LEARNING VERBS UNDER REFERENTIAL UNCERTAINTY: THE ROLE OF REFERENTIAL AND SYNTACTIC CONTEXTS

BY

ROSE MARIE SCOTT

DISSERTATION
Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology in the Graduate College of the University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:
Professor Cynthia Fisher, Chair
Professor Renée Baillargeon
Assistant Professor Sarah Brown-Schmidt
Professor Gary S. Dell
Associate Professor Susan M. Garnsey
ABSTRACT

Previous research on early verb learning has focused largely on how children acquire verbs in highly informative contexts. However, children often hear verbs in a variety of less informative circumstances. They hear verbs in the presence of multiple potential referents, and they hear verbs in the absence of relevant scenes. This dissertation investigated whether children can learn anything about verbs under these circumstances of referential uncertainty. The first two sets of experiments (Chapters 2 and 3) examined situations in which children heard a novel verb in a sentence that was consistent with more than one referent within a scene. Experiments 1 and 2 (Chapter 2) showed that 31-month-olds can encode information about the set of potential referents for a novel verb under these circumstances. They can then integrate this information with additional scene information, using cross-situational consistency to identify the verb’s referent. Experiments 3 and 4 (Chapter 3) showed that 28-month-olds can also identify a verb’s referent by integrating information across multiple sentence structures. Thus, children who heard a novel verb used in the causal alternation realized that the verb had to describe an action on an object that produced an effect (e.g., caused-motion). The final set of experiments (Chapter 4) focused on situations in which children hear a verb in the absence of a relevant referential. 28-month-olds acquired useful combinatorial information about a novel verb simply by listening to its use in sentences and later retrieved this information to map the verb to an appropriate referent.

Taken together, these studies show that children can form lexical entries for verbs, even when they are uncertain about the verb’s referent. They can then attach to this entry facts about the set of potential referents for the verb, the sentences structures in which the verb occurs, and facts about the verb’s arguments. Children identify a verb’s referent by integrating information
across different referential scenes, integrating information across sentence structures, and integrating previous sentential information with later referential information.
ACKNOWLEDGMENTS

I would like to thank my advisor Cynthia Fisher for her guidance throughout my graduate career. Not only did her comments and criticism play a crucial role in the development of this research, but I am certain this dissertation would not have been possible without her persistent encouragement and enthusiasm. I would also like to thank my committee members – Renée Baillargeon, Sarah Brown-Schmidt, Gary Dell, and Susan Garnsey – for their many thoughtful questions, comments, and discussions.

I would like to express my gratitude to the many research assistants who have worked in the Language Acquisition Lab for their invaluable assistance with data collection and coding. I especially wish to thank Amélie Bernard, Yael Gertner, Cynthia Lukyanenko, Zhenghan Qi, and Sylvia Yuan for many helpful discussions and for their friendship and support. Many thanks as well to the many parents and children who participated in this research.

I would like to thank my parents, Cynthia Scott and Ian Scott, my aunt, Patricia Fitzpatrick, and my friend Nicholas Pasquale, for always being there when I needed them, whether it was to share my joy or my frustration. Finally, I would like to offer special thanks to Derek Devnich for his tireless love and support, for his many helpful discussions of this research, and for starring in the stimulus events for Experiment 4.
# TABLE OF CONTENTS

CHAPTER 1 THE PROBLEM OF VERB LEARNING .................................................................1

CHAPTER 2 2.5-YEAR-OLDS USE CROSS-SITUATIONAL CONSISTENCY TO LEARN VERBS UNDER REFERENTIAL UNCERTAINTY ...............................................................7

CHAPTER 3 USING MULTIPLE SENTENCE FRAMES TO LEARN VERBS ......................31

CHAPTER 4 2-YEAR-OLDS USE DISTRIBUTIONAL CUES TO INTERPRET TRANSITIVITY ALTERNATING VERBS ..................................................................................67

CHAPTER 5 CONCLUSIONS ..............................................................................................97

FIGURES .............................................................................................................................102

TABLES ..................................................................................................................................110

REFERENCES .........................................................................................................................121

APPENDIX .............................................................................................................................134
CHAPTER 1
THE PROBLEM OF VERB LEARNING

Verbs are the syntactic and semantic heart of sentences: they determine which phrases can occur in a sentence and specify the semantic relationship between those phrases. Learning verbs is thus a fundamental part of learning to comprehend and produce one’s native language. However, learning verbs is a difficult task. Whenever children encounter a new word, be it verb or noun, the accompanying referential scene offers many potential interpretations. In order to learn the word, children need to somehow eliminate the erroneous interpretations. All words pose these problems, but verbs pose a number of additional difficulties. Verbs typically describe actions and relations between objects, dimensions of experience that are often abstract and therefore do not lend themselves well to direct observation. Their referents also tend to be fleeting: a child who hears the word *run* may turn to find a scene that contains jumping, laughing, and talking, but that no longer contains running. Finally, verbs do not simply label events – they adopt a particular perspective on an event. For instance, the verbs *chase* and *flee* both refer to events involving pursuit. In order to determine whether the speaker was referring to chasing or fleeing, the child would need some way of identifying the speaker’s perspective on the event.

The difficulty posed by verbs is attested to in cross-linguistic studies of children’s early vocabularies. Across languages as disparate as Korean, English, and Hebrew, children’s early vocabularies contain a larger proportion of nouns than verbs (e.g., Bornstein et al., 2004; Nelson, 1974). Nevertheless, studies also show that children acquire verbs quite early – verbs are included in children’s first 50 words (Bloom, Tinker, & Margulis, 1993; Nelson, 1973;
Tomasello, 1992). Some of the earliest verbs children learn are the highly abstract verbs that ought to be the most difficult to acquire (e.g., go; Bloom et al., 1993; Dale & Fenson, 1996).

Verb learning thus presents an intriguing puzzle that has inspired several decades of research. Given the apparent difficulty of verb learning, how is it that children acquire verb meanings at an early age? The assumed answer to this question is that children have access to sources of information that help them to overcome the challenges posed by verbs. In particular, a great deal of research has focused on what constraints children might use to narrow their interpretation of a new verb when they encounter it. Most studies have investigated this question by presenting children with highly informative contexts: children hear a verb in the presence of a relevant referential scene and the available constraints pick out a single referent. The findings from such studies have shown that children use a variety of non-linguistic and linguistic cues, together with the referential scene, to constrain verb interpretation (e.g., Fisher, 1996; Fisher, 2002; Gleitman, 1990; Landau & Gleitman, 1985; Naigles, 1990; Naigles, & Kako, 1993; Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009; Tomasello & Barton, 1994; Yuan, Fisher, & Snedeker, in prep.)

One cue that children use to narrow their interpretation of a verb is the structure of the sentence in which that verb occurs (i.e. syntactic bootstrapping; Fisher, 1996; Fisher, 2002; Gleitman, 1990; Landau & Gleitman, 1985; Naigles, 1990; Naigles & Kako, 1993; Yuan et al., in prep.). Naigles (1990) presented 25-month-olds with scenes that contained both a caused motion (e.g., a duck pressed down on a bunny’s head, causing him to bend at the waist) and a simultaneous activity (e.g., the duck and the bunny moved their arms in a circle). Each scene was accompanied by a novel verb used in either a transitive sentence (e.g., The duck is gorping the bunny) or a conjoined-subject intransitive sentence (The duck and the bunny are gorping). At
test, the two components of the complex scene were separated and presented on different screens; children were then asked to Find gorping. Children who had heard the verb used in a transitive sentence looked significantly longer at the caused-motion event than children who heard the verb used in the conjoined-subject intransitive sentence. Thus, by 25 months, children recognize that transitive sentences refer to relationships between two participants.

Yuan et al. (in prep) have shown that even 19-month-olds can use surface properties of a sentence to make inferences about a verb’s meaning. Children watched two simultaneous events: one depicted a two-participant relation (a boy pushing on another boy’s shoulders, causing him to bend forward) and the other depicted a solo action (a boy making arm motions as in jumping jacks). While viewing these events, children heard a novel verb in a transitive sentence (e.g., He’s gorping him) or an intransitive sentence (e.g., He’s gorping). Children who heard the verb used in a transitive sentence looked longer at a two-participant relation than did children who heard the verb used in the intransitive sentence.

Young children can also use learned language-specific sentence cues when interpreting a new verb. Gertner, Fisher, and Eisengart (2006) showed 21-month-olds two videos: a boy spinning a girl on a chair and a girl causing a boy to bend. While watching these videos, children heard either “The boy is gorping the girl” or “The girl is gorping the boy.” In both cases, children looked reliably longer at the video in which the subject of the sentence played an agent’s role. A follow-up experiment showed that when children instead heard a sentence such as Who’s gorping the boy, they looked longer at the video in which the object of the sentence played a patient’s role. Together, these results suggest that by 21 months, children have learned to use the order of nouns in a sentence to infer whether a participant plays an agent or a patient role. They
can then restrict their interpretation of a novel verb to events that are consistent with the role assignments conveyed in the sentence.

These studies suggest that in highly informative contexts, children can use cues such as sentence structure to interpret a new verb. However, in a natural learning context, children often encounter verbs in much less informative contexts than those used in the studies described above. This is because sentences and scenes are independent sources of information about verb meaning and they do not always occur in useful conjunction. For instance, children could encounter a verb accompanied by a relevant referential scene, but without useful sentence information. This could occur either because the children failed to understand the sentence that the verb occurred in, or because the sentence was consistent with multiple interpretations of the present referential scene. To illustrate, consider a child who is watching a group of girls playing: One girl is pushing another girl on a swing and a third girl is hopping. If the child heard the intransitive sentence *She’s dacking* while viewing this scene, this child would have no way of knowing whether the verb described the solo action (e.g., *She’s hopping*) or a component of the two-participant relation (*She’s pushing* or *She’s swinging*; Yuan et al., in prep.). Children also often encounter verbs in sentences in the absence of a relevant referential scene (Gleitman, 1990). Imagine a child who hears her mother say “Let’s see if Granny’s home” while watching her mother dial the phone to call Grandma. This child has access to syntactic facts about the verb *see* (e.g., it occurs with sentential complements), but little useful referential information about the event that the verb describes.

Can children learn anything about a verb’s meaning from these types of observations? While the two scenarios described above involve uncertainty about the verb’s referent, they still contain potentially useful information about the verb’s meaning. In order to be able to take
advantage of this information, however, children would need to be able to do at least two things. First, children must be capable of encoding information about the sentences and scenes that occur with a verb, even when these two information sources do not converge on a single present referent. If children hear a sentence that is consistent with multiple referents within a scene, they must be able to encode information about the set of potential referents for the verb without knowing which part of the referential scene the verb describes. If children hear a verb used in the absence of a relevant referential scene, they need to encode facts about the sentences in which the verb occurs, even though they have no referential information about the event described by the verb. In both cases, children need to be able to form a proto-lexical entry for a verb – based on the verb’s phonological form – and attach to it facts about how the verb is used. Second, because observations of this kind leave the verb’s meaning ambiguous, children would need to be able to integrate the information gathered from these observations with that provided by additional observations of sentences and/or scenes. This would enable them to update and refine their interpretation of the verb.

This dissertation examined whether children could overcome these difficulties and learn verbs under referential uncertainty. The experiments in Chapters 2 and 3 investigated whether children could learn in situations where the available sentence information was consistent with multiple potential referents. In Chapter 2, we held the sentence context constant (children always heard a verb used in the same sentence structure) and asked whether children could identify a verb’s referent by taking into account additional referential contexts. Children encountered a series of sentence-scene pairs, each of which was ambiguous because the sentence was consistent with two candidate event-referents. Across trials, however, each verb consistently occurred with only one of the event-referents. If children can keep track of the co-occurrence of events and
verbs under referential uncertainty, then over time they should be able to identify the correct referent for the verb.

In Chapter 3, we held the referential scene constant (children only saw a single set of referential options accompanying the verb) and examined children’s ability to integrate information across multiple sentence contexts. While a single sentence structure is likely to be consistent with many interpretations of a scene, encountering a verb in multiple sentence structures should provide greater constraint on a verb’s meaning. To test this possibility, we presented children with verbs that occurred in both transitive and intransitive sentences. When considered individually, the transitive and intransitive sentences provided little constraint on the verb’s meaning. If, however, children could consider both sentence frames simultaneously, then this would provide them with enough information to identify the verb’s correct referent.

The experiments in Chapter 4 examined whether children could learn useful information about a verb in the absence of a relevant referential scene. Children saw two women carrying on a conversation in which they repeatedly used a made-up verb in both transitive and intransitive sentences. The surface properties of the sentence (e.g., animacy, lexical overlap) identified the verbs as belonging to one of the two transitivity alternations examined in Chapter 3. After viewing these dialogues, children saw two referential options and heard the verb in a transitive sentence that was consistent with both referential options. If children could encode the sentence information present in the dialogue, then they should be able to use that information to determine which of the two referents the verb described.
Whenever children encounter a new verb, the referential scene offers many potential interpretations. For instance, imagine a child who is watching a group of girls playing on a playground. One girl is pushing another girl on a swing, a third girl is sliding down a slide, and a fourth is climbing on a jungle-gym. While viewing this scene, the child hears her mother say, “Oh look, she’s pimming!” Based on this sentence alone, the child has no way of knowing whether pimming refers to pushing, swinging, sliding, or climbing. Is the child able to learn anything about the meaning of the verb pim, given this referential ambiguity?

One way that children might overcome this type of ambiguity is to take into account additional referential contexts in which the verb occurs. Across occurrences of a verb, events (or event components) that are not part of a verb’s meaning should vary more than those that are central to the verb. If children could detect which elements consistently co-occurred with a verb across referential contexts, then this would help them isolate the verb’s meaning. Although researchers have long assumed that cross-situational consistency plays a role in verb learning (e.g., Fisher, Hall, Rakowitz, & Gleitman, 1994; Pinker, 1984), the process is clearly far from trivial. In order to benefit from cross-situational consistency in verb learning, children would have to establish a lexical entry for a verb and attach to that entry information about the verb’s potential referents without knowing which aspect of the scene the verb described. They would then have to update this information as they encountered the verb in additional referential contexts.
Recent experiments have shown that, despite these challenges, adults and 12-month-old children can use cross-situational consistency to learn novel *nouns* that refer to objects (Smith & Yu, 2008; Yu & Smith, 2007). Here we adapt this established technique to ask whether 2-year-olds can use information derived from cross-situational consistency to identify the referents of verbs. We begin by reviewing the recent work with nouns and what would be required for children to perform similar learning with verbs. We then present experiments in which 2-year-olds encounter novel intransitive (Experiment 1) and transitive verbs (Experiment 2) under referential uncertainty and ask whether they can benefit from cross-situational consistency in verb learning. We will find that 2-year-olds do benefit from this source of information, but we also find significant limits in their ability to do so. These findings help to define the role of cross-situational observation in verb learning.

Smith and Yu (2008) tested 12-month-olds’ ability to use cross-situational observation to learn novel nouns. Infants saw a series of training trials in which pictures of two novel objects were presented side by side; each pair of objects was accompanied by two novel labels. On each trial, it was impossible to tell which novel label went with which novel object. Across trials, however, each word consistently occurred with only one object. In each test trial, infants heard one label and saw its referent and a distractor object. Infants looked longer at the referent object, suggesting that they had learned the labels by keeping track of which word consistently appeared with each picture. Thus, infants were able to attach some information about potential object referents to each noun’s lexical entry, even when the referential context contained multiple possible referents. They were then able to update this information across trials and identify correct object-noun pairings.
Can children also use cross-situational observation to learn verbs? In order to do so, they must overcome several additional challenges that verbs pose.

First, there is strong evidence that observation of the extralinguistic context is less useful for learning verbs than it is for learning nouns. In experiments using the “human simulation” paradigm, adults tried to guess the words mothers said to their toddlers based on observing a series of referential scenes in which each word was uttered. Adults were much better able to guess the mothers’ nouns than their verbs under these circumstances (Gillette, Gleitman, Gleitman, & Lederer, 1999; Snedeker & Gleitman, 2004); similar results were obtained with 7-year-olds (Piccin & Waxman, 2007). Further results suggested that the trouble with learning verbs via scene observation is that verb referents tend to be more abstract than noun referents, and therefore verb referents are naturally harder to observe. Adults correctly identified concrete verbs (e.g., throw, hammer) much more often than abstract verbs in this procedure (e.g., think, know; Gillette et al., 1999; Snedeker & Gleitman, 2004); similar results were obtained for concrete (e.g., spoon) versus abstract nouns (e.g., thing; Gillette et al., 1999). These results tell us that cross-situational observation is likely to be most useful for learning concrete verbs, those whose referents are relatively easy to observe. For this reason, we focus here on children’s ability to use cross-situational observation to learn verbs that describe concrete actions.

Second, encoding referential information under uncertainty and updating that information across referential contexts is likely to be more difficult for verbs than for nouns. In order to learn the nouns in Smith and Yu’s (2008) experiment, infants had to encode information about a set of static objects that occurred with the words on each trial. In contrast, consider a situation where multiple candidate verb referents are available, such as the playground scenario described above. In order to learn from this observation, the child must encode information about multiple,
causally-unrelated actions (e.g., *swinging, climbing, sliding*). Thus, the information that children need to attach to lexical entries is more complex for verbs than for nouns. To date there is no evidence that children can attach referential information to a verb’s lexical entry when multiple candidate referents are present.

If children can encode information about potential referents, they must then update this information based on subsequent observations. The infants in Smith and Yu’s experiment saw the same six object tokens repeated across both the training and test trials. In order to keep track of the co-occurrence patterns, they only needed to note whether a particular label had occurred with a particular object token in the past. While children do encounter repeated object tokens in real life (e.g., their own toys, their own bottle), they see repeated *action* tokens only on video. This is because actions vary when instantiated by different actors and objects, and even the same actor is unlikely to perform an action the same way twice. In order to learn verbs via cross-situational observation, children must be able to compare distinct event tokens and abstract across irrelevant variation.

The evidence is mixed regarding children’s ability to abstract across irrelevant variation when learning verbs. Some studies show that when presented with a series of training exemplars, each of which is clearly labeled with the same verb, children can compare distinct event tokens to draw inferences about verb meaning (e.g., Behrend, 1995; Childers, 2008). Childers (2008) showed 31-month-olds a series of training events depicting a consistent manner of action (with varying result) or a consistent result (with varying action); each event was enacted in isolation, and accompanied by the same novel verb (e.g. “I’m dacking it!”). Children were then given a set of objects and asked to enact the action named by the novel verb. Children’s enactments tended to preserve elements that had been consistent in the training events. This suggests that children
assumed that the event component that consistently occurred with a verb was central to the verb’s meaning, while components that varied were not. However, other studies have found that young children sometimes fail to identify consistent components across events due to superficial variation. For instance, Maguire et al. (2008) found that 31-month-olds who learned a novel verb by seeing four distinct event tokens had more difficulty extending that verb to relevant new events than those trained with a single repeated event token. Taken together, then, these two sets of studies suggest that while children are capable of detecting consistency in the face of irrelevant variation, this process is difficult for them. Thus, it is possible that they would be unable to detect consistency when faced with the additional challenge of encoding and comparing events under referential uncertainty.

**Experiment 1**

In Experiment 1, we asked whether 31-month-old children could attach referential information to novel intransitive verbs when multiple potential referents were present and then use cross-situational consistency to identify each verb’s correct referent. To examine this question, we adapted Smith and Yu’s (2008) paradigm for use with novel verbs. On each trial, two novel one-participant actions, each performed by a different actor, were presented along with two novel intransitive verbs (e.g., "She's pimming ... and she's nading."; Figure 1). Each of the potential referents was a concrete action. The sentence structure identified the new words as intransitive verbs, but provided no information as to which verb described which action. The only cue to each verb's intended referent was the consistency with which it accompanied the same action across trials.

Unlike in Smith and Yu’s (2008) experiment, we did not include separate test trials. Instead, we employed a continuous study-test procedure. Within each trial, we examined where
children looked when they heard a novel verb. This allowed us to see how children’s interpretation of a verb changed over time. We predicted that if children could keep track of which event consistently occurred with each verb, then when they encountered a verb, they would look longer at the matching referent than at the non-matching referent for that verb.

Method

Participants

Thirty-six 31-month-olds participated in the experiment (mean = 31.2, range 30.1-33.1, 18 male, 18 female). All were native speakers of English. An additional 4 children were tested but eliminated because they had a side bias (1), became fussy (1), or failed to complete the experiment (2). Children’s productive vocabulary was measured using the short form of the Bates MacArthur CDI, Level 2 (Fenson et al., 2000). Vocabulary scores ranged from 42 to 100 with a median of 80.5. Half of the children were assigned to verb order 1 and half to verb order 2. An equal number of boys and girls was assigned to each verb order.

