

## Action Speaks Louder Than Words: Young Children Differentially Weight Perceptual, Social, and Linguistic Cues to Learn Verbs

Amanda C. Brandone  
*University of Michigan*

Khara L. Pence  
*University of Virginia*

Roberta Michnick Golinkoff  
*University of Delaware*

Kathy Hirsh-Pasek  
*Temple University*

This paper explores how children use two possible solutions to the verb-mapping problem: attention to perceptually salient actions and attention to social and linguistic information (speaker cues). Twenty-two-month-olds attached a verb to one of two actions when perceptual cues (presence/absence of a result) coincided with speaker cues but not when these cues were placed into conflict (Experiment 1), and not when both possible referent actions were perceptually salient (Experiment 2). By 34 months, children were able to override perceptual cues to learn the name of an action that was not perceptually salient (Experiment 3). Results demonstrate an early reliance on perceptual information for verb mapping and an emerging tendency to weight speaker information more heavily over developmental time.

Understanding the process of verb learning is crucial to any theory of language acquisition. Verbs are the architectural centerpiece of a sentence. They allow us to talk about, not only people and things, but, crucially, the relations between them. Verbs are necessary to express everything from the simplest to the most complex events; without a verb, there is no sentence.

Abundant research suggests that verb learning is complicated; in many languages it lags behind and is more complex than noun learning (Bornstein et al., 2004; Gentner, 1982; but see Tardif, 1996). As Gentner (1982) suggested, verbs do not directly label actions in the same way that nouns label objects. Nouns commonly refer to objects that are naturally perceived as distinct units. Verbs, however, refer to relations within events, and any event can be conceptualized in terms of a multitude of different components, including, but not limited to, *path* (the trajectory of an action with respect to some reference point, e.g., approach, enter),

*manner* (how an action is carried out, e.g., walk, swagger, stroll), *result* (e.g., open, fill), and *instrument* (e.g., hammer, shovel) (Talmy, 1985). It is the task of the verb learner to decide which relation or relations in an event is the verb referent. Deciphering which aspect of an event is being labeled is often difficult, even for adults (Gillette, Gleitman, Gleitman, & Lederer, 1999). Further, there is much cross-linguistic variation as to which elements of motion events are most likely to be verb referents (Talmy, 1985). Thus, learning verbs involves disentangling a variety of simultaneously occurring components and deciding between a number of possible meanings.

Proliferating evidence from a variety of languages confirms that verb learning is difficult (Hirsh-Pasek & Golinkoff, 2006; but see Tardif, 1996). Verbs are hard to learn, even for older children (Forbes & Farrar, 1995), and even in languages such as Japanese where verbs can appear in isolation and in salient sentence-final position (Bornstein et al., 2004; Imai, Haryu, & Okada, 2002; Imai et al., 2006). However, despite the difficulty in learning verbs, words for actions or events appear among children's earliest receptive and production vocabularies (Bates, Bretherton, & Snyder, 1988; Bloom, 1993; Bloom, Tinker, & Margulis, 1993; Bowerman, 1974; Gopnik & Meltzoff, 1986; Hampson, 1989; Huttenlocher, Smiley, & Charney, 1983; Lieven, Pine, & Barnes, 1992; McCune-Nicolich, 1981; Nelson, 1973;

---

The research described here was supported by grants #SBR601306 and SBR 9615391 from the National Science Foundation to the third and fourth authors. We are grateful to all of the children who participated in these studies and to their parents. We extend special thanks to Dede Addy, Rachel Pulverman, Sara Salkind, Rebecca Seston, and Jenny Sootsman Buresh for their contribution to the study design and their assistance in data collection and coding. We also thank Letitia Naigles and two anonymous reviewers for very constructive comments on earlier versions of this manuscript.

Correspondence concerning this article should be addressed to Amanda Brandone, Department of Psychology, University of Michigan, Ann Arbor, MI 48109. Electronic mail may be sent to Brandone@umich.edu.

