ELSEVIER

Cognition 58 (1996) 221–251

COGNITION

The use of multiple frames in verb learning via syntactic bootstrapping

Letitia R. Naigles*

Department of Psychology, Yale University, P.O. Box 208205, New Haven, CT 06520-8205, USA

Received 26 November 1993, final version accepted 5 June 1995

Abstract

Following the original Syntactic Bootstrapping proposal of Landau and Gleitman (1985), this study investigated whether young 2-year-old children (mean age = 28 months) can use multiple syntactic frames, in addition to the extralinguistic scene, to help focus on the meaning of a novel verb. The multiple frames tested were combinations of transitive and intransitive frames in two alternation patterns, Causative and Omitted Object. By hypothesis, the Causative alternation would be more predictive of actions involving physical causation and the Omitted Object alternation more predictive of actions involving repeated physical contact without causation. Subjects were presented with videos depicting both actions, together with a novel verb. The actions were subsequently separated, and the children were asked to select which action was the referent of the novel verb. The novel verb was presented either in transitive and intransitive frames in the Causative alternation (CS: *The duck is sebbing the frog, the frog is sebbing*) or the Omitted Object alternation (OO: *The duck is sebbing the frog, the duck is sebbing*), or in intransitive frames only (IO: *The duck is sebbing*), or without a frame (FF: *Sebbing*!). In the CS, IO, and FF conditions, children preferred the causative action as the referent of the verb. However, the girls in the OO condition showed a significantly different preference, and looked more toward the contact actions than their peers in the other conditions did. This study thus provides the first experimental evidence that young 2-year-old children can use *multiple* syntactic frames to help determine the meaning of a novel verb.

* E-mail: naigles@yalevm.ycc.yale.edu.

SSDI 0010-0277(95)00681-8

1. The use of multiple frames in verb learning via Syntactic Bootstrapping

The theory of Syntactic Bootstrapping holds that young children in the process of acquiring the meaning of a verb will exploit the syntactic frames in which that verb appears, in addition to the extralinguistic scenes coincident with the verb's use (Landau & Gleitman, 1985; Gleitman, 1990). One reason that the syntactic information seems necessary for verb acquisition involves the plurality of interpretations allowed by the extralinguistic scene alone (e.g., Quine, 1960). For example, the child who hears "bring" while holding a doll and walking towards her parent could conjecture that "bring" refers to *bringing*, *carrying*, *walking*, *holding*, *playing*, or many other actions or relations. Syntactic frames may serve to constrain the set of possible interpretations for the verb learner because many frames have more or less specific semantic implications for the verbs that appear in them. In the above example, if the child heard "Are you bringing me the doll?", the double object dative frame *V NP NP* suggests that *bring* is a verb of transfer and thus eliminates the meanings of *hold*, *carry*, *walk*, and *play* from contention (for more discussion, see Bowerman, 1982a, 1982b; Jackendoff, 1983, 1990; Fisher, Gleitman, & Gleitman, 1991; Levin, 1985; Levin & Rappaport, 1991; Naigles, Fowler, & Helm, 1992; Naigles, Gleitman, & Gleitman, 1993; Pinker, 1989). In sum, a major tenet of Syntactic Bootstrapping is that children will exploit correlations between syntax and verb semantics in verb acquisition.

Experimental validations of the Syntactic Bootstrapping hypothesis have focused on the ability of children to distinguish the meaning of a novel verb on the basis of a single frame in which the verb is presented, in combination with a visual scene (Naigles, 1990; Naigles & Kako, 1993; Fisher, Hall, Rakowitz, & Gleitman, 1984). The demonstrations so far have primarily exploited the correlation between transitive syntactic frames and causative verb meanings. For example, Naigles and Kako presented young 2-year-olds with side-by-side videos, each depicting two animals engaged simultaneously in two distinct types of actions. One of the actions was noncausative and Synchronous; e.g., a duck and a rabbit were making arm circles in perfect synchrony. The other action was either Causative (e.g., the duck forces the rabbit to bend over) or involved Contact without causation (e.g., the duck contacts the rabbit's head with splayed fingers). These videos were paired with audios in which a novel verb was presented, either in isolation (e.g., "Blicking!"), in a transitive frame (e.g., "The duck is blicking the bunny!"), or in an intransitive frame with conjoined subject (e.g., "The duck and the bunny are blicking!"). The two actions were subsequently separated, such that one video screen showed the Causative or Contact action while the other showed the Synchronous action. These separated-action videos were paired with the audio "Find blicking"; the dependent measure was the children's looking preferences to each action. Naigles and Kako found that, relative to their peers' performance with the verb in isolation, the children

who had heard the verb in the transitive frame looked longer at the Causative or Contact actions when asked to "find blicking". That is, the transitive frame enabled the children to select the Causative or Contact action over the Synchronous one as the referent for the novel verb.

Fisher et al. (1994) further showed that preschoolers will use the syntactic frame in which a novel verb is placed to distinguish between aspects of a single action or event. In that study, one video depicted a rabbit spooning food into an elephant. If the frame was transitive (e.g., "The bunny is blicking the elephant"), then the children translated the novel verb as *feeding*; if the frame was intransitive (e.g., "The elephant is blicking"), the children translated the novel verb as *eating*. In sum, recent research supports Syntactic Bootstrapping as a procedure for verb learning, as several studies have demonstrated that young children will use the syntactic frame in which a novel verb appears as a clue to that verb's meaning.

While this evidence is quite compelling, the Syntactic Bootstrapping procedure as just presented is incomplete, both theoretically and empirically (e.g., Landau & Gleitman, 1985; Gleitman, 1990; Naigles & Kako, 1993; Pinker, 1994). One theoretical issue concerns the fact that there does not exist a single, distinct, semantically revealing syntactic frame for every distinct verb in a language. Instead, the thousands of verbs in any one language are limited to appearing in a relatively small number of distinct syntactic frames (probably less than 50; see Fisher et al., 1991; Naigles & Hoff-Ginsberg, 1995, for more discussion). The implication of this fact for Syntactic Bootstrapping is that a single syntactic frame may *not* be sufficient to distinguish among the verb meaning possibilities allowed by the extralinguistic scene. For example, recall the scene described earlier in which the child was holding the doll and walking towards her parent. If the parent had used the prepositional dative construction "Are you bringing the doll to me?" instead of the double object dative "Are you bringing me the doll?" then either *bring* or *carry* would be good candidates for the meaning of "bring", because the prepositional dative frame is not limited to verbs of transfer (Pinker, 1989). Similarly, hearing *think* followed by a sentence complement (e.g., "I think THAT SUE IS WEARING A GREEN DRESS") suggests that it is a mental state verb, and so distinguished from such motion verbs as *go* and *come*. However, the sentence complement construction would not be sufficient to distinguish the meaning of *think* from that of *know*, as "know" can also appear with the same complement frame, and *know* might also be a consonant with the extra-linguistic situation. Further, while the link between transitive frames and causative verbs seems quite strong (Bowerman, 1982, 1983; Naigles et al., 1993), noncausative verbs can also appear in transitive frames (e.g., *I SEE the house*, *I WANT some cookies*, *I TOUCHED the ball*). How would a child who hears only "You're X-ing the cat" while stroking a responsive cat determine whether the verb referred specifically to the action performed on the cat (*stroke*), or to the resulting action the cat produced (*cataltituding* (Hall, 1984))?

This theoretical issue, concerning the generality of verb meaning imparted by a single syntactic frame, manifested itself with particular empirical force in the third experiment of Naigles and Kako (1993), the focus of which concerned distinguishing specific Contact and Causative actions. In that experiment, the two actions that the two characters were initially engaged in were Causative (the duck making the frog bend over) and Contact-without-causation (the duck contacting the frog's head with splayed fingers). Children first saw these two actions together paired with a single novel verb either in isolation (e.g., "Blocking!") or in a transitive frame (e.g., "The duck is blicking the frog!"), and then saw the two actions separated. When asked for the referent of the novel verb, children in the isolation condition strongly preferred the Causative action, while children in the transitive condition showed a similar, although non-significant preference for the same action. Thus, the transitive frame alone did not succeed in drawing the children's attention to the Contact action, as it had in the earlier experiment when the contrasting action was noncausative and Synchronous. And perhaps it shouldn't have, given that both contact and causative actions can be described by transitive verbs. These results led Naigles and Kako to the same question asked above: given (1) that children are probably faced with numerous scenes that allow both contact and causative interpretations (e.g., stroking a responsive cat), (2) that their preference without syntactic information appears to be 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 are they to acquire verbs involving contact without causation? This question, motivated by both a theoretical issue and an empirical result, was the instigation for the current research.

One very simple answer is that contact verbs are acquired after the confounding causative verbs have been learned, perhaps by means of a strategy of Contrast. For example, a child faced with an ambiguous scene involving contact and causation may realize that she already has a word for the causative aspect of this action, and so conjectures that a new verb must refer to some other aspect, such as contact (see Clark, 1987). Data from children's production and comprehension of verbs indicates, though, that verbs involving contact, (e.g., *touch*, *pat*, *rub*) are among the earliest learned, and show no delay in development relative to causative verbs (Goldin-Meadow, Seligman, & Gelman, 1975; Tomasello, 1992). Clearly, another approach is warranted.

