructed not to watch the videos and was given a visor to wear over her eyes while in the testing room.

The stimulus displays were color videotapes, recorded with stimuli (described below) arranged against a white wall. Intertrial intervals were created by inserting 3 sec of black tape. Each tape was prepared as one of
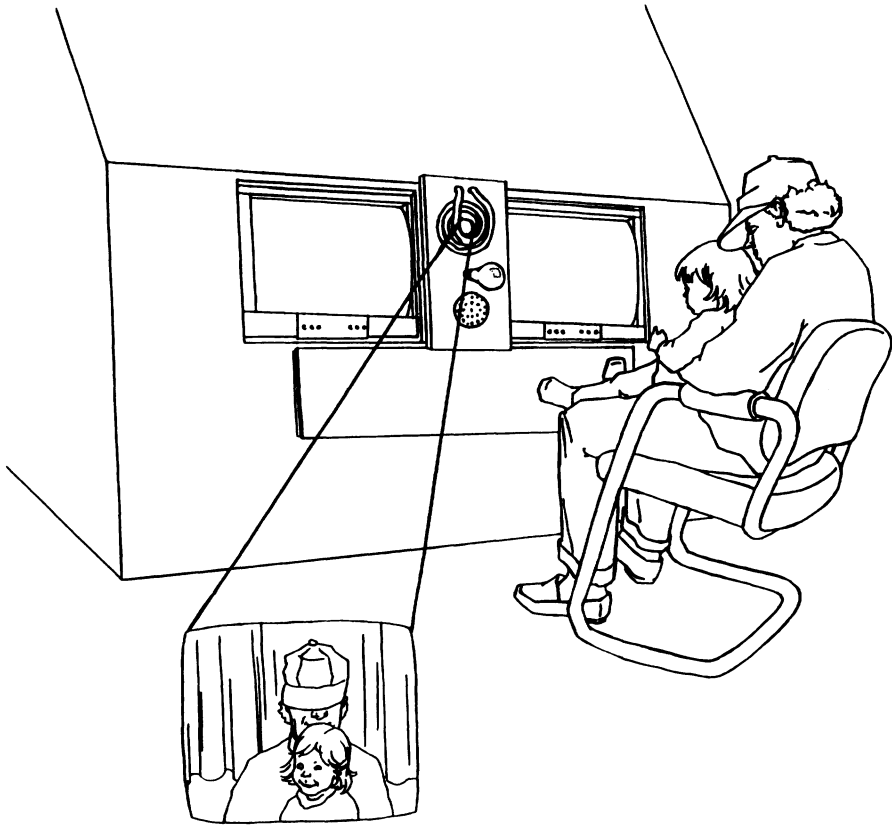
Fig. 1.—Experimental set-up of the preferential looking paradigm

a pair. Thus, tapes in a pair had the identical structure, down to the number of frames per episode. This level of precision enabled the synchronous operation of tapes, such that neither tape began before the other in a pair, and both tapes had an equal number of episodes of action per trial. The auditory stimuli were then dubbed onto one channel of the final version of the tapes. On the other channel, a 1 kHz tone was dubbed on to coincide with the trial segments to permit a computer to record trials. While the child was watching the videos, the tone was copied onto the single auditory channel of the videotape made of the child's face. These tones were read by a tone decoder that interfaced with an IBM PC computer.

Stimuli and Design

The stimulus videotapes were the same ones used by Naigles (1990); the following description is drawn in large part from that paper. A layout of the videotapes and the "Frameless" audio is presented in Table 2. The left and right columns indicate videos, while the center column indicates the audio.

The audio of a female voice using infant-directed speech intonation (Fernald & Simon, 1984) was first heard during 3 sec of black tape before each presentation of a pair. The audio was then repeated as the scenes were presented for 6 sec each. First, an introductory passage familiarized the subjects with the situation and the characters (see App. A), and then the novel actions and verb were presented. The crucial sequence was as follows. In trial 1, one screen presented a multiple-action scene—two actions going on simultaneously, performed by the same two actors. One of the actions was Causative (one character causes the other to move in some way; e.g., the duck forces the rabbit into an odd bending position), and the other was non-Causative and Synchronous (the same characters each generate their own actions; e.g., the duck and rabbit make identical arm gestures in synchrony). This multiple-action scene is shown in the top panel of Figure 2; in the video, both actions were iterative and were repeated the same number of times. The accompanying audio presented a novel verb—*gorp*—in isolation as a gerund (e.g.,

TABLE 2

STRUCTURE OF THE STIMULUS VIDEOTAPES FOR THE VERB "GORP," EXPERIMENT 1

	Tape 1	Audio	Tape 2
1.	The duck is forcing the rabbit to bend over; both are making arm gestures	Look! Gorping!	Black
2.	Black	See? Gorping!	Duck is forcing rabbit to bend over; both are making arm gestures
3.	Duck is forcing rabbit to bend over; both are making arm gestures	Wow! Gorping!	Duck is forcing rabbit to bend over; both are making arm gestures
4.	Duck is forcing rabbit to bend over	Oh! They're different now!	Duck and rabbit are making arm gestures
5.	Duck is forcing rabbit to bend over	Where's gorping now?	Duck and rabbit are making arm gestures
6.	Duck is forcing rabbit to bend over	Find gorping!	Duck and rabbit are making arm gestures

NOTE.—From L. Naigles (1990), "Children use syntax to learn verb meanings," *Journal of Child Language*, 17, 374. Copyright by Cambridge University Press. Adapted by permission.

"Look! Gorping!"). This simultaneous presentation of two novel actions and one novel verb (repeated twice: once on the other screen and once on both screens, trials 2 and 3) comprised the "teaching" phase of the experiment.

Next, the two actions were separated into single-action scenes (trial 4 in Table 2): One screen showed only the Causative action of the duck forcing the rabbit into the odd bending position, while the other screen, displayed at the same time, showed only the Synchronous action of the duck and the rabbit making the arm gestures. This familiarization trial (formerly called the control trial; the audio is "Oh! they're different now!") permitted the child to inspect the two alternatives visually before the directive audio was introduced.

Finally, the test trials (trials 5 and 6) were presented. The single-action scenes appeared again, paired with the test audios, "Where's gorping?" or "Find gorping now!" These trials tested what the children learned during the teaching phase.

This pattern of teaching and testing was repeated across four different nonsense verbs for each of the subjects, for a total of eight test trials per subject, with a duration of 6 sec per trial. The actions chosen were likely to be novel to young 2-year-olds because most are performed only in the context of aerobics workouts, and none can be de-

scribed by a single lexical item. In all three experiments reported here, each verb was paired with only one set of actions (i.e., there was no counterbalancing across verbs; see App. B for the specific actions). The side of the matching screen was counterbalanced across subjects, by varying the placement of the two tapes in the videotape decks. The side of the matching screen was also counterbalanced within subjects so that the match occurred equally on the left and right sides.

Procedure

Subjects and their parents were first interviewed in the playroom by the experimenter for an average of 15 min. At this time, the parent filled out a brief language questionnaire (Rescorla, 1991) and was given the visor to wear. The experimenter then escorted the parent and child to the testing room. The mother was told to place the child on the center of her lap facing forward. The experimenter then activated the presynchronized videotapes and left the testing room. During a trial, children generally scanned back and forth across both monitors, sampling the contents of both screens. After the tapes had ended, the experimenter returned to the testing room and replayed the videotapes for the parent.