Apparatus

Children sat on a parent’s lap facing two 20-inch television screens placed about 30 inches away. The screens were 12 inches apart and about at the children's eye level. Soundtracks were played from a concealed central speaker. A camera hidden between the two screens recorded the children’s eye movements during the experiment. Parents wore opaque sunglasses, preventing them from biasing their children’s responses.

Materials and Procedure

Materials consisted of color videos of four women performing novel solo actions. There were four novel actions, each paired with a novel intransitive verb (see Table 1). Each woman performed three of the four novel actions for a total of 12 unique event tokens. Tokens were
presented in synchronized pairs and accompanied by a soundtrack recorded by a native English speaker. Left-right positioning of event tokens was counterbalanced within verb-order.

Children were tested with a continuous study-test procedure consisting of 12 trials. On each trial, a pair of event tokens, each performed by a different actor, was presented for 8 seconds. While viewing the events, children heard two intransitive sentences, each containing a novel verb (see Figure 1). All sentences were of the form, “She’s verbing.” The use of the pronoun subject she ensured that the sentences were ambiguous and could refer to either event. The second sentence of each trial was recorded with contrastive stress on both the subject pronoun and the verb. The presentation of the sentences was timed such that the onset of the first verb occurred approximately 1 second into the trial and the onset of the second verb occurred approximately 5 seconds into the trial. Trials were separated by a 3-second blank-screen interval.

The experiment was divided into two blocks of six trials. Each of the 12 event tokens occurred once in the first block and once in the second block. Within each block, children saw all possible combinations of actions once (e.g., the tilting-torso event in Table 1 paired once with squatting, once with arm-stretching, and so forth). Event tokens were paired and presented in the same sequence for all participants. This sequence was randomized with the following constraints: the same actor never appeared on both screens at once, no action occurred more than two trials in a row, and each action appeared an equal number of times on the right and left screens. Across the experiment, each action was labeled six times (three times in each block). The action on the left screen was labeled first in half the trials, and each action was labeled first in half the trials in which it occurred.

Children were assigned to one of two verb orders that differed in which verb occurred first in each trial. For instance, in trial 1, children in verb order 1 heard “She’s pimming...And
she’s nading.” while children in verb order 2 heard “She’s nading…And she’s pimming.” Thus, for any given trial, half of the children heard one event labeled first, while the other half heard the other event labeled first.

Analyses

We coded where children looked (left-screen, right-screen, away) frame by frame from silent video. To assess reliability, 8 children’s data were independently coded by a second coder. The first and second coders agreed on the children’s direction of gaze for 98% of coded video frames.

Within each trial, we examined children’s visual fixations within two 2.5-second observation windows that began at the onset of each novel verb. Observation windows in which the children looked away from the screen for more than 67% of the 2.5-second interval were dropped from the analyses (30 observation windows, 3.5% of the 864 possible observation windows). For the remaining 834 observation windows, we performed a preliminary analysis of variance (ANOVA) on the time children spent looking away from the two screens, with within-trial observation window (1, 2) and trial (1-12) as within-subject factors. This analysis revealed that children tended to look away more during the second observation window in each trial, $F(1, 20) = 4.983, p < .05$ (window 1: $M = .10, SD = .12$; window 2: $M = .14, SD = .14$). There was also a marginal effect of trial, suggesting that children looked away more in later than in earlier trials, $F(11, 220) = 1.658, p = .08$ (trial 1: $M = .06, SD = .15$; trial 12: $M = .20, SD = .35$). Because look-away times varied within and across trials, we conducted our main analyses on raw looking times to the matching and non-matching events rather than on a proportion of the time spent looking to the matching event out of the time spent looking at the two events.
For each observation window, we calculated a target-advantage score by subtracting the time spent looking at the non-matching event (i.e. the distractor event) from the time spent looking at the correct referent for the novel verb (i.e. the target event). We omitted the first observation window for each verb because the first time children encountered each verb they could not have any cross-situational information about the verb’s meaning\textsuperscript{1}. For each verb, the target-advantage scores were averaged across the second through sixth presentations of the verb. Preliminary analyses of target-advantage scores revealed no effects of sex or verb order. These factors were not examined further.

Results and Discussion

In order to succeed in the present task, children needed to look at one event during the first half of the trial, and then switch to the other event upon encountering the second verb in the second half of the trial. The extent to which they did so provides a measure of the degree to which children understood the task. Figure 2 displays the proportion of fixations over time to each of the two events (verb1 match, verb2 match) averaged across observations 2 through 6 for all verbs. Fixations are plotted separately for children whose vocabulary scores were above the median (high-vocabulary group; Figure 2A) and children whose vocabulary scores were below the median (low-vocabulary group; Figure 2B). At the onset of the trial, children were equally likely to fixate each of the two events. At the average onset of the first verb, there was an increase in the proportion of fixations to the matching referent for the first verb. This increase occurred in both vocabulary groups, though the shift occurred slightly later in the low-

\textsuperscript{1} In principle, children could have had cross-situational information about the meanings of the verbs that were introduced for the first time on the second and third trial. To illustrate, consider the second trial shown in Figure 1. Because children already had some evidence that pimming referred to the torso-tilting motion, they could have used this information, along with assumptions of mutual exclusivity, to deduce that tazzing must refer to the girl lifting her knee. However, preliminary analyses confirmed that children’s target-advantage scores on the first observation for each verb did not differ significantly from 0, even for the verbs introduced for the first time in the second and third trials, all \( p > .2 \) (Pim \( M = -.31s \), Nade \( M = .09s \), Tazz \( M = .37s \), Rivv \( M = .02s \)).
vocabulary group. Looks to the verb1 match persisted throughout the first observation window. Approximately 200ms prior to the average onset of the second verb, children in the high-vocabulary group shifted their attention to the matching referent for the second verb. This shift occurred slightly later within the low-vocabulary group. Thus, both vocabulary groups demonstrated the same overall pattern of focusing their attention on the verb1 match during the first observation window and then shifting their attention to the verb2 match during the second observation window. This pattern suggests that children understood the structure of the task, and therefore were in a good position to benefit from cross-situational consistency.

Next, we analyzed children’s target-advantage scores. We first asked whether children’s vocabulary level predicted their ability to learn the verbs in this experiment. Children’s overall target-advantage scores were averaged across verbs and analyzed using an ANOVA with vocabulary group (high, low) as a between-subjects factor. As might be expected from the time-course data shown in Figure 2, the effect of vocabulary was not significant, $F(1, 34) < 1$. Vocabulary showed no significant correlation with overall target-advantage scores, $r = .22, p = .19$. We therefore collapsed across vocabulary group in subsequent analyses.

Table 2 shows the target-advantage scores averaged across all verbs, as well as separately for each verb. Averaged across all four verbs, children’s target-advantage scores were significantly greater than 0, $t(35) = 3.010, p < .01$. Furthermore, the means for all four verbs were positive, showing a consistent preference for the target over the distractor (see Table 2). Despite the referential ambiguity of each trial, children were able to gather information about the events that occurred with each verb and update that information across trials to identify the verb’s referent.
Similar results emerged when we analyzed block 1 alone: Children’s overall target-advantage scores were again significantly greater than 0, \( t(35) = 2.813, p < .01 \), and each verb showed a mean preference for the target over the distractor. Recall that for each verb, there were three unique event tokens, each featuring a different actor. Each event token occurred once in the first block of trials, and again in the second block. As a result, during the second and third presentation of each verb, children already had some opportunity to collect referential information, but had not yet seen any repeated event-tokens. Thus, the appearance of this effect in the first block indicates that despite variability in the actor performing each action, children were capable of using cross-situational consistency to identify the referents for the novel verbs.

**Experiment 2**

In Experiment 1, 31-month-olds were able to use cross-situational consistency to identify the referents of novel intransitive verbs under referential uncertainty. In Experiment 2, we sought to explore the limits of this ability by testing whether children could also identify the referents of transitive verbs under these circumstances. Transitive verbs present several additional sources of potential difficulty for children. First, because transitive sentences contain an additional argument, they might impose a greater language processing load than the simple intransitive sentences used in Experiment 1. Second, in order to learn a transitive verb, children need to abstract not only across actions performed by different actors, but also across variation in the object acted upon and variation in the details of the action when enacted on different objects. This additional variability may make detecting cross-situational consistency more challenging. Thus, this experiment should provide an initial hint as to how well children’s ability to attach referential information to verbs under uncertainty scales up to more difficult learning conditions.
On each trial, children saw two different actors, each performing a novel action on a distinct object. The events were accompanied by two novel transitive verbs (e.g., “She’s nading her toy…and she’s pimming her toy.”; Figure 3). Across trials, the actor and object performing the action varied. As in Experiment 1, the only cue to the verb’s referent was the consistent pairing of the verb with a particular action. If children can use cross-situational comparison to identify the referents of transitive verbs under uncertainty, then when presented with a verb, they should look longer at the event that consistently co-occurred with that verb across trials.

**Method**

**Participants**

Thirty-six 31-month-olds participated (mean = 30.9, range 29.9-33.1, 18 male, 18 female). All were native speakers of English. An additional 3 children were tested but eliminated because they were fussy (1), distracted (1), or because their overall target-advantage score was more than 3.5 SD from the mean (1). Vocabulary scores ranged from 15 to 100 with a median of 82.5. Equal numbers of children were randomly assigned to each of four combinations of experimental list and verb order (see procedure below). An equal number of boys and girls was assigned to each of these four combinations.

**Apparatus**

The apparatus was identical to that used in Experiment 1.

**Materials and Procedure**

Materials consisted of color videos of four women performing novel actions on one of four different toys; the pairing of woman and toy was consistent across trials. There were four novel actions, each paired with a novel transitive verb. The pairing of verb and action varied across experimental lists (see Table 3). Each woman performed three of the four novel actions.
for a total of 12 unique event tokens. Tokens were presented in synchronized pairs accompanied by a soundtrack recorded by a native English speaker. Left-right positioning of event tokens was counterbalanced within verb order and list.

The procedure was identical to that of Experiment 1 with two exceptions. First, the continuous study-test procedure was preceded by two display trials. On the first display trial, children saw two of the female actors, each holding her toy, and heard a soundtrack calling attention to each actor and her toy (e.g., “Oh look! She has a toy! And she has a toy!”). After a 3-second blank-screen interval, children saw a second display trial in which the other two female actors and their toys were introduced in the same manner. Each display trial lasted 8 seconds. The purpose of these trials was to familiarize children with the actors, their toys, and the fact that the objects would be referred to as “toys” throughout the experiment. Second, during the continuous study-test phase, children heard the novel verbs used in transitive rather than intransitive sentences. All sentences were of the form She’s verbing her toy; the second sentence in each pair was recorded with contrastive stress on the subject, verb, and possessive pronoun (e.g., She’s pimming her toy. And she’s nading her toy!). The use of the pronoun “she” and the generic noun “toy” ensured that the sentences were ambiguous and could refer to either event.

Analyses

As in Experiment 1, we coded where children looked (left-screen, right-screen, away) frame by frame from silent video. 16 children’s data were independently coded by a second coder. The first and second coders agreed on the children’s direction of gaze for 98% of coded video frames.

Within each trial, we examined children’s fixations within two 2.5-second observation windows that began at the onset of each novel verb. Observation windows in which the children
looked away from the screen for more than 67% of the 2.5-second interval were dropped from the analyses (20 observation windows, 2.3% of the 864 possible observation windows). A preliminary ANOVA on time spent looking away with observation window (1, 2) and trial (1-12) as within-subject factors revealed significant interaction of window and trial, $F(11, 231) = 2.146, p < .025$; the main effects of window and trial were not significant, both $Fs < 1.4, ps > .2$. Because looking times varied across windows and trials, we again conducted our analyses on raw target-advantage scores.

For each observation window, we calculated a target-advantage score (time to target – time to distractor). These scores were averaged across the second through sixth presentations for each verb. Preliminary analyses revealed no effects of sex, list, or verb order. These factors were not examined further.

**Results and Discussion**

Figure 4 shows the proportion of fixations over time to each of the events (verb1 match, verb2 match) averaged across observations 2 through 6 for all verbs. Fixations are plotted separately for children whose vocabulary scores were above the median (high-vocabulary group; Figure 4A) and children whose vocabulary scores were below the median (low-vocabulary group; Figure 4B). Children in the high-vocabulary group performed similarly to the children in Experiment 1. Within the high-vocabulary group, fixations to the verb1 match began to increase about 800 ms after the onset of the first verb. Looks to the verb1 match persisted until about 400 ms before the onset of the second verb, at which point there was an increase in the proportion of fixations to the verb2 match. Children in the high-vocabulary group then maintained a preference for the verb2 match throughout the remainder of the trial. Thus, the high-vocabulary children appeared to understand the structure of the task.
The low-vocabulary group showed a different pattern. About 800 ms after the onset of the first verb, the proportion of fixations to the verb1 match increased slightly in the low vocabulary group. However, this preference for the verb1 match vanished by 1000 ms after verb1 onset, only returning a second later (2s after verb1 onset), at which point the low vocabulary group again shifted their attention to the verb1 match. This preference for the verb1 match was then maintained until 2s after the onset of the second verb, when there was a shift in fixations to the verb2 match. This pattern does not strongly suggest that the low-vocabulary children identified the referent for either verb within the trial, or even that they learned that each event was labeled on each trial.

Table 4 shows the target-advantage scores, separately by vocabulary group. Overall target advantage scores were significantly higher for high- than for low-vocabulary children, $F(1, 34) = 8.943, p < .01$. We also found a reliable correlation between vocabulary and overall target-advantage scores, $r = .38, p < .025$. Given these differences, we conducted separate analyses for the high- and low-vocabulary groups.

Analysis of the data for children whose vocabulary was above the median revealed a pattern similar to that found in Experiment 1. Averaged across all verbs, high-vocabulary children’s target-advantage scores were significantly above 0, $t(17) = 3.639, p < .005$. Examination of the individual verbs shows that all means were positive, demonstrating a consistent preference for the target over the distractor (see Table 4). Furthermore, this effect emerged within the first block: averaged across all verbs, high-vocabulary children’s target-advantage scores were significantly above 0, $t(17) = 3.109, p < .01$, and the means were again positive for all verbs. This suggests that the high-vocabulary children were able to encode
information about the referents for transitive verbs under referential uncertainty, despite variation in both the actor and the object in the events.

Performance was much worse in the low-vocabulary group. Averaged across all verbs, low-vocabulary children’s target-advantage scores did not differ significantly from 0, $t < 1$. In fact, the overall mean, as well as that of 3 of the 4 verbs, was negative, with children looking longer at the distractor than at the target event. The low-vocabulary children failed to identify the referents of the transitive verbs in this experiment, suggesting that this experiment was much more difficult than Experiment 1. The added difficulty may have resulted from the increased linguistic difficulty of the present task. By adding an extra argument to each sentence, we may have overwhelmed the low-vocabulary children’s limited language-processing abilities, making it harder for them to process each verb. Another possibility is that their difficulty was not uniquely linguistic, but instead arose from limitations in other systems that are correlated with vocabulary level. Recent work by Marchman and Fernald (2008) has shown that vocabulary size is positively correlated with working memory, for example. If the low-vocabulary children in this experiment also tended to have lower working memory capacity, then they may have had difficulty encoding, retrieving, and comparing the more complicated sentences and events they encountered in Experiment 2.

**General Discussion**

When children encounter a new verb, the referential scene is likely to offer many potential interpretations. Researchers have long assumed that children can constrain these interpretations by noting which scene elements consistently co-occur with the verb across referential contexts (i.e. cross-situational consistency). However, using cross-situational consistency to interpret verbs is not an easy task: children must attach complex event information
to a lexical entry under referential uncertainty, and then update that information when they encounter that verb again in the future. In doing so, children must be able to detect consistencies across referential contexts despite irrelevant variation, such as changes in the actors and objects involved in the events.

Here we investigated whether children could overcome these difficulties and use cross-situational observation to identify the referents of novel verbs. In two experiments, children repeatedly encountered pairs of events accompanied by two novel verbs. While on any given trial either novel verb could refer to either event, across the experiment, each verb consistently occurred with only one action. Experiment 1 tested children’s ability to identify the referents of intransitive verbs that described solo actions. Results showed that when children encountered a verb, they looked longer at the target referent than expected by chance. This effect emerged within the first block, before children encountered any repeated event tokens, suggesting the children were able to identify which referent consistently occurred with the verb even when they had to abstract across irrelevant variation in the actor performing the action.

Experiment 2 examined whether children could identify the referents for transitive verbs that described actions on objects. Here the results revealed an effect of vocabulary. Children whose vocabulary was above the median looked longer at the target event than expected by chance, and this pattern again emerged in the first block, before children saw any repeated event tokens. Thus, the high vocabulary children were able to identify the referents for the transitive verbs, even when they had to abstract across the objects and actors involved in the events. Children whose vocabulary was below the median failed to learn the verbs in Experiment 2. This failure suggests that the increase in complexity of the sentences and the scenes made it too difficult for them to keep track of the pairings between verbs and actions.
The fact that children succeeded in these tasks is striking. Setting aside the low-vocabulary children in Experiment 2, 31-month-old children were able to use cross-situational consistency to identify the referents of both intransitive and transitive verbs, despite the fact that they never received a single unambiguous labeling trial. While previous experiments have shown that children can use cross-situational consistency to identify the referents of novel nouns under referential uncertainty (Smith & Yu, 2008), the present experiments are the first to demonstrate that children can also do so with novel verbs. Our results show that children can encode information about potential event referents that occurred with a verb, without knowing which event the verb described. They were then able to update that information across multiple ambiguous trials in order to identify the verb’s referent. In addition, children in both experiments were able to successfully locate the correct referents within the first block of trials, demonstrating that they can compare across observations even when the actors and objects change across referential scenes.

These results complement several recent studies showing that under referential uncertainty, children can encode facts about the sentences in which verbs occur (Arunachalam & Waxman, 2010; Scott & Fisher, 2009; Yuan & Fisher, 2009). For instance, Yuan and Fisher (2009) presented 28-month-old children with dialogues in which two people repeatedly used a made-up verb in transitive sentences (e.g., A: ‘Anna blicked the baby!’ B: ‘Really, she blicked the baby?’) or intransitive sentences (A: ‘Anna blicked!’ B: ‘Really, she blicked?’). The children later heard the verb used in isolation (‘Find blicking!’) while they viewed test videos depicting two referential options: a two-participant causal action (one girl lifted and lowered another girl’s leg) and a one-participant action (a girl made arm-circles). Children who had previously heard the verb used in transitive sentences looked longer at the two-participant event than did children
who had heard the verb used in intransitive sentences. These results suggest that even in the absence of any referential information about a verb’s meaning, children can form a lexical entry and attach to that entry information about the types of sentences that the verb occurs in. Children can then integrate these facts with later scene information in order to identify the verb’s referent.

Together with the present results, these studies suggest that children routinely encode facts about a new verb, even when they have little to no information about the verb’s referent. This paints a picture of a developing lexicon in which partial representations of verb meanings gather up information about how the verb is used, even before the child could be said to know the verb in the usual sense. These experiments add to this picture by showing that children can encode information about a verb’s potential referents, even before they know which aspect of a referential scene a verb describes.

However, our results also suggest that there are clear limits on children’s ability to encode and compare potential referents under uncertainty. The small increase in complexity of the sentences and scenes used in Experiment 2 overwhelmed the low vocabulary children, preventing them from showing learning of any of the verbs in this experiment. Note that this difficulty occurred, despite the fact that we attempted to simplify the learning situation in several ways. The referential uncertainty in our task was much lower than in a typical language-learning situation because children saw only two referents per trial. We also attempted to make it easier for children to compare across trials by making the different tokens of each action quite similar, as well as making the four actions very distinct from one another so that they were not confusible.

In addition, the presence of two events and two verbs in each trial served as a source of implicit contrast information. Although children were not explicitly told that one verb applied to
one event and the other applied to the other event, this probably became clear from the structure of the task. The fact that children in Experiment 1 and the high-vocabulary children in Experiment 2 reliably switched from one referent to the other within a trial suggests that they noticed this contrast. Providing children with a positive exemplar and a contrasting event, as our task effectively did, has been shown to facilitate word learning. While 3-year-olds failed to extend a newly learned verb to a novel instance of an action when trained with only a single exemplar (Imai, Haryu, & Okada, 2005; Piccin & Waxman, in prep.), they succeeded if they were also shown a contrasting event to which they were told the verb does not apply (e.g., *Uh-oh, he’s not larping that*; Piccin & Waxman, in prep.; see Waxman & Klibanoff, 2000 for evidence that contrast also supports the extension of novel adjectives). The presence of implicit contrast in each trial of the current experiments should have made it easier for children to extend the verbs to new event tokens involving different actors and objects.