The solution that is the focus of this paper is the Multiple Frames hypothesis. As Landau and Gleitman (1985) hypothesized in their original proposal of Syntactic Bootstrapping, children may exploit the *multiple* frames in which a verb can be presented in their acquisition of that verb's meaning. The idea is that while a single frame may be insufficient to accurately constrain a verb's meaning given a particular extra-linguistic context, multiple frames may provide enough information to do so. For example, in Landau and Gleitman's investigation of the acquisition of *look*

by a blind child named Kelli, they suggested that it was the set of syntactic frames in which *look* appeared that enabled Kelli to establish this as a verb of perception. To Kelli, *look* was a verb of haptic perception rather than visual perception (i.e., it meant *explore with the hands* rather than *explore with the eyes*). Part of this meaning could be attributed to the fact that Kelli's mother only used *look* in reference to nearby things and not far-away things. However, she also used such verbs as *have*, *give*, *play*, and *put* in reference to nearby things; that is, the spatial context did not distinguish *look* from the nonperceptual verbs. Interestingly, the syntactic frames in which *look* was used did provide this distinction: *look* differed from *have* in allowing PPs (compare *Kelli looked AT THE BALL* with **Kelli has AT THE BALL*), it differed from *put* in not allowing frames containing three NPs (compare *KELLI put THE BALL in THE BASKET* with **KELLI looked THE BALL in THE BASKET*), and it differed from *go* in allowing sentence complements (compare *Look HOW TO DO IT* with **Go HOW TO DO IT*). Thus, Landau and Gleitman conjectured that it was the *conjunction* of these frames in Kelli's input, plus the accompanying spatial contexts, which enabled Kelli to establish *explore haptically* as the meaning of *look*.

More direct evidence for the usefulness of multiple syntactic frames in verb learning comes from a study by Naigles and Hoff-Ginsberg (1993, 1995). They investigated the syntactic uses of 25 verbs in the speech of 57 mothers to their children when those children were just beginning to combine words (Time 1), and tabulated the verb production of the children 10 weeks later (Time 2). Forty-six different post-verb frames emerged from a syntactic parsing of the mothers' speech, and many of these frames helped to distinguish among semantic subclasses within the set of 25 verbs (e.g., sentence complements followed internal state verbs almost four times more often than they followed non-internal state verbs). With regard to the Multiple Frames Hypothesis, there were two major findings: first, over three-quarters of the targeted verbs in maternal speech at Time 1 were used multiple times within a single morning, and over three-quarters of these multiple-use verbs were used in multiple frames over the same time period. Thus, the input that mothers provided verb-learning children illustrated the diversity of frames in which many verbs appear. Second, the greater the number of different syntactic frames in which a given verb was presented at Time 1, the greater the frequency with which children used that verb at Time 2, controlling for maternal frequency of verb use at Time 1. In sum, preliminary analyses suggest that the presentation of verbs in multiple syntactic frames in the maternal input significantly facilitated the children's subsequent verb use.

The purpose of this paper is to move the investigation of the use of multiple frames in syntactic bootstrapping from the naturalistic to the experimental realm. The focus of the research was on how children might use multiple frames to distinguish actions involving direct physical causation,

which I am calling Causative actions, from actions involving repeated physical contact without causation, which I am calling Contact actions. The hypothesis that will be investigated in this study is that two different types of *transitive-intransitive frame alternations* can allow children to distinguish some kinds of Causative actions from some kinds of Contact ones. These alternations present different ways in which transitive and intransitive frames containing the same verb relate to each other; these relations are described in more detail below.

1.1. The alternations and their semantic correlates

As described by Levin (1993), in the Causative (CS) alternation of transitive and intransitive frames the object of the transitive frame and the subject of the intransitive frame are the same:

(1) The girl dropped the ball.
(2) The ball dropped.

Note that, while transitive and intransitive sentences differ on the component of causality (in (1) the girl is the causer of the ball's dropping, while (2) does not express an external causer), they can refer to the same event. Thus, if (1) occurred then (2) also occurred. Not all verbs that involve causation can participate in this alternation (e.g., *kill*); however, all verbs that participate seem to involve some notion of causation. According to Dixon (1991), the types of verbs that can participate in the CS alternation are those in which a "Causer could be responsible for the event happening in a natural manner" (p. 295). Levin (1993) lists several subclasses of English verbs that participate in causative alternations, including change-of-state verbs such as *break*, *fold*, and *crumble*, change-of-position verbs such as *roll*, *bounce*, and *drop*, induced-action verbs such as *run*, *jump*, and *race* (e.g., *The rider jumped the horse over the fence/The horse jumped*), and verbs of spatial configuration such as *hang*, *lean*, and *sit*.

The CS alternation is here contrasted with what I am calling the Omitted Object (OO) alternation, in which the subject of the transitive frame and the subject of the intransitive frame are the same:

(3) The cat was scratching the door.
(4) The cat was scratching.

As with the two sentences of the CS alternation, both (3) and (4) can refer to the same event. The semantic correlate(s) to this alternation are not as transparent or clearcut as those to the CS alternation. In one discussion of the types of verbs that participate in OO alternations, Dixon (1991) focuses on two factors: (1) the activity of the verb, which typically involves a specific target or product that is pursued for a longish period of time, and

which thus appears in the progressive (e.g., *saw*, *knit*, *rake*); (2) the ability of the verb's omitted object to be inferred by the listener from earlier discourse, the current situation, or common ground (e.g, *see*, *help*, *choose*). Two of Levin's (1993) subclasses of English verbs that can omit their objects also seem to reflect these factors: activity verbs whose object is unspecified but assumed to be present to some degree (e.g., *bake*, *chop*, *sketch*), and verbs that describe a characteristic property of the agent of the action (e.g., *bite*, *pinch*, *scratch*); the omitted object is interpreted as something like "people". Notice that while the classes of Contact verbs are not co-extensive with the classes of verbs that undergo OO alternations, many of the verbs that do participate appear to involve physical contact of a simple or complex sort (e.g., *bite*, *kick*, *scratch*, *prick*, *dust*, *mend*, *sketch*, *saw*, *rake*, *polish*, *knit*). The hypothesis tested in this study is that the different alternation patterns of the transitive and intransitive frames (OO and CS) help young verb learners select actions involving contact without causation from those involving causation.

1.2. The semantic correlates and the specific actions under investigation

The method used was the preferential-looking paradigm developed by Golinkoff and Hirsh-Pasek (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1991); this method has been successful in revealing effects of syntax on children's novel verb learning (Naigles, 1990; Naigles & Kako, 1993). The novel actions presented to the children were the same as those in the third experiment of Naigles and Kako (1993; see Table 1). In this section, I discuss how these actions manifest the semantic correlates described above.

In each of the four Causative actions listed in Table 1, Character 1 grasps a body part of Character 2, and by exerting force causes Character 2 to change its bodily position or spatial configuration. Following Levin's (1993) categorization, these actions resemble those of the verbs in both the induced-action subclass and the spatial configuration subclass. Like the induced-action verbs *jump* and *leap*, these caused actions are accompanied by Character 1, and Character 2 (the causee) is both animate and volitional. Like the spatial configuration verbs *lean* and *sit*, these actions could be internally controlled (i.e., in *sebbing* the duck could be bending over by itself), but are in these scenes caused by another. In their manifestations of accompanied causation and change of position/spatial configuration, these actions are akin to those whose verbs are among the earliest overgeneralized on the Causative alternation by child learners (e.g., "Daddy go me around" [=spin, turn], produced by Christy at 2;8, from Bowerman, 1983, p. 14). The proposal here is that children who hear the novel verbs in the CS alternation will chose these novel causative actions as the referents of the verbs.

In each of the four Contact actions listed in Table 1, Character 1 performs

Table 1
Actions associated with each novel verb[a]

VERB	Contact	Causative
SEB	Frog contacts duck's head with (frog's) fingers	Frog holds duck's shoulder, forces duck into bending position
TIG	Duck rubs frog's leg with duck's foot	Duck hold frog's arms, makes arms flex
LORP	Frog moves foot up and down, stepping on duck's foot	Frog holds onto duck's shoulder, making duck's head and shoulders tilt
PIM	Duck taps frog's stomach with duck's wing	Duck holds frog's elbow makes frog "bonk" frog's head with hand

[a] Adapted from "First contact in verb learning: Defining a role for syntax" by L. Naigles and E. Kako, 1993, *Child Development*, *64*, 1665–1687. Copyright held by University of Chicago Press. Reprinted by permission.

an action with its hand, wing, or foot such that this body part comes into repeated contact with some part of Character 2. These actions may not be those of prototypical transitive activity verbs (which usually involve some inanimate instrument); however, they do encompass activities in a broader sense (cf. Vendler, 1967). That is, these actions are not discrete, nor do they involve a particular result or accomplishment. Moreover, like the verbs in Levin's "characteristic property of agent" subclass, these novel actions may be considered to be idiosyncratic to the particular Character 1 engaged in them (e.g., just as dogs bite, ducks tap with their wings). Finally, because these actions are being performed in full view, the role of Character 2 is easily inferred from context, so that it may be considered permissible to omit mention of Character 2. In sum, the proposal is that children who hear the novel verbs in the OO alternation will choose these novel contact actions as the referents of the verbs.