Nelson, Hampson, & Shaw, 1993; Pine, 1992; Tomasello, 1992; Tomasello & Todd, 1983).

A central question, then, is how children first attach verbs to their referents. To address this question, we use the Emergentist Coalition Model (ECM) of word learning (Hirsh-Pasek, Golinkoff, & Hollich, 2000; Hollich, Hirsh-Pasek, & Golinkoff, 2000) as a framework. The ECM argues that children have at their disposal a range of cues (perceptual, social, and linguistic) that they can utilize to attach a novel label to a novel referent. This paper considers how children make use of two possible solutions to the mapping problem for verbs: attention to perceptually salient actions in the environment and attention to social and linguistic information provided by a speaker. The current experiments ask: (a) are young children biased to rely on perceptual salience as a cue to what verbs label and (b) if so, can they override these perceptual preferences in favor of social and linguistic cues that might more precisely alert children to the package of information to be encoded in the verb?

#### *Solution 1: Attention to Perceptually Salient Actions*

One possible solution to the referential ambiguity of verbs is that children may assume that words map onto the most perceptually salient actions in their environment. Children may simply map a novel verb onto the action or event that “stands out” in their world. Research in the domain of noun learning suggests that, in its earliest stages, children’s interest in objects guides word-to-world mapping (Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006). Could perceptual information offer a toehold into verb learning?

Naigles and Kako (1993) explored this question by showing young 2-year-olds ambiguous, compound events paired with bare novel verbs. For example, children saw a duck making a rabbit bend over while simultaneously both were circling their arms. Accompanying this scene, children heard, “Look! Gorp!” The actions were then separated with bending on one screen and circling on the other, and children were asked to select the action that was the referent of the verb. Results suggest that children come to the task of verb learning with initial preferences for particular actions or aspects of actions that appear to be based on perceptual salience. When causative actions (one character forces another to move in some way) were paired with synchronous actions (both characters move simultaneously), children preferred to map the verb to the *synchronous* actions. When causative actions were then paired with contact actions (one character simply touches the other), children pre-

ferred to map the verb to the *causative* actions. The three kinds of actions tested can be rank ordered such that synchronous actions are the most preferred, followed by causative, and finally contact actions. Naigles and Kako interpreted this pattern of results in terms of perceptual salience resulting from the *size* of the action. A follow-up study, in which adults were asked to rate the size of the action in each of the Naigles and Kako videos, revealed that toddlers’ preferences could indeed be predicted by action size. In the absence of further information, children preferred to map a label to the largest and most perceptually salient of the actions presented.

Maguire (2004) reported a similar finding. When 2- and 2½-year-olds were shown an ambiguous event composed of both a manner and a path (e.g., an animated starfish doing jumping jacks in an arc below a stationary ball in the center of the screen) accompanied by a novel verb (e.g., “Starry’s moding!”), children preferred to map the verb to the *path* rather than to the manner of the action. As in Naigles and Kako’s (1993) study, results may be attributed to the relative size of each action component. Whereas manners were defined as local movements of the character’s limbs, paths were larger trajectories pervading the entire screen. Thus, Maguire’s findings suggest that children preferred to map the label to the largest available salient action component.

Beyond action size, Lakusta and Landau (2005) demonstrated that children and adults have a robust bias to linguistically encode goal paths over source paths in events. When asked to describe manner of motion, change of possession, change of state, and attachment/detachment events, participants consistently included the goal path but not the source path (e.g. “The bird flew *into the pitcher*” not “The bird flew out of the bucket”). Lakusta and Landau interpreted these results as support for a goal-biased perspective on events, whereby conceptual endpoints are preferred over conceptual starting points. The source of this asymmetry can be attributed to perceptual salience. A model proposed by Regier (1996, 1997), for example, suggests that perceivers weight the endpoint of an event more strongly than the starting point. This early preference to attend to perceptual endpoints provides an additional toehold into the resolution of the verb-mapping problem.