Thus, our results suggest that while children can attach referential information to a verb’s lexical entry under referential uncertainty, this mechanism is limited, even in 2.5-year-old children. In situations of greater referential uncertainty, as might be expected in natural language learning, this mechanism may require support from additional sources of information. One potentially useful source of information is prior experience with the sentences and nouns that occur with the verb. If children have heard the verb used in sentences in the past, then this might make it easier for them to process future sentences in which the verb occurs. This should free up processing resources, making it easier for children to encode potential referents. Knowledge of the sentence structures that the verb occurs in may also enable children to form a conjecture about the verb’s meaning (e.g., the verb describes a relationship between two participants). Children could then direct their attention to subsets of the referential scene that were consistent
with their conjecture. This should ease processing difficulties by limiting the number of potential referents that children need to encode and later retrieve. We intend to explore this possibility in future experiments.

A question of mechanism: Two conjectures

Although our results demonstrate that children can use cross-situational consistency to learn novel verbs under referential uncertainty, it remains unclear how they do so. Here we outline two broad hypotheses about how cross-situational consistency might operate.

The first possibility, true cross-situational learning, is cross-situational comparison as it has traditionally been construed – a process of intersection discovery (Pinker, 1984; Siskind, 1996; Smith & Yu, 2008). According to this view, when children encounter a new verb, they note whatever co-occurred with that verb. The next time they encounter the verb, they compare the current set of co-occurrences with the previous set; any element that is not consistent across the two sets is rejected as a plausible component of the verb’s meaning. Co-occurrence information could come in the form of simple associations between event-components and words (e.g., Smith & Yu, 2008), or it could consist of a set of candidate interpretations that children generate using other cues such as syntactic structure (e.g., Siskind, 1996). Regardless of precisely what children are tracking, this model of cross-situational comparison makes at least two predictions about learning. First, children should gain information about a verb’s meaning from each encounter (although some observations may be more informative than others). Second, learning should be an incremental process in which children gradually converge on a verb’s correct meaning.

An alternative possibility is that children are not capable of storing and comparing all possible sets of potential referents, and thus are unable to engage in true cross-situational
learning as described above. Instead, children may engage in a form of *approximate* or *conjecture-based cross-situational learning* (Medina, Trueswell, Snedeker, & Gleitman, 2009; Smith, Smith, & Blythe, 2009). When children encounter a new verb, children may generate a conjecture or guess about that verb’s meaning. It is this conjecture that is stored and refined (or falsified) over subsequent observations. Children may also store some amount of information about the potential alternative referents, depending on their available resources. If they later discover that their guess was wrong (e.g., they encounter the verb without the hypothesized referent), then they must form a new conjecture based on the information in the present scene and any information they have retained from previous observations. This mechanism also generates at least two predictions. First, it predicts that in order to benefit from a particular observation, children must be able to generate a conjecture about the verb’s meaning. Second, while this mechanism does make use of cross-situational comparison, it is approximate in the sense that not all information is assumed to be retained across observations (in some cases, only the children’s conjecture is retained). As a result, children may sometimes entertain interpretations of a verb that are inconsistent with the previous referential contexts in which that verb has occurred.

Although our results cannot distinguish between these two alternatives, several recent experiments support the approximate cross-situational learning mechanism. In one experiment, Medina et al. (2009) tested adults’ ability to learn novel nouns using the human-simulation paradigm. Participants saw a series of silent video clips of parent-child interactions and had to guess which word the parent had just uttered. For each noun, adults saw five clips that varied in their informativeness: one clip was highly informative, while the remaining four were low on informativeness (informativeness was established in a separate norming study; informative clips
were those on which participants guessed the word correctly at least 50% of the time). Across participants, the sequence of clips varied such that the informative clip could be either the first, third, or fifth observation for a particular noun. If participants were using true cross-situational observation, then they should not have been affected by the order of observations: all participants saw the same clips, so they should be equally good or bad at guessing the word’s meaning. Instead, participants who viewed the highly informative clip first performed significantly better than participants who saw the informative clip in the middle or at the end of a sequence. This suggests that participants needed to see the highly informative clip first so that they could generate a conjecture about the word’s meaning. This enabled them to make use of later less informative observations.

Smith et al. (2009) also compared the true and approximate cross-situational learning strategies in the context of novel noun learning. Adults saw a set of objects on a screen and heard a novel label (e.g., *kupo*). They were then shown an array of 15 objects and asked to click on the object that went with that label (e.g., *Click on the kupo*). Two different difficulty factors were manipulated within participant. First, different words were learned under different levels of referential uncertainty: labels could be presented with 3, 6, or 9 objects on the screen. Second, some words were presented serially (all three observations of the word in a row) and others were interleaved (other words intervened between observations of a given word). The patterns of each participant’s responses were then examined in order to determine which learning strategy they used. Results suggested that participants used true cross-situational learning only at the lowest level of difficulty (e.g., 3 possible referents, serial presentation). At all higher levels of difficulty, participants reverted to an approximate cross-situational strategy.
These results suggest that even adults may not be able to use true cross-situational observation in a noun-learning task. If this is the case, then it seems highly unlikely that children, who have more limited computational resources (e.g., working memory) and far less linguistic knowledge would be able to use this type of mechanism to acquire verbs. Instead, it seems likely that for children, all learning situations are more like the high referential uncertainty conditions of Smith et al.’s (2009) experiment and that they must therefore rely on an approximate cross-situational learning mechanism. In future experiments, we hope to test this possibility by increasing the number of potential referents, thus enabling us to model responses. Hopefully this will shed light on which type of mechanism children are using.
CHAPTER 3

USING MULTIPLE SENTENCE FRAMES TO LEARN VERBS

For any given utterance of a verb, the referential scene offers a wide array of potential interpretations. The syntactic bootstrapping hypothesis (Landau & Gleitman, 1985) maintains that children could constrain these interpretations by exploiting systematic links between syntactic structure and verb meaning. Many experiments provide support for this hypothesis. For example, children interpret a novel verb in a transitive sentence (but not an intransitive sentence) as referring to a relationship between two participants (Fisher, 2002; Naigles, 1990; Yuan & Fisher, 2009) and use knowledge of English word order to map a novel verb to an event in which the participants play the correct roles (Gertner et al., 2006).

However, a single sentence provides only highly abstract information about the meaning of a new verb. To illustrate, consider encountering a novel transitive verb, as in “John blicked the doctor.” Occurrence in the transitive frame indicates that blick describes a relationship between two participants. A wide range of meanings is consistent with this constraint: blicked could mean anything from killed or murdered to loved or visited. Intransitive sentences pose similar problems. Occurrence in an intransitive sentence like John blicked indicates only that the verb describes some state or activity that the subject of the sentence is engaged in. It specifies neither the nature of this state/activity, nor whether the subject of the sentence is an actor (e.g., John cooked) or an undergoer (e.g., John fell). Because many diverse meanings can map onto transitive and intransitive sentence frames, occurrence in either leaves the meaning of blick ambiguous.

Of course, observing the referential scenes that accompany a verb reduces this ambiguity somewhat. If children hear the sentence John blicked the doctor in the context of John visiting a
doctor, it seems unlikely that they will entertain ‘murder’ as a potential meaning for *blick*, even though it is consistent with the structure of the sentence. However, there are still many things that *blick* could mean within this referential scene (e.g., see, like, visit).

One way that children might reduce this residual ambiguity is by considering the full set of syntactic frames in which a verb occurs (Landau & Gleitman, 1985). Many verbs occur in more than one sentence structure, and each structure that a verb occurs in provides additional constraint on its meaning. Consider the verb *explain*.

1) Mary explained that he left.
2) Mary explained the problem to John.

As shown in (1), *explain* can occur with sentential complements, which typically indicates that the verb takes an argument with propositional content (Landau & Gleitman, 1985). *Explain* can also occur in the three-argument structure seen in (2), where it takes both a direct and an indirect object. This structure is consistent with transfer or motion towards a goal (e.g. *give, put*; Fisher, Gleitman, & Gleitman, 1991; Levin, 1993). The combination of these two frames strongly constrains the interpretation of *explain*: it refers to the transfer of propositional content (Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005).

Landau and Gleitman (1985) argued that because the information provided by individual frames is highly ambiguous, the use of multiple frames could play a major role in verb learning. Here we sought to test whether (a) children can actually use the set of frames that a verb occurs in to draw appropriate inferences about its meaning and (b) whether the set of frames provides stronger constraint on verb meaning than a single sentence frame. In order to explore these questions, we chose to examine sets of sentence frames that children would be likely to comprehend. Previous evidence suggests that children draw appropriate inferences about novel
verbs presented in either transitive or intransitive sentence frames (Naigles, 1990; Naigles & Kako, 1993; Yuan et al., in prep.; Yuan & Fisher, 2009). Given these findings, we asked whether children could appropriately interpret verbs that occurred in both of these sentence frames—verbs that alternate in characteristic ways between transitive and intransitive uses. As described above, occurrence in either a transitive or intransitive sentence can leave the meaning of a verb ambiguous. Does encountering a verb in both of these sentence frames help children reduce this ambiguity?

Previous research on this question (Naigles, 1996, 1998) has focused on two specific types of transitivity alternations: the causal (3) and unspecified-object alternations (4). Verbs that participate in the causal alternation describe an action on an object that causes a salient effect on the object (e.g., an externally-caused motion such as bounce or a change of state such as break); both the action and the resulting change are part of the verb’s meaning. The actor’s action on the object can be described by a transitive sentence with the actor as the subject and the recipient of the action (the undergoer) as the object (3a). The effect on the undergoer can be described separately in an intransitive sentence with the undergoer as the subject (3b). Verbs that participate in the unspecified-object alternation also describe an actor acting on an object and appear in transitive sentences (4a). However, unspecified-object verbs typically describe ongoing activities and include only the actor’s action in their meanings, without reference to any particular effect on the object acted upon. For instance, while poking someone could have an effect (e.g., the person poked becomes irritated), only the action is described by the verb poking; this action would still be considered poking even in the absence of an effect. Consistent with the intuition that these verbs do not include an effect in their meanings, they cannot appear in undergoer-subject intransitive sentences. Instead, the actor’s activity can be described separately,
without any reference to the object acted upon, in an intransitive sentence with the actor as the subject (4b). Unspecified-object verbs vary more in their meanings than do causal verbs, but many describe contact actions (e.g., dust, poke, push).

3) a. Anne broke the lamp.
   b. The lamp broke.

4) a. Anne dusted the lamp.
   b. Anne dusted.

If children were to encounter a new verb in one of these alternations, how might they use the set of sentence frames to draw conclusions the meaning of that verb? Researchers have suggested two distinct mechanisms by which sets of frames could assist verb learning: a category-mediated mechanism and a direct-inference mechanism.

Proponents of the category-mediated mechanism suggest that children could use the set of sentence frames that a verb occurs in to assign meaning via previously-learned verb categories (Pinker, 1989, 1994). The causal and unspecified-object alternations define distinct categories of verbs that share syntactic and semantic similarities. While comparable verb categories are found across the world’s languages, the particular set of verbs that can participate in each alternation is constrained by language-specific semantic restrictions and thus the categories must be learned (Pinker 1989). For instance, the verb bounce (It bounced/She bounced it), along with other verbs of manner of motion, participates in the causal alternation in English, but the verb fall (It fell/*She fell it), along with other verbs of inherently-directed motion, does not. Different restrictions apply in other languages.

Once children have created a verb class that has both syntactic and semantic properties, they can use it to make inferences about new words (e.g., Brooks & Tomasello, 1999; Gropen,
Pinker, Hollander, Goldberg, & Wilson, 1989; Pinker, Lebeaux, & Frost, 1987). For instance, children could learn that in English, some verbs that can be both transitive and intransitive share similar distributional properties: the same nouns tend to occur as both the transitive objects and the intransitive subjects of these verbs (i.e. causal alternation verbs, see 3 above). Children may also note that this class of transitivity-alternating verbs shares the meaning “action culminating in a noteworthy change.” Upon encountering a new verb that demonstrates the distributional properties of the causal alternation, children could extend the meaning of the category to this new verb. A similar process could be used to assign meaning to unspecified-object verbs, though here the inferences that children could draw are less constrained. Verbs in the unspecified-object category share the semantic property of not entailing a specific effect, but the meanings of these verbs are diverse, as they describe a broad range of activities. Thus, upon encountering a verb with the distributional properties of the unspecified-object alternation, children could infer that the verb described an activity and not an effect, but they may not be able to determine the nature of that activity.

Other researchers have proposed that children could use a direct-inference mechanism to interpret sets of sentence frames (e.g., Fisher et al., 1994; Fisher, 1996). This account assumes that some properties of sentences are inherently meaningful, allowing children to draw inferences about verb meaning directly without reference to previously-learned verb classes. For instance, children might possess an unlearned assumption that the number of nouns in a sentence corresponds to the number of participants involved in the event that the sentence describes (Fisher, 1996; Lidz, Gleitman, & Gleitman, 2003). Using this assumption of one-to-one mapping between nouns and participant roles, children could infer that both the causal and unspecified-object alternations describe events that could be construed in terms of either an action on an
object (e.g., 3a, 4a) or the action or state of a single individual (e.g., 3b, 4b). Children could gain further insight into the verb’s meaning by considering the semantic role assigned to the subject of the intransitive sentence in each alternation. The subject of (3b) is an undergoer, indicating that the sentence describes something that happened to the lamp. Taken together with the transitive sentence, this suggests that the verb describes an action on an object that causes some sort of change or effect in that object. In the case of the unspecified-object alternation, the subject of the intransitive sentence (4b) is an actor. This suggests that the verb describes an action on an object where the actor’s activity is sufficiently interesting to deserve description in a separate sentence.

Note that in order to use this direct-inference mechanism to interpret the causal and unspecified-object alternations, children must be able to identify the semantic roles of participants in sentences. If children only considered the surface forms of these sentences, then the direct inferences that they could draw about the two alternations would be identical: they both involve a transitive and intransitive sentence, indicating that they describe both a two-participant relation and the action of a single individual. In order to differentiate between the two, children need to be able to determine the semantic role assigned to the subject of the intransitive sentence of each alternation. One potential source of role-relevant information is the referential scene. If both sentences of an alternation occur in the presence of a relevant referential scene (see Figure 5), then children could use the overlap between the two sentences and the referential scene to infer that both sentences probably describe the same set of participants (e.g., the same girl and the same boy) and that the roles played by individual participants do not change from one sentence to the next. For the scenario depicted in Figure 5, this would lead
children to infer that the subject of the second sentence, *The boy is pimming*, is an undergoer rather than an actor.

Prior experimental evidence suggests that both category-mediated and direct-inference mechanisms play a role in word learning. Older children and adults possess learned categories for a variety of verb alternations, and these categories affect how novel verbs are produced and interpreted (e.g., Ambridge, Pine, Rowland, & Young, 2008; Brooks & Tomasello, 1999; Gropen et al., 1989; Pinker et al., 1987). While no one has yet examined the presence of these language-specific verb classes in children under three, there is evidence that category-mediated inference plays a role in learning other classes of words from early in infancy. For instance, 14-month-olds use learned distributional categories to determine whether a novel word refers to an object (e.g., “This is a blicket!”) or a property (e.g., “This is a blickish one!”; Booth & Waxman, 2003). Similarly, 2.5-year-olds can use surface properties of sentences, such as whether the sentence is transitive or intransitive, to make direct inferences about verb meaning (i.e. whether the verb describes a two-participant relation or a solo action; Yuan & Fisher, 2009).

Given these findings, we do not attempt here to determine which mechanism children use to interpret sets of sentence frames, as it seems likely that children draw on a combination of these two mechanisms. Instead, we focus on predictions that follow from both of these mechanisms about how the use of multiple sentence frames should operate. Regardless of which mechanism children use, encountering a verb in the causal alternation should tell children that the verb describes an action on an object that produces a salient effect. Returning to the scenario in Figure 5, this would mean that hearing “*The girl is pimming the boy. The boy is pimming.*” should direct children’s attention towards the caused-motion event, as this is the only referent with the proper action + result structure. In contrast, the unspecified-object alternation should
focus children’s attention on an actor’s activity rather than on an effect. If children were to hear “The girl is pimming the boy. The girl is pimming.” while viewing the events in Figure 5, they would be able to conclude that the verb described either the girl’s dusting action in the contact-activity event or her pushing action in the caused-motion event. While these two alternations could potentially direct children’s attention to the same referent within this scene (i.e. the caused-motion event), they result in different interpretations of the verb: the causal alternation should result in an interpretation like “cause to bend”, which includes the effect in the verb’s meaning, while the unspecified-object should result in an interpretation like “push”, which does not include the effect.

As can be seen from this example, the causal alternation should provide greater constraint on verb meaning than the unspecified-object alternation because the causal alternation picks out a very restricted set of events as candidate referents. Thus, the causal alternation should be informative across a range of referential contexts, consistently directing children’s attention towards an event with an action + result structure. In contrast, the unspecified-object alternation is consistent with a wider array of interpretations, and its informativeness is likely to vary as a function of the referential options with which it occurs. If children encounter a verb in the unspecified-object alternation while viewing a referential scene that contains multiple activities, they may not be able to determine which activity the verb describes.

Two previous studies have tested children’s interpretation of the causal and unspecified-object alternations (Naigles, 1996, 1998). Naigles (1996) showed 28-month-olds training scenes that contained both a caused-motion and contact-activity action. For instance, one training scene showed a frog pushing down on a duck’s shoulder with one hand, causing him to bend (caused motion), while simultaneously patting the duck on the head with the other hand (contact
activity). Each scene was accompanied by a novel verb in one of 4 different sentence structures (see 5-8 below). After viewing each scene, children saw test trials in which the caused-motion and contact-activity actions were separated and presented on different screens. Children were then asked to find the referent of the novel verb (e.g., *Find sebbing*).

5) Causal alternation: The frog is sebbing the duck. The duck is sebbing.

6) Unspecified-object alternation: The frog is sebbing the duck. The frog is sebbing.

7) Isolated verb: Look! Sebbing!

8) Actor-subject intransitive: The frog is sebbing.

If children used the set of sentences that the verb occurred in to draw conclusions about which component of the training event the verb described, then children who heard the verb used in the causal alternation (5) should have inferred that the verb referred to an action + result and they should therefore have interpreted the verb as describing the caused-motion action. In the test trial, these children should have looked longer at the caused-motion event than the children in the isolated verb condition (7), who provided a baseline measure of children’s preference for the two events. In contrast, children who heard the unspecified-object alternation (6) should have inferred that the verb focused on an agent’s activity without entailing a particular result. Because this inference is consistent with both actions, these children should have looked relatively more at the contact-activity event than children in the causal condition. Children who heard only the actor-subject sentence (8) should not have had enough information to determine which event the verb described and should therefore have demonstrated looking patterns similar to those of the children in the isolated verb condition.

The results showed that children in the isolated verb condition looked significantly longer at the caused-motion than at the contact-activity event, suggesting that children had a strong
baseline preference for the caused-motion event. Children in the causal and actor-subject intransitive conditions also looked longer at the caused-motion than at the contact-activity event, and the magnitude of their preference did not differ from that demonstrated by children in the isolated verb condition. Thus, contrary to what was predicted, children who heard the verb used in the causal alternation did not look longer at the caused-motion event than children who heard an isolated verb. Interestingly, the children in the unspecified-object condition did differ from children in the isolated verb condition. Specifically, girls (but not boys) who heard the verb used in the unspecified-object alternation looked equally at the two events, and this pattern differed significantly from that demonstrated in the other three conditions. While this effect only emerged for the girls in Naigles (1996), this effect was later replicated with boys who were 33 months old (Naigles, 1998).

These results suggest that hearing a novel verb used in the unspecified-object alternation influenced children’s interpretation of that verb, causing them to look relatively longer at the contact-activity event than they would have in the absence of useful sentence information (i.e. the isolated verb condition). Furthermore, children in the unspecified-object condition demonstrated a different pattern from those in the actor-subject intransitive condition. This provides evidence that hearing a verb used in unspecified-object alternation results in a different interpretation of that verb than hearing the verb in just one of the alternation’s component frames. However, hearing a verb used in the causal alternation did not appear to influence children’s interpretation of the verb in this experiment, as the causal and isolated verb conditions did not differ.

While these results provide the first evidence that hearing a verb used in a set of sentence frames can provide greater constraint than hearing a verb used in a single sentence structure, the
particular pattern of results in this experiment is puzzling. If children were using either the
category-mediated or the direct-inference mechanism to interpret the sentences in this
experiment, then the causal alternation should have directed children’s attention to the caused-
motion event, since only that event had the proper action + result structure. In contrast,
unspecified-object alternation could have referred to either the contact-activity event or the
contact portion of the caused-motion event (e.g., pushing on the duck’s shoulders). Thus, the
causal alternation should have been more likely to differ from the baseline than the unspecified-
object alternation.