2. Method

2.1. Subjects

One-hundred and twenty children completed the experiment. One hundred and ninety-seven children were tested, but 77 were excluded due to the following: inability of 2 out of 5 coders to agree on the child's fixations (21), refusal of the child to watch the videos for more than half of the trials (13), parent watching video along with the child (13), fixation bias of more than 75% to a single side, averaged across test trials (19), and experimenter error

Table 2
Detailed subject information

Condition	N	Words[a]	Verbs[b]	Age[c]	Multi-word utterances[d]	Parental education[e]
Causative (CS)	32	297.13 (74.69)	35.06 (9.51)	28.02 (0.75)	84.38	3.3
Omitted Object (OO)	32	323.06 (31.98)	37.39 (4.54)	28.60 (0.60)	95.00	4.3
Intransitive Only (IO)	24	288.04 (73.86)	32.46 (10.94)	28.15 (0.68)	91.6	2.9
Frameless	32	304.19 (49.69)	35.59 (5.99)	27.92 (0.74)	89.00	3.7

[a] Mean number of words (*SD*) used spontaneously by parental report, out of 354 on Rescorla et al.'s (1993) checklist.
[b] Mean number of verbs (*SD*) used spontaneously by parental report, out of 40 on Rescorla et al.'s (1993) checklist.
[c] Mean age (*SD*) 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.

(11). The remaining children, half boys and half girls, ranged in age from 26 to 30 months. Demographic and linguistic information concerning the subjects by audio condition is listed in Table 2; the excluded subjects did not differ from these on any of the listed variables. 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.

2.2. Apparatus

The apparatus, stimuli, and procedure were very similar to that in Naigles and Kako (1993); the following description is drawn in large part from that paper.

The basic set-up is shown in Fig. 1. The child was seated on the parent's (usually mother's) lap and observed two different, simultaneously presented video events on two side-by-side video monitors. Between the two monitors, which were separated by 12 inches, 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

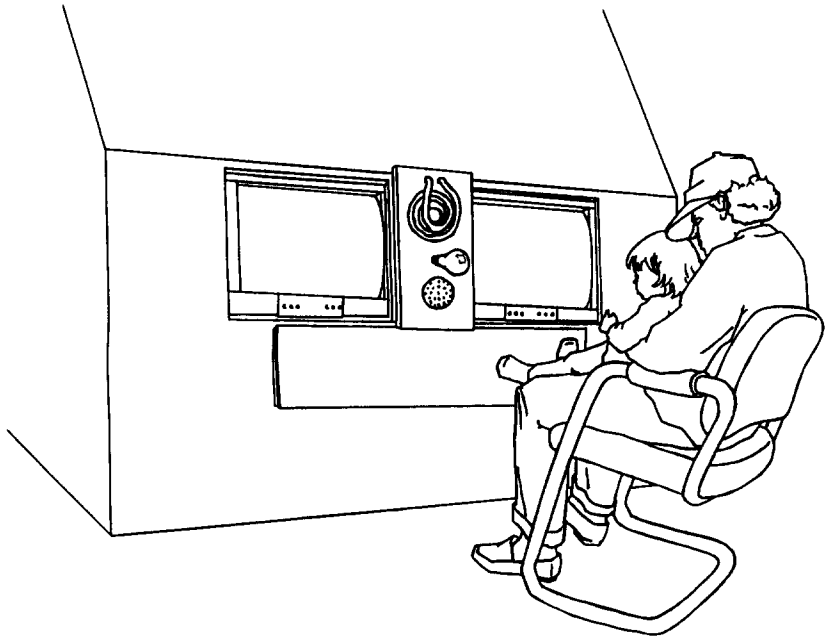

Fig. 1. Experimental set-up of the preferential looking paradigm. From "First contact in verb acquisition: Defining a role for syntax" by L. Naigles and E. Kako, 1993, *Child Development*, *64*, 1665–1687. Copyright held by University of Chicago Press. Reprinted by permission.

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 beige curtains. Inter-trial intervals were created by inserting three seconds of black tape. Each tape was prepared as one of 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.

2.3. Stimuli and design

The stimulus videotapes were the same ones used by Naigles and Kako (1993), Experiment 3. A layout of the videotapes and the audio with the

verb in the OO alternation is presented in Table 3. 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 was heard during 3 seconds of black tape before each presentation of a pair, as well as during each 6-second scene. First, an introductory passage familiarized the subjects with the situation and the characters (trials 1–9), and then the novel actions and verb were presented. The crucial sequence was as follows. In trial 10, one screen presented a Multiple Action Scene: two actions going on alternately, 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 frog into an odd bending position), and the other was noncausative and involved Contact (one character contacts the other in a novel way; e.g., the duck contacts the frog's head with splayed fingers). Both actions were iterative, and occurred the same number of times per trial. The accompanying audio presented a novel verb (e.g., *seb*) three times: once when both screens were blank and then twice when the pictures were presented. Depending on the audio condition (see below), the novel verb was presented in isolation or in three different patterns of syntactic frames. This simultaneous presentation of two novel actions and one novel verb (repeated twice: once on the other screen and once on both screens, trials 11 and 12) comprised the "teaching" phase of the experiment.

Next, the two actions were separated into Single Action Scenes (trial 13 in Table 3): one screen showed *only* the Causative action of the duck forcing the frog into the bending position, while the other screen, displayed at the same time, showed *only* the Contact action of the duck contacting the frog's head with splayed fingers. This familiarization 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 14 and 15) were presented. The Single Action Scenes appeared again, paired with the test audios, "Where's sebbing?" or "Find sebbing now!" These trials tested what the children learned during the teaching phase.

This pattern of teaching and testing was presented across four different pairs for Causative and Contact actions with four different nonsense verbs for each of the subjects (see Table 1). There was thus a total of eight test trials per subject. 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 described by a single lexical item in English. Each verb was paired with only one set of actions (i.e., there was no counterbalancing across verbs). 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.

There were four between-subjects audio conditions, as shown in Table 4.

Table 3
Structure of the stimulus videotapes. Character familiarization sequence and the verb "seb"

Tape 1	Audio	Tape 2
BLACK	There's the duck!	BLACK
1 Duck waves	See the duck?	Black
BLACK	There's the frog!	BLACK
2 Black	See the frog?	Frog waves
BLACK	There's the duck!	BLACK
3 Duck waves	See the duck?	Black
BLACk	There's the frog!	BLACK
4 Black	See the frog?	Frog waves
BLACK	Oh, look now!	BLACK
5 Duck waves	Here they are!	Frog waves
BLACK	Where's the duck?	BLACK
6 Duck waves	Find the duck!	Frog waves
BLACK	Where's the frog?	BLACK
7 Duck waves	See the frog?	Frog waves
BLACK	Where's the frog?	BLACK
8 Duck waves	Find the frog!	Frog waves
BLACK	Where's the duck?	BLACK
9 Duck waves	Find the duck!	Frog waves
BLACK	The duck is sebbing the frog!	BLACK
10 The duck is grasping the frog's shoulder with one hand, forcing it to bend over, and is contacting the frog's head with its other hand	The duck is sebbing the frog! The duck is sebbing!	Black
BLACK	The duck is sebbing the frog!	BLACK
11 Black	The duck is sebbing the frog! The duck is sebbing!	Duck is grasping frog's shoulder with one hand, forcing it to bend over, and is contacting the frog's head with its other hand
BLACK	The duck is sebbing the frog!	BLACK
12 Duck is grasping frog's shoulder with one hand, forcing it to bend over, and is contacting the frog's head with its other hand	The duck is sebbing the frog! The duck is sebbing!	Duck is grasping frog's shoulder with one hand, forcing it to bend over and is contacting the frog's head with its other hand
BLACK	Oh! They're different now!	BLACK
13 Duck is forcing frog to bend over	See? They're different now!	Duck is contacting frog's head
BLACK	Where's sebbing now?	BLACK
14 Duck is forcing frog to bend over	Find sebbing now!	Duck is contacting frog's head
BLACK	See sebbing now?	BLACK
15 Duck is forcing frog to bend over	Look at sebbing!	Duck is contacting frog's head

Table 4
Teaching trial audios by condition

Frameless	Hey, sebbing! See, sebbing! Look, sebbing!
Causative	The duck is sebbing the frog! The duck is sebbing the frog! The frog is sebbing!
Omitted Object	The duck is sebbing the frog! The duck is sebbing the frog! The duck is sebbing!
Intransitive Only	The duck is sebbing! The duck is sebbing! The duck is sebbing!