Coding and Dependent Variables

Visual fixation was coded from the videotapes. An observer, who could not hear
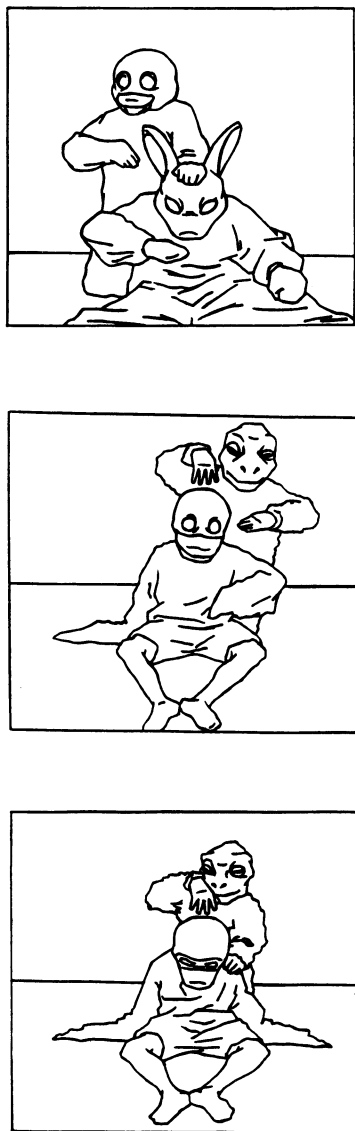
FIG. 2.—Pictorial representation of multiple-action scenes for Experiment 1 (top panel), Experiment 2 (middle panel), and Experiment 3 (bottom panel).

the stimulus audio and so could not know where the child was "supposed" to look, pressed buttons for the duration of the child's fixation to the left or right screens, to the center, or entirely away from the screens (e.g., up or down). Data from the button presses were collected and tabulated by the computer.

Each trial was coded for two dependent variables: (1) Total visual fixation time to the matching and nonmatching screens (mea-

sured in hundredths of a second). (2) Direction of first look to the matching versus nonmatching screen. Because the single-action scenes for the test trials appeared on the same screens as they had during the familiarization trial, and because the audio was presented both while the screens were blank and when the videos appeared, we expected that once children had heard the test audio the first time, they might direct their looking toward the screen where they anticipated the matching scene would be. This should hold even for the very first test trial, because this pattern of same scene–same side was established during the preceding character-familiarization trials (see App. A). Therefore, all measures were coded from the point at which the infant looked at the central light, when the screens were blank, for more than 0.3 sec. Trials where the child had not looked at the center light for a minimum of 0.3 sec were excluded. In addition, trials where the child had not looked at either screen (once the pictures appeared) for a minimum of 0.3 sec were excluded. Using these criteria, an average of 8% of the trials were excluded from further analysis across all conditions of the three experiments (range 6%–12% of trials).

The video of each child was initially viewed by two different coders. If the coders did not agree on visual fixation for the test trials for each verb (within approximately .5 sec), a third (and sometimes fourth and fifth) coder viewed the video, until two coders were found to be in agreement. An average of 3.27 coders was necessary to reach agreement on each subject (range 2.63–3.52 coders across all conditions). The chronologically earlier of these two coders-in-agreement was selected as the one whose coding was included in the analysis ("Chosen Coder"). Reliability was calculated between these two coders and also between the Chosen Coder and the first coder of all who viewed that child's video. If the Chosen Coder was the first coder, then the second reliability calculation was performed between the Chosen Coder and the second coder. The reliability comparisons were based on the difference between each child's mean fixation times to the two screens across test trials. Reliability for all 104 children in the three experiments was high, with an average r of .95 (range .87–.99 for individual audio conditions). Reliability between coders was also calculated for the first look measure, based on the difference between each child's total number of first
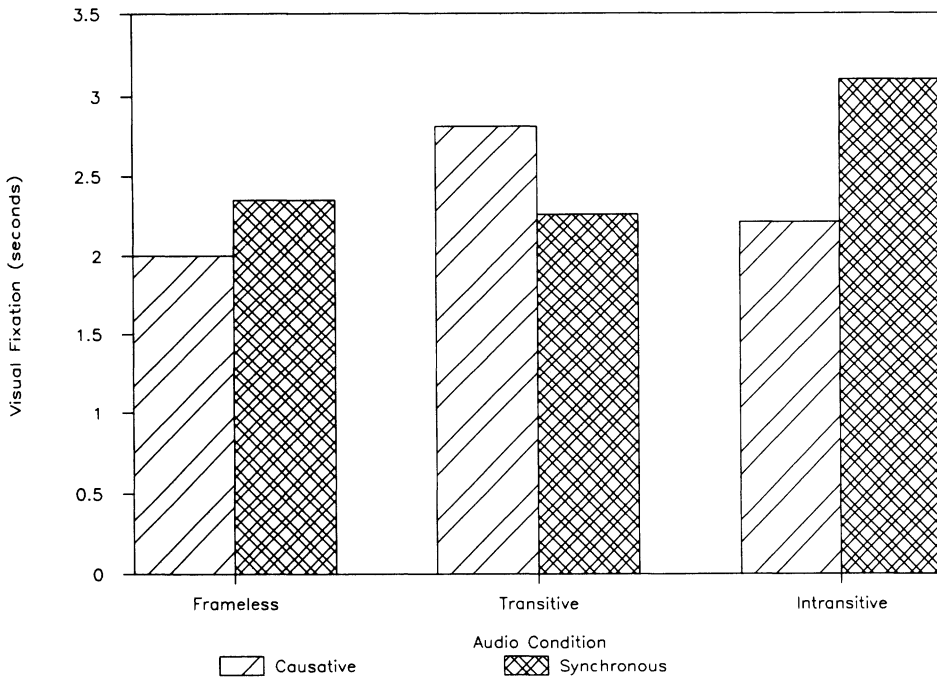
Fig. 3.—Mean visual fixation during the test trials to the Causative and Synchronous actions, for the Frameless condition, Experiment 1, and for the Transitive and Intransitive conditions, Naigles (1990). Right two panels from L. Naigles (1990), "Children use syntax to learn verb meanings," *Journal of Child Language*, 17, 366. Copyright held by Cambridge University Press. Adapted by permission.

looks to each screen across test trials. Although first look was not considered when selecting our coders-in-agreement, reliability between coders on the first look measure was still quite high, with an average *r* of .85 (range .73–.94 for individual audio conditions).

RESULTS AND DISCUSSION

The results were organized so as to address two questions. The first question concerned the children's action focus without the syntactic frames, so the data were analyzed to see whether the children showed a significant looking preference for the Causative or the Synchronous actions during the test trials. The second question asked whether such an initial focus could be altered by presenting the verb in a syntactic frame. To investigate this, we compared the results from Experiment 1 with those obtained from the children in Naigles (1990), who saw the same videos but heard the verb presented in either a transitive or an intransitive syntactic frame.

Frameless Audio

After hearing the novel verb presented in the Frameless audio, the children's mean visual fixations to the Causative and Synchronous screens was 2.0 and 2.35 sec, respectively (SD = .53 and .82, respectively), as shown in the left-most panel of Figure 3. Out of the 6 sec of each trial, the children spent an average of 4.35 sec looking at the screens (SD = .79); the rest of the time was spent centering or looking away. A three-way ANOVA (sex × verb [*gorp, blick, krad, dax*] × screen) revealed no significant main effect of screen, suggesting that the children had no preference for either action when the verb was presented in isolation. A significant interaction between verb and screen was obtained, $F(3, 42) = 4.06$, $p < .05$. Post hoc Tukey's tests for each verb revealed that only *dax* elicited a significant screen preference, which was for the Synchronous action ($p < .01$). Ten of the 16 subjects preferred to watch the Synchronous screen overall, while six preferred the Causative screen ($p > .10$ using a binomial distribution). Likewise, while the subjects tended to look first to the Synchronous screen rather than to the Causative screen (see Table 3), a two-way ANOVA (sex × screen) yielded no significant effects. The above analyses were performed collapsing across the first and second test trials for each novel verb. When the data

TABLE 3

NUMBER OF FIRST LOOKS TO EACH SCREEN BY EXPERIMENT AND
AUDIO CONDITION

	Causative	Synchronous	p
Experiment 1:			
Frameless	3.75	3.63	NS
	(1.45)	(1.71)	
Transitive[a]	3.25	3.50	NS
	(1.53)	(1.71)	
Intransitive[a]	3.50	3.67	NS
	(1.44)	(1.43)	

	Contact	Synchronous	
Experiment 2:			
Frameless	2.50	4.50	**
	(.94)	(1.17)	
Transitive	3.15	4.25	NS
	(1.68)	(1.72)	
Intransitive	3.13	4.38	*
	(1.05)	(1.16)	

	Causative	Contact	
Experiment 3:			
Frameless	4.94	2.38	**
	(1.14)	(.93)	
Transitive	4.25	3.05	+
	(1.47)	(1.20)	

[a] These data were gathered in the study by Naigles (1990); however, the scores for this dependent measure were not reported in that paper. Standard deviations in parentheses.
+ $p < .10$.
* $p < .05$.
** $p < .01$.

from the first and second trials were considered separately, the same pattern of effects was obtained.