Finally, in a series of experiments, preschool children were taught novel verbs that labeled events that included both an action and a result (Behrend, 1989, 1990; Forbes & Farrar, 1993; Forbes & Poulin-Dubois, 1997). Children between 2 and 5 years of age demonstrated a result bias. That is, children were least likely to extend a newly learned verb to an event in which

the result was different from the event originally used to teach the verb. These findings can be understood in terms of perceptual salience: children's preference to map a label to an action's result over its manner may be due, at least in part, to increased attention to perceptually salient results in an action. A result bias also helps constrain children's initial hypotheses about what a novel verb means.

Despite the apparent importance of action size, perceptual endpoint, and results in children's initial interpretations of novel verbs, none of these explanations capture the whole story of how young children learn the meaning of verbs. Children learn names not only for salient, perceptually available actions, but also for mundane actions (e.g., *stand*), actions without obvious results (e.g., *touch*), and even unperceivable actions (e.g., *think*, *love*). Verb learning requires more than default perceptual biases since the meaning of a verb depends upon more than its appearance. Children need to learn not only where to attend, but also how to package different elements into the verb referent.

#### *Solution 2: Attention to Linguistic and Social Information*

As an alternative to the perceptual salience account, to discover how a given language community packages events and how specific verbs encode event components, children may rely on information from the speaker. Such information includes *linguistic data* imparted in syntactic frames (e.g., syntactic bootstrapping; Fisher, 2002; Fisher, Hall, Rakowitz, & Gleitman, 1994; Gleitman, 1990; Gleitman & Gillette, 1995; Naigles, 1990, 1996) as well as *social-pragmatic information* conveyed by the extralinguistic context (e.g., comprehension of a speaker's attentional focus and communicative intent (Baldwin, 2000; Childers & Tomasello, 2002, 2006; Tomasello 1995; Tomasello, Strosberg, & Akhtar, 1996) and comprehension of actor intent (Behrend & Scofield, 2006; Poulin-Dubois & Forbes, 2002). Such cues in the social and linguistic context surrounding a verb constrain possible verb meanings.

The strongest data on the difficulty of determining verb meaning from perceptual cues alone come from studies by Gleitman and colleagues (Gillette et al., 1999; Snedeker & Gleitman, 2004). In their human simulation paradigm, adults viewed a series of video clips of a mother and child playing. The video clips were silent except for a beep that coincided with a missing verb. Participants guessed which verb the speaker might have used in place of the beep. The findings were dramatic: the proportion of correct guesses was a paltry 15% (as compared to 45% in

the case of nouns). In fact, if only responses for mental verbs were considered, the proportion of correct guesses dropped to zero! Although performance was best for actions that were perceptually available, performance for all actions increased when participants were given additional linguistic information (e.g., the nouns used in the sentence surrounding the verb and/or the frames in which the verbs occurred). From these results, Gleitman, Cassidy, Nappa, Papafragou, and Trueswell (2005) concluded that to determine the meaning of "hard words" (e.g., verbs—especially abstract verbs of perception, communication, and thinking), learners need supplementary linguistic evidence. Sophisticated linguistic knowledge is therefore required for children to figure out which aspects of the environment are coded in a verb.

Research is accumulating to suggest that linguistic information can alter preliminary perceptual preferences (e.g., Behrend, Harris, & Cartwright, 1995; Echols & Marti, 2004; Fisher, 2002; Imai et al., 2006; Maguire, 2004; Naigles & Kako, 1993). For example, Naigles and Kako (1993) demonstrated that before the age of 30 months, children can use syntactic information, particularly that encoded in the transitive verb frame, to redirect or refocus their attention on a less favored action (e.g., a causative action), thereby leading to a successful mapping. Similarly, when Maguire (2004) offered her participants multiple exemplars (e.g., Starry spinning over, under, past, around the ball) and additional syntactic information in the form of a prepositional phrase (e.g., "Starry's moding *over the ball*"), they also redirected their attention to the less-favored action component (manner). Behrend et al. (1995) showed that children even adjusted their result bias as a function of how the novel verb was inflected during training. Specifically, preschoolers were more likely to make manner interpretations of novel verbs presented with a progressive *-ing* ending than those presented with the past *-ed* ending.