Two of these have already been introduced: the causative (CS) condition, in which the novel verb was presented during the Multiple Action Scenes in transitive and intransitive frames in the Causative alternation, and the Omitted Object (OO) condition, in which the novel verb was presented during the Multiple Action Scenes in transitive and intransitive frames in the Omitted Object alternation. When both screens were black, both conditions presented the verb in the identical transitive frame ("Look, the duck is sebbing the frog!"). The prediction is that the children in the CS condition will prefer the Causative action as the referent of the verb, while the children in the OO condition will prefer the Contact action. The other two conditions served as controls: in the Frameless condition (FF), the verb was presented throughout the teaching trials without a syntactic frame (e.g., "Look! sebbing!"). This condition was designed to reveal the children's baseline (i.e., presyntactic) preference for labelling each action. In the Intransitive Only (IO) audio condition, the verb was presented throughout the teaching trials in the same intransitive frame as had appeared in the OO condition ("Look! The duck is sebbing!"). This condition thus tested whether significant effects in the OO condition could in fact be attributed to the combined transitive and intransitive frames, or just to the presence of the intransitive frame itself.

2.4. Procedure

Subjects and their parents were first interviewed in the playroom by the experimenter for an average of 15 minutes. At this time, the parent filled out a brief language questionnaire (Rescorla, Hadicke-Wiley, & Escarce, 1993), 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. After the tapes had

ended, the experimenter returned to the testing room and replayed the videotapes for the parent.

2.5. Coding and reliability

Total visual fixation time to the matching and non-matching screens (measured in hundredths of a second) was coded from the videotapes.[1] An observer, who could not hear the stimulus audio and so could not know here 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 recorded and tabulated by the computer. Visual fixation was calculated from the point at which the infant looked at the central light, when the screens were blank, for more than 0.3 seconds. Trials where the child did not look at the center light for a minimum of 0.3 seconds were excluded. In addition, trials where the child had not looked at either screen (once the pictures appeared) for a minimum of 0.3 seconds were excluded. Using these criteria, an average of 5% of the trials were excluded from further analysis across all conditions of the experiment (range 3–10% of trials).

The method used for determining coder reliability was the same as that used by Naigles and Kako (1993) and Naigles and Gelman (1995). This is a conservative procedure in which the video of each child was initially viewed by two different coders. If the coders did not agree on visual fixation for each pair of test trials within 0.5 seconds (approximately the duration of a saccade), a third (and sometimes fourth and fifth) coder viewed the video. Such multiple viewing was necessary because the children often had restless periods while watching the videos, when they would scan both screens with great rapidity and so be difficult to follow. When two coders would found to be in agreement on all pairs of test trials for a given subject, the chronologically earlier of these was selected as the one whose coding was included in the analysis ("Chosen Coder"; the other was designated as the "Pair"). Across all conditions, an average of 3.36 coders (range 2.94–3.71) was necessary to reach agreement on each subject. The correlations between the codings of the Chosen Coder and those of the Pair ranged from .90 and .99; thus, we can be confident that the codings of the Chosen Coder are reliable. In sum, corroboration of the Chosen Coder's judgements was required for every subject, instead of relying on a reliability calculation based on a subset of subjects.

[1] Some studies using the preferential-looking paradigm also report the number of trials that the children looked first at the matching screen (i.e., Naigles and Gelman, 1995). This measure will not be reported for the present study because previous research with single-frame Syntactic Bootstrapping (Naigles & Kako, 1993) and preliminary analyses of the present data indicated that the First Look measure is considerably more variable than Visual Fixation, and is not sensitive to the effects of syntax.

3. Results

The results will be reported in two parts. The Within-audio analyses investigated the screen preferences of the children in each audio condition separately; for example, did the CS audio lead children to prefer the Causative action? The Between-audio analyses compared the screen preferences across audio conditions; for example, did the OO audio elicit significantly different preferences from the CS audio?

3.1. Within-audio analyses

The results from the within-audio analyses are presented in Table 5. The screen preferences were assessed via separate two-way ANOVAs (Gender (Boys vs. Girls) × Screen (Causative vs. Contact)) for each audio condition with two-tailed tests of significance. The top rows show the data from the Frameless condition. When the actions were paired with a novel verb in isolation, the children looked longer at the Causative actions than the

Table 5
Action preferences (in seconds) by gender, all conditions Screen

Condition	Gender	Contact	Causative
Frameless	Girls	1.97	2.84
		(0.51)	(0.74)
	Boys	1.95	2.63
		(0.50)	(0.53)
	Total	1.96	2.74
		(0.51)	(0.65)
Causative	Girls	2.17	2.72
		(0.49)	(0.55)
	Boys	1.91	2.43
		(0.65)	(0.65)
	Total	2.04	2.57
		(0.62)	(0.59)
Omitted Object	Girls	2.39	2.44
		(0.59)	(0.48)
	Boys	1.90	2.76
		(0.61)	(0.50)
	Total	2.14	2.60
		(0.65)	(0.52)
Intransitive Only	Girls	2.02	2.79
		(0.44)	(0.51)
	Boys	1.54	2.44
		(0.71)	(0.51)
	Total	1.78	2.62
		(0.54)	(0.62)

Contact ones; the ANOVA yielded a significant effect of screen ($F(1, 30) = 20.41$, $p < .01$) and no other significant effects. The second set of rows shows the data from the CS condition. When the actions were paired with transitive and intransitive frames in the CS alternation pattern, the children looked longer at the Causative actions than the Contact ones; the ANOVA yielded a significant effect of screen ($F(1, 30) = 8.95$, $p < .01$), a significant effect of gender ($F(1, 30) = 4.95$, $p < .05$), and no significant interaction. The source of the gender effect was that the girls looked longer at the screens overall than the boys did (M (girls) = 4.88 second, $SD = .47$, M (boys) = 4.34 seconds, $SD = .79$).

The third set of rows shows the data from the OO condition. When the actions were paired with transitive and intransitive frames in the OO alternation pattern, these children looked somewhat longer at the Causative actions than the Contact ones; the ANOVA yielded a significant effect of screen ($F(1, 30) = 6.62$, $p < .05$), and a significant interaction of gender and screen ($F(1, 30) = 5.35$, $p < .05$). Post hoc Tukey's tests revealed that the boys showed a significant preference for the Causative actions ($p < .05$) but the girls did not. Instead, the girls' looking times to the screens were almost identical. These was no main effect of gender.

The fourth set of rows shows the data from the IO condition. When the actions were paired only with intransitive frames, these children also looked longer at the Causative actions than the Contact ones; the ANOVA yielded a significant effect of screen ($F(1, 22) = 17.53$, $p < .01$), a significant effect of gender ($F(1, 22) = 10.26$, $p < .01$), and no other significant effects. As in the CS condition, the source of the gender effect was that the girls looked longer at the screens overall than the boys did (M (girls) = 4.90 seconds, $SD = .42$, M (boys) = 3.98 seconds, $SD = .86$).

In sum, in each audio condition children showed a significant preference for the Causative actions over the Contact ones. The significant Causative preference in the FF condition replicated that found in Naigles and Kako (1993) when there were two presentations of the verb during the teaching trials instead of the three in this study. The significant Causative preference in the CS condition was as predicted by the Multiple Frames hypothesis; however, the boys' significant Causative preference in the OO condition was *not* as predicted by the hypothesis of Multiple Frames. The girls in the OO condition were the only children not to show a significant Causative preference; they were essentially equivocal in their screen preferences. The significant Causative preference in the IO condition was perhaps the most surprising, because while the audio is directing the children towards the duck's motions (for *seb* and *pim*; for *tig* and *lorp* it is directing them towards the frog's actions), they evidently preferred to watch its Causative actions rather than its Contact activities. Finally, in all four conditions the girls tended to pay more attention to the screens overall than the boys did, and in two of the four conditions (CS and IO), this gender difference reached significance.

3.2. Between-audio analyses

The critical tests for the Multiple Frames hypothesis are the between-audio analyses, because these can test the extent to which the preferences elicited by the multiple frame audios (CS and OO) were significantly different from each other, as had been predicted. Moreover, the between-audio analyses test whether the preferences elicited by the CS and OO audios were significantly different from those elicited by the baseline audio (FF), and whether the preferences elicited by the OO audio were significantly different from those elicited by the Intransitive Only (IO) audio.

To carry out these analyses a three-factor ANOVA was set up, with Audio condition (FF vs. CS vs. OO vs. IO) and Gender as between-subjects factors and Screen (Causative vs. Contact) as the within-subjects factor. Four planned comparisons were of interest: between the OO and CS conditions, between the OO and FF conditions, between the OO an IO conditions, and between the CS and FF conditions. Because gender effects and interactions were found in the within-audio analyses, each planned comparison was carried out separately for the boys and girls. Because the overall preference in all conditions was for the Causative action, the tests of significance were one-tailed: did any audio condition show movement *away* from this overall preference?

The planned comparison between the OO and CS conditions with the girls' preferences revealed a significance difference ($F(1, 112) = 3.19$, $p < .05$). As shown in the middle panels of Fig. 2, the girls in the CS condition preferred the Causative actions while the girls in the OO condition showed essentially no preference. In contrast, the planned comparison between the OO and CS conditions with the boys' preferences revealed no significance difference ($F < 1$). As shown in the middle panels of Fig. 3, the boys in both conditions preferred the Causative actions.