Frameless Audio versus Transitive and Intransitive Audios

In these analyses, the children's fixations to a screen were compared across audio conditions: Frameless versus Transitive and Intransitive. The question asked was whether the children's screen preferences after hearing the verb presented in a syntactic frame differed from their peers' screen preferences after hearing the verb presented in isolation.

The 24 children (12 in each audio condition) in Naigles (1990) were somewhat younger than those in Experiment 1 and had on average a slightly smaller vocabulary (see Table 1 for details). The stimuli, procedure, and design were exactly the same as in Experiment 1, with the following exceptions:

(1) The verbs in the teaching trials were presented in either transitive (e.g., "Look! The duck is gorping the bunny!") or intransitive with conjoined subject (e.g., "Look! The duck and the bunny are gorping!") syntactic frames; (2) The children's visual fixations were not filmed; instead, these were recorded on-line by an observer hidden behind the apparatus, who was unaware of the experimental condition.

Visual fixation.—As shown in Figure 3, the screen preferences during the test trials in the Transitive condition appear different from those in the Frameless condition, whereas the preferences in the Intransitive condition resemble those in the Frameless condition. The mean visual fixations to the Causative and Synchronous screens, respectively, were 2.82 and 2.16 (SDs = .70 and .66) in the Transitive condition, and 2.21 and 3.10 (SDs = .50 and .51) in the Intransi-

tive condition. A four-way ANOVA (audio [Frameless, Transitive, Intransitive] × sex × verb × screen) yielded a significant interaction between audio and screen, $F(2, 34) = 6.16$, $p < .01$. Planned comparisons revealed that the screen preferences in the Transitive condition were significantly different from those in the Frameless condition, $F(1, 34) = 5.36$, $p < .05$; however, the screen preferences in the Intransitive and Frameless conditions did not differ significantly. Post hoc Tukey's tests within each syntactic audio condition indicated that the children in the Transitive condition significantly preferred the Causative screen, while those in the Intransitive condition significantly preferred the Synchronous screen ($p < .05$). There was also a significant interaction between verb and screen, $F(3, 102) = 3.41$, $p < .05$. Post hoc Tukey's tests revealed that the first verb, *gorp*, elicited significantly different preferences from the second and fourth verbs, *blick* and *dax* ($p < .05$ and $p < .01$, respectively). The difference was that *gorp* tended overall to elicit longer visual fixations for the Causative screen while *blick* and *dax* (and to a lesser extent *krad*) tended to elicit longer visual fixations for the Synchronous screen. There was no significant interaction of audio, verb, and screen. Finally, there was a main effect of audio, $F(2, 34) = 5.38$, $p < .01$, where the children in the Transitive and Intransitive conditions spent 5.12 and 5.22 sec, respectively (out of a total of 6 sec), looking at the video monitors, while those in the Frameless condition spent only 4.36 sec looking at the video monitors. In sum, hearing a novel verb in a transitive frame yields a significantly different looking patterns from hearing the verb in isolation; however, hearing a verb in an intransitive frame yields basically the same looking pattern as hearing it in isolation.

First look.—With both the Transitive and Intransitive audios the children tended to look first to the Synchronous screen (see Table 3). Because this was also the children's tendency in the Frameless condition, the three-way ANOVA (audio × sex × screen) revealed no significant interactions between audio and screen, $F(2, 34) < 1$, and no other significant effects.

These comparisons indicate that the frame in which a novel verb was presented did bring about differences in children's conjectures as to what action that verb might refer to. When the verb was presented in isolation, as with the Frameless audio of Experiment 1, the children showed no real

preference for Synchronous or Causative actions. When the verb was presented in a transitive frame (by Naigles, 1990), the children's preferences were different: They looked significantly longer at the Causative actions. Thus, we have demonstrated an effect of syntactic frame across audio conditions. Finally, when the verb was presented in an intransitive frame (by Naigles, 1990), children's preferences were more strongly for the Synchronous actions. This preference was in the same direction as that of the Frameless condition, though, and so did not demonstrate an effect of syntax.

Experiment 2

This experiment assessed whether Syntactic Bootstrapping facilitates the acquisition of Contact verbs, as it has already been shown to do with Causative verbs. That is, can transitive frames lead children to attend to Contact actions in a multiple-action scene? To investigate this, we presented multiple-action scenes containing novel Contact and Synchronous actions, paired with nonsense verbs in transitive or intransitive frames. We continued to explore the effect of initial focus by presenting the videos without a syntactic frame as well.

METHOD

Subjects

Fifty-two children completed this experiment. Seventy-three children were tested, but 21 children were eliminated from the final sample for the following reasons: Fixation bias of more than 75% to a single side (six), parent watching video along with child (nine), refusal of the child to watch the videos for more than half of the trials (three), inattention to the centering light (two), and inability of two out of five coders to agree on child's fixations (one). The final sample included 26 males and 26 females with a mean age of 2 years, 3 months ($SD = 29.8$ days). Children were assigned to the Frameless condition ($n = 16$, half males and half females), the Transitive condition ($n = 20$, half males and half females), or the Intransitive condition ($n = 16$). Table 1 presents the detailed subject information for each audio condition. The 21 children who were eliminated did not differ significantly from the final subject pool in any of the listed variables.

Stimuli and Design

The stimulus videotapes used in both Experiment 2 and Experiment 3 were recorded with a video camera filming against
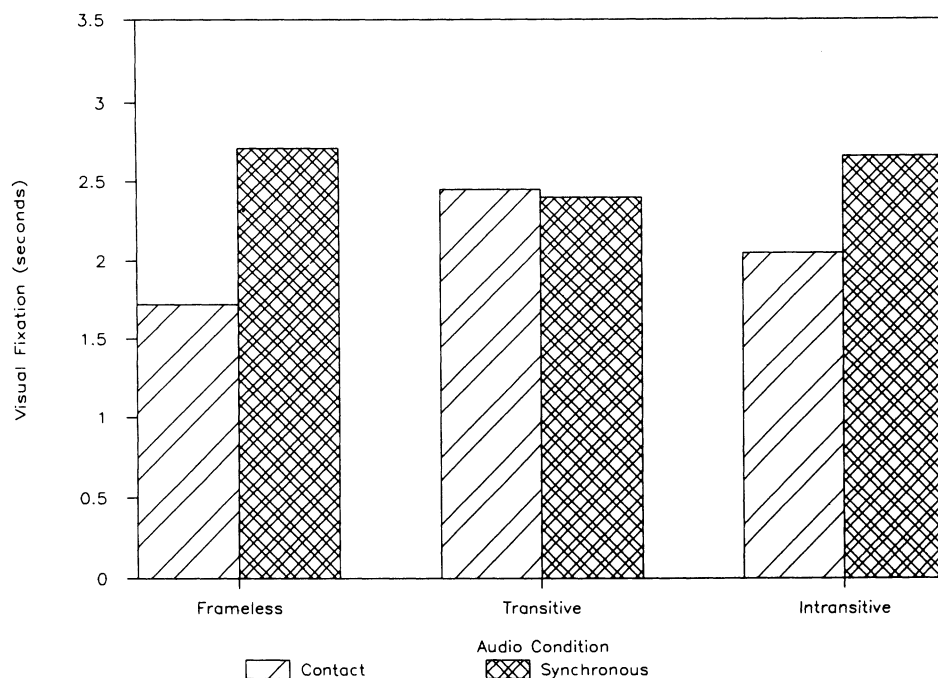
FIG. 4.—Mean visual fixation during the test trials to the Contact and Synchronous actions for the Frameless, Transitive, and Intransitive audio conditions, Experiment 2.