One possible reason why Naigles’ (1996, 1998) results differed from this predicted
pattern is that 28-month-olds do not yet comprehend the causal alternation. Relative to the
unspecified-object alternation, the causal alternation has a more complex sentence structure (i.e.
the intransitive sentence involves a non-standard mapping of an undergoer to subject position)
and a more complex semantic structure (i.e. it describes both actions and their associated results).
However, while the causal alternation may be more complex than the unspecified-object
alternation, it is still composed of two very simple sentence frames (i.e. transitive and
intransitive). Other sets of sentence frames, such as the ditransitive and sentential-complement
structures shown in (1-2), are more complex. Thus, the causal alternation is one of the simplest
sets of sentence frames that could play a role in verb learning. If 28-month-olds cannot interpret
this simple set of sentence structures, then this does not bode well for the utility of multiple
frames in early verb learning.

An alternative possibility is that 28-month-olds do understand the causal alternation, but
this understanding was masked by various properties of Naigles’ (1996, 1998) task. The pattern
of looking times could have resulted from a ceiling effect. That is, children in the causal
condition may have interpreted the verb as referring to the caused-motion event, but the set of frames could not influence their looking behaviors to any greater degree. This prevented the causal condition from differing from the baseline, which was already skewed towards the caused-motion event. It may also be that children in the causal condition did look longer at the caused-motion action than children in the baseline condition, but this effect emerged during the training period. Since both actions were presented on a single screen during the training phase, a correct interpretation during this phase cannot be assessed using children’s looking behavior.

The Present Research

The present research sought to extend the evidence for the use of multiple sentence frames in early verb learning by examining children’s understanding of the causal alternation. We began by asking whether children could appropriately interpret verbs presented in this alternation. We investigated this question by making several changes to the basic procedure used by Naigles (1996, 1998). First, we eliminated the composite training phase. Children first heard the verb in a set of test trials in which they saw a caused-motion event on one screen and a contact-activity event on the other. Because children always encountered the events on different screens, we were able to assess children’s interpretation of the verb from its first presentation.

Second, we sought to create sets of events in which the contact-activity event was at least as interesting as the caused-motion event\(^2\). By moving the baseline towards the contact-activity event, we hoped to reduce the possibility of problematic ceiling effects, making it easier to observe any effect of the causal alternation on verb interpretation.

We next asked whether, as predicted by Landau and Gleitman (1985), considering the full set of sentence frames that a verb occurs in provides children with more information about

\(^2\) This manipulation turned out to be more than effective, resulting in a baseline preference for the contact-activity event instead.
that verb’s meaning than considering just a single sentence frame. We focus here on whether the
causal alternation provides more constraint on verb interpretation than its intransitive frame (e.g.,
The boy is pimming in Figure 5; henceforth referred to as the boy-subject sentence). Intransitive
sentence frames simply indicate that the verb describes some action or activity of the subject,
without specifying whether the subject of the sentence is an actor (e.g., 4b) or an undergoer (e.g.,
3b). Intransitive sentences are thus consistent with a broad range of events. The causal
alternation picks out a restricted subset of these events (i.e. those with an action + result) and
should therefore provide more information about a verb’s meaning.

Previous findings suggest that when interpreting a sentence, children and adults consider
both the constraint provided by the sentence and the properties of the referential scene.
Specifically, several studies have shown that when a sentence is consistent with more than one
interpretation within a scene, children assume that the sentence describes the most prominent or
noteworthy\(^3\) of those potential interpretations (e.g., Fisher & Song, 2006; Gropen, Pinker,
Hollander, & Goldberg, 1991; Hollich, Hirsh-Pasek, & Golinkoff, 2000). If, however, a sentence
is consistent with only one interpretation within the scene, then the relative noteworthiness of
scene participants has much less of an effect on children’s interpretation of the sentence (e.g.,
Fisher & Song, 2006). Thus, children’s interpretation of weakly constraining sentence frames
(i.e. sentences that are consistent with many potential referents) is more strongly influenced by
properties of the referential scene than their interpretation of strongly constraining sentence
frames.

Based on these findings, we tested the relative constraint provided by the causal
alternation and boy-subject intransitive sentences by conducting two experiments using two

\(^3\) While prominence is somewhat complex and difficult to define, people tend to find components of the scene that
involve animate, brightly-colored, and moving objects to be more prominent than those involving inanimate, drab,
and stationary objects.
different sets of test events. In both experiments, children saw a caused-motion event (a girl pushing down on a boy’s shoulders, causing him to bend) and a contact-activity event (a girl dusting a boy’s back with a feather duster). In Experiment 3, the boy remained stationary while being dusted (Figure 5); in Experiment 4, he marched in place (Figure 6). Thus, in Experiment 3 the only noteworthy role for the boy was his role in the caused-motion event (see Figure 5), while in Experiment 4 the boy also had a noteworthy solo-action role in the contact-activity event (see Figure 6). We predicted that because the boy-subject intransitive sentence is weakly constraining, children’s interpretation of this sentence should be affected by the particular set of referential options provided at test. In other words, the boy-subject intransitive sentence should result in different interpretations of the verb across the two experiments. In contrast, hearing a verb used in the causal alternation should consistently direct children’s attention to the caused-motion event in both experiments, regardless of how noteworthy the boy is in the contact-activity event.

**Experiment 3**

In this experiment we asked whether, given a set of events very similar to those used in Naigles (1996), children could demonstrate understanding of the causal alternation. 28-month-old children saw the caused-motion and contact-activity events described above (see Figure 5). We added an instrument to the contact-activity event to make that event more attractive, thus moving children’s baseline preference away from the caused-motion event. These events were presented simultaneously in two test trials and accompanied by the novel verb *pimming* used in the causal alternation (*The girl is pimming the boy. The boy is pimming.*), the unspecified-object alternation (*The girl is pimming the boy. The girl is pimming*), a boy-subject intransitive sentence (*The boy is pimming*), or a girl-subject intransitive sentence (*The girl is pimming*). A fifth group
of children heard only neutral audio (e.g., *Look! What’s going on?*), allowing assessment of children’s baseline preference for the two events.

If children can assign appropriate interpretations to verbs presented in these two alternations, then children who heard the causal alternation should look longer at the caused-motion event. In comparison to the causal alternation condition, children who heard the unspecified-object alternation should look relatively longer at the contact-activity event. We predicted that children who heard only intransitive sentences should interpret the verb as describing the event in which the subject of the sentence was most noteworthy. In order to assess the noteworthiness of the boy and the girl in the two test events, we showed 12 adults the stimulus videos and asked them to choose the video where the boy (or the girl; the order of these questions was counterbalanced) captured their attention more. When asked about the boy, 10 out of the 12 adults selected the caused-motion event. When asked about the girl, 9 out of 12 adults selected the contact-activity event. Based on these judgment data, we predicted that children who heard the boy-subject intransitive sentence should look longer at the caused-motion event, while those who heard the girl-subject intransitive sentence should look longer at the contact-activity event.

**Method**

**Participants**

Forty 28-month-old children participated (mean = 28.0, range 27.0-29.9, 20 male, 20 female). All children were native speakers of English. Three additional children were tested but excluded from analyses due to a side bias (1) or because the average difference between their looking time to the caused-motion and contact-activity events was more than 2.5 SD away from the mean of their condition (2). Children’s productive vocabulary was measured using the short
form of the Macarthur CDI, Level II (Fenson et al., 2000). Vocabulary scores ranged from 23 to 96 with a median of 74. Eight children were randomly assigned to each of five conditions: causal, unspecified-object, boy-subject, girl-subject, and neutral. An equal number of boys and girls were assigned to each condition.

**Apparatus**

Children sat on a parent’s lap facing two 20-inch television monitors located about 30 inches away. The screens were separated by a 12 inch gap and placed at eye level. Soundtracks were presented from a central speaker. A camera hidden between the two screens recorded the children’s eye movements throughout the experiment. Parents wore opaque sunglasses, preventing them from biasing their child’s responses to the videos.

**Materials and Procedure**

Children viewed a synchronized pair of videos involving a boy and a girl. The left-right positioning of events was counter-balanced within each condition. The videos were accompanied by a soundtrack recorded by a native English speaker. The video sequence (see Figure 7) consisted of three phases: character-familiarization, practice, and test. The character-familiarization and practice phases were intended to familiarize the children with the actors, and to teach them that the soundtrack matched one event during each trial.

In the character-familiarization phase, a female actor was shown waving on one screen (4s) and was labeled twice (e.g., “There’s a girl!”) while the other screen remained blank. Following a 2-s blank-screen interval, the male actor was introduced on the other screen in the same manner (e.g., “There’s a boy!”). This was followed by two 4-s trials, separated by 2-s blank-screen intervals. In each trial, the girl appeared on one screen and the boy appeared on the
other. In the first trial, children were asked to “Find the boy!” and in the second trial they were asked to “Find the girl!”

The character-familiarization phase was followed by a brief practice phase involving two familiar intransitive verbs. The practice phase began with a 4-s interval in which both screens were blank. During this interval, the first practice action was labeled in the future tense (“The boy’s gonna jump!”). Children then viewed an 8-s trial in which the boy was jumping on one screen and pretending to sleep on the other; the boy’s jumping action was labeled three times (“The boy is jumping!”). This was followed by a 4-s blank-screen interval, during which the boy’s action was labeled in the past tense (“The boy jumped!”) and children were instructed to “Find jumping.” The children then viewed the jumping and sleeping events again (8s) and were asked to “Find jumping.” The second set of practice trials was identical to the first set except that children now viewed the girl clapping on one screen and the girl eating on the other and were asked to “Find clapping!”

The test phase began with a 7-s blank-screen interval, during which children heard either neutral audio, or the novel verb *pimming* used in one of four different sets of experimental sentences (see Table 5). This was followed by two 8-s test trials involving the two novel action events depicted in Figure 5. Both actions were repeated throughout the entire test trial. The trials were separated by 7-s blank-screen intervals in which children heard sentences in the past tense. In each condition, the set of sentences was presented four times and in three different tenses (past, present, and future).

**Coding**

We coded where children looked (left screen, right screen, away) frame by frame from silent video. To assess reliability, 10 children’s data were independently coded by a second
coder. The two coders agreed on the children’s direction of gaze for 98% of coded video frames. Any 8-s trial in which the child away for at least 50% of the trial was eliminated (n = 1).

The amount of time children spent looking away from the two video screens during the test trials was analyzed by an analysis of variance (ANOVA) with trial (1, 2) as a within-subject factor and condition (causal, unspecified-object, boy-subject, girl-subject, neutral) as a between-subjects factor. The analysis revealed a significant of trial, $F(1, 34) = 6.746, p < .025$; there was no effect of condition and no trial by condition interaction, both $F$s < 1. Children tended to look away from the screens more on the second trial ($M = .48$, $SD = .56$) than on the first trial ($M = .23$, $SD = .15$). Due to differences in time spent looking away across the two trials, we conducted our analyses on raw looking times to the caused-motion and contact-activity events rather than on the proportion of time spent looking to one event out of the total time spent looking at either event.

Preliminary analyses of test-trial performance revealed no effects of sex or whether vocabulary or practice performance was above or below the median. These factors were not analyzed further.

**Results and Discussion**

Table 6 shows the average time spent looking to the caused-motion and contact-activity events during each test trial, separately by condition. For each of the comparisons reported below, separate analyses were performed on looking times to the caused-motion event and looking times to the contact-activity event.

We first asked whether children assigned different interpretations to the novel verb when it was presented in the causal and unspecified-object alternations. The amount of time children spent looking at each event was analyzed using an ANOVA with trial (1, 2) as a within-subject
factor and condition (causal, unspecified-object) as a between-subjects factor. Both analyses showed a significant main effect of condition (caused-motion: $F(1, 14) = 7.761, p < .025$; contact-activity: $F(1, 14) = 9.303, p < .01$). No other effects were significant, all $Fs < 1.1, ps > .3$. Children who heard the verb used in the causal alternation looked significantly longer at the caused-motion event, and significantly less at the contact-activity event, than did children who heard the verb used in the unspecified-object alternation. This result demonstrates that children interpreted these alternations in appropriately different ways, recognizing that while the causal alternation systematically picks out events with an action + result, the unspecified-object alternation does not.

We next examined each alternation separately and asked whether (a) the alternation differed from baseline, (b) whether the relevant intransitive sentence differed from baseline, and (c) whether the alternation differed from its component intransitive frame. To address these questions, we divided the conditions into a causal alternation group (causal, boy-subject, and neutral conditions) and an unspecified-object alternation group (unspecified-object, girl-subject, and neutral conditions) and analyzed each group separately.

**Causal alternation group.** The means in Table 6 show that children who heard the causal alternation or the boy-subject intransitive looked longer at the caused-motion than at the contact-activity event. In contrast, children in the neutral condition preferred the contact-activity event. To test these differences, the amount of time children spent looking at each event was analyzed using an ANOVA with trial (1, 2) as a within-subject factor and condition (causal, boy-subject, neutral) as a between-subjects factor. Both analyses yielded a significant main effect of trial (caused-motion: $F(1, 20) = 4.445, p < .05$; contact-activity: $F(1, 20) = 6.673, p < .025$), and a significant main effect of condition (caused-motion: $F(2, 20) = 5.878, p < .025$; contact-activity:
The trial by condition interaction was not significant (caused-motion $F(2, 20) = 5.948, p < .01$). The trial effect appears to have resulted from the fact that children in all three conditions looked somewhat longer at the caused-motion event (and less at the contact-activity event) on the second trial than on the first trial. However, this effect did not interact with the effect of condition. We therefore performed planned condition comparisons on the average of the two test trials (see Table 6). Planned, two-tailed t-tests revealed that children who heard the verb used in the causal alternation looked significantly longer at the caused-motion event, and significantly less at the contact-activity event, than children who heard only neutral audio (caused-motion: $t(14) = 3.487, p < .005$; contact-activity $t(14) = -3.460, p < .005$). This suggests that hearing a verb used in the causal alternation biased children to interpret the verb as referring to the caused-motion event, providing evidence that children can use the causal alternation to draw inferences about verb meaning.

We next examined whether children’s interpretation of the boy-subject sentence was influenced by the relative salience of the boy in the two test events. The adult judgment data suggested that the boy was more noteworthy in the caused-motion event. Correspondingly, children who heard the novel verb used in the boy-subject sentence looked significantly longer at the caused-motion event, and significantly less at the contact-activity event, than children who heard neutral audio (caused-motion: $t(14) = 3.046, p < .01$; contact-activity: $t(14) = -2.939, p < .025$). Thus, as predicted, children interpreted the boy-subject sentence as describing the event in which the subject of the sentence was most noteworthy.

Comparison of the causal and boy-subject intransitive conditions revealed no significant effects, both $ts < 1$, suggesting that in this referential context both the causal alternation and the
boy-subject intransitive sentence alone directed children’s attention towards the caused-motion event. This does not mean, however, that children in the causal and boy-subject conditions arrived at the same interpretation of the novel verb. It seems likely that children who heard the causal alternation attended both to the causal relationship between the girl and the boy and the effect on the boy (i.e. interpreting the verb as meaning “cause to bend”) while children who heard the boy-subject sentence simply attended to the boy’s motion (i.e. interpreting the verb as “bend” or “squat”; see also Bunger, 2006; Bunger & Lidz, 2004, 2006). We return to this point in Experiment 4.

*Unspecified-object alternation group.* The means in Table 6 suggests that the unspecified-object, girl-subject, and neutral conditions all demonstrated a preference for the contact-activity event over the caused-motion event and the magnitude of this preference was similar across the three conditions. This was confirmed by analyzing the amount of time children spent looking at each event using an ANOVA with trial (1, 2) as a within-subject factor and condition (unspecified-object, girl-subject, neutral) as a between-subjects factor. There were no significant effects, all $F$s < 1, indicating that the three conditions did not differ.

The fact that the unspecified-object condition did not differ from baseline may have resulted from children’s baseline preference for the contact-activity event. Given this baseline preference, any additional biasing toward the contact-activity event by the unspecified-object alternation may have been difficult to measure. Alternatively, this finding may reflect the fact that the unspecified-object alternation is less constraining than the causal alternation. Regardless of whether children use the category-mediated or direct-inference mechanism to interpret this alternation, the unspecified-object alternation tells them only that the verb describes an action on an object and that the emphasis is on the actor’s activity. Since both of our test events depicted
an actor (the girl) performing an ongoing activity (e.g., pushing, dusting), children who heard the unspecified-object alternation could have construed the verb as referring to aspects of either test event.

Based on adults’ judgments, we predicted that the girl-subject sentence would bias children towards the contact-activity event. While the looking patterns are consistent with this prediction, this condition did not differ from baseline. This may have resulted from the fact that, as mentioned above, the neutral baseline was skewed towards the contact-activity event, making any additional effect of the girl-subject sentence difficult to measure.

Note that as with the causal group, we found no difference between the alternation and its intransitive frame. In this case, children in both the unspecified-object and the girl-subject conditions preferred the contact-activity event. This finding is not consistent with Naigles (1996, 1998), who reported a difference between the unspecified-object and actor-subject conditions in her experiment. We return to this issue in the General Discussion.

**Experiment 4**

The results of Experiment 3 showed that children can appropriately interpret verbs presented in the causal alternation. However, because the causal condition did not differ from the boy-subject intransitive condition, we cannot conclude based on Experiment 3 alone that the causal alternation provides greater constraint than its intransitive frame. In fact, one might argue that the children in the causal condition were not actually interpreting the set of sentence frames. In principle, they could have ignored the transitive sentence and interpreted the verb solely based on the intransitive sentence.

Experiment 4 thus had two goals. First, we sought to replicate the finding from Experiment 3 that children interpret a verb presented in the causal alternation as describing a
caused-motion rather than a contact-activity event. Second, we tested whether hearing a verb used in the causal alternation resulted in a different, more constrained interpretation of the verb than hearing it used in just an intransitive sentence. If the causal alternation is more constraining than the boy-subject intransitive, then children’s interpretation of the causal alternation should be less affected by properties of the referential scene than their interpretation of the boy-subject intransitive. In particular, changes in the relative noteworthiness of the boy across the two test events should alter children’s interpretation of the boy-subject sentence without affecting their interpretation of the causal alternation. If, however, children cannot interpret the causal alternation and instead simply interpret its intransitive sentence frame, then any changes to the referential scene should affect both the boy-subject condition and the causal condition.

To investigate these questions, we created a new set of caused-motion and contact-activity events (see Figure 6). The caused-motion event was like the one used in Experiment 3 except that it featured a different girl and boy. The contact-activity event was altered in several ways. First, the positioning of the girl and boy was changed to be more similar to the configuration in the caused-motion event. Rather than facing to the side, both the boy and the girl now faced forward. The boy was centered and slightly in front, while the girl stood partly behind the boy. Second, the boy now marched in place while being dusted rather than standing still.

The goal of these changes was to increase the noteworthiness of the boy’s role in the contact-activity event. We assessed the effectiveness of these manipulations by again showing the stimulus videos to 12 adults and asking them to choose the video in which the boy or the girl (asked in counterbalanced order) attracted their attention more. When asked about the boy, 10 out of the 12 adults selected the contact-activity event. This is the opposite of what was found for in Experiment 3, where 10 out of 12 adults stated that the boy was more attention-getting in the
caused-motion event. This suggests that our manipulation successfully affected the relative noteworthiness of the boy in the test events. These changes also affected the noteworthiness of the girl in the two events. When asked where the girl grabbed their attention, only 5 out of 12 adults chose the contact-activity event (as compared to 9 out 12 in Experiment 3). It seems that by increasing the noteworthiness of the boy’s role in the contact-activity event, we reduced the noteworthiness of the girl’s role in that event. As a result, the girl’s role was about equally noteworthy in the two events.

While viewing these events, children heard either a transitive sentence or one of the sentence conditions from Experiment 4. We added a transitive condition in this experiment to insure that despite the changes to the contact-activity event, children still viewed it as a plausible referent for the transitive sentence *The girl is dackling the boy*. If children considered both events equally plausible referents, then their looking behaviors in this condition should be similar to those of the children in the neutral baseline condition.

We predicted that both the boy-subject and girl-subject conditions should be affected by the changes in the test events. Because the boy’s role was now more noteworthy in the contact-activity event, children in the boy-subject condition should look longer at the contact-activity than at the caused-motion event. The reduction in salience of the girl’s role in the contact-activity event should make children relatively more likely to consider her action in the caused-motion event as a referent for the girl-subject sentence. In contrast, if children can appropriately interpret the causal alternation, then their interpretation of this alternation should not be affected by the change in test events. The boy in the contact-activity event was now more interesting, but his action in that event was not caused by the girl’s action. Thus, children in the causal condition should again interpret the verb as referring to the caused-motion event. If, however, children who
heard the causal alternation in Experiment 3 responded solely based on the intransitive sentence frame, then the noteworthiness manipulations should affect children’s interpretation of the causal alternation in this experiment.

Unlike the causal alternation, children’s interpretation of the unspecified-object alternation may be affected by the change in test events. Because the unspecified-object alternation focuses children more on the actor’s activity than on an effect, children’s interpretation of this alternation may be affected by the relative noteworthiness of the actor’s actions within the two test events. It is possible that by reducing the noteworthiness of the girl in the contact-activity event that we made that event a less good referent for the unspecified-object alternation. If this is the case, then we might expect that children who heard the verb used in the unspecified-object alternation would look relatively more at the caused-motion event in this experiment than they did in Experiment 3.