The planned comparison between the OO and FF conditions with the girls' preferences revealed a significance difference ($F(1, 112) = 5.58$, $p < .01$). As shown in the far left panel of Fig. 2, the girls' Causative preference in the FF condition was significantly different from their peers' lack of preference in the OO condition. In contrast, the planned comparison between the OO and FF conditions with the boys' preferences revealed no significance difference ($F < 1$). As shown in the far left panel of Fig. 3, the boys in both conditions preferred the causative actions.

The planned comparison between the OO and IO conditions with the girls' preferences revealed a significance difference ($F(1, 112) = 2.81$, $p < .05$). As shown in the far right panel of Fig. 2, the girls' Causative preference in the IO condition was significantly different from their peers' lack of preference in the OO condition. In contrast, the planned comparison between the OO and IO conditions with the boys' preferences revealed no significance difference ($F < 1$). As shown in the far right panel of Fig. 3, the boys in both conditions preferred the Causative actions.

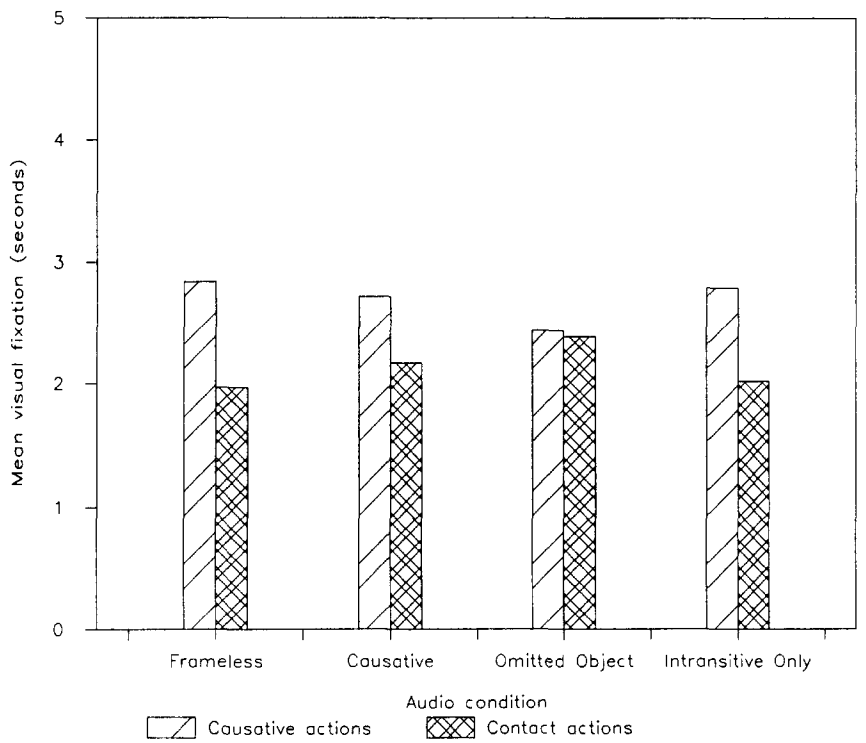

Fig. 2. Mean visual fixation of the Contact and Causative actions for each audio condition, girls only.

The planned comparison between the CS and FF conditions with the girls' preferences revealed no significance difference ($F < 1$). As shown in the left two panels of Fig. 2, the girls in both conditions preferred the Causative actions to more or less the same extent. Similarly, the planned comparison between the CS and FF conditions with the boys' preferences revealed no significance difference ($F < 1$). As shown in the left two panels of Fig. 3, the boys in both conditions preferred the Causative actions to more or less the same extent.

In sum, these analyses revealed that the girls in the OO condition showed significantly different action preferences from their peers in either the CS, IO, and FF conditions; unlike these girls, those who heard the OO audio did *not* prefer the Causative actions as the referent of the novel verb. Thus, the multiple-framed audio in the OO condition seems to have made a difference in the girls' looking preferences. Notice, though, that both the girls' and boys' preferences in the CS condition were statistically identical to those of their peers in the FF condition: this multiple-framed audio did not succeed in changing their baseline preference. Finally, the effect of multiple frames

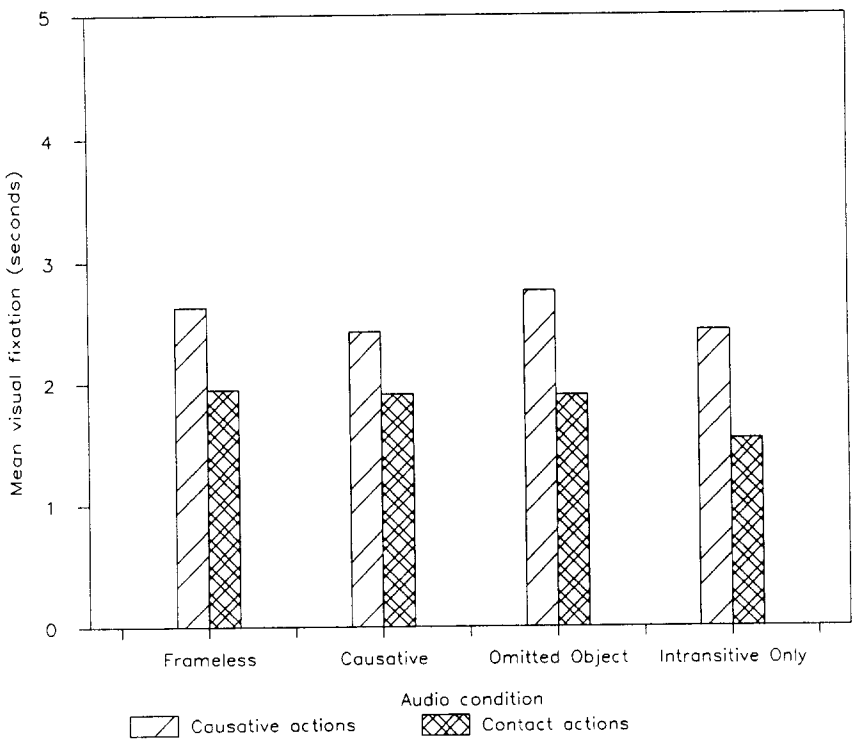

Fig. 3. Mean visual fixation of the Contact and Causative actions for each audio condition, boys only.

held only for the girls: these analyses demonstrated that the boys in all conditions showed statistically similar preferences for the Causative actions.

This significance shift in action preference (away from the Causative action) that was exhibited by the girls who heard a novel verb in the OO alternation was obtained using data pooled across the four verb–action pairs. This next analysis investigates the extent to which these four items yielded consistent preferences in the same direction. Fig. 4 presents the degree of Causative action preference (the amount of looking to the Causative screen minus the amount of looking to the Contact screen) by item (in order of presentation) for the girls in the FF, CS, OO, and IO audio conditions. As the figure shows, all four items in the FF condition yielded strong Causative action preferences, as did three of the four items in the IO condition. More importantly, for the first three items, the OO audio consistently yielded the *least* preference for the Causative action. In fact, the verb *lorp* in the OO audio actually elicited a small preference for the Contact action. In comparison, the girls' visual fixations with these three items in the CS condition were directed more at the Causative action, on the

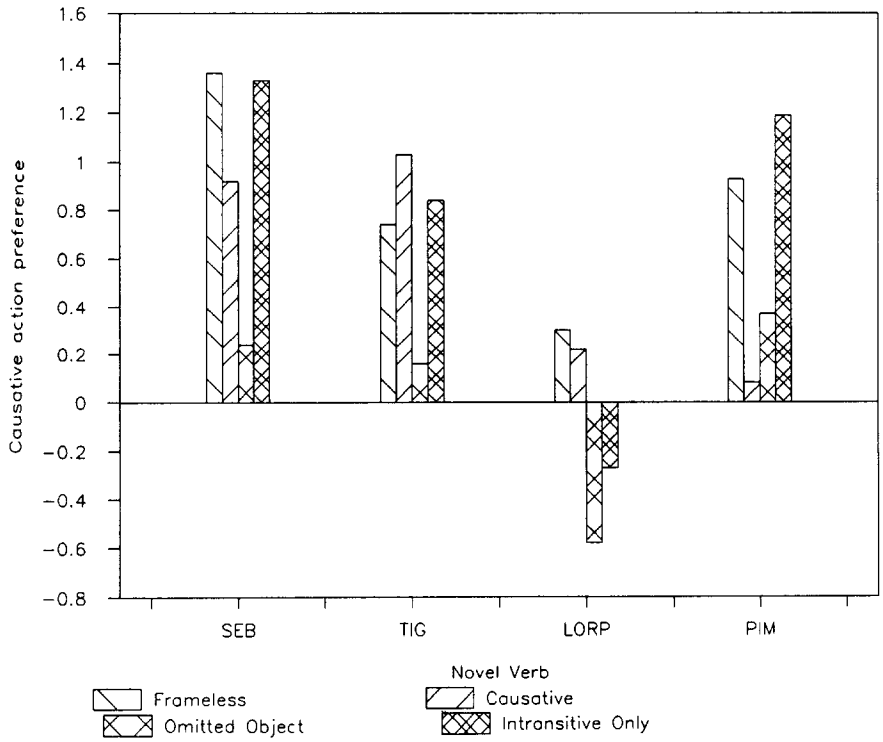

Fig. 4. Mean Causative action preference (visual fixation to Causative screen minus fixation to Contact screen) for each novel verb and audio condition, girls only.

order of 0.68 to 0.87 seconds. Only the fourth item (associated with the verb *pim*) failed to follow this pattern; the girls in both the CS and OO conditions showed no significant preferences for either action. A two-factor ANOVA (Audio (FF vs. CS vs. OO) × Screen (Causative vs. Contact) that was performed by items rather than by subjects exhibited the consequences of this failure, as the planned comparison between the FF and OO conditions and between the IO and OO conditions only approached significance ($F(1, 24) = 3.46$, $p < .10$ and $F(1, 24) = 3.26$, $p < .10$, respectively), and the planned comparison between the CS and OO conditions was not significant ($F(1, 24) = 1.93$, $p > .10$). Thus, the girls in the OO condition showed considerable but not complete consistency across items in their lack of preference for the Causative action. While the results would have been stronger had all four items yielded the same effect, the item difference that have been observed invite interesting consideration in terms of their actions' suitability for the CS and OO alternations; this will be addressed in the Discussion.