beige curtains. The structure of the video-tapes was exactly the same as in Experiment 1 (that is, three teaching trials, one familiarization trial, and two test trials). The actual scenes were different: The multiple-action scene was composed of a Synchronous action (e.g., the duck and frog [note new character] making identical arm gestures) and a Contact action (one character makes physical contact with the other, yielding no resulting movement; e.g., the frog contacting the duck's head with grasping finger motions). An example of this multiple-action scene is presented in static picture form in the middle panel of Figure 2; in the video, both actions were iterative and were repeated the same number of times. There were three accompanying audio conditions, all of which were presented in infant-directed speech intonation. Children heard the novel verbs either without syntactic frames (Frameless), in transitive frames (Transitive), or in intransitive frames with conjoined subjects (Intransitive).

There were again four different nonce verbs and sets of novel actions; see Appendix B for the specific actions. The procedure was exactly the same as in Experiment 1.

RESULTS AND DISCUSSION

As in Experiment 1, the first analysis considered the initial focus question, investigating whether the children in the Frameless condition had a preference for the Contact or the Synchronous actions. The second analysis compared the children's preferences across the audio conditions, to investigate the effect of syntax.

Frameless Audio

In the Frameless audio condition, the children's mean fixation times to the Contact and Synchronous screens were 1.76 and 2.65 sec, respectively (SD = .77 for both means), as shown in the left-most panel of Figure 4. The children spent an average of 4.42 sec per trial (out of a total of 6 sec) looking at the screens (SD = .69). A three-way ANOVA (sex $\times$ verb $\times$ screen) revealed a significant main effect of screen, $F(1, 14) = 9.31$, $p < .05$, and no other significant main effects or interactions. Eleven of the 16 subjects preferred to watch the Synchronous screen overall, while five preferred the Contact screen ($p < .10$ using a binomial distribution). The subjects also tended to look first at the Synchronous screen rather than at the Contact screen (see Table 3). This pattern is

in the same direction as that of the visual fixation measure, and as happened with that measure, it reaches significance, $F(1, 14) = 19.58$, $p < .001$.

Frameless versus Transitive and Intransitive Audios

Visual fixation.—As shown in Figure 4, the screen preferences are different in the Frameless and Transitive conditions: There is a large preference for the Synchronous screen in the Frameless condition and essentially no preference for either screen in the Transitive condition ($M = 2.45$ sec to the Contact screen and 2.40 sec to the Synchronous screen, SD = .59 and .56, respectively). By contrast, the preferences in the Frameless and Intransitive conditions are quite similar; for the Intransitive condition, the mean visual fixations were 2.05 sec to the Contact screen and 2.66 sec to the Synchronous screen (SD = .71 and .56, respectively). Total looking times to the screens averaged 4.86 sec per trial in the Transitive condition (SD = .65) and 4.78 sec per trial in the Intransitive condition (SD = .58). A four-way ANOVA (audio [Frameless, Transitive, Intransitive] × sex × verb × screen) revealed a significant main effect of screen, $F(1, 46) = 10.13$, $p < .01$, a significant interaction between audio and screen, $F(2, 46) = 3.76$, $p < .05$, and no other significant effects or interactions. Planned comparisons revealed that the screen preferences in the Transitive condition were significantly different from those in the Frameless condition, $F(1, 46) = 7.28$, $p < .01$; however, there was no significant difference between the screen preferences in Frameless and Intransitive conditions. Post hoc Tukey's tests within each syntactic audio condition yielded no significant preferences for either screen in the Transitive condition and only a weak preference for the Synchronous screen in the Intransitive condition ($p < .10$).

First look.—In all three audio conditions, the children tended to look first to the Synchronous screen (see Table 3). Overall, then, there was a significant main effect of screen, $F(1, 46) = 15.00$, $p < .001$, and no significant interaction between audio and screen.

The above analyses were performed collapsing across the first and second test trials for each novel verb. When the data from the first and second trials were considered separately, the same results were obtained, although the effects were at higher levels of

significance for the first trial than for the second.

In sum, these results demonstrate that the frame in which a novel verb is presented made a difference to children's conjectures as to what action that verb might refer to. When the verb was presented in isolation, as in the Frameless condition, the children showed a significant preference for Synchronous over Contact actions. When the verb was presented in a transitive frame, the children's preferences were significantly different: They looked longer at the Contact actions than did their peers in the Frameless condition. In contrast, like the results of Experiment 1, the presentation of the novel verb in the intransitive frame did not make a difference to children's conjectures as to what action that verb might refer to, relative to their peers' demonstrated preference for the Synchronous action when the verb is presented in isolation.

Experiment 3

In the preceding two experiments, we have demonstrated that transitive frames can lead 2-year-old children to focus more on Causative actions (Experiment 1) and on Contact actions (Experiment 2). The purpose of the present experiment was to compare the relative strengths of the transitive-causative and transitive-contact linkages. If both actions were present in the environment paired with a verb in a transitive frame, which would the child choose? As before, we also addressed the initial focus issue, investigating which action would be chosen if the verb was presented without a syntactic frame.

METHOD

Subjects

Thirty-six children completed the experiment. Sixty children were tested, but 24 children were eliminated from the final sample for the following reasons: Parent watching video along with child (seven), refusal of the child to watch the videos for more than half of the trials (three), inability of two out of five coders to agree on child's fixation (three), fixation bias of more than 75% to a single side (six), and equipment failure (five). The final sample included 18 males and 18 females with a mean age of 2 years, 3 months (SD = 28.6 days). Children were assigned either to the Frameless condition ($n = 16$, half males and half females) or the Transitive condition ($n = 20$). Table 1 pre-
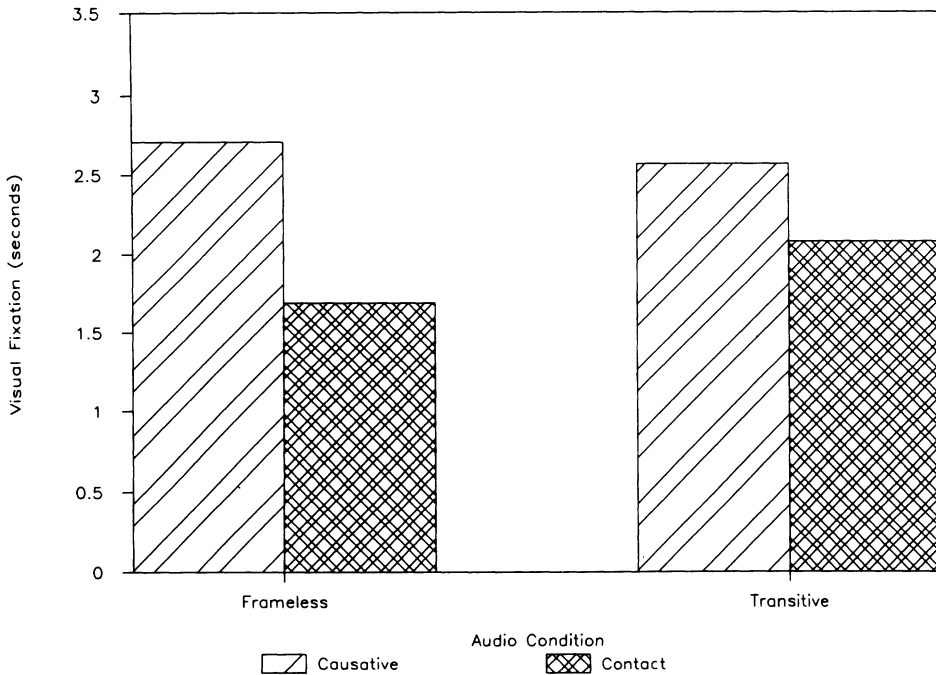
Fig. 5.—Mean visual fixation during the test trials to the Contact and Causative actions for the Frameless and Transitive audio conditions, Experiment 3.

sents the detailed subject information for each condition separately. The 24 children who were eliminated did not differ significantly from the final subject pool in any of the listed variables.