Method

Participants

Sixty 28-month-old children participated (mean = 28.2, range 27.1-30.1, 30 male, 30 female). All children were native speakers of English. Five additional children were tested but excluded from analyses because they were distracted (2), failed to complete the experiment (1), or because the average difference between their looking times to the caused-motion and contact-activity events more than 2.5 SD away from the mean of their condition (2). Children’s productive vocabulary was measured using the short form of the Macarthur CDI, Level II (Fenson et al., 2000). Vocabulary scores ranged from 20 to 100 with a median of 62. Ten children were randomly assigned to each of six conditions: causal, unspecified-object, boy-
subject, girl-subject, neutral, and transitive. An equal number of boys and girls were assigned to each condition.

*Apparatus*

The apparatus was identical to that used in Experiment 3.

*Materials and Procedure*

As in Experiment 3, the materials consisted of color videos of a boy and girl accompanied by a soundtrack recorded by a native English speaker. The videos for the present experiment featured a different boy and girl. Other than the difference in actors, the character-familiarization, practice, and caused-motion test events were as described in Experiment 3. The contact-activity test event differed in structure, however. The girl and boy now stood in a similar configuration to that used in the caused-motion event: the boy and girl faced forward, and the girl stood behind and slightly to the left of the boy. The girl dusted the boy’s right arm and shoulder while he marched in place. The dusting and marching action were not enacted in synchrony, making it less likely children would think the dusting somehow caused the boy to march.

The procedure was identical to that of Experiment 3 with three exceptions. First, because the contact-activity event (with its added marching action by the boy) was more complicated than the contact-activity event in Experiment 3, the test phase began with two preview trials. These trials gave children more time to inspect the events before the test trials. In the first preview trial, children saw the caused-motion event presented on one screen for 8 seconds, accompanied by neutral audio (e.g., "Look here! See this?"), while the other screen remained blank. In the second preview trial, the contact-activity event was previewed in the same way on the other screen. The two trials were separated by a 4-s blank-screen interval. Second, the novel
verb *pimming* was replaced with the novel verb *dacking*. Third, during the test trials in the alternation conditions, children heard the transitive sentence prior to the intransitive sentence. This was done to insure that the novel verb did not conspicuously align with the onset or peak of an action in one event but not the other.

**Coding**

We coded where children looked (left screen, right screen, away) frame by frame from silent video. To assess reliability, 11 children’s data were independently coded by a second coder. The two coders agreed on the children’s direction of gaze for 98% of coded video frames. Any 8-s trial in which the child looked away for at least 50% of the trial was eliminated (n = 3).

Look-away times during the test trials were analyzed by an ANOVA with trial (1, 2) as a within-subject factor and condition (causal, unspecified-object, boy-subject, girl-subject, neutral, transitive) as a between-subjects factor. The analysis revealed a significant of trial, $F(1, 51) = 13.101, p < .005$, no effect of condition and no trial by condition interaction, both $Fs < 1.1, ps > .39$. As in Experiment 3, children tended to look away from the screens more on the second trial ($M = .55, SD = .73$) than on the first trial ($M = .23, SD = .28$), regardless of their condition. Due to differences in time spent looking away across the two trials, we again conducted our analyses on raw looking times to the caused-motion and contact-activity events.

Preliminary analyses of test-trial performance revealed no effects of sex or whether vocabulary or practice performance was above or below the median. These factors were not analyzed further.

**Results**

Table 7 shows the average time spent looking to the caused-motion and contact-activity events during each test trial, separately by condition. For each set of comparisons reported
below, separate analyses were performed on the looking times to the caused-motion event and the looking times to the contact-activity event.

As in Experiment 3, we began by examining children’s interpretation of the causal and unspecified-object alternations. Inspection of the means in Table 7 suggests that children in both alternation conditions looked longer at the caused-motion than at the contact-activity event. This was confirmed by analyzing the amount of time children spent looking at each event using an ANOVA with trial (1, 2) as a within-subject factor and condition (causal, unspecified-object) as a between-subjects factor. There was no effect of trial (caused-motion: $F < 1$; contact-activity: $F(1, 18) = 2.634, p = .12$) or condition (both $F$s < 1); there was also no interaction of trial and condition (both $F$s < 1). Thus, unlike in Experiment 3, the two alternations did not result in different patterns of looking behavior. This is because children in the unspecified-object condition demonstrated a different pattern from their counterparts in Experiment 3: they interpreted the verb as describing the caused-motion event rather than the contact-activity event.

One possible explanation for this pattern is that children found the altered contact-activity event to be an unacceptable referent for the transitive sentence in each alternation. If this were true, it would cause children in both alternation conditions to interpret the verb as describing the caused-motion event. However, the results of the transitive condition do not support this explanation. Children who heard the verb used in just a transitive sentence showed a preference for the contact-activity event on both trials. The transitive condition did not differ from the neutral condition on either trial (caused-motion: trial 1 $t < 1$, trial 2 $t(16) = -1.380, p = .19$; contact-activity: trial 1 $t < 1$, trial 2 $t(16) = 1.664, p = .12$), or on the average of the two trials, (caused-motion: $t < 1$; contact-activity: $t(18) = 1.163, p = .26$). This suggests that children who heard a transitive sentence alone found both events to be equally suitable referents for the verb.
The more likely explanation for why the two alternation conditions did not differ is that, as we suspected, children’s interpretation of the unspecified-object alternation was based on the relative noteworthiness of the actor’s action in the test events. By reducing the relative noteworthiness of the girl’s action in the contact-activity event, we caused children to see the caused-motion event as the better referent. The fact that the unspecified-object condition, but not the causal condition, was affected by the change in the referential options is consistent with the prediction that the unspecified-object alternation should provide less constraint on verb interpretation than the causal alternation.

We next examined each alternation separately and asked whether (a) the alternation differed from baseline, (b) whether the intransitive sentence differed from baseline, and (c) whether the alternation differed from its component frames. To address these questions, we divided the data into a causal alternation group (causal, boy-subject, neutral, and transitive conditions) and an unspecified-object alternation group (unspecified-object, girl-subject, neutral, and transitive conditions) and analyzed each group separately.

**Causal alternation group.** The means in Table 7 suggest that, as in Experiment 3, children in the causal condition preferred the caused-motion event. In contrast, children in the boy-subject, transitive, and neutral conditions looked longer at the contact-activity event. The time children spent looking at each event was analyzed using an ANOVA with trial (1, 2) as a within-subject factor and condition (causal, boy-subject, neutral, transitive) as a between-subjects factor. The two event analyses showed similar patterns, but differed in the strength of those patterns. Analysis of the looking times to the caused-motion event revealed no significant effects of trial, $F < 1$, or condition, $F(3, 33) = 1.222, p = .32$; there was also no interaction of trial and condition, $F(3, 33) = 1.996, p = .13$. Analysis of looking times to the contact-activity event
also revealed no significant effects of trial, $F < 1$, or condition, $F (3, 33) = 2.133, p = .12$, but did show a significant trial by condition interaction, $F (3, 33) = 3.139, p < .05$.

To investigate the trial by condition interaction, we analyzed looking times to the contact-activity event on each trial separately using an ANOVA with condition (causal, boy-subject, neutral, transitive) as a between-subjects factor\(^4\). On trial 1, there was no effect of condition, $F < 1$. On trial 2, there was a main effect of condition, $F (3, 34) = 3.742, p < .025$. To investigate this effect of condition, condition comparisons were performed on the second test trial. These comparisons revealed that children in the causal condition looked significantly less at the contact-activity event than did children in the neutral condition, $t(18) = -3.846, p < .005$. This suggests that children’s interpretation of the causal alternation was not influenced by the change in events: they still interpreted the verb as describing the caused-motion event. This replicates the finding from Experiment 3, demonstrating that the causal alternation consistently directs children’s attention towards an event with an action + result structure.

Unlike in Experiment 3, the boy-subject condition did not differ from the neutral condition, $t < 1$. This suggests that the increased noteworthiness of the boy’s role in the contact-activity event caused children to view the contact-activity event as a better referent for the boy-subject sentence. Furthermore, children in the causal condition looked significantly less at the contact-activity event than did children in the boy-subject condition, $t(18) = -2.414, p < .05$. This demonstrates that children in the causal condition must have been interpreting the verb based on both sentence frames. If they had been interpreting the verb based on the intransitive sentence alone, there would have been no difference between the causal and boy-subject conditions.

\(^4\) Because the ANOVA on caused-motion looking times failed to show any significant effects, we only performed condition comparisons on the looking times to the contact-activity event. However, for each of these comparisons, the looking times to the caused-motion event showed the same pattern.
The difference between the causal and boy-subject conditions emerged on the second trial, which suggests that despite the additional event previews, the events in this experiment were more difficult for children to process than those used in Experiment 3. The test events in this experiment may have posed particular difficulties for children in the causal condition. Both of the test events in this experiment involved a girl acting on a boy and a boy acting. Upon hearing the verb used in a causal alternation, children had to examine each event closely in order to determine whether the boy’s action was causally linked to the girl’s action, and thus whether the event was an appropriate referent for the sentences they had just heard.

While the causal condition differed from the boy-subject condition on the second test trial, it did not differ from the transitive condition, $t(18) = -1.168$, $p = .26$. Close inspection of the transitive condition showed that the looking times in this condition were highly variable. On the second test trial, looking times to the caused-motion event ranged from 0s to 6.90s (out of 8s) and looking times to the contact-activity event ranged from .87s to 7.43s (out of 8s). No other condition showed this degree of variability. Children in the transitive condition also spent more time looking away from the two screens: 2 of the 10 children in this condition looked away for more than 50% of the second trial, resulting in dropped trials. Taken together, these two patterns suggest that because both events were equally suitable referents for the transitive sentence, children in this condition responded less systematically than children in other conditions. This may have made it difficult to see a difference between the transitive condition and other sentence conditions.

**Unspecified-object alternation group.** The amount of time children spent looking at each event was analyzed using an ANOVA with trial (1, 2) as a within-subject factor and condition (unspecified-object, girl-subject, transitive, neutral) as a between-subjects factor. Both analyses
revealed a significant main effect of condition (caused-motion: $F(3, 34) = 3.899, p < .025$; contact-activity: $F(3, 34) = 4.710, p < .01$). The effect of trial was not significant (caused-motion: $F < 1$; contact-activity: $F(1, 34) = 2.558, p = .12$), and there was no interaction of trial and condition (caused-motion: $F(3, 34) = 1.517, p = .23$; contact-activity: $F(3, 34) = 1.890, p = .15$). We therefore performed condition comparisons on the average of the two test trials. These comparisons revealed that children in the unspecified-object condition looked significantly longer at the caused-motion event, and significantly less at the contact-activity event, than did children in the neutral condition (caused-motion: $t(18) = 3.161, p < .01$; contact-activity: $t(18) = -3.505, p < .005$). Children in the girl-subject condition looked significantly longer at the caused-motion event, and significantly less at the contact-activity event, than children in the neutral condition (caused-motion: $t(18) = 2.681, p < .025$; contact-activity: $t(18) = -3.103, p < .01$). This suggests that our manipulation of the test events affected not only the unspecified-object condition, but also the girl-subject condition. By reducing the noteworthiness of the girl’s action in the contact-activity event, we caused children to see the caused-motion event as a better referent for both the unspecified-object alternation and the girl-subject sentence.

The unspecified-object condition again did not differ from the girl-subject condition, both $ts < 1$. Children in the unspecified-object condition did look marginally longer at the caused-motion event than children in the transitive condition, $t(18) = 1.915, p = .07$; however, these two groups did not differ in their looking times to the contact-activity event, $t(18) = -1.546, p = .14$.

**General Discussion**

Previous research suggests that children can use the structure of a sentence to constrain their interpretation of a new verb when they encounter it. However, a single sentence structure can be consistent with many meanings, leaving considerable ambiguity in verb interpretation.
Landau and Gleitman (1985) proposed that children could resolve this ambiguity by considering the full range of sentence frames in which a verb occurred. Here we investigated this multiple frames hypothesis by testing children’s interpretations of two transitivity alternations: causal and unspecified-object verbs. Verbs that occur in the causal alternation describe actions on objects that cause an ensuing effect (i.e. caused motion, change of state). Verbs that occur in the unspecified-object alternation also describe actions on objects, but they focus only on the actor’s activity rather than an effect.

Previous evidence suggests that children can interpret verbs presented in the unspecified-object alternation (Naigles, 1996, 1998). Here we sought to extend these findings by showing that children can also interpret verbs presented in the causal alternation. In two experiments, 28-month-olds encountered a novel verb in either the causal alternation, the unspecified-object alternation, or one of the individual sentence frames from these alternations (i.e. boy-subject intransitive, girl-subject intransitive, transitive). In both experiments, this verb was accompanied by a caused-motion event and a contact-activity event. The nature of the contact-activity event varied across the two experiments. In Experiment 3, the girl dusted the boy while he remained stationary. In Experiment 4, the girl dusted the boy while he marched in place.

The results showed that hearing the novel verb in the causal alternation consistently directed children’s attention towards the caused-motion event. This effect occurred in both experiments, despite the fact that in Experiment 4 the boy’s individual role was more noteworthy in the contact-activity than in the caused-motion event. This suggests that children interpreted the causal alternation as describing the caused-motion event because only that event had the correct action + result structure. As a result, their interpretation of the verb was not influenced by the relative noteworthiness of the boy’s role in the two test events. In contrast, when children heard
only the intransitive sentence of the causal alternation (i.e. the boy-subject sentence), their interpretation of the verb was influenced by the noteworthiness of the boy in the test events. In Experiment 3, children who heard the boy-subject sentence preferred the caused-motion event. In Experiment 4, they preferred the contact-activity event. Taken together, the results from these two conditions suggest that children can appropriately interpret verbs that occur in the causal alternation, and that hearing a verb in the causal alternation provides children with a more constrained interpretation of the verb than hearing it in an intransitive sentence.

The results of the two experiments also showed that children interpret the causal and unspecified-object alternations differently. In Experiment 3, children who heard the verb used in the unspecified-object alternation looked longer at the contact-activity than at the caused-motion event. While the unspecified-object condition did not differ from the neutral condition, it did differ from the causal condition. This pattern reversed in Experiment 4: children who heard the unspecified-object alternation looked longer at the caused-motion than at the contact-activity event, and this condition differed from the neutral condition but not the causal condition. This suggests that while children do not systematically interpret the unspecified-object alternation as describing an action with an effect (as they do the causal alternation), they recognize that the alternation can sometimes describe actions that produce effects (much as pushing and poking can have effects on objects). Moreover, the pattern across the two experiments suggests that the noteworthiness of the actor’s action in the events affected whether children viewed the event as a suitable referent for the unspecified-object alternation. When the girl was more noteworthy in the contact-activity event, children assumed that the unspecified-object alternation described the contact-activity event (Experiment 3). When the noteworthiness of the girl in the contact-activity
event was reduced, children were more likely to interpret the alternation as describing the caused-motion event (Experiment 4).

Unlike Naigles (1996), we failed to find a difference between the unspecified-object and girl-subject conditions in either experiment. This was probably because children saw both events as equally suitable referents for the transitive sentence of the alternation. This meant that in order to select a referent for the verb, children had to decide which event was a better match for the intransitive sentence (i.e. in which event is the girl’s activity most noteworthy). However, Naigles (1996) events were also designed to be equally suitable referents for a transitive sentence. Why then did she find a difference between the unspecified-object and actor-subject conditions when we did not?

This discrepancy may have resulted from the structure of Naigles’s (1996) task. Because both actions were on the same screen, it is impossible to assess children’s initial interpretations of the verb. It could be that children in the unspecified-object and actor-subject conditions actually preferred the same action (i.e. the contact-activity) during the composite training phase, but the two conditions differed in how long this preference persisted. Children in actor-subject condition may have finished processing the verb’s meaning more quickly because they heard only an intransitive sentence. Once they finished processing the verb, they may have chosen to look at the more interesting event instead (i.e. the caused-motion event). If children in the unspecified-object condition took slightly longer to encode the verb’s referent, and thus continued looking at the contact-activity event longer, then this would explain why the two conditions differed during the test phase. Suggestive support for this interpretation comes from the fact that the difference between these two conditions only emerged for two of the four items in Naigles’s experiment.
How do our findings compare with the predictions generated by the category-mediated and direct-inference mechanisms? The pattern across the two experiments is precisely what one would expect if children were using either (or both) of these mechanisms to interpret sets of sentence frames. The causal alternation consistently directed children towards the caused-motion event, while the unspecified-object alternation did not. The fact that the unspecified-object condition varied across the two experiments is consistent with the prediction that this alternation should provide less constraint than the causal alternation. Thus, these experiments provide support for Landau and Gleitman’s (1985) multiple frames hypothesis by demonstrating that children possess a mechanism for interpreting multiple frames, be it the category-mediated or direct-inference mechanism, and that they can use this mechanism to interpret both causal and unspecified-object verbs.
CHAPTER 4

2-YEAR-OLDS USE DISTRIBUTIONAL CUES TO INTERPRET TRANSITIVITY

ALTERNATING VERBS

Languages contain systematic relationships between sentence structure and word meaning (e.g., Bloom, 1970; Carlson & Tanenhaus, 1988; Dowty, 1991; Levin & Rappaport-Hovav, 2005). Thus, verbs that describe one participant acting on another tend to occur in transitive sentences, with two noun-phrase arguments (e.g., “She tickled her.”), while verbs that describe one-participant actions tend to be intransitive, with one argument (e.g., “She laughed.”).

Children use these systematic relationships between structure and meaning to constrain the interpretation of new verbs in a process known as syntactic bootstrapping (Landau & Gleitman, 1985). For example, when they encounter a novel verb in a simple transitive sentence, 2-year-olds infer that the verb refers to a relationship between two participants rather than to an action by a single participant (e.g., Fisher, 2002; Naigles, 1990; Naigles & Kako, 1993). Young children also use word order in a transitive sentence to relate a new verb to an event in which the subject of the sentence plays an agent’s role (Gertner et al., 2006). Somewhat older children infer that an unknown verb has a mental-content meaning (e.g., believe) when it occurs with sentence-complement syntax (e.g., “Matt corps that his grandmother is under the covers;” Papafragou, Cassidy, & Gleitman, 2007).

In the preceding examples, children used structural elements of a single sentence to guide their interpretation of a novel verb. But syntactic bootstrapping is not limited to inferences from a single sentence. The syntactic bootstrapping view holds that children also gain information

---

5 Chapter 4 is reprinted from an article in its entirety from *Language and Cognitive Processes, 24, 777-803*. The copyright belongs to the Taylor & Francis Group, which has granted permission to reproduce the article.
about the meaning of a verb from the set of sentence structures in which it occurs (Landau & Gleitman, 1985).

To illustrate, many verbs occur in multiple sentence structures (e.g., transitive and intransitive, see examples 9 and 10). Verbs that occur in the same set of structures tend to have similar meanings (Levin, 1993; Pinker, 1989). For instance, verbs that occur in the causal alternation shown in (9) describe actions that cause a salient change in the object acted upon (e.g., an externally-caused motion such as bounce or a change of state such as break); both the action and resulting change are part of the meaning of the verb. Verbs that occur in the unspecified-object alternation shown in (10) are similar to causal verbs in that they describe actions on objects. However, unspecified-object verbs include only the action in their meanings, without reference to any particular result. For instance, while poking someone could result in some effect (e.g., the person poked becomes irritated), only the physical action is described by the verb poking; this action would still be considered poking even in the absence of an effect (see Wagner, 2006, for discussion). Unspecified-object verbs vary more in their meanings than do causal verbs, but many describe contact actions (e.g., dust, poke, push).

9)  a. Anne broke the lamp.
    b. The lamp broke.

10) a. Anne dusted the lamp.
    b. Anne dusted.

Experimental evidence suggests that children can use sets of frames such as those in (9-10) to draw appropriate inferences about verb meaning (Naigles, 1996, 1998; Scott & Fisher, 2007). For example, in one study, Scott and Fisher (2007) presented 28-month-old children with a caused-motion event (a girl causing a boy to bend at the knees by pressing down on his
shoulders) and a contact-activity event (a girl dusting a boy’s back with a feather duster; see Figure 5). These two events were shown simultaneously and were accompanied by a soundtrack in which a novel verb occurred either in the causal or in the unspecified-object alternation. Children who heard the novel verb used in the causal alternation looked longer at the caused-motion event than did those who heard it used in the unspecified-object alternation.