3.3. Analyses of the preferences of individual subjects

The final analyses compared the action preferences of the individual subjects in the CS and OO conditions. Information about subject preferences can reveal the extent to which individual children in a particular audio condition followed the pattern set by the group as a whole. The first analysis tabulated the number of children who looked longer overall at the Causative actions and found little difference between the conditions. In the CS condition, 11 of the 16 girls and 11 of the 16 boys looked longer at the Causative actions than the Contact ones. In the OO condition, 9 of the 16 girls and 14 of the 16 boys showed the same preference for the Causative actions.

A more in-depth analysis compared the number of children who preferred the Causative actions at the two levels of preference: a large preference (a difference of more than 0.65 seconds between screens) that could indicate more certainty about which action is the referent for the verb, and a small preference (a difference of less than 0.65 seconds between screens) that could indicate less certainty. A difference of 0.65 seconds was chosen as the cutoff because it was the median preference for both girls and boys in both audio conditions. At the level of small preferences (less than 0.65 seconds), 60% of boys and 63% of girls in the CS condition preferred the Causative screen, and 78% of both boys and girls in the OO condition preferred the Causative screen. Thus, there appeared to be no differences across gender or audio condition in the preferences of the children who had small looking time differences. At the level of large preferences, though, the girls in the OO condition again stood out: while 83% of boys and 75% of girls in the CS condition, and 100% of boys in the OO condition, preferred the Causative screen, only 28.5% of girls in the OO condition showed the same preference. That is, 71.5% of the girls in the OO condition who showed a large preference preferred to look at the Contact actions rather than the Causative ones. Thus, the effect of the multiple frames in the OO condition seems due to the substantial number of girls who showed large (more than 0.65 seconds) preferences for the Contact actions.

4. Discussion

This study was demonstrated that young children can make use of multiple frames in verb learning. When presented with a scene containing an action involving physical causation and an activity involving physical contact without causation, young 2-year-olds made different conjectures about the meaning of an accompanying novel verb depending on the syntactic frames in which the verb was heard. When the verb was heard without a frame, or in an intransitive frame, they conjectured it referred to the Causative action.

However, when the verb was presented in transitive and intransitive frames in the Omitted Object (OO) alternation pattern (*The duck sebs the frog, the duck sebs*), the girls behaved significantly differently, and showed no consistent preference for the Causative action. Instead, they looked equally overall at both actions, with considerable (75%) consistency across items. Over 70% of the girls showing large screen preferences preferred the Contact actions. This effect could not be due to the mere presence of any combination of transitive and intransitive frames, because children who heard the verb in transitive and intransitive frames in the Causative (CS) alternation (*The duck sebs the frog, the frog sebs*) behaved significantly differently from the girls who heard the OO audio; those who heard the CS audio again looked longer at the Causative action. In sum, these results provide the first direct experimental support for the multiple frames tenet of the Syntactic Bootstrapping hypothesis: when a single frame is insufficient to constrain the meaning of a novel verb in a given extralinguistic context, young children can exploit the semantic implications of a *pair* of frames to narrow the focus of that verb's meaning.

These positive results must be accompanied by three qualifications. First, the effect of the OO audio was found for the girls only and not the boys. Second, while the girls in the OO condition behaved significantly differently from their peers in the other conditions, and did not show a preference for the Causative actions, they also did not look significantly *toward* the Contact action as they had been predicted to do. Third, while the children in the CS condition showed a significant preference for the Causative action, as had been predicted, this preference was not significantly different from that of their peers in the baseline (Frameless) condition. I consider each of these qualifications in turn.

4.1. Girls show greater linguistic sophistication than their male agemates

Only the girls seemed able to take advantage of the semantic implications of the OO alternation pattern, such that they did not choose the Causative actions as the referents for the novel verbs in the OO condition. These are not the first girls to demonstrate greater linguistic sophistication in comprehension than their male peers. Gender differences have been observed in other investigations of language development using the preferential-looking paradigm, and in each case the girls have been the ones to demonstrate earlier syntactic or lexical comprehension than the boys (e.g., Golinkoff et al., 1987, Hirsh-Pasek & Golinkoff, 1991; Fernald, McRoberts, & Herrera, 1993). Moreover, Katz, Baker, and Macnamara (1974) showed that 1-year-old girls were more sensitive than boys to the semantic implications of noun-phrase syntax. Scrutiny of the children's language production measures in the OO condition revealed that the girls in this condition were reported to produce more words overall than the boys (M (girls) = 333.06 words (SD = 16.29), M (boys) = 314.63 words (SD = 40.66), $t(30) = 1.73$,

$p < .05$ on a one-tailed test). The girls' and boys' measures on number of verbs and percent of three-word combinations did not differ. In sum, it seems likely that the gender effect observed here is yet another indication of the well-known linguistic advantage girls demonstrate over their male age-mates (see also Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991).

4.2. Perceptual and semantic contributions to syntactic bootstrapping

Recall that the initial hypotheses was that the CS alternation pattern would lead children to look at the Causative actions, while the OO alternation pattern would lead children to pick out the Contact actions. The results supported this hypothesis, in that the girls in the OO condition did look significantly less at the Causative actions than their peers in the CS, IO, and FF conditions did, and the difference between the OO and other audios could be observed for three of the four verb–action pairs tested. However, the girls in the OO condition, as a group, did not look longer at the Contact actions; instead, they looked equally at both. Moreover, while the girls in the CS condition did look significantly longer at the Causative actions than the Contact ones, this preference did not differ from that of their peers in the FF condition, when the verb was presented in isolation. As detailed below, I believe that these two qualifications to the basic result are interrelated to a certain extent, and derive from both perceptual and semantic factors.

Perceptual factors: the salience of causative actions

The children in the FF condition showed a significant and consistent preference for the Causative action as the referent of the novel verb (see Figs. 2 and 4). It is likely that this presyntactic focus reflects the perceptual salience of the causative actions; Naigles and Kako (1993) reported that adult subjects viewing these videos consistently rated the Causative action as larger than the Contact action (i.e., the duck's and frog's movements displayed greater displacement). Furthermore, causal relations have been shown to elicit more attention from prelinguistic infants than noncausal ones (Cohen & Oakes, 1993). Fisher et al. (1994) also found the causal component of a scene to be quite powerful in influencing the interpretation of an associated verb (in isolation), in that events in which one character was giving food to another to eat were more likely to be described as *feeding* than as *eating*. This well-documented presyntactic salience of causative actions may have contributed to the findings described above in two ways. With respect to the girls' equivocal preference in the OO condition, it is likely that they were experiencing a conflict between the salience of the Causative action and their understanding that the verb in the OO alternation referred to the Contact action. As detailed in the Results, girls with large preferences appeared more influenced by the latter factor and girls with small preferences appeared more influenced by the former. Given the

"pull" of the Causative actions, perhaps the best that these girls could do was to look more or less equally at both screens.

The presyntactic salience of the Causative actions may have also contributed to the absence of a significant shift in focus in the CS condition. Given that the children's preference without syntax was already likely to be strongly for the Causative actions, a syntactically based preference for the Causative actions in the CS condition would not have much leeway for expression. A similar phenomenon was observed in the second experiment for Naigles and Kako (1993), when the two contrasting actions were Contact and Non-causative/Synchronous. In that experiment, the children's preferences in the Frameless condition of that experiment were strongly for the Synchronous actions, and their peers who heard a novel verb in an intransitive frame showed statistically equivalent preferences for the two actions.