In addition to the influence of linguistic knowledge is the influence of social-pragmatic information on verb learning. Prior to the time when they start adding verbs to their lexicon, young children already have an extensive understanding of social-pragmatics (Carpenter, Akhtar, & Tomasello, 1998; Gergely, Nadasdy, & Csibra, 1995; Leslie & Keeble, 1987; Meltzoff, 1995; Woodward, 1998). Children are able to apply two forms of social-pragmatic knowledge to the learning of novel verbs: (a) inferring the speaker's attentional focus and communicative intent, and (b) inferring the intent of the actor.

For example, Akhtar and Tomasello (1996) demonstrated that by 24 months of age, children are able to

infer the intent of a speaker to name a novel action, even when they have never seen that action performed in the presence of that label. Here the experimenter told children that they were going to *meeek* Big Bird. After searching, the experimenter informed the child that she could not find Big Bird. The target action was then performed with other objects but never labeled. As a test, children were asked to *meeek* a new character, Cookie Monster. Two-year-old children were able to produce the action with the novel object at the same rate as a group of children who heard the label while the action was performed. These results demonstrate children's ability to use their understanding of a speaker's communicated intent to resolve the problem of referential ambiguity and interpret the meaning of a novel verb.

In order to refine their understanding of verb meaning, children must also attend to the intent of the actor performing a labeled action. For example, Poulin-Dubois and Forbes (2002) found that 27-month-olds, but not 21-month-olds, could use social cues when distinguishing between novel actions that looked quite similar except for barely detectable social information. Specifically, 27-month-olds attended to the subtle cue of eye gaze when determining whether a verb meant something like *topple* or *knock over*. Children seem to understand that actions may be perceptually similar, but have distinct labels because of the intentions of the actor (see also Behrend & Scofield, 2006). These results hint at a sophisticated ability on the part of the verb learner to mine the social context and determine the meaning of a novel verb.

Thus, although support exists for children's use of preferences based on perceptual salience to determine the referential focus of a novel verb, research also suggests that these preferences are not sufficient. In order to successfully map an action or event onto a verb, children must rely on some combination of perceptual cues, linguistic information, and social intent. As of yet, it is unclear how young children coordinate these cues.

### *The Present Experiments*

Here we investigate young children's use of conflicting or coinciding perceptual, social, and linguistic information to determine how children discover the referent for a novel verb. Importantly, our purpose is not to study the separate use of social and linguistic cues, which often, though not always, occur in tandem as speaker information; rather, our goal is to tease apart the use of perceptual information from what we will refer to as *speaker information*—a com-

plex of social and linguistic cues provided by a speaker. As in the noun-learning work of Hollich et al. (2000) and Pruden et al. (2006), to arrive at the differential weightings of perceptual and speaker information early in the verb-learning process, we created a situation in which these cues could either coincide or conflict. By pitting these cues against one another, the current research addresses whether young children differentially weight perceptual salience and speaker information during verb learning.

Three experiments manipulated the perceptual and speaker information available to young children as they engaged in an active verb-learning task. We focused on action verbs *with* and *without* results. These verbs were selected for four reasons. First, verbs encoding action are among children's earliest verbs (Bloom, 1978). Second, research suggests that early in the verb-learning process, events are represented in terms of movements and outcome states—manners and results (Huttenlocher et al., 1983; Smiley & Huttenlocher, 1985; Wagner & Carey, 2005). Thus, basic action verbs with and without results serve as the simplest actions for young participants to comprehend and perform. Third, evidence suggests that results are an extremely salient aspect of an event (Behrend 1989, 1990; Forbes & Farrar 1993, 1995), and toddlers prefer to perform actions that create results (Gibson, 1969). Finally, these actions were selected based upon their experimental manipulability: the same action might or might not produce a result (e.g., pressing a button to produce a tone or not). Thus, the use of actions with and without results allowed us to manipulate the factor of perceptual salience, creating a situation in which perceptual cues could either *coincide* or *conflict* with speaker information.