The present research asked how children use the alternations shown in (9) and (10) to assign different meanings to causal and unspecified-object verbs. The two alternations are syntactically identical: Both involve a transitive and an intransitive sentence. Moreover, the two alternations assign similar roles to the subject and object arguments of the transitive sentence of each pair (9a, 10a): both types of verbs assign the agent of action to the subject position and the recipient of action (the undergoer) to object position. The difference between the two alternations lies in the semantic roles assigned to subject position in the intransitive sentence of each pair. In intransitive sentences, unspecified-object verbs assign the agent to the subject position (10b), just as they do in transitive sentences. In contrast, causal verbs assign the undergoer to this position (9b). In order to assign appropriately different interpretations to novel verbs presented in these two alternations, children must be able to distinguish actor-subject from undergoer-subject intransitive sentences. The two syntactic cues that have been shown to influence young children's verb interpretation, verb transitivity and word order in transitive sentences (e.g., Naigles, 1990; Gertner et al., 2006), would not help children differentiate these two types of intransitive sentences. The finding that 28-month-olds interpret these alternations appropriately (Naigles, 1996, 1998; Scott & Fisher, 2007) suggests that they have access to another source of information, one that reflects the underlying role difference between the two alternations.
One likely source of role information is the referential scene. In experiments with adults, Knoeferle, Crocker, Scheepers, and Pickering (2005) showed that when an ambiguous sentence is accompanied by a relevant referential scene, the apparent roles of the scene participants constrain the adults’ assignment of semantic roles to the noun phrases in the sentence. It has long been assumed that children's sentence comprehension exploits the constraints of the referential context in much the same way (e.g., Chapman & Miller, 1975; Clark, 1973; Huttenlocher, 1974; Shatz, 1978). If children encountered the alternations in (9-10) in the presence of relevant referential scenes, as they did in the experiments by Naigles (1996, 1998) and by Scott and Fisher (2007), then in principle they could use the referential scene to assign likely semantic roles to the noun phrases within each alternation. For instance, in Scott and Fisher’s (2007) study, children in the unspecified-object condition might have concluded that “The girl is pimming” was an actor-subject intransitive sentence because the girl played an agent role in the accompanying test events.

A second source of role-relevant information that could be useful even in the absence of an informative referential setting comes from the distributional properties of the sentences in the input. In an examination of text from the Wall Street Journal, Merlo and Stevenson (2001) found that causal and unspecified-object verbs systematically varied on a number of distributional features, and that a machine-learning model could reliably use these features to distinguish between the two verb types. Merlo and Stevenson’s model relied primarily on three features: each verb’s frequency of occurrence in the transitive structure (transitivity), its rate of subject-noun animacy, and the degree of lexical overlap between the nouns in its subject and object positions.
How might these features help identify causal versus unspecified-object verbs? Unspecified-object verbs (10) assign the same role (agent) to subject position regardless of transitivity. These verbs have mostly animate subjects (because all their subjects are agents, which tend to be animate), and they rarely have the same lexical items in subject and object position (because all their subjects are agents and their objects are undergoers). In contrast, causal verbs (9) assign the undergoer role to the object of the transitive sentences and to the subject of the intransitive sentences. These verbs have fewer animate subjects (because their intransitive subjects are undergoers, which tend to be inanimate) and are more likely to have the same lexical items in subject and object position (because their objects and intransitive subjects are both undergoers).

Could young children use the distributional cues identified by Merlo and Stevenson (2001) to distinguish causal and unspecified-object verbs? In order to so, children must be able to a) detect and keep track of all three distributional cues, and b) use these distributional cues to draw inferences about verb meaning. There are encouraging hints that young children possess these abilities.

Previous evidence suggests that both adults (e.g., Wonnacott, Newport, & Tanenhaus, 2008) and children (e.g., Gordon & Chafetz, 1990; Kidd, Lieven, & Tomasello, 2006; Snedeker & Trueswell, 2004; Yuan & Fisher, 2006, 2009) keep track of the sentence structures in which a verb has appeared. For example, in recent work by Yuan and Fisher (2006, 2009), 28-month-old children watched videotaped dialogues in which two people repeatedly used a made-up verb in either only transitive sentences (e.g., A: “Anna blicked the baby!” B: “Really, she blicked the baby?”) or only intransitive sentences (A: “Anna blicked!” B: “Really, she blicked?”). The children later heard the verb used in isolation (“Find blicking!”) while they viewed test videos
depicting two referential options: a two-participant causal action (one girl lifted and lowered another girl’s leg) and a one-participant action (a girl made arm-circles). Children who had previously heard the verb used in transitive sentences looked longer at the two-participant event than did children who had heard the verb used in intransitive sentences. This and other results suggested that the children learned whether the verb was transitive or intransitive simply by hearing it used in sentences, even though they did not yet know the verb’s semantic content. When later presented with the verb in a referential context, children retrieved this syntactic-distributional information and used it to select an appropriate referent for the verb.

Previous work also suggests that children keep track of role-relevant information about the nouns that occur with verbs. Children keep track of the animacy of nouns in various grammatical positions. For example, 2-year-olds better comprehend sentences with animate than inanimate subjects, suggesting some sensitivity to the tendency for subject noun phrases to be animate (Childers & Echols, 2004; Corrigan, 1988; Lempert, 1989). Two-year-olds also use knowledge of the animacy of likely argument role-fillers for a particular verb in sentence comprehension: for example, they inferred that an unfamiliar noun must refer to an animal if it was the object of feed (“Mommy’s feeding the ferret!”; Goodman, McDonough, & Brown, 1998; see also Corrigan & Stevenson, 1994; Fernald, 2007). Children’s ability to track the distributions of particular lexical items relative to one another is also well documented. Infants track syllable co-occurrences in artificial grammar-learning experiments (e.g., Gómez, 2002; Saffran, Aslin, & Newport, 1996), and early sentence production reflects word combinations that are frequent in the input (e.g., Rowland, 2007).

Can young children use distributional features to assign meanings to novel verbs that occur in the causal and unspecified-object alternations? We examined this question in two parts.
In Experiment 5, we sought to verify that the features identified by Merlo and Stevenson (2001) reflected the semantic roles assigned by causal and unspecified-object verbs in a corpus of child-directed speech. Having determined that relevant distributional cues occur in child-directed speech, in Experiment 6 we used the dialogue-training method described earlier to examine 2-year-olds’ ability to use these distributional cues to infer the meaning of a novel alternating verb. In the General Discussion, we will speculate about two possible mechanisms by which young children might use distributional cues to draw such inferences.

**Experiment 5**

The *Wall Street Journal* corpus that Merlo and Stevenson (2001) analyzed differs in many ways from casual speech to children; for example, child-directed speech consists of much shorter sentences, is more repetitive, and employs a simpler vocabulary (e.g., Bard & Anderson, 1994; Newport, Gleitman, & Gleitman, 1977). Moreover, the same common verbs are almost certainly used in different senses in newspaper text and in casual speech to children (e.g., *The corporation folded* versus *Let's fold the laundry*). Such verb-sense differences across different discourse registers are important because they can cause striking differences in the frequency with which the same verb occurs in particular sentence structures (e.g., Roland & Jurafsky, 2002). Given these differences, it cannot be assumed that transitivity, subject animacy, and lexical overlap would identify the argument-structure patterns of the causal and unspecified-object alternations in child-directed speech just as they did in Merlo and Stevenson’s (2001) newspaper text.

To get at this issue, we examined the distributions of transitivity, subject animacy, and lexical overlap for causal and unspecified-object verbs in a large sample of child-directed speech.
We then used an unsupervised learning algorithm, k-means clustering, to determine whether the distributions of these features differentiated the two groups of verbs.

**Method**

**Materials**

We selected the following part-of-speech tagged corpora from the CHILDES database (MacWhinney, 2000): Bloom 1970, Brown, Clark, Demetras Working, Higginson, Kuczaj, New England, Post, Suppes, and Warren-Leubecker. From these corpora we selected all transcripts in which the target child was 32 months old or younger (range 13.5-32 months). These transcripts contained 112,000 parental utterances.

We used the CLAN program to search the morphological tier of all parental utterances for the “v|” tag$^6$, to obtain a list of all the verbs used by the parents. From this list, we selected all causal and unspecified-object verbs (classification based primarily on Levin, 1993) that occurred at least 30 times in total, and that occurred in at least 5 of the 10 corpora. The 29 verbs that met these criteria appear in Table 8. These verbs were used in 12,521 parental utterances across all the selected corpora.

**Coding**

The author hand-corrected the part-of-speech tagging of all 12,521 parental utterances containing the 29 selected verbs. The corrected utterances were coded with the CLAN program using the search heuristics described below. All coding was hand-checked by the author.

**Transitivity.** Utterances containing the target verbs were coded as transitive if the verb was immediately followed by a noun, a pronoun, a determiner (e.g., *a, the*), or any of a set of quantifiers (*some, any, all, much*, and *more*). An utterance was coded as intransitive if the verb was immediately followed by a punctuation mark, a conjunction (e.g., *and, or*), a preposition, a

---

$^6$ We used part-of-speech tagging in corpora downloaded in February 2005.
locative phrase (e.g., *here*), another verb, or a filler (e.g., *uh-oh*). Utterances containing phrasal verbs (e.g., “He tore up the paper.”) were hand-corrected to be transitive. Utterances in which the verb was followed by a part of speech other than those described above could not be reliably classified using machine-coding heuristics and were excluded. The final data set, after transitivity coding, consisted of 11,748 utterances.

**Animacy.** Following Merlo and Stevenson (2001), we used pronouns as a machine-extractable approximation of animacy. Pronoun arguments are very frequent in child-directed speech (e.g., Laakso & Smith, 2007) and 2-year-olds more readily understand transitive sentences when subjects and objects are pronouns marked for animacy (e.g., *He [verb] it*; Childers & Tomasello, 2001). This suggests that children know the meanings of many pronouns from an early age and might be able to use them to track the animacy of each verb’s arguments.

An utterance was coded as having an animate subject if the verb was preceded by *he, she, we, I, you, let’s* or *who*, permitting one intervening auxiliary verb. Imperatives were also coded as having animate subjects. Inanimate subjects were *it, that, this, that one, this one, or what*. These heuristics captured 75% of the subjects in the data set.

**Lexical overlap.** To calculate the lexical overlap between subject and object noun-phrases for each verb, we followed the procedure outlined by Merlo and Stevenson (2001): We extracted the subject and object noun-phrases of each sentence containing a target verb. The subject-noun set for each verb contained all nouns and pronouns that occurred as subjects of the verb in either transitive or intransitive sentences; the object-noun set contained all nouns and pronouns that occurred as objects of the verb. We then found all nouns that occurred in both the subject- and object-noun set for a particular verb. We selected the set in which the noun was more frequent

---

7 If imperatives are removed from the analyses, the patterns of significance described in the results section remain the same.
and added its number of occurrences in that set to the overlap set for the verb. For example, if *break* had the subject set \{pencil, Adam, it, it, baby\} and the object set \{pencil, it, it, glasses\}, the overlap set would be \{pencil, it, it, it\}. We then divided the number of items in the overlap set by the total number of subjects and objects for that verb (in this example, 4/10 = 40% lexical overlap). In determining the overlap set, noun number (e.g., *pencil* vs. *pencils*) and case (*he* vs. *him*) were ignored.

**Measures**

We assigned to each verb a score for each of four variables: transitivity (the proportion of coded utterances that were transitive), overall subject animacy (the proportion of coded subjects that were animate), intransitive-subject animacy (the proportion of coded intransitive subjects that were animate), and lexical overlap as defined above. Intransitive-subject animacy was examined separately because of its predicted importance in distinguishing causal and unspecified-object verbs. Transitive-subject animacy was not examined separately because 99% of transitive subjects were animate and this did not vary by verb class (causal mean = 99%; unspecified-object mean = 99%).

**Classification analyses**

In their classification analyses, Merlo and Stevenson’s (2001) machine-learning algorithm learned to classify the verbs via discriminant analysis based on explicit feedback about the correct classification of a training subset of verbs. Supervised learning procedures of this type are generally considered a poor model for ordinary language acquisition, as children receive no direct feedback about the proper classification of verbs they have learned. To approximate this aspect of language acquisition, we chose to classify the verbs using k-means clustering (MacQueen, 1967), an algorithm that does not receive direct feedback about correct
classification. Note that while an unsupervised learning algorithm is more similar to the circumstances of language acquisition than a supervised algorithm, we do not presume that language learners use a process like k-means clustering to acquire and classify verbs. Rather, this clustering analysis is simply a tool for determining whether particular cues could be used to detect category divisions in the input.

K-means clustering takes scores on \( p \) variables (i.e. the scores described above) for \( n \) objects (i.e. verbs) and attempts to organize those objects into \( k \) clusters. The algorithm forms a random initial division of the \( p \)-dimensional similarity space into \( k \) clusters and then iteratively reorganizes the objects to minimize the sum, over all clusters, of the within-cluster distance between each object and the center of its cluster. Once reorganization of the objects can no longer reduce the total within-cluster distance, the algorithm stops. K-means clustering is sensitive to the initial random partitioning of the data, sometimes returning solutions that are not the best fit for the data. The standard procedure for avoiding suboptimal fits is to repeat the analysis with different random initial divisions of the data and to select the solution with the lowest total within-cluster distance as the best fit for the data. For each analysis reported here, we performed 100 replications.

The k-means clustering algorithm requires that the number of clusters, \( k \), be specified in advance. For all the analyses reported here, 2 clusters were used. Since children do not know a priori how many classes of verbs they are likely to encounter in the input, this aspect of the analyses is a poor approximation of natural language acquisition. However, pilot analyses with other numbers of clusters (3 and 4) produced poorly separated clusters. This hints that, within the set of transitivity-alternating verbs, children might naturally arrive at 2 groups of verbs, as
exploratory attempts with other numbers of groups would fail to produce cohesive clusters of verbs.

Cluster evaluation

Two measures were used to evaluate the results of each clustering analysis. We first calculated the proportion of verbs correctly classified (accuracy). That is, each verb was counted as correctly classified if its a priori categorization as shown in Table 8 matched that of the majority of verbs in its cluster. Because this accuracy score sometimes overestimates the quality of a clustering solution, we also calculated the Adjusted Rand Index ($R_{adj}$; Hubert & Arabie, 1985), which measures the overall quality of a clustering solution by taking into account both correct classifications and incorrect classifications. The index ranges from 0 (random grouping) to 1 (perfect classification), with occasional negative values for extremely poor solutions.

To provide a baseline for the obtained accuracy and $R_{adj}$ scores, we created a reference distribution of randomly-obtained scores via Monte Carlo sampling. That is, we created 5,000 random permutations of the data (by randomizing the assignment of transitivity, animacy, and lexical-overlap scores to particular verbs), submitted them to the k-means clustering algorithm, and obtained Accuracy and $R_{adj}$ scores for the resulting clusters. Separate reference distributions were created in this way for each analysis reported below. $P$-values were calculated as the proportion of scores in the reference distribution that were as or more extreme than the score obtained in the experimental analysis.

Results

Table 9 shows the mean scores for each variable, separately by verb class. As expected, unspecified-object verbs had animate subjects more often than did causal verbs, $t(27) = 3.197, p < .01$. This difference was greater for intransitive-subject animacy than for subject animacy.
overall: only 44% of the intransitive sentences containing causal verbs had animate subjects, whereas 97% of the intransitive sentences containing unspecified-object verbs had animate subjects \( t(27) = 6.913, p < .001 \). Causal verbs also exhibited marginally higher lexical overlap than unspecified-object verbs, \( t(27) = 1.971, p = .061 \). The two verb groups did not differ in transitivity in this sample \( t < 1 \).

We performed five k-means clustering analyses, one using each variable alone, and one using all variables. Table 10 shows the accuracy and \( R_{adj} \) scores for each cluster analysis. The \( p \)-values shown in the Table were obtained from reference distributions created through Monte Carlo sampling, as described in the Methods section. The clustering solution based on all variables, shown in the bottom row of Table 10, yielded highly accurate verb classification, grouping 24 of the 29 verbs correctly. Its \( R_{adj} \) of .41 represents a significant improvement over the randomly-generated baseline. Neither transitivity nor overall subject animacy yielded classifications that differed from the random baseline when considered alone. Intransitive-subject animacy yielded the best classification; with 24 out of 29 verbs correctly classified, it performed as well as the analysis using all variables. The analysis based on lexical overlap classified 20 of the 29 verbs correctly, with a significant \( R_{adj} \) of .12.

**Discussion**

The results of Experiment 5 show that child-directed speech contains distributional features that discriminate causal from unspecified-object verbs. The individual feature analyses provide some insight into which of these features would be most useful for making this distinction.

In our corpus, unlike in Merlo and Stevenson’s (2001), both verb types were equally likely to be transitive, causing transitivity to perform poorly in classification analyses. Such
differences between corpora are generally unsurprising and, as mentioned above, can be partially attributed to differences in the distribution of verb senses in different corpora (Roland & Jurafsky, 2002). For instance, in our data the verb fold was used primarily in the context of doing laundry (e.g. “Let’s fold it [the towel] nice and neat.”) and therefore the uses of this verb were predominantly transitive (73%). It is unlikely that the Wall Street Journal contains the sense of fold that pertains to laundry. Instead, in the Wall Street Journal, fold is probably used to refer to events such as the collapse of a corporation (e.g. “After serious financial trouble, the company folded.”), resulting in a much lower rate of transitivity (23%; Merlo & Stevenson, 2001).

Although we did not code for verb sense, it seems likely that differences in discourse context and style between the two corpora led to many sense differences like those seen with fold, and that these differences contributed to the different transitivity patterns.

The high transitivity rate in our corpus also affected the animacy results. Because both causal and unspecified-object verbs assign an agent to the transitive subject position, the transitive subjects were mostly animate, regardless of verb class. For the overall subject animacy measure, this large number of animate transitive subjects obscured any differences in subject animacy between the two verb types. As a result, overall subject animacy failed to classify the verbs in our data set.

Once the uniformly animate transitive subjects were removed, however, the difference between the two groups became apparent. Intransitive-subject animacy classified most of the verbs correctly, performing as well as a model considering all of the predictors together. Merlo and Stevenson’s (2001) lexical overlap measure also performed well in our analyses. Thus, linking animacy with syntactic positions and tracking lexical overlap between syntactic positions
are both useful cues for distinguishing causal and unspecified-object verbs in corpora as diverse as the *Wall Street Journal* and casual speech to children.

These results suggest that the input contains distributional features that would allow children to distinguish between causal and unspecified-object verbs. Many questions remain about how children would use these features to discover verb categories in the input. For instance, in the analyses presented here, only relevant features were presented to the classification algorithm. In natural language acquisition, learners must determine both how many categories are present in the input and which cues are relevant to each. In principle, the set of cues could be unbounded, resulting in an intractably large search space for useful distributionally-defined categories. Proposed solutions to this problem typically adopt some combination of two theoretical tactics. First, children’s initial search space for discovering grammatical categories and subcategories is assumed to be constrained because children are predisposed to attend to a subset of the conceivable cues, the grammatically-relevant ones (e.g., Gleitman, 1990; Pinker, 1989). Second, as children learn about words and structures, the apparent meanings of those words and structures provides semantic feedback for learning. This feedback, while noisy, could help children to determine which distributional cues are diagnostic of meaning differences (e.g., Maratsos & Chalkley, 1980; Pinker, 1989).

In the present case, the cues that we used – verb transitivity, subject noun-phrase animacy, and lexical overlap between subject and object positions – are clearly grammatically relevant features. In addition, the usefulness of these cues is not limited to telling apart causal and unspecified-object verbs. Cues based on the nouns and pronouns that appear in various argument positions are informative for discriminating a number of meaningfully distinct classes of verbs (e.g., Joanis, Stevenson, & James, 2008; Laakso & Smith, 2007; Pinker, 1989; Resnik,
1996). We will return to these issues in the General Discussion, and propose two mechanisms whereby children could use the distributional cues examined here to draw inferences about verb meaning.

**Experiment 6**

The results of Experiment 5 showed that child-directed speech contains at least two useful distributional cues to the differences between causal and unspecified-object verbs: intransitive-subject animacy and the lexical overlap between nouns in subject and object position. In Experiment 6, we asked whether 28-month-old children could use these distributional cues to guide the interpretation of an unknown verb. To isolate the contribution of distributional cues, we presented the new verb’s distributional properties separately from its possible event referents using the dialogue-training technique introduced by Yuan and Fisher (2006, 2009). This task is a modification of the standard looking-preference comprehension task, which relies on the well-established tendency of children and adults to look at scenes related to sentences they hear (e.g., Tanenhaus, Spivey, Eberhard, & Sedivy, 1995).

Children first watched a video of two women using the novel verb *dacking* in a conversation. Depending on the dialogue condition, the conversation exhibited the animacy and lexical overlap patterns of either the causal or the unspecified-object alternation (Figure 8). All children then viewed the two test events shown in Figure 8 and heard, “The girl is dacking the boy! Find dacking.” The causal event showed a girl causing a boy to squat by pressing down on his shoulders. The contact-activity event showed the girl brushing the boy's back with a feather duster.