The behavior of two of the four verb–action pairs seen by the girls is consonant with both of these explanations. Both *seb* and *tig* elicited strong Causative preferences in the FF and CS conditions and no action preferences in the OO condition (see Fig. 4). The rationale that the salience of the Causative actions distracted the girls' attention from the Contact actions in the OO condition is also supported by the behavior of the verb–action pair associated with *lorp*. Recall that this verb was the only one to elicit a small Contact action preference in the OO condition, and that the FF and CS conditions generated even smaller Causative preferences. It seems likely that the Causative action associated with this verb was *not* particularly salient, and as a consequence, the OO alternation could exert its influence in the direction of the Contact action more readily. Thus, the behavior of *lorp* shows that if the paired Causative action is not too diverting, the OO alternation can actually direct young 2-year-old girls' attention *to* Contact actions. However, the "salience of causal actions explanation" cannot account for the behavior of *lorp* in the CS condition: If the Causative action associated with *lorp* was not particularly salient, and thus did not yield a large Causative preference in the FF condition, then why did *lorp* not elicit a strong Causative preference in the CS condition? This explanation also cannot account for the results with the verb *pim*, which elicited essentially equivocal preferences in both the CS and OO conditions. To fully explain the girls' behavior with these verbs, we must also consider some semantic factors.

Semantic factors: the match between the actions and the alternations' correlates

A supplementary explanation for the absence of a significant Contact preference in the OO condition, and for a Causative preference in the CS condition that was not greater than that shown in the FF condition, may be rooted in the types of Contact and Causative actions used. That is, these empirical qualifications may reflect inexact matches between the semantic

correlates commonly associated with the alternations, and aspects of the stimulus actions. I will discuss the Contact actions first. As presented in the introduction to this paper (and see Table 1), the Contact actions the children saw embodied some of the correlates often associated with the OO alternation; for example, they involved activities rather than accomplished actions. However, they differed from many of the English verbs that participate in the OO alternation in that Character 2, the goal or theme of the contacting action, was animate. Some analyses (e.g., Levin, 1993) have suggested that the prototypical verb that participates in the OO alternation takes an *inanimate* direct object (e.g., the cat is scratching the door, the man is sweeping the floor). Thus, perhaps the girls in the OO condition did not look significantly toward the Contact actions because these actions did not embody all of the semantic correlates that would be expected for a verb in the OO alternation. Their preferences may have been less equivocal if Character 1 had been contacting a box instead of a duck or frog.

The Causative actions used in the current study may have suffered from a similar problem, as they too involved an animate Character 2. According to some analyses (Dixon, 1991; Levin, 1993), causative actions that participate in the CS alternation also prototypically involve an animate causer an an inanimate causee; moreover, they are more likely to involve changes of state than the induced actions or changes in spatial configuration that were shown in the videos (Levin & Rappaport, Hovav, 1994). Perhaps the children in the CS condition did not look even more toward the Causative actions because these actions did not manifest these particular attributes. Levin (1993; Levin & Rappaport Hovav, 1994) has further suggested that when induced action verbs do participate in the CS alternation, the transitive frame also requires a locative phrase (e.g., *The rider jumped the horse OVER THE FENCE*), which was not presented in the audio stimuli in this study.[2] Perhaps if the children had heard "The frog is sebbing the duck to the floor/downwards/over" during the teaching trials they might have shown a greater preference for the Causative action during the test trials.

None of these factors yet explains the peculiar behavior of *lorp* and *pim* in the CS condition, though, because the issues of Character 2 animacy, induced action, and locative phrases are equally applicable to all four items. To explain why *lorp* and *pim* did not elicit significant Causative action preferences in the CS condition, we must consider another semantic correlate to the CS alternation and look more closely at the particulars of the causative actions used with these verbs. Recall Dixon's (1991) requirement that verbs that can participate in CS alternations are those in which a

[2] There is some evidence that children as young as 3 are sensitive to this property of the CS alternation: when Naigles et al. (1993) asked preschoolers and adults to enact (ungrammatical) transitive sentences containing the verbs *come*, *go*, *fall*, and *stay*, they found that all subjects were more likely to perform causative actions when the sentences contained locative prepositional phrases than when they were "bare" transitives (see also Naigles et al., 1992).

"Causer could be responsible for the event happening *in a natural manner*" (p. 295, italics added). In a recent analysis of causative alternations, Levin and Rappaport Hovav (1994) have elaborated on this requirement for one particular subclass of verbs that participates in CS alternations. Specifically, they suggest that nonagentive motion verbs (i.e., *roll*) are inherently intransitive but can become transitive if the motion can be externally controlled. Scrutiny of the causative action associated with *lorp* reveals that it meets this requirement of Dixon's and Levin and Rappaport Hovav's less than perfectly. Whereas the causative actions of *seb* and *tig* can be easily performed *either* under internal or external control (i.e., it is easy to bend over by oneself, and it is easy for another to force the bending over), the sideways tilt of *lorp* may not be so easily caused. In fact, *lorp* seems more similar to the actions of *lurch* or *sway*, both of which would be rather difficult to control externally. Thus, perhaps the girls in the CS condition did not look longer at the causative action associated with *lorp* because it did not embody this semantic correlate of the CS alternation.

The causative action associated with *pim* may have been less than optimal in a different way. Here, the causative action involves the duck holding onto the frog's elbow and moving that elbow to make the frog "bonk" himself (repeatedly) on the head with his hand. Thus, the causative action has the duck forcing the arm movement of the frog, an action that can occur both spontaneously and under external control. However, this action is actually more complicated than the other causative actions, as it involves two results instead of one: the arm movement of the frog in general and the "bonking" on the frog's head in particular. This latter action, in fact, resembles the contact activities of the other scenes, and here may lie the reason for the children's equivocal performance with *pim* with both the OO and CS audios: because both of the duck's actions during the teaching trials involve repeatedly contacting the frog, the children may have interpreted either or both of them to refer to "pimming". Thus, the verb in the OO alternation (The duck is pimming the frog/The duck is pimming) could really refer to either action. Furthermore, while there is a causative referent for the verb in the CS alternation, the stimulus sentences may not have targeted it appropriately (The duck is pimming the frog/The frog is pimming): the duck was really pimming the frog's *arm*, and the frog's *arm* was doing the pimming.

In sum, the qualifications to the main result of this study have actually been quite revealing concerning young 2-year-olds' knowledge and use of the relations between verb syntax and verb semantics. First, the basic result still holds: girls who heard a novel verb in two syntactic frames in the Omitted Object alternation showed significantly different action preferences when asked for the referent of that verb than their peers who heard the verb either with no frame, in a single intransitive frame, in a single transitive frame (cf. Naigles & Kako, 1993), or in two frames in the Causative alternation. Second, this effect was found even though the actions that were

to embody the semantic correlates of the CS and OO alternations may have lacked some important components (e.g., inanimate patients); therefore, it may be conjectured that the ability to use the transitive/intransitive alternations in the service of verb acquisition is fairly robust. Third, because the causal interpretation of a scene is so compelling to young verb learners (i.e., as seen in the FF condition), multi-frame syntax (in this case, the OO alternation) appears to be an important aid in redirecting their attention to other aspects of the scene (e.g., activities involving contact). Fourth, the item differences suggest that young 2-year-old girls understand already that the basic CS alternation (N1 V N2/N2 V) does not target all causal relations: those that occur easily (naturally) under external control are preferred (compare *seb* and *tig* with *lorp*). And finally, these child subjects showed considerable sophistication in rejecting acceptable causative actions that were not well described by the surface sentences (e.g., involving *pim*). In short, by exploring how young children exploit multiple-frame syntax to target the meanings of novel verbs, we have also shed some light on children's early understanding of the complex relations between verbs, their instantiations in scenes, and the sentence structures they can appear in.

4.3. Multiple frames in language acquisition

To step back, this study has demonstrated that a verb's appearance in diverse syntactic frames can help children target that verb's meaning. This work thus provides important experimental evidence for a growing theoretical stance in the literature that holds that cross-sentential observation has a critical role to play in language acquisition (Fisher & Tokura, in press; Gerken, Jusczyk, & Mandel, 1994; Gillette & Gleitman, 1995; Gleitman, 1990; Hoff-Ginsberg, 1985; Landau & Gleitman, 1985; Morgan, Meier, & Newport, 1987; Naigles & Hoff-Ginsberg, 1993, 1995; Rispoli, 1989, 1995). The proposals thus far have been primarily suggestive, showing how cross-sentential observation and analysis have the potential to address some knotty problems in acquisition. For example, one arena where cross-sentential observation has been deemed important is concerned with the role of prosody in the acquisition of syntax. Briefly, to the extent that the prosodic units of utterances map onto their syntactic units, prosodic cues (e.g., pauses, melodic contours) may be used by infants to delimit syntactic constituents. Gerken et al. (1994) have presented evidence that 9-month-olds are sensitive to the prosodic cues that distinguish lexical NP subjects from their predicates in declarative utterances as well as those that distinguish pronominal subjects from their predicates in interrogative utterances. However, both theory and data (see Gerken et al., 1994) have indicated that pronominal subjects in declarative sentences cannot be distinguished from their predicates by prosodic information. Thus, Gerken et al. propose that infants perform some sort of comparison of the two valid ways to adduce subclausal constituents, and then make the inference that

pronominal subjects in declarative sentences also form such constituents. Addressing the same issue, Fisher and Tokura (in press) suggest that the reduction of subject NPs across utterances, whether by a decrease in prosodic stress, by pronominalization, or by ellipses, might serve to highlight the accompanying VP as a syntactic unit. In both cases, the infant would be led to the acquisition of a new structure (the subject/predicate boundary) via a set of individually indeterminate but jointly converging sentence types.