In the present studies, children were presented with a pair of actions and were taught the label for one of those actions using a combination of perceptual and speaker cues. Cues to verb meaning in our experiments were not subtle. Perceptual salience was determined by whether or not the action produced an interesting result. If a result was produced, we inferred that the action was more salient than if it had no result. Because children were able to perform the actions themselves and therefore produce the results themselves, the perceptual salience of the action (or lack thereof) was clear. Speaker cues to verb meaning were also transparent. We did not require children to rely on subtle social cues, such as eye gaze, to determine the referent of the novel label; rather, the action performed during training was the action being labeled. Linguistic cues were provided by the grammar of the training sentences. Thus, social and

linguistic cues to verb reference were available through attention to the speaker during the time of labeling.

### Experiment 1: To What Extent Do Young Children Rely on Perceptual Salience as a Cue to Verb Meaning?

Each child was exposed to a pair of actions. Half of the children received a label for the action with a result (the coincident condition), whereas the other half received a label for an action without a result (the conflict condition). One action produced a salient result; the other action did not produce a result. If children weight perceptual salience more heavily than speaker cues in verb mapping, it should be easy for them to learn the name of the action with a result (because perceptual cues coincide with speaker information) and more difficult for them to learn the name of the action without a result (because perceptual cues and speaker information conflict). On the other hand, if children weight speaker cues more heavily than perceptual cues in verb learning, they should be able to learn the name of the action regardless of whether it produces a result. Based on prior work documenting the difficulty of verb learning (e.g., Hirsh-Pasek & Golinkoff, 2006) and children's early perceptual preferences (e.g., DeLoache, 2004; Smiley & Huttenlocher, 1985), we predicted that young children would encounter difficulty overcoming the lure of perceptual salience in a verb-learning task.

Verb mapping was assessed after training by presenting the actions on video in the Intermodal Preferential Looking Paradigm (IPLP; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996). Children were tested with a visual preference task rather than an act-out task to remove such performance obstacles as motivation and preference for a particular action.

#### Method

##### Participants

Participants included 32 native English-speaking toddlers (16 male, 16 female) ranging in age from 21.01 to 24.07 months ( $M = 22.48$  months,  $SD = 0.99$  months). Participants were predominantly Caucasian and of middle- to upper-middle class. The data from an additional 10 participants were excluded for the following reasons: child ended the study ( $n = 6$ ), child watched less than 65% of the video ( $n = 2$ ), and experimenter/technical error ( $n = 2$ ). Upon arrival, parents completed a MacArthur-Bates Communica-

tive Development Inventory Toddler Short Form (Fenson et al., 2000). Children's expressive vocabularies ranged from 4 to 99 words ( $M = 45.35$ ,  $SD = 25.75$  words).

##### Materials

We presented children with a 14 by 8 inch metal box. Affixed to the box were a light switch accompanied by a small red light and a Morse code key with a black button (see Figure 1). Flipping the light switch illuminated the red light and pressing the Morse code key produced an audible tone. A battery hidden inside the box powered the light and the tone. Importantly, either or both effects could be disabled by disconnecting the battery source.

##### Procedure

Children sat at a table on their parent's lap facing the experimenter. We instructed parents to refrain from speaking or assisting their child in any way during the course of the experiment.

*Demonstration phase.* The experimenter began by covering one side of the box with a transparent plastic container. Then the experimenter gave a brief demonstration of the action on the exposed side (see Table 1). Children were asked to perform that same action using a neutral request (e.g., "Can you do that?"). The transparent container was then moved to cover the other side of the box. The experimenter demonstrated the other action and gave the child an opportunity to perform that same action. All children were able to perform both of the actions.