Stimuli and Design

The structure of the videotapes was exactly the same as in Experiments 1 and 2. The multiple-action scenes were composed of pairs of the four Causative and Contact actions used in Experiments 1 and 2 (with the exception of "pimming"; see App. B). The bottom panel of Figure 2 provides an example of one multiple-action scene in static-pictorial form; in the videos, both actions were iterative and were repeated the same number of times. The accompanying nonce verbs were presented either without a syntactic frame (Frameless condition) or in a transitive frame (Transitive condition). The procedure was exactly the same as in Experiments 1 and 2.

RESULTS AND DISCUSSION

As in Experiments 1 and 2, the first analysis considered the data collected in the Frameless condition, while the second analysis compared the children's preferences across the audio conditions.

Frameless Audio

The children's mean fixations to the Contact and Causative screens were 1.69 and 2.74 sec, respectively (SD = .82 and .71, respectively), as shown in the left-most panel of Figure 5. Total looking times to the screens averaged 4.43 sec per trial (SD = 1.08). A three-way ANOVA (sex × verb × screen) revealed a significant main effect of screen, $F(1, 14) = 14.67$, $p < .01$, and no other significant main effects or interactions. Fourteen of the 16 subjects preferred to watch the Causative screens overall, while two preferred the Contact screens ($p < .01$ using a binomial distribution). The subjects also directed their first looks overwhelmingly to the Causative screen (see Table 3). This pattern is in the same direction as that from the visual fixation measure, and as happened with that measure, it reaches significance, $F(1, 14) = 24.37$, $p < .001$.

Frameless versus Transitive Audios

Visual fixation.—As shown in Figure 5, the looking preferences are in the same direction in both the Frameless and Transitive conditions: In the latter condition, the mean visual fixations were 2.57 sec to the Causative screen and 2.08 sec to the Contact screen (SD = .70 and .80, respectively). Total looking times to the screens during the

Transitive condition averaged 4.68 sec (SD = .82). A four-way ANOVA (audio [Frameless vs. Transitive] sex × verb × screen) revealed a significant main effect of screen, $F(1, 32) = 14.31, p < .001$. Although the preference for the Causative screen is numerically greater in the Frameless condition than in the Transitive condition, there was no significant interaction between audio and screen, and no other significant effects. A post hoc Tukey's test within the Transitive condition revealed no significant preference for either screen.

First look.—In both the Transitive and Frameless audio conditions, the children tended to look first to the Causative screen (see Table 3). Overall, there was a significant main effect of screen, $F(1, 32) = 30.01, p < .001$, and no significant interaction between audio and screen.

As in the preceding experiments, the above analyses were performed collapsing across the first and second test trials for each novel verb. When the data from the first and second trials were considered separately, the same pattern of results was obtained.

The most striking finding in Experiment 3 is that the significant effect of the transitive frame found in Experiments 1 and 2, in yielding a different pattern of looking from the Frameless condition, was not found in this experiment. Instead, the significant preference for the Causative action in the Frameless condition did not differ from the nonsignificant preference for the same action found in the Transitive condition.

General Discussion

Our results demonstrate that 2-year-old children do exhibit initial action preferences upon their first exposure to two novel actions and a single "frameless" novel verb. These preferences can be altered, though, when the verb is presented in a transitive syntactic frame. Specifically, the children's preferences for the Causative actions in Experiment 1 and for the Contact actions in Experiment 2 were greater in the Transitive conditions than were those of their counterparts in the Frameless conditions. In the discussion which follows, we consider these findings with regard to the methodological advances made over Naigles's (1990) work and then address the following two questions: (1) What is the underlying basis for the initial preferences revealed in Experiments 1, 2, and 3? (2) What is the status now of the linking rule which allowed the transi-

tive frame to indicate or afford either Causative or Contact meanings?

Methodological Issues

One of the methodological differences between this and earlier research (e.g., Naigles, 1990) is the use of the first look measure in addition to the visual fixation measure. The findings of first look corroborate those of visual fixation for the children's initial focus on particular actions; however, the effects of syntax were obtained from the visual fixation measure alone. That is, while the amount of visual fixation to a screen changed depending on whether the novel verb was presented in isolation or in a transitive frame, the direction of first look did not. We can rule out reliability as a factor in these divergent findings because the intercoder reliability for both measures was quite high (r's of .87 and .95 for first look and visual fixation, respectively).

Differences in the variability of the two measures could also result in a lack of correspondence, because greater variability within conditions could attenuate between-condition effects. An F test comparing the grand variances of the two measures confirmed that the first look variance was significantly greater than that of visual fixation ($s^2 = 2.43$ and 0.60, respectively; $F(103, 103) = 4.02, p < .05$). Interestingly, individual F tests reached significance in the Transitive conditions of Experiments 2 and 3, $F(19, 19) = 8.77$ and 3.33, respectively, $p < .05$, and in the Frameless condition of Experiment 1, $F(15, 15) = 5.92, p < .05$, but not in the Frameless conditions of Experiments 2 and 3. It appears that the first look variance is significantly greater than that of visual fixation in the two between-condition comparisons where first look failed to corroborate the visual fixation findings. Finally, while the variability of the measures is clearly implicated in their divergent findings, properties of the measurement scale may also be a contributing factor. Unlike the continuous visual fixation measure, the first look variable is a discrete one. Discrete variables are generally less sensitive than continuous ones, so preferences or subject size would have to be greater (or less varied) to achieve significance.

A second methodological innovation involved our use of the Frameless audio to provide a baseline condition from which the effects of syntax could be assessed. The importance of this condition can be clearly

seen in our interpretation of the results of Experiments 2 and 3. In Experiment 2, the children showed no significant screen preferences within the Transitive condition, and an almost-significant preference for the Synchronous action over the Contact one within the Intransitive condition (see Fig. 4). Without the Frameless comparison, we might have concluded that only the intransitive frame had an effect. Likewise in Experiment 3, the children within the Transitive condition looked somewhat longer at the Causative action than the Contact one (see Fig. 5). The two experiments together might have been taken to indicate that the transitive frame implicates specifically Causative actions. However, the Frameless conditions demonstrate that these looking patterns toward the Synchronous action in the Intransitive condition of Experiment 2, and toward the Causative action in the Transitive condition of Experiment 3, were not due to the presentation of a syntactic frame. As a result, the effects of syntax come from the transitive frames in Experiments 1 and 2, and we conclude that the transitive frame directs attention to both Causative and Contact actions.

This Frameless condition was an advance over the use of the within-subject familiarization trials as the basis of comparison in Naigles (1990). The change was made because we conjectured that looking preferences for these trials were unavoidably influenced by the immediately preceding teaching audio. This conjecture is supported by data from the current set of experiments, in that, for the Transitive and Intransitive conditions in Experiment 2 and the Transitive condition in Experiment 3, the action preferences during the familiarization trials were similar to those exhibited during the following test trials, and sometimes different from those obtained during the test trials in the corresponding Frameless conditions. Thus, the within-subject shifts between familiarization and test trials found by Naigles (1990) were not replicated in these experiments. Such inter- and intraexperiment variability with the familiarization trials suggests that they are not reliable indicators of children's initial action preferences, and so the earlier results bear reevaluation. Because the transitive frame effect obtained by Naigles (1990) was replicated here with the Frameless condition as the baseline, the validity of this effect is strengthened. In contrast, the intransitive frame effect obtained by Naigles (1990) was not replicated here,

either in Experiment 1 or 2; therefore, this effect must be viewed with some skepticism.