Within the realm of verb acquisition, other researchers have also highlighted the potential importance of multi-sentence analysis. Rispoli (1989, 1995) has suggested that cross-sentential observation may be especially critical for Japanese verb learners. Because of unrestricted zero anaphora, Japanese input does not appear to distinguish agentive from non-agentive verbs by surface syntax (i.e., transitivity); in Rispoli's corpus, less than half of the agentive verbs appeared with a direct object or accusative case marker. At the very least, then, it should be incumbent upon Japanese verb learners to compare verb uses across input utterances before deciding whether a verb is agentive or not. However, Japanese also provides more consistent grammatical cues to verb agentivity in the form of consistently produced agentive morphemes (e.g., the desiderative, prohibitional, permissive, and request postpositions), as well as pragmatic cues that are instantiated by the absence of animate patients. Rispoli's proposal is that a child who was able to pool these sources of information across utterances would successfully subcategorize Japanese agentive and non-agentive verbs into two different classes.

Similarly, Hoff-Ginsberg (1985) found that structural changes across American English mothers' utterances (e.g., addition or deletion of constituents) during mother–child play periods had a facilitating effect on their 2-year-olds' verb production 2 months later, and suggested that "the child may benefit from the presentation of pairs of utterances that are syntactically similar, but not identical, because he compares the structural properties of adjacent utterances in input, and makes use of information which derives from that comparison" (1985, p. 380). Finally, on the grounds that any single syntactic frame may be insufficient to fully target a verb's meaning in context, Naigles and Hoff-Ginsberg (1993, 1995; see also Landau & Gleitman, 1985) analyzed the entire set of verb frames used in a corpus of maternal speech. They found that for 30 common verbs, the set of frames in which each verb appeared was sufficient to distinguish it from all of the others. Moreover, they showed that the more frames in which mothers used particular verbs at Time 1, the more frequently their children produced those verbs 10 weeks later.

The common thread running through all of this research is that diverse sentence forms in the input can, if analyzed together, provide children with important information concerning subclausal constituents or verb meanings. The current study not only supplies the first experimental evidence that

young language learners can take advantage of such cross-sentential information, but goes one step further: in the bulk of the studies reviewed above, the contribution of each sentence form may be interpreted as additive. Each form provides another piece of the puzzle (i.e., the verb representation), or strengthens the probability that the working configuration is correct (i.e., the subject/predicate boundary). Here, in contrast, the pair of sentence frames (the transitive/intransitive alternation) makes a specific prediction concerning the meaning of the verb that appears in it. That is, hearing a verb in a transitive frame can inform a child that the verb's action affects an object, possibly in a causative relation (Fisher et al., 1994; Gropen, Pinker, Hollander, & Goldberg, 1991; Naigles, 1990; Naigles et al., 1993; Naigles & Kako, 1993), and so directs the child to focus on these types of actions in the accompanying scene. Hearing a verb in an intransitive frame suggests, among other things, that the verb's action is spontaneous or internally controlled, so that the child's attention is directed to these types of actions (Fisher et al., 1994; Naigles, 1990; Naigles et al., 1993). Hearing the same verb in both frames may inform the child that both types of actions are feasible for this verb, but may not (yet) provide enough direction for the child's attention in a given scene of set of scenes. However, hearing the verb in both frames *in the Omitted Object alternation* provides information that the verb's action is not causative, and that it refers to an ongoing activity in the scene (in this study, this was one involving contact). In sum, the information provided by the two frames in the OO alternation in this study would seem to be more specific than that provided by sentence forms being considered in combination but not in relation to each other. Both kinds of cross-sentential information are likely to be useful to the language learner, though, and both should be investigated in more detail.

In conclusion, making connections between multiple frames (i.e., Syntactic Bootstrapping) appeared to be necessary for the children in this experiment to determine that the novel verb referred to the Contact rather than Causative action. Notice, though, that the information provided by the syntax was not applied on a separate, non-interacting track from other sources of information. Instead, the children viewed the scenes, heard the frames, and used the frames to help specify the relevant aspects of the scene for each particular verb.

Acknowledgements

This research was supported by the NIH FIRST Award HD26595. I thank Melissa Highter, Karen Lin, Kristina Hanson, Nicole Anker, Rudy Mendoza-Denton, Nancy McGraw, Jo Pichanick, Jenn Cunningham, and all the volunteers of the Yale Infant Language Lab for their assistance in data collection and analysis. Thanks also go to Cynthia Fisher, Lila R. Gleitman, Erika Hoff-Ginsberg and two anonymous reviewers for their helpful insights

and commentary on earlier versions of this paper. Portions of this research were presented at the 17th Annual Boston University Conference on Language Development, January 1994.

References

Bowerman, M. (1982). Starting to talk worse: clues to language acquisition from children's late speech errors. In S. Strauss (Ed.), *U-shaped behavioral growth* (pp. 101–145). New York: Academic Press.

Bowerman, M. (1983). Evaluating competing linguistic models with language acquisition data: implications of developmental errors with causative verbs. *Quaderni di Semantica, 3*, 5–66.

Clark, E. (1987). The principle of contrast. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (p. 3–20). Hillsdale, NJ: Erlbaum.

Cohen, L., & Oakes, L. (1993) How infants perceive a simple causal event. *Developmental Psychology, 29*, 421–433.

Dixon, R.M.W. (1991). *A new approach to English grammar, on semantic principles*. Oxford: Oxford University Press.

Fernald, A., McRoberts, G., & Herrera, C. (in press). *Effects of prosody and word position on lexical comprehension in infants. Journal of Experimental Child Psychology.*

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. (1994). When it is better to receive than to give: syntactic and conceptual constraints on vocabulary growth. *Lingua, 92*, 333–375.

Fisher, C., & Tokura, H. (in press). Prosody in speech to infants: direct and indirect cues to syntactic structure. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to syntax in early acquisition*. Hillsdale, NJ: Erlbaum.

Gerken, L., Jusczyk, P., & Mandel, D. (1994). When prosody fails to cue syntactic structure nine-month-olds' sensitivity to phonological vs. syntactic phrases. *Cognition, 51*, 237–265.

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

Gleitman, L., & Gillette, J. (1995). The role of syntax in verb learning. In P. Fletcher and B. MacWhinney (Eds.), *The handbook of child language* (pp. 413–427). London: Blackwell.

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.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: the role of lexical semantics in the acquisition of argument structure. *Cognition, 41*, 153–195.

Hall, R., & Friends (1984). *Sniglets (snig' let): Any word that doesn't appear in the dictionary, but should*. New York: Macmillan.

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. (1985). Some contributions of mothers' speech to their children's syntactic growth. *Journal of Child Language, 12*, 167–185.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–247.

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

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

Katz, N., Baker, E., & Macnamara, J. (1974). What's in a name? A study of how children learn proper and common names. *Child Development, 45*, 469–473.

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* (Lexicon Project Working Papers No. 1) (pp. 1–62). Cambridge, MA: MIT Center for Cognitive Science.

Levin, B. (1993). *English verb classes and alternations*. Chicago: University of Chicago Press.

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

Levin, B., & Rappaport Hovav, M. (1994). A preliminary analysis of causative verbs in English. *Lingua*, *92*, 35–77.

Morgan, J., Meier, R., & Newport, E. (1987). Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, *19*, 498–550.

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., Fowler, A., & Helm, A. (1992). Developmental changes in the construction of verb meanings. *Cognitive Development*, *7*, 403–427.

Naigles, L., & Gelman, S.A. (1995). Overextensions in production and comprehension, revisited: a study of *dog*, *cat*, and *cow*. *Journal of Child Language*, *22*, 19–46.

Naigles, L., Gleitman, H., & Gleitman, L.R. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and development* (pp. 104–140). Norwood, NJ: Ablex.

Naigles, L., & Hoff-Ginsberg, E. (1993). *Input to verb learning: verb frame diversity in mother's speech predicts children's verb use*. Poster presented at the Biennial meetings of the Society for Research in Child Development, New Orleans, LA.

Naigles, L., & Hoff-Ginsberg, E. (1995). Input to verb learning: evidence for the plausibility of Syntactic Bootstrapping. *Developmental Psychology*, *31*, 827–837.

Naigles, L., & Kako, E. (1993). First contact: biases in verb learning with and without syntactic information. *Child Development*, *64*, 1665–1686.

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

Pinker, S. (1994). How could a child use verb syntax to learn verb semantics? *Lingua*, *92*, 377–410.

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

Rescorla, L., Hadicke-Wiley, M., & Escarce, E. (1993). Epidemiological investigation of expressive language delay at age two. *First Language*, *13*, 5–22.

Rispoli, M. (1989). Encounters with Japanese verbs. *First Language*, *9*, 57–80.

Rispoli, M. (1995). Missing arguments and the acquisition of predicate meanings. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: The acquisition of verbs*. Hillsdale, NJ: Erlbaum.

Tomasello, M. (1992). *First verbs*. Cambridge University Press.

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca: Cornell University Press.