Because both dialogue groups heard the same transitive sentence while viewing the test events, any differences between the groups in the interpretation of the novel verb could be
attributed to the distributional information manipulated in the dialogues. If children can use verbs’ distributional properties as a source of role-relevant information, then those who heard the causal dialogue should assign a causal interpretation to the verb and thus look longer at the caused-motion event; those who heard the unspecified-object dialogue should assign an activity interpretation and thus look longer at the contact-activity event.

We also included control conditions in which children heard one of the *dacking* dialogues, but heard a different verb when the test events were presented (“The girl is pimming the boy. Find pimming.”). Since the children in these control conditions had not encountered *pimming* in the dialogues, they should treat the dialogues as irrelevant to the test trials. We therefore predicted that in the control conditions, children in the two dialogue groups would not differ in their looking patterns.

**Method**

**Participants**

Forty-eight 28-month-olds participated (mean 28.1, range 27.0-30.0, 24 male, 24 female). All were native speakers of English. Five additional children were excluded due to parental interference (1), failure to complete the experiment (1), or because they moved out of camera range (3). Children’s productive vocabulary was measured using the short form of the Bates-Macarthur CDI, Level 2 (Fenson et al., 2000). Vocabulary scores ranged from 15 to 100 with a median of 70. Twelve children were randomly assigned to each of the four combinations of the dialogue (causal, unspecified-object) and test-verb (same-verb, different-verb) conditions.

**Apparatus**

Children sat on a parent’s lap facing two 20-inch television screens placed 30 inches away. The screens were 12 inches apart and about at the child's eye level. Soundtracks were
played from a concealed central speaker. A camera hidden between the two screens recorded the children’s eye movements during the experiment. Parents wore opaque sunglasses, preventing them from biasing their children’s responses.

Materials and Procedure

Stimulus materials were color videos of dialogues between two women and of test events depicting actions involving a girl and a boy. Test events were shown in synchronized pairs and accompanied by a soundtrack recorded by a native English speaker. The left-right positioning of the familiarization, practice, and test events was counter-balanced with dialogue and test-verb condition.

The procedure had three phases: character-familiarization, practice, and test. The character familiarization and practice phases were designed to acquaint the children with the actors in the test events and with the structure of the experiment.

In the character-familiarization phase, a female actor was shown waving on one screen (4s) and was labeled twice (e.g., “There’s a girl!”) while the other screen remained blank. Following a 2-s interval, a male actor was introduced on the other screen in the same manner (“There’s a boy!”). This was followed by two 4-s trials, separated by 2-s blank-screen intervals. In each trial, the girl appeared on one screen while the boy appeared on the other. In the first trial, children were asked to “Find the boy!”; in the second trial they were instructed to “Find the girl!”

In the practice phase, two familiar intransitive verbs were presented. Children first watched two women using the verb jump in 8 sentences in a conversation (e.g., “Matt jumped!” “Yeah, he was jumping”). The conversation was presented in two dialogue video-clips (each 16 - 18s long), each of which contained 4 sentences, separated by a 3-s blank-screen interval. The
same dialogue video appeared on both screens simultaneously. Next, during a 7-s blank-screen interval, children heard “Look! Jumping!” Children then saw a pair of 8-s videos, one showing the boy jumping and the other showing the boy pretending to sleep; the soundtrack asked children to “Find jumping.” After a 3-s interval, this 8-s video pair was presented again, and children were again asked to “Find jumping.” Following a 4-s blank-screen interval, this procedure was repeated with a second practice verb (clap); the videos showed the girl clapping and the girl pretending to eat.

Finally, the novel verb dack was introduced. Children first encountered the verb in three dialogue clips (each 18 - 22 s), each of which presented 4 sentences containing the novel verb. In both dialogue conditions, children heard the novel verb in 12 sentences, 6 transitive and 6 intransitive (the complete dialogues appear in the Appendix). Depending on dialogue condition, dacking was presented in either the causal or the unspecified-object alternation (see Figure 8). The two dialogue conditions differed in their patterns of animacy and lexical overlap. In the causal dialogues, 67% (4/6) of the intransitive subjects were inanimate, and the nouns that served as objects of transitive sentences were always re-used as subjects of intransitives. In the unspecified-object dialogue, none of the subjects were inanimate and the nouns that served as subjects of transitive sentences were re-used as the subjects of intransitives. The nouns used in these dialogues did not obviously label the actors in either test event.

Next, during a 7-s blank-screen interval following the third novel-verb dialogue, children in the same-verb groups heard, “Watch! The girl is gonna dack the boy!” The children then saw a pair of 8-s videos, one depicting a caused-motion and the other a contact-activity event (Figure 5). These events were accompanied by the sentence, “The girl is dacking the boy. Find dacking.”
As in the practice phase, the pair of videos was presented twice, separated by a 3-s interval. These two 8-s trials tested children's interpretations of the novel verb.

The different-verb groups received the same dialogues and the same test trials as the same-verb groups, with one key difference: The transitive sentences that accompanied the novel-verb test events contained the verb *pim* instead of *dack* (“The girl is gonna pim the boy!”).

**Analyses**

We coded where children looked (left-screen, right-screen, away) frame by frame from silent video. To assess reliability, 12 children’s data were independently coded by a second coder. The first and second coders agreed on the children’s direction of gaze for 96% of coded video frames.

The amount of time children spent looking away from the two video screens during the test trials was analyzed by means of a 2 x 2 x 2 mixed-model analysis of variance (ANOVA) with dialogue (causal or unspecified-object) and test verb conditions (same-verb or different-verb) as between-subjects factors, and test-trial as a within-subjects factor. No effect was significant, all \( F_s < 2.5 \) and all \( p_s > .1 \), suggesting that the children in the two experimental and two control groups tended to look away about equally, and equally briefly, during the test trials (same-verb: causal M = .40 s, SD = .50, unspecified-object M = .43 s, SD = .49; different-verb: causal M = .37 s, SD = .32; unspecified-object M = .44 s, SD = .43). Given the uniformity of time spent looking away, we conducted our main analyses on a single measure, the proportion of looking time to the caused-motion event, out of the total time spent looking at either test event during each test trial. Analyses based on absolute looking times to the caused-motion or to the contact-activity event revealed the same pattern of significant effects as the main analyses reported below.
Preliminary analyses of children’s looking time performance in the test trials revealed no interactions of dialogue and test-verb with sex or with whether the child’s vocabulary or performance in the practice trials was above or below the median, all $F$s < 1. These factors were not examined further.

Results and Discussion

Table 11 shows the proportion of looking time to the caused-motion event, out of the total time spent looking at either test event during each test trial, separately by dialogue and test-verb condition. As predicted, dialogue type strongly affected looking preferences in the test trials, but did so only for children in the same-verb condition. Inspection of Table 11 suggests that this effect emerged on the second trial.

A 2 (dialogue: causal, unspecified-object) by 2 (test-verb: same-verb, different-verb) by 2 (trial) mixed-model ANOVA revealed a main effect of dialogue ($F(1,44) = 4.792, p < .05$), marginal interactions of dialogue and test-verb condition ($F(1,44) = 3.985, p = .052$), trial and dialogue ($F(1,44) = 3.823, p = .057$), and trial and test-verb ($F(1,44) = 3.179, p = .081$), as well as a reliable three-way interaction of dialogue, test-verb, and trial ($F(1,44) = 7.242, p = .01$). To further investigate the 3-way interaction involving trial, we performed separate ANOVAs for each test-verb condition.

An analysis of the different-verb condition revealed no significant effects of dialogue or trial and no significant interaction of these two factors, all $F$s < 1. The absence of an effect of dialogue in the different-verb condition suggests that children used the presentation of the verb in the test trial as a cue to retrieve what they knew about this verb. When the verb in the test trials was entirely new, the information from the dialogue was treated as irrelevant.
Analysis of the same-verb condition revealed a significant effect of dialogue, $F(1,22) = 8.216, p < .01$, a marginal effect of trial, $F(1,22) = 3.085, p = .093$, and a significant interaction of dialogue and trial, $F(1,22) = 12.102, p < .005$. As shown in Table 11, during the first trial children in the same-verb condition looked about equally at the two test events regardless of dialogue condition ($t < 1$). In the second trial, children who heard the causal dialogue looked significantly longer at the caused-motion event than did those who heard the unspecified-object dialogue, $t(22) = 4.231, p < .001$.

During the test portion of the 2-phase procedure, all children encountered the verb in a transitive sentence. If children’s attention to the test events were influenced only by that sentence, then the two dialogue groups should have shown the same looking patterns. The effect of dialogue in the same-verb condition suggests that children encoded useful distributional information about the new verb during the preceding dialogues, retrieved that information when they encountered the verb again in the test trials, and used it to select an appropriate referent event. The emergence of this effect on the second trial suggests that the task was not easy for the children. The apparent difficulty of this task is not surprising: When the test events appeared, children had to inspect the two novel events to understand their structure, identify the novel verb in the transitive test sentence, and retrieve what they had learned about this verb during the dialogue. Any or all of these steps could contribute to the relatively slow appearance of the dialogue effect during the test trials in the same-verb group.

**General Discussion**

Previous experiments have shown that children assign appropriately different interpretations to verbs presented in the causal and unspecified-object alternations (Naigles, 1996, 1998; Scott & Fisher, 2007). Here we explored how they might do so. Causal and
unspecified-object verbs are syntactically similar in that they can be either transitive or intransitive, but they differ in the semantic roles they assign to the subjects of intransitive sentences. Causal verbs assign an undergoer to this position (The lamp broke), while unspecified-object verbs assign an agent to this position (Anna dusted). In order to tell apart verbs presented in these two alternations, children must detect this difference in role assignments. In the present study, inspired by computational work by Merlo and Stevenson (2001), we examined distributional features of the input as a potential source of information about this role difference.

The results of Experiment 5 showed that child-directed speech contains distributional cues that reflect the underlying argument-structure differences between causal and unspecified-object verbs. Specifically, both intransitive-subject animacy and the lexical overlap between nouns in subject and object position proved useful for discriminating the two verb types. The results of Experiment 6 suggest that children can encode some combination of these two cues, and later use them to identify an appropriate referent for novel causal and unspecified-object verbs. These findings broaden what we know about syntactic bootstrapping in three ways.

First, the results of Experiment 6 confirmed the key findings of Yuan and Fisher (2006, 2009). As in Yuan and Fisher's study, children in the two dialogue groups had access to the same syntactic and visual information during the test phase. Thus, the difference between the two dialogue conditions in the same-verb groups indicates that the children gathered distributional facts about the verb while listening to the dialogues, retrieved this information during the test phase when they heard the verb again, and used it to identify an appropriate referent for the verb. The finding that the different-verb groups did not show an effect of dialogue during the test phase provides evidence that children attached distributional facts to a particular new verb and
used the reappearance of the same verb during the test trials as a cue to retrieve that
distributional knowledge.

Second, our results extend Yuan and Fisher’s (2006, 2009) findings by showing that what
children learn from listening experience is not limited to the number of noun-phrase arguments
that occur with a new verb. In Experiment 6, children had to gather role-relevant information
about the new verb’s arguments from the dialogues in order to successfully identify the referent
of the verb in the test phase. The dialogues contained two role-relevant cues: intransitive-subject
animacy and lexical overlap between the transitive and intransitive sentences. The finding that
the children in the same-verb groups interpreted the new verb appropriately suggests that they
extracted from their listening experience information about the lexical items that occurred in the
verb’s argument slots and/or a coarse semantic encoding of its arguments (i.e. animate vs.
inanimate). Because the dialogue conditions in Experiment 6 differed in both animacy and
lexical overlap, we cannot determine which cue children used to succeed in the task. One cue
may have been more useful than the other, or children may have required both cues to succeed.
Future experiments will disentangle these two information sources, estimating their relative
weight in verb interpretation.

Third, our results add to a growing body of evidence that young children represent verbs
somewhat independently of the sentential context in which they occur (Fernandes et al., 2006;
Naigles, Bavin, & Smith, 2005; Yuan & Fisher, 2006, 2009). In the present case, children heard
the new verb in a transitive sentence and as a bare gerund (“Find dacking”) at test. In order to
select an appropriate referent for the verb at test, however, children had to retrieve information
about previous encounters with the same verb in an intransitive sentence. Thus, the results of
Experiment 6 suggest that 2-year-olds can use knowledge of a verb observed in one syntactic structure to sensibly interpret that verb in a different structure.

**Mechanisms**

By what mechanism did children use the distributional information present in the dialogues to constrain their interpretation of the verb? We can envision two classes of mechanisms for the cues we provided, which for ease of discussion we will refer to as the category-mediated and direct-inference mechanisms.

**Category-mediated.** One possibility is that the children used the distributional cues in the dialogues to assign meaning via previously-learned verb categories. The causal and the unspecified-object alternations define categories of verbs that share both syntactic and semantic similarity. Similar categories are found across languages, but the set of particular verbs that participate in each syntactic alternation appears to be constrained by semantic restrictions that are subtle and language-specific, and thus clearly learned (e.g., Pinker, 1989). For instance, the verb *bounce* (*It bounced/She bounced it*), along with other verbs of manner of motion, participates in the causal alternation in English, but the verb *fall* (*It fell/*She fell it*), along with other verbs of inherently-directed motion, does not. Once a class that has both syntactic and semantic properties is created, it can be used to make inferences about new words (e.g., Brooks & Tomasello, 1999; Gropen, et al., 1989; Pinker et al., 1987).

Existing experimental evidence for such class-based inferences in verb learning comes only from studies with older children (e.g., Ambridge et al., 2008; Brooks & Tomasello, 1999; Gropen et al., 1989); however, learners draw inferences of the same kind in learning other categories of words from a very early age. For example, Booth and Waxman (2003) showed that 14-month-olds had different expectations for the meanings of novel words presented as count
nouns ("This is a blicket!") versus adjectives ("This is a blickish one!"). Slightly older children
use morphological context to discriminate count nouns from proper names ("This is Blicket!"; Hall & Lavin, 2004). The creation of grammatical categories and sub-categories based on
semantic and distributional learning, and their use to guide new word-learning, has long been
assumed as a core mechanism of syntax acquisition (e.g., Brown, 1957; Maratsos & Chalkley,
1980; Waxman & Markow, 1995).

Such category-mediated inferences could explain the results of Experiment 6. For instance, the children may have previously learned that in English, some verbs that can be both
transitive and intransitive share the meaning “action culminating in a noteworthy change.” They
might also note that the verbs in this class share distributional features: their intransitive subjects
tend to be inanimate and they often have the same nouns in subject and object positions. Upon
encountering a new verb that demonstrates these distributional properties, children could extend
the meaning of the category to the new verb.

The utility of such an inference in our task does not require that 28-month-olds have
already worked out the refined semantic restrictions that help determine which verbs can
participate in the causal and unspecified-object alternations in English (i.e. the intricacies that
Pinker (1989) termed narrow-range restrictions). Rather, children could succeed in our task via
this category-mediated mechanism as soon as they have established a rough category of
meanings associated with each alternation. Although this category could be too broad, including
meanings that do not participate in the alternation in English, it would still permit children to
make useful inferences about new verbs (see Ambridge et al., 2008, for evidence of a broad
causal alternation category in 5-year-olds).
Direct-inference. Alternatively, children may have used the distributional cues to infer facts about the new verb directly, without using a previously-learned class. Consider first lexical overlap. Particular nouns differ in their tendency to occur in different grammatical positions. Given such lexical asymmetries, learners could develop expectations about which nouns make good subjects and objects (Resnik, 1996). In principle, this information could be represented both verb-generally and relative to particular verbs (e.g., *juice* generally makes a better object than a subject, and makes a good object for *drink*, but not for *break*), and could be used to infer roles in sentences. For instance, upon encountering the sentences “Matt dacked the pillow” and “The pillow dacked” in Experiment 6, children may have tended to assign the *pillow* an undergoer role in the intransitive sentence. They could have done so because: (a) in their experience with English, *pillow* has more often occurred as an object than a subject, so they assumed it played an object-like role, even when it occurred in subject position; or because (b) when they encountered *pillow* as the object of the transitive sentence, they added it to the set of *dackable* things and extended this interpretation to its occurrence in the intransitive sentence. By either route, lexical overlap between syntactic positions could yield direct information about likely role assignments for noun-phrases in intransitive subject position.

Turning to animacy, children could again have used this cue to infer role assignments directly rather than via an established syntactic/semantic category of verbs. A number of researchers have proposed that the linguistic system has a built-in tendency to align particular semantic roles with different levels of animacy (e.g., animate $\rightarrow$ agent, inanimate $\rightarrow$ patient; Aissen, 1999; Dowty, 1991). The correspondence between semantic roles and animacy could also arise from children’s strong expectations about the capacities of animate and inanimate entities. Toddlers know that animate entities can move on their own while inanimate things
cannot, for example (Golinkoff, 1975; Massey & Gelman, 1988; see Luo, Kaufman, & Baillargeon, 2009, for a similar distinction in 5-month-old infants). These two routes differ in whether the link between animacy and agency is part of a unique endowment for language acquisition (i.e. a Universal Grammar), or is an effect of conceptual knowledge on language interpretation. However, either of these routes would enable children to infer facts about likely semantic role assignments directly from animacy cues.

Prior evidence suggests that somewhat older children use animacy via one of these routes to infer role assignments. Gelman and Koenig (2001) found that 5-year-olds used animacy to assign roles to the subjects of intransitive sentences containing the verb *move*. When the subject was animate (“The dog moved”), children assumed it was the agent of its own motion; when the subject was inanimate (“The cup moved”), children were more likely to assume the object was undergoing externally-caused motion. Relatedly, Becker (2007) proposed that preschoolers use subject animacy to distinguish two syntactically confusable classes of verbs, known as control verbs (e.g., *want* in *The sheep wanted to disappear* vs. *The hay wanted to disappear*) and raising verbs (*seem* in *The sheep seemed to disappear* vs. *The hay seemed to disappear*). The underlying intuition here is that *seem* cannot have a meaning similar to *want* if it freely occurs with inanimate subjects. Similarly, the children in Experiment 6 could have inferred that the animate subjects of intransitive sentences in the dialogues were likely to be agents, but that the inanimate subjects (e.g., *pillow*), which could not initiate action on their own, were better suited to an undergoer role.

The direct-inference mechanism is interesting for two reasons. First, it would provide a route whereby the distributional cues we manipulated, subject animacy and lexical overlap, could be inherently meaningful to children without the prior establishment of a semantic/syntactic
subcategory of verbs, and thus could guide early word learning. Second, this mechanism is intriguing because it hints that children may have assigned a partial interpretation to the sentences containing the novel verbs while they listened to the dialogues, even though no referential scene was provided. One way to draw inferences about what types of roles particular entities play is to speculate about what those roles might be, using the ordinary processes of sentence comprehension. Adults watching our dialogue videos have the experience of wondering what it means for the pillow (or Matt) to dack. Children may do the same, using the tentative role assignments that result from their attempts to comprehend the dialogues to interpret the novel verb when they encounter it again in the test phase.

At present, we cannot determine which mechanism the children in Experiment 6 used to interpret the distributional cues provided in the dialogues. As noted above, however, there is clear evidence that older children and adults use both mechanisms to interpret sentences (Ambridge et al., 2008; Brooks & Tomasello, 1999; Gelman & Koenig, 2001; Gropen et al., 1989; Pinker et al., 1987), that category-mediated inferences guide word interpretation even in infancy (e.g., Booth & Waxman, 2003), and that 2-year-olds interpret sentences in part based on the plausible roles that the nouns in various argument positions could play (Chapman & Kohn, 1978). On such grounds, we might speculate that the 28-month-olds in our task already relied on a combination of the two mechanisms. We anticipate that the dialogue and test method used here can provide a powerful route for investigating these questions, in future experiments that vary the distributional cues offered in the dialogues, and the referential options provided at test.

The present results provide new evidence of the 2-year-old’s prowess in distributional learning, and the relationship of that learning to the creation of a meaningful lexicon. Ordinary child-directed speech contains distributional cues that can be used to tell apart syntactically
confusable classes of verbs. Two-year-olds can detect and retain these useful distributional cues from listening experience alone. The listener who hears “Matt dacked the pillow! And the pillow dacked!” can draw, at best, only highly abstract conclusions about what it means to *dack*. Despite this referential uncertainty, children encode distributional facts relevant to the characteristics of the nouns that fill its argument slots.
CHAPTER 5

CONCLUSIONS

Previous research on early verb learning has shown that when children hear a verb in the presence of a relevant referential scene, they use a variety of linguistic and non-linguistic cues to identify the verb’s likely referent. This research has largely focused on highly informative contexts, in which the constraints available to the child pick out a single potential referent from the accompanying scene. However, children frequently encounter verbs in less informative learning situations. They hear verbs in sentences that are consistent with multiple potential interpretations within the referential scene, and they hear verbs in the absence of relevant referential scenes. In order to learn from these types of observations, children must be able to form a proto-lexical entry for a verb, even when the verb’s meaning is uncertain, and then attach to it information about how the verb is used. They must then integrate that information with information gathered from later observations in order to learn the verb’s meaning.