The absence of an intransitive frame effect in our experiments could be due to theory-based factors such as an inadequate linkage between intransitive frames and noncausative meanings, at least in our 2-year-old subjects, or to methodological factors. While the first cannot be ruled out, we believe the second is more likely. Naigles, Gleitman, and Gleitman (1993) demonstrated that 2-year-olds, given ungrammatical intransitive sentences to enact (e.g., "The zebra brings to Noah"), altered the (usually Causative) verbs' interpretations to fit the syntactic frame. Their predominant enactment of such sentences was non-Causative (e.g., The zebra moving alone to Noah), indicating that the syntax-semantics linkage is in use. A related possibility is that the Synchronous actions in Experiments 1 and 2 were not encoded or conceptualized as non-Causative actions by the children, so that the intransitive-noncausative correlation did not appear to apply. From the available data we cannot determine whether or not this was the case; what is needed is an assessment of how broad or narrow is the 2-year-old's category of non-Causative actions (and for that matter, how broad or narrow are the 2-year-old's categories of Causative and Contact actions).

The explanation we think most likely is that because the children's nonsyntactic focus was on the Synchronous action, and the designated referent for the Intransitive audio was also the Synchronous action, any effect of the intransitive frame was masked by the effect of the verb in isolation. This is especially telling in Experiment 2, where the preference for the Synchronous screen over the Contact one in the Frameless condition almost reached 1 sec (see Fig. 4). In order to demonstrate a significant difference in the Intransitive condition, the mean looking preference would have had to be on the order of 2 sec; however, mean preferences on this order of magnitude have not been obtained in any of the published language-comprehension studies using the preferential-looking paradigm (Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1991; Naigles, 1990). Our conjecture, then, is that the absence of significant effects for the intransitive frame is due to ceiling effects.

Even without yielding a significant comparison with the Frameless condition,

the Intransitive conditions served another crucial function in this research: They demonstrated that the children's preferences during the Transitive conditions were in fact due to the structural properties of the transitive frame. If only one syntactic condition had been presented, the observed effects could have been due simply to the presence of a full syntactic frame of any configuration, or to the presence of a syntactic frame containing two noun phrases (NPs) referring to the characters in the scene. The syntactic frames in the Intransitive conditions included both of these properties, yet they did not succeed in eliciting different preferences concerning the referent of the novel verb. The syntactic frames in the Transitive condition included these properties plus a particular arrangement of NPs: The agent NP was in subject position and the patient/goal NP was in object position. This configuration did succeed in eliciting different selections of action referents. Therefore, we can conclude that it was indeed the particular structure of the transitive frame that was influencing the subjects' looking preferences.

Taken together, the results of these three experiments fortify findings that emerged from earlier research on syntactic bootstrapping, showing that young children use syntactic information to constrain or alter the meanings of novel verbs. The only other study comparable in goals and approach is that of Fisher et al. (in press), in which preschoolers were shown videotapes of familiar actions and taught novel verbs either in isolation or in syntactic frames. As in our study, action preferences were obtained when the verbs appeared in isolation, preferences which were different when the verbs appeared in a syntactic frame. Our results extend Fisher et al.'s (in press) in two ways: We have demonstrated an effect of syntax with younger children (mean age = 2 years, 3 months) and with novel actions. The actions we used are not easily described by a single (and so possibly familiar) lexical item in English; therefore, inasmuch as our subjects probably did not have single words for the actions presented, our situation provides a closer simulation of the actual verb-learning scenario.

WHY THE INITIAL PREFERENCES?

We found initial (i.e., extrasyntactic) preferences in children's looking at some of the actions presented; however, the preferences differed according to what the alternative actions were. That is, Synchronous actions appear to be preferred overall, significantly when paired with Contact actions and nonsignificantly when paired with Causative actions. When Causative and Contact actions were paired, the children preferred to watch the Causative actions. Thus, the three kinds of actions can be rank ordered, where Synchronous ones are most preferred, Causative next, and Contact least. At this point, one might wonder how representative of the classes of Synchronous, Causative, and Contact actions were the ones utilized in these experiments. The actions used here could not be prototypical ones, because such actions are quite common and their labels probably would been known to our subjects. Furthermore, the typical Causative actions have inanimate patients, while ours required animate ones so that the Intransitive audio could include the same NPs as the Transitive audio. Thus, the true representativeness of our actions is difficult to judge. The paucity of individual verb/action effects throughout the experiments suggests, though, that within a class all four actions were treated similarly. We next consider two possible bases for the obtained ranking of action preferences.

Lexicalization.—There have been several recent demonstrations of young children's initial biases to lexicalize particular aspects of events (e.g., Behrend, 1990; Gentner, 1978; and Gropen, Pinker, Hollander, & Goldberg, 1991a, 1991b). For example, Behrend (1990) has shown that, given a (frameless) novel verb paired with an action which involves both a manner and a result or end state, preschoolers' first guess is that the verb refers to the result rather than the manner (e.g., *cut* refers to the result of cutting, not the particular manner of cutting, which can be sawing, "scissoring," etc.). Might the results of the present set of studies be interpreted analogously, through the argument that given a (frameless) novel verb paired with both Synchronous and Contact or Causative actions, 2-year-olds tend to assume that the novel verb refers to the Synchronous action? There is some supporting evidence for this; Huttenlocher, Smiley, and Charney (1983) have shown that 1-year-olds understand the English labels for spontaneous motion events (e.g., The ball *bounced*) before they understand the same label used for caused events (e.g., The boy *bounced* the ball), and our Synchronous actions can be

viewed as spontaneous ones. However, in order to make this claim conclusively, at least two more kinds of evidence are needed. First, it must be demonstrated that the children did in fact view the novel lexical items in the Frameless audio as verbs or at least action words (see Tomasello, 1991). Second, it must be clear that the novel verb elicits different preferences than those exhibited when no verb is presented at all. We believe that there is sufficient support for the first point, but not for the second.

To adults, the "-ing" inflection we suffixed to our nonce words would be sufficient to suggest that they are verbs. Brown (1957) has shown that children as young as 3 years of age also assume that a novel word ending in "-ing" refers to an action, but there is no such clear-cut data for children under the age of 3. However, the "-ing" progressive form is one of the first inflections acquired in production, with ages of acquisition ranging from 1-10 to 2-10 (Brown, 1973), so it is not unlikely that our 2-year-olds had learned it. Although there also exist nouns in English which end in syllabic "-ing" (e.g., *wedding, carpeting, dressing*), these are of much lower frequency than the progressive or gerund forms of many common verbs (Francis & Kucera, 1982) and so are probably not confounding to such young children. Thus, the chances seem fairly good that the children considered the novel lexical items to be, at least, action words.

Presentation of the visual stimuli without a novel lexical item is a common control in studies of noun acquisition; it demonstrates that the specific presence of the noun changes the child's conjectures concerning the stimuli. This control has not been used in many verb acquisition studies, including this one; however, it is clearly critical for any claims concerning initial lexical biases in verb acquisition (see Naigles, Eisenberg, & Kako, 1992). Lacking this latter piece of evidence, we cannot establish that the action preferences in the Frameless conditions are lexicalization ones, and so we turn to a second possible explanation for the results in the Frameless conditions.

Perceptual salience.—Perhaps the actions are ranked according to perceptual salience; that is, perhaps the children choose to watch the scene which is most interesting or visually salient. What operates in order to make an action more or less perceptually salient? According to Gestalt psychology, movements in tandem like the Synchronous

actions in these experiments are basic ones for perceptual organization (Kohler, 1947), and infants have been shown to be sensitive to such Gestalt principles (Bornstein, 1988; Gibson & Spelke, 1983). Perhaps this "tandemness" is what makes the Synchronous actions so interesting. Alternatively, the Synchronous actions differed somewhat in size from the other two kinds of actions (i.e., in the degree of displacement of the different body parts or in the actual size of the body part performing the action; e.g., moving a torso vs. moving an arm). It is also widely recognized that larger actions are more interesting to children than smaller ones (Gibson & Spelke, 1983). Clearly, there could be several perceptual bases for the observed preferences. A full investigation into this issue is beyond the scope of this paper; however, we performed one follow-up study to see to what extent the size of the actions could predict the toddlers' preferences.

In this follow-up, eight Yale undergraduates were asked to rate the single-action scenes in all three pairs of videotapes for the size of the action, on a scale of 1 (small) to 10 (big). The subjects saw each scene separately, so no pairwise comparisons could be made on-line. The four single-action scenes on a given tape were presented in series; however, the six tapes (three pairs of tapes) were presented in random order. Each subject saw 24 scenes and made 24 ratings.

The raw (nonnormalized) mean ratings are presented in Table 4, averaged across subjects and across verbs. The first thing to note is that the Synchronous, Causative, and Contact actions from different sets of tapes were rated very similarly; this indicates that the individual actions had been performed similarly even though they had been, in the cases of the Synchronous and Causative actions, performed by different actors. More important, the ratings do appear to differ by type of action: The Synchronous actions received the highest ratings, the Contact actions the lowest, and the Causative actions fell in between. The critical comparisons are between the actions that were paired in the presentation to the children; t tests were performed on standardized z scores to see if these pairs of ratings differed significantly. As the table shows, the ratings of size of action did not differ significantly for the Causative and Synchronous actions in Experiment 1; however, the ratings of size of action for both the Synchronous and Contact actions in Experiment 2, and for the Causative and Contact actions in Experiment 3, were

TABLE 4

RATINGS OF SINGLE ACTION SCENES BY ADULTS

| | ACTIONS | | | |
SOURCE	Synchronous	Causative	Contact	t
Experiment 1	5.88 (1.40)	5.47 (1.20)	...	2.19, NS
Experiment 2	5.84 (1.38)	...	3.09 (.80)	80.66**
Experiment 3	...	5.59 (1.10)	3.44 (1.03)	37.04**

NOTE.—Rating scale was from 1 to 10, where 1 was designated for very small actions. Standard deviations are in parentheses. The t tests were performed on standardized z scores created for each rater.
** $p < .01$.

significantly different, $t(7) = 80.66, p < .001$ and $t(7) = 37.04, p < .001$, respectively. This pattern of significant differences matches that obtained from the children in the Frameless audio conditions. Thus, we have provided some evidence that these initial preferences for particular actions may be guided to a large extent by perceptual factors.

DEFINING A ROLE FOR SYNTAX

In the first two experiments, the screen preference the children exhibited when they heard the novel verb in the Frameless audio was different from that of their agemates who heard the verb presented in the transitive syntactic frame. Specifically, in Experiment 1 there was a change across conditions from no action preference in the Frameless condition to a significant Causative action preference in the Transitive condition. In Experiment 2, the change was from a significant Synchronous action preference in the Frameless condition to no significant action preference in the Transitive condition. Essentially, the looking time toward the Contact action in this condition increased, relative to that in the Frameless condition, and the looking time for the Synchronous action decreased. In contrast, in Experiment 3, the children in both conditions looked longer at the Causative actions, although the pattern within the Transitive condition did not reach significance. Had this Transitive looking pattern been significantly different from that in the Frameless condition, the absence of a within-condition preference might have indicated realization on the part of the children that a Contact action could also be the referent of a transitive verb. This was the expected result, given the dual implications of the transitive frame and the results from Experiments 1 and 2 suggesting that 2-year-olds know that Contact and Causative actions can be described by a verb in the transitive frame. However, given the strong preference already shown for the Causative action in the Frameless condition, and the fact that it is still a valid choice when the verb is in the transitive frame, the continued tendency toward the Causative action in the Transitive condition was perhaps inevitable. The Contact action, in comparison, had only the transitive syntactic frame supporting its selection, and the semantic implications of the frame were not as unambiguous as they had been in Experiment 2. The data provide one small indication that the audio stimuli in the Transitive condition did contribute to disproportionate confusion in our subjects, but because this comes from subject elimination data, it must be interpreted with caution. The Transitive condition in Experiment 3 experienced the highest percentage of subject mortality across all three experiments (47%, as compared with a mean of 26% [range 20%–28%] for the other five conditions in the three experiments). These subjects were eliminated due sometimes to factors such as side bias and inattention to the screens or the centering light, which directly reflect child inattention to the screens or the audio, and other times to factors such as parental video watching and coder disagreement, which can be indirect reflections of child inattention. While certainly not conclusive, this high elimination rate might be an indication of confusion which could be attributed to the ambiguous implications of the transitive frame.

We have found that the transitive frame facilitated children's action preferences toward both Causative actions (Experiment 1 comparison with Naigles, 1990) and Contact actions (Experiment 2). We have thus demonstrated that the procedure of syntactic bootstrapping can affect the acquisition of a second class of verbs, namely, Contact verbs like *touch* and *pat*. The use of syntax in verb learning is evident now for two classes of actions that are lexicalized as verbs in English. Moreover, these results suggest that the rule linking the NP arguments of the transitive frame with verb semantics is general rather than specific. That is, the component(s) of verb semantics to which the syntax is related appear to be general, encompassing (at least) both Causative and Contact actions. Gropen et al. (1991b) have proposed that the link with the arguments of the transitive frame involves object affectedness. That is, they point out that in some classes of transitive verbs, the direct object represents the noun most affected by the verb's action. For example, when you *load a wagon with hay*, the *wagon* is the object most salient or affected, while *loading hay into a wagon* highlights the affectedness of the *hay*. This notion is easily extended to both Causative and Contact verbs. In both classes, the object is the entity most affected by the verb's action; it is that which is *closed* or that which is *scratched*. Object affectedness, then, appears to capture the similarity between these two classes of verbs. If we recast this notion in terms of the semantics of the verb, we might say that transitive verbs are those that affect some other entity.

This discussion has led us, however, to a second conundrum: Given (1) that the child is probably faced with numerous scenes that allow both Contact and Causative interpretations (e.g., stroking a responsive cat), (2) that her preference without syntactic information, at least with the stimuli in Experiment 3, is to focus on the Causative action, and (3) that presentation of the verb in a single transitive frame does not sufficiently override this preference, how is she to acquire the accompanying Contact verbs? We suggest, following Landau and Gleitman (1985), that additional syntactic frames would be informative in this regard. That is, while transitive frames may be the canonical ones for Causative and Contact verbs, these verbs can often also appear in intransitive frames. Crucially, though, the entailment relations of the NP arguments in those intransitive frames differ between the two classes

of verbs (see Dowty, 1990; Levin, 1985; see also Naigles, 1990; Naigles, Gleitman, & Gleitman, 1993). For example, for Causative verbs such as *drop,* the object (patient) of the transitive sentence must become the subject (experiencer) of the intransitive sentence in order to generate the same reading:

1. The girl dropped the ball.
2a. The ball dropped.
2b. The girl dropped.

Thus, 2a is entailed by 1; that is, they describe the same event. Although 2b is also a grammatical sentence, it does not entail, nor is it entailed by, 1 or 2a; it describes a different event, where the girl, instead of the ball, moves downward. In contrast, for Contact verbs such as *scratch,* the subject of the transitive sentence must remain the subject of the intransitive sentence in order to generate the same reading:

3. The cat scratched the dog.
4a. The dog scratched.
4b. The cat scratched.

Here, the transitive sentence 3 entails the intransitive 4b; only these two describe the same event. Again, 4a is a grammatical sentence which describes a different event, in which the dog, not the cat, is doing the scratching.