The studies reported in this dissertation examined whether children could learn anything about verb meaning under the circumstances of referential uncertainty described above. The findings from these studies shed light on what kinds of information children can attach to a verb’s proto-lexical entry, as well as what types of observations children can compare and integrate when learning verbs.

Chapter 2 investigated situations where the available sentence information was consistent with multiple potential referents and asked whether children could identify a verb’s referent by taking into account additional referential contexts. Results showed that 31-month-olds were able to encode information about the potential referents that occurred with a verb, despite the fact that they never received a single unambiguous labeling trial. This is the first evidence that children
can attach referential information to a verb’s proto-lexical entry under circumstances of referential uncertainty. Children were also able to update information across multiple ambiguous trials in order to identify the verb’s referent. While previous experiments have shown that children can use cross-situational consistency to identify the referents of novel nouns under referential uncertainty (Smith & Yu, 2008), the present experiments are the first to demonstrate that children can also do so with novel verbs. In addition, children were able to successfully locate the correct referents within the first block of trials, demonstrating that they can compare across observations even when the actors and objects change across referential scenes.

These impressive abilities, however, are not without their limits. Experiment 1 tested children’s ability to learn intransitive verbs for solo actions, whereas Experiment 2 tested children’s ability to learn transitive verbs for actions on objects. While children were able to identify the verbs’ referents in Experiment 1, only the high-vocabulary children were able to do so in Experiment 2. Children whose vocabulary scores fell below the median failed to learn any of the verbs in Experiment 2. The increase in complexity of the sentences and/or scenes in Experiment 2 made it too difficult for the low-vocabulary children to keep track of the pairings of verbs and actions. Thus, our results suggest that while children can attach referential information to a verb’s lexical entry under referential uncertainty, this mechanism is limited, even in 2.5-year-old children. In situations of greater referential uncertainty, as might be expected in natural language learning, this mechanism may require support from additional sources of information.

Chapter 3 also examined situations in which a sentence was consistent with multiple potential referents; in this case, we presented children with additional sentence structures and asked whether children could identify a verb’s referent by considering the set of sentence frames
in which a verb occurred. These studies focused on two sets of sentence frames: the causal alternation and the unspecified-object alternation. Our findings showed that by 28 months, children can draw appropriate inferences about verbs that occur in either of these alternations. Furthermore, considering the full set of sentence frames allows them to arrive at a more refined interpretation of the verb than if they had considered just one of the two sentence frames in the alternation. Children who heard the intransitive sentence *The boy is pimming* were only able to infer that the verb described something interesting that the boy was doing (and thus their interpretation of this sentence varied as a function of the referential context). In contrast, children who heard the causal alternation were able to draw a more specific inference about the verb’s meaning: it described an action on an object that produced an effect (i.e. caused-motion). Thus, the results of Chapter 3 complement those of Chapter 2 by showing that children can identify a verb’s referent by either integrating referential information across multiple scene observations, or integrating sentence information across multiple sentence observations.

Chapter 4 investigated children’s ability to learn from informative sentences that occurred in the absence of a relevant referential scene. In Experiment 5, we showed that child-directed speech contains distributional cues that reflect the underlying argument-structure differences between causal and unspecified-object verbs (i.e. intransitive-subject animacy and lexical overlap between nouns in subject and object position). In Experiment 6, children first watched a video of two women using the novel verb *dacking* in a conversation; the conversation exhibited the animacy and lexical overlap patterns of either the causal or the unspecified-object alternation. Children then viewed a caused-motion and contact-activity event and heard, “The girl is dacking the boy! Find dacking.” Children who heard the causal dialogues looked longer at the caused-motion event than those who heard the unspecified-object dialogues.
These results confirm previous findings (Yuan & Fisher, 2009) that when children hear a verb in the absence of a relevant referential scene, they can encode facts about the sentences in which the verb occurs. Our results extend these findings by showing that what children learn from listening experience is not limited to the number of noun-phrase arguments that occur with a new verb. In Experiment 6, children had to gather role-relevant information about the new verb’s arguments from the dialogues in order to successfully identify the referent of the verb in the test phase. The finding that the children interpreted the new verb appropriately in the test phase suggests that while listening to the dialogues, they encoded information about the lexical items that occurred in the verb’s argument slots and/or a coarse semantic encoding of its arguments (i.e. animate vs. inanimate). In addition, these findings provide additional evidence for children’s ability to integrate across observations: children integrated syntactic facts gathered at one time point with referential facts gathered at a later time point in order to identify the verb’s referent.

In summary, this dissertation explored children’s ability to learn a verb under referential uncertainty. The findings presented here show that children can form proto-lexical entries for verbs, despite uncertainty about the verb’s referent. Having formed a proto-lexical entry, they can attach to it facts about the set of potential referents for a verb, the sentence structures in which the verb occurs, and role-relevant information about the verb’s arguments. Children can identify a verb’s referent by integrating information across different referential scenes, integrating information across sentence structures, and integrating previous sentence information with later scene information. These studies thus provide evidence that children can learn verbs under referential uncertainty. They do so by routinely encoding information about the scenes, sentence structures, and nouns that accompany new verbs, even when they do not yet know the
verb’s meaning. They can then bring this information to bear on future observations, allowing them to refine their interpretation of a verb over time.
FIGURES

Trial 1: “She’s pimming. And she’s nading.”

Trial 2: “She’s tazzing. And she’s pimming.”

Trial 3: “She’s nading. And she’s rivving.”

*Figure 1.* Three sample trials from Experiment 1 with accompanying audio.
Figure 2. Proportion of fixations to verb1 match, verb2 match, and away, separately for high-vocabulary (A) and low-vocabulary (B) groups. Vertical lines indicate the average onset of verb1 and verb2.
Trial 1: “Look! She’s nading her toy…and she’s pimming her toy.”

Figure 3. Sample trial from Experiment 2 with accompanying audio.
Figure 4. Proportion of fixations to verb1 match, verb2 match, and away, separately for high-vocabulary (4A) and low-vocabulary (4B) groups. Vertical lines indicate the average onset of verb1 and verb2.
Figure 5. Caused-motion event (left) and contact-activity event (right) used in Experiment 3.
Figure 6. Caused-motion event (left) and contact-activity event (right) used in Experiment 4.
Character-Familiarization

**Familiarization 1:** Find the boy! Where’s the boy? Look at the boy. (4s)

**Familiarization 2:** Find the girl! Look at the girl. Where’s the girl? (4s)

Practice

Blank screen: The boy is gonna jump.

**Practice 1:** The boy is jumping. The boy is jumping. See? The boy is jumping! (8s)

Blank screen: The boy jumped. Find jumping.

**Practice 2:** The boy is jumping. Find jumping. Find jumping (8s)

Blank screen: The girl is gonna clap.

**Practice 3:** The girl is clapping. The girl is clapping. See? The girl is clapping. (8s)

Blank Screen: The girl clapped. Find clapping.

**Practice 4:** The girl is clapping. Find clapping. Find clapping (8s)

Test

Blank screen: The girl is gonna pim the boy. The boy is gonna pim.

**Test 1:** The boy is pimming. The girl is pimming the boy. See? (8s)

Blank screen: The girl pimmed the boy. The boy pimmed. Find pimming.

**Test 2:** The boy is pimming. The girl is pimming the boy. See? (8s)

*Figure 7.* Video sequence for Experiment 3 (Test trials show Causal audio).
Figure 8: Training and test phases for the novel verb (Experiment 6). Test events: Contact-activity event (left) and caused-motion event (right).

**Causal dialogue**
A: Matt dacked the pillow.
B: Really? He dacked the pillow?
A: Yeah. The pillow dacked.
B: Right. It dacked.

**Unspecified-object dialogue**
A: Matt dacked the pillow.
B: Really? He dacked the pillow?
A: Yeah. He dacked.
B: Right. He dacked.

**Contact-activity test event**

**Causal test event**

Same-verb: “The girl is dacking the boy. Find dacking.”

Different-verb: “The girl is pimming the boy. Find pimming.”
**TABLES**

Table 1

_Novel Verbs and Actions Used in Experiment 1._

<table>
<thead>
<tr>
<th>Verb</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pimming</td>
<td>Tilting torso side to side</td>
</tr>
<tr>
<td>Nading</td>
<td>Squatting</td>
</tr>
<tr>
<td>Tazzing</td>
<td>Lifting knee</td>
</tr>
<tr>
<td>Rivving</td>
<td>Stretching arms over head</td>
</tr>
</tbody>
</table>
**Table 2**

*Mean (SD) Overall and Block 1 Target-advantage Acroses (in Seconds) for Experiment 1.*

<table>
<thead>
<tr>
<th>Verb</th>
<th>Block 1</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pimming</td>
<td>.09 (.28)</td>
<td>.10 (.87)</td>
</tr>
<tr>
<td>Nading</td>
<td>.37 (.12)</td>
<td>.26 (.88)</td>
</tr>
<tr>
<td>Tazzing</td>
<td>.10 (.06)</td>
<td>.16 (.69)</td>
</tr>
<tr>
<td>Rivving</td>
<td>.48 (.36)</td>
<td>.29 (.84)</td>
</tr>
<tr>
<td>Overall</td>
<td>.28 (.59)</td>
<td>.21 (.41)</td>
</tr>
</tbody>
</table>
Table 3

Novel Verbs and Actions Used in Experiment 2, Separately by List.

<table>
<thead>
<tr>
<th>Verb</th>
<th>List 1</th>
<th>List 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pimming</td>
<td>Sliding in an S on table-top</td>
<td>Spinning on table-top</td>
</tr>
<tr>
<td>Nading</td>
<td>Lifting and lowering in one hand</td>
<td>Tossing from hand to hand</td>
</tr>
<tr>
<td>Tazzing</td>
<td>Tossing from hand to hand</td>
<td>Lifting and lowering in one hand</td>
</tr>
<tr>
<td>Rivving</td>
<td>Spinning on table-top</td>
<td>Sliding in an S on table-top</td>
</tr>
</tbody>
</table>
Table 4

*Mean (SD) Overall and Block 1 Target-advantage Scores (in Seconds) for Experiment 2, Separately by Vocabulary Group.*

<table>
<thead>
<tr>
<th>Vocabulary group</th>
<th>Verb</th>
<th>Block 1</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pimming</td>
<td>.07 (.90)</td>
<td>.26 (.52)</td>
</tr>
<tr>
<td>High Vocabulary</td>
<td>Nading</td>
<td>.55 (1.04)</td>
<td>.38 (.72)</td>
</tr>
<tr>
<td></td>
<td>Tazzing</td>
<td>.31 (1.43)</td>
<td>.06 (.94)</td>
</tr>
<tr>
<td></td>
<td>Rivving</td>
<td>.98 (1.00)</td>
<td>.43 (.80)</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>.45 (.62)</td>
<td>.28 (.32)</td>
</tr>
<tr>
<td></td>
<td>Pimming</td>
<td>.76 (1.09)</td>
<td>.31 (.69)</td>
</tr>
<tr>
<td>Low Vocabulary</td>
<td>Nading</td>
<td>-.56 (1.24)</td>
<td>-.41 (.75)</td>
</tr>
<tr>
<td></td>
<td>Tazzing</td>
<td>.16 (1.29)</td>
<td>-.10 (.71)</td>
</tr>
<tr>
<td></td>
<td>Rivving</td>
<td>.20 (1.74)</td>
<td>-.07 (1.08)</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>.15 (.67)</td>
<td>-.07 (.37)</td>
</tr>
</tbody>
</table>
Table 5

*Audio Conditions for Experiment 3.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sentence pairs for Test Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal</td>
<td>“The boy is pimming. The girl is pimming the boy.”</td>
</tr>
<tr>
<td>Unspecified-object</td>
<td>“The girl is pimming. The girl is pimming the boy.”</td>
</tr>
<tr>
<td>Boy-subject</td>
<td>“The boy is pimming. The boy is pimming.”</td>
</tr>
<tr>
<td>Girl-subject</td>
<td>“The girl is pimming. The girl is pimming.”</td>
</tr>
<tr>
<td>Neutral</td>
<td>“Look! Isn’t that fun? See?”</td>
</tr>
</tbody>
</table>


Table 6

Mean (SD) Looking Time to the Caused-motion and Contact-activity Events (in seconds) During Each Test Trial in Experiment 3, Separately by Condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 1 Caused-motion</th>
<th>Trial 1 Contact-activity</th>
<th>Trial 2 Caused-motion</th>
<th>Trial 2 Contact-activity</th>
<th>Average Caused-motion</th>
<th>Average Contact-activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal</td>
<td>3.94 (.67)</td>
<td>3.76 (.68)</td>
<td>4.49 (1.50)</td>
<td>3.07 (1.55)</td>
<td>4.22 (.80)</td>
<td>3.42 (.77)</td>
</tr>
<tr>
<td>Unspecified-object</td>
<td>2.78 (1.39)</td>
<td>5.00 (1.37)</td>
<td>2.64 (1.39)</td>
<td>4.99 (1.34)</td>
<td>2.71 (1.31)</td>
<td>4.99 (1.24)</td>
</tr>
<tr>
<td>Boy-subject</td>
<td>3.34 (1.32)</td>
<td>4.49 (1.36)</td>
<td>5.10 (1.14)</td>
<td>2.41 (1.36)</td>
<td>4.16 (.91)</td>
<td>3.53 (.86)</td>
</tr>
<tr>
<td>Girl-subject</td>
<td>3.53 (.80)</td>
<td>4.25 (.87)</td>
<td>2.75 (1.28)</td>
<td>4.66 (1.39)</td>
<td>3.14 (.86)</td>
<td>4.46 (.95)</td>
</tr>
<tr>
<td>Neutral</td>
<td>2.98 (.84)</td>
<td>4.80 (.86)</td>
<td>3.02 (1.56)</td>
<td>4.45 (1.46)</td>
<td>3.00 (.57)</td>
<td>4.63 (.61)</td>
</tr>
</tbody>
</table>
Table 7

_Mean (SD) Looking Time to the Caused-motion and Contact-activity Events (in Seconds) During Each Test Trial in Experiment 4, Separately by Condition._

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 1 Caused-motion</th>
<th>Trial 1 Contact-activity</th>
<th>Trial 2 Caused-motion</th>
<th>Trial 2 Contact-activity</th>
<th>Average Caused-motion</th>
<th>Average Contact-activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal</td>
<td>3.77 (1.78)</td>
<td>4.04 (1.79)</td>
<td>4.03 (1.56)</td>
<td>3.00 (1.50)</td>
<td>3.90 (1.31)</td>
<td>3.52 (1.23)</td>
</tr>
<tr>
<td>Unspecified-object</td>
<td>4.15 (1.10)</td>
<td>3.68 (1.05)</td>
<td>4.52 (2.20)</td>
<td>2.99 (2.07)</td>
<td>4.34 (1.14)</td>
<td>3.34 (1.04)</td>
</tr>
<tr>
<td>Boy-subject</td>
<td>4.10 (1.82)</td>
<td>3.60 (1.98)</td>
<td>2.80 (1.88)</td>
<td>4.74 (1.71)</td>
<td>3.34 (1.60)</td>
<td>4.26 (1.59)</td>
</tr>
<tr>
<td>Girl-subject</td>
<td>3.46 (.90)</td>
<td>4.27 (.80)</td>
<td>4.49 (1.85)</td>
<td>2.91 (2.02)</td>
<td>3.98 (.92)</td>
<td>3.59 (.94)</td>
</tr>
<tr>
<td>Neutral</td>
<td>3.25 (1.19)</td>
<td>4.58 (1.23)</td>
<td>2.32 (1.33)</td>
<td>5.40 (1.28)</td>
<td>2.79 (1.06)</td>
<td>4.99 (1.07)</td>
</tr>
<tr>
<td>Transitive</td>
<td>2.98 (1.39)</td>
<td>4.70 (1.70)</td>
<td>3.47 (2.17)</td>
<td>4.02 (2.20)</td>
<td>3.26 (1.38)</td>
<td>4.28 (1.62)</td>
</tr>
</tbody>
</table>
Table 8

*Verbs Used in Experiment 5.*

<table>
<thead>
<tr>
<th>Verb class</th>
<th>Selected verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unspecified-Object</td>
<td>bite, draw, drink, eat, hit, play, pull, push, read, see, throw, tickle, try, wash, write</td>
</tr>
<tr>
<td>Causal</td>
<td>bounce, break, change, close, fold, move, open, pop, roll, shut, slide, spill, tear, turn</td>
</tr>
</tbody>
</table>
Table 9

*Mean (SD) for Each Variable in Experiment 5, Separately by Verb Class.*

<table>
<thead>
<tr>
<th></th>
<th>Causal</th>
<th>Unspecified-object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitivity</td>
<td>.71 (.20)</td>
<td>.67 (.22)</td>
</tr>
<tr>
<td>Overall subject animacy</td>
<td>.87 (.12)</td>
<td>.98 (.05)</td>
</tr>
<tr>
<td>Intransitive-subject animacy</td>
<td>.44 (.27)</td>
<td>.97 (.09)</td>
</tr>
<tr>
<td>Lexical overlap</td>
<td>.44 (.14)</td>
<td>.29 (.26)</td>
</tr>
</tbody>
</table>
Table 10

Accuracy (Proportion of Verbs Classified Correctly) and Adjusted Rand Index ($R_{adj}$) Scores for k-means Clustering Solutions Based on Each Variable, and on all Variables.

<table>
<thead>
<tr>
<th>Variable used</th>
<th>Accuracy</th>
<th>$R_{adj}$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitivity</td>
<td>.59</td>
<td>-.001</td>
<td>.43</td>
</tr>
<tr>
<td>Overall subject animacy</td>
<td>.62</td>
<td>.04</td>
<td>.17</td>
</tr>
<tr>
<td>Intransitive subject animacy</td>
<td>.83</td>
<td>.41</td>
<td>.0004</td>
</tr>
<tr>
<td>Lexical overlap</td>
<td>.69</td>
<td>.12</td>
<td>.04</td>
</tr>
<tr>
<td>All variables</td>
<td>.83</td>
<td>.41</td>
<td>.0004</td>
</tr>
</tbody>
</table>
Table 11

*Mean (SD) Proportion of Looking Time to the Caused-motion Event During Each Test Trial in Experiment 6, Separately by Dialogue and Test-verb Condition.*

<table>
<thead>
<tr>
<th>Test verb</th>
<th>Dialogue</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same-verb</td>
<td>Causal</td>
<td>.45 (.24)</td>
<td>.77 (.18)</td>
<td>.61 (.17)</td>
</tr>
<tr>
<td></td>
<td>Unspecified-object</td>
<td>.44 (.22)</td>
<td>.33 (.31)</td>
<td>.39 (.21)</td>
</tr>
<tr>
<td>Different-verb</td>
<td>Causal</td>
<td>.47 (.25)</td>
<td>.38 (.23)</td>
<td>.43 (.19)</td>
</tr>
<tr>
<td></td>
<td>Unspecified-object</td>
<td>.43 (.15)</td>
<td>.40 (.33)</td>
<td>.41 (.17)</td>
</tr>
</tbody>
</table>
REFERENCES


Piccin, T., & Waxman, S.R. (in prep.). Children use contrast and cross-situational exposure to learn verbs.


## APPENDIX

Dialogues Used in Experiment 6, Separately by Dialogue Condition

<table>
<thead>
<tr>
<th>Dialogue</th>
<th>Causal</th>
<th>Unspecified-object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialogue 1:</td>
<td>A: Matt dacked the pillow.</td>
<td>A: Matt dacked the pillow.</td>
</tr>
<tr>
<td></td>
<td>B: Really? He dacked the pillow?</td>
<td>B: Really? He dacked the pillow?</td>
</tr>
<tr>
<td></td>
<td>B: Right. It dacked.</td>
<td>B: Right. He dacked.</td>
</tr>
<tr>
<td>Dialogue 2:</td>
<td>A: Kelly is gonna dack Adam.</td>
<td>A: Kelly is gonna dack Adam.</td>
</tr>
<tr>
<td></td>
<td>B: Hmm…she’s gonna dack Adam?</td>
<td>B: Hmm…she’s gonna dack Adam?</td>
</tr>
<tr>
<td></td>
<td>B: Great. He’s gonna dack.</td>
<td>B: Great. She’s gonna dack.</td>
</tr>
<tr>
<td>Dialogue 3:</td>
<td>A: Jessica is gonna dack the flower.</td>
<td>A: Jessica is gonna dack the flower.</td>
</tr>
<tr>
<td></td>
<td>B: Wow! She’s gonna dack the flower?</td>
<td>B: Wow! She’s gonna dack the flower?</td>
</tr>
<tr>
<td></td>
<td>A: Yeah. The flower is gonna dack.</td>
<td>A: Yeah. She’s gonna dack.</td>
</tr>
<tr>
<td></td>
<td>B: Yeah. It’s gonna dack.</td>
<td>B: Yeah. She’s gonna dack.</td>
</tr>
</tbody>
</table>