Thus, the pattern of entailment relations in the transitive-intransitive frame alternation can distinguish Causative and Contact verbs syntactically, in the adult language. Our proposal is that this difference in syntactic patterning could also help the child distinguish the two classes of verbs. Preliminary evidence suggests that such multiple-frame information is available to toddlers. Using Hoff-Ginsberg's (1991) database of the video-recorded interactions of 57 mothers and their toddlers during mealtime, dressing, and free play with toys, Naigles and Hoff-Ginsberg (1993) have found that almost half of the 30 verbs investigated were used more than once within a single day. Of these multiple-use verbs, almost three-quarters were used in more than one syntactic frame during the same time period. Preparations are also underway to test this multiple-frame hypothesis directly, by presenting novel verbs in two syntactic frames in succession, paired with the scenes in Experiment 3. Note that this is not a "riddle" procedure as suggested by Pinker (1989, p. 263), because the child is viewing the scene while hearing the syntactic

frames. That is, the child does not hear the two frames and then try to guess what kind of action could fit both of them. Instead, the child views the scenes, hears the frames, and uses the frames to pull out the relevant aspects of the scene (see also Fisher et al., 1991; Gleitman, 1990; Naigles, 1991).

CONCLUSIONS

Children have initial preferences for particular actions or aspects of actions, pref-

erences that appear to be based on the perceptual salience of those actions. However, before they reach the age of 2½ years, they can use syntactic information, particularly that encoded in the transitive frame, to redirect or refocus their attention on those actions that were not initially favored. The transitive frame facilitates lexicalization of both Causative and Contact actions, indicating that the meaning with which it is correlated is a more general one, perhaps akin to object affectedness.

Appendix A

TABLE A1

STRUCTURE OF THE STIMULUS VIDEOTAPES FOR THE CHARACTER FAMILIARIZATION SEQUENCES

Tape 1	Audio	Tape 2
Duck waves	There's the duck!	Black
Black	There's the bunny!	Rabbit waves
Duck waves	See the duck?	Black
Black	See the bunny?	Rabbit waves
Duck waves	Here they are!	Rabbit waves
Duck waves	Where's the duck?	Rabbit waves
Duck waves	Where's the bunny?	Rabbit waves
Duck waves	Find the bunny!	Rabbit waves
Duck waves	Find the duck!	Rabbit waves

Appendix B

TABLE B1

ACTIONS ASSOCIATED WITH EACH NOVEL VERB

Verb	Causative	Synchronous
Experiment 1:[a]		
Gorp	Duck forces rabbit into bending position	Duck and rabbit flex own arms
Blick	Rabbit makes duck pat duck's head	Rabbit and duck lift own legs
Krad	Rabbit makes duck tilt duck's head	Duck and rabbit make arm circles
Dax	Duck lifts rabbit's leg	Duck and rabbit cover own eyes with own arm

	Contact	Synchronous
Experiment 2:		
Seb	Frog contacts duck's head with (frog's) fingers	Duck and rabbit flex own arms
Tig	Duck taps frog's stomach with duck's wing	Rabbit and duck lift own legs
Pim	Duck rubs frog's leg with duck's foot	Duck and rabbit make arm circles
Lorp	Frog steps on duck's foot	Duck and rabbit cover own eyes with own arm

TABLE B1 *(Continued)*

Verb	Causative	Synchronous
	Contact	Causative
Experiment 3:		
Seb	Frog contacts duck's head with (frog's) fingers	Frog forces duck into bending position
Tig	Duck taps frog's stomach with duck's wing	Duck makes frog pat frog's head
Lorp	Frog steps on duck's foot	Frog makes duck tilt duck's head
Pim	Duck rubs frog's leg with duck's foot	Duck makes frog's arms flex

[a] Experiment 1 actions from L. Naigles (1990), "Children use syntax to learn verb meanings," *Journal of Child Language*, 17, 365. Copyright held by Cambridge University Press. Reprinted by permission.

References

Behrend, D. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development,* 61, 681–696.

Bornstein, M. (1988). Perceptual development across the life cycle. In M. Bornstein & M. Lamb (Eds.), *Developmental psychology: An advanced textbook: Vol. 2. Perceptual, cognitive, and linguistic development* (pp. 151–204). Hillsdale, NJ: Erlbaum.

Bowerman, M. (1974). Learning the structure of causative verbs: A study in the relationship of cognitive, semantic and syntactic development. *Proceeding on Research in Child Language Development,* 8, 142–178.

Bowerman, M. (1982). Evaluating competing linguistic models with language acquisition data: Implications of developmental errors with causative verbs. *Quaderni di Semantica,* 3, 5–66.

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology,* 55, 1–5.

Brown, R. (1973). *A first language.* Cambridge, MA: Harvard University Press.

Dowty, D. (1990). Thematic proto-roles and argument selection. *Language,* 67, 547–619.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology,* 20, 104–113.

Fisher, C., Gleitman, H., & Gleitman, L. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology,* 23, 331–392.

Fisher, C., Hall, G., Rakowitz, S., & Gleitman, L. (in press). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua,* to appear.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Gibson, E., & Spelke, E. (1983). The development of perception. In J. Flavell & E. Markman (Eds.), P. Mussen (Series Ed.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 1–76). New York: Wiley.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition,* 1, 3–55.

Gleitman, L. R., Gleitman, H., Landau, B., & Wanner, E. (1987). Where learning begins: Initial representations for language learning. In F. Newmeyer (Ed.), *The Cambridge Linguistic Survey* (Vol. 3). Cambridge: Cambridge University Press.

Goldin-Meadow, S., Seligman, M., & Gelman, R. (1976). Language in the two-year-old. *Cognition,* 5, 189–202.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language,* 14, 23–45.

Grimshaw, J. (1990). *Argument structure* (Linguistics Inquiry Monograph No. 18). Cambridge, MA: MIT Press.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991a). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language,* 18, 115–152.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991b). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition,* 41, 153–196.

Hirsh-Pasek, K., & Golinkoff, R. (1991). Language comprehension: A new look at some old

themes. In N. Krasnegor, D. Rumbaugh, M. Studdert-Kennedy, & R. Schiefelbusch (Eds.), *Biological and behavioral aspects of language acquisition*. Hillsdale, NJ: Erlbaum.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development, 62,* 782–796.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). The emergence of action categories in the child: Evidence from verb meaning. *Psychological Review, 90,* 72–93.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Kohler, W. (1947). *Gestalt psychology*. New York: Liveright.

Landau, B., & Gleitman, L. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.

Levin, B. (1985). Lexical semantics in review: An introduction. In B. Levin (Ed.), *Lexical semantics in review* (pp. 1–62; Lexicon Project Working Papers No. 1). Cambridge, MA: MIT Center for Cognitive Science.

Levin, B., & Rappaport Hovav, M. (1991). Wiping the slate clean: A lexical semantic exploration. *Cognition, 41,* 123–152.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17,* 357–374.

Naigles, L. (1991). [Review of S. Pinker's *Learnability and cognition*]. *Language and Speech, 34,* 63–79.

Naigles, L., Eisenberg, A., & Kako, E. (1992). Acquiring a language-specific lexicon: Motion verbs in English and Spanish. In preparation.

Naigles, L., Gleitman, H., & Gleitman, L. R. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and development*. Norwood, NJ: Ablex.

Naigles, L., & Hoff-Ginsberg, E. (1993). *Input to verb learning: Verb frequency and frame diversity in mother's speech predicts children's verb use*. Unpublished manuscript, Yale University.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rescorla, L. (1991). Identifying expressive language delay at age 2. *Topics in Language Disorders, 11,* 14–20.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics, 11,* 145–166.

Slobin, D. (1985). Cross-linguistic evidence for the language-making capacity. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition: Vol. 2. Theoretical issues*. Hillsdale, NJ: Erlbaum.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science, 12,* 49–100.

Taylor, M., & Gelman, S. (1988). Adjectives and nouns: Children's strategies for learning new words. *Child Development, 59,* 411–419.

Tomasello, M. (1991). *First verbs*. Cambridge: Cambridge University Press.

Waxman, S., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1,* 139–156.

Waxman, S., & Kosowski, T. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development, 61,* 1461–1473.