*Free play and language training phase.* Following the demonstration phase, a 20-second free play trial was conducted in which the clear container was removed from the box. Children were permitted to perform



Figure 1. Toy used for the training portion of the study.

Table 1  
*Demonstration and Training Phases of Experiments 1–3*

Trial type	Experimenter actions	Experimenter directions	Participant actions
Demonstration of first action	Flips light switch up and down (button on Morse code key is covered with a transparent box)	“Hey, look what I can do on the box!”	
	Passes box to child	“Can you do that on the box?”	Performs requested action
	Retrieves box	“Yes, you can! Thank you.”	
Demonstration of second action	Presses button on Morse code key (light switch is covered with a transparent box)	“Now look what I can do on the box!”	
	Passes box to child	“Can you do that on the box?”	Performs requested action
	Retrieves box	“Yes, you can! Thank you.”	
Free play trial 1	Removes transparent container. Passes box to child.	“What can you do on the box?”	Free play with either or both actions
First series of labels	Retrieves box from child, then flips light switch up and down while looking back and forth between box and child	“I’m going to feb on the box.” Then, “Look! I’m febbing on the box” (5 times).	
Free play trial 2	Passes box to child	“What can you do on the box?”	Performs either or both actions
Second series of labels	Retrieves box from child, then flips light switch up and down while looking back and forth between box and child	“I’m going to feb on the box.” Then, “I’m febbing on the box” (5 times).	
Free play trial 3	Passes box to child	“What can you do on the box?”	Performs either or both actions

*Note.* The order in which the actions were performed was counterbalanced across participants as was the action labeled by the experimenter. In this example, in the Coincident Condition, flipping the light switch (the labeled action) produces a result; in the Conflict Condition, pressing the Morse code key (the unlabeled action) produces a result; in the Equal Salience Condition, both actions (labeled and unlabeled) produce a result.

either or both of the actions after a neutral request (e.g., “What can you do on the box?”).

The experimenter then announced, “I’m going to feb on the box,” as she removed the box from the child’s reach. The experimenter provided five additional labels in the present progressive sentence frame as she simultaneously demonstrated the labeled action (e.g., “I’m febbing on the box”). The experimenter looked back and forth between the child and the box throughout the labeling phase, maintaining joint attention. Children were then allowed another 20 seconds to explore the box after a neutral request from the experimenter (e.g., “What can you do on the box?”). A second series of labels and another 20-second play period followed, presented in the same manner as the first. Children were thus permitted to perform the actions and witness their results or lack thereof both before and after the selected action was labeled. Importantly, labeling occurred only when the box was in the experimenter’s possession and only the labeled action was performed by the experimenter during the training phase. The action labeled by

the experimenter was counterbalanced across participants.

In a between-subjects design, participants were randomly assigned to one of two conditions, defined by whether the speaker produced a label for the salient action with a result or for the nonsalient action without a result. In the *coincident* condition, perceptual cues were aligned with speaker information such that the labeled action produced a salient result (i.e., flipping the light switch illuminated a red light or pressing the Morse code key sounded a tone), whereas the unlabeled action did not produce a result. In the *conflict* condition, perceptual cues were *not* aligned with speaker information, such that the labeled action did *not* produce a salient result, whereas the unlabeled action did.

Speaker information was held constant across experimental conditions. Because the experimenter performed the labeled action as she named it, speaker cues always coincided with the labeled action. Speaker information consisted of a complex of social and linguistic cues. Social cues included maintaining

joint attention and using eye gaze alternation between the box and the child to draw attention to the action being demonstrated and labeled. Linguistic cues included announcing the intention to perform an action beforehand ("I'm going to feb on the box"; Tomasello & Kruger, 1992) and using the present progressive morpheme *-ing* ("I'm *febbing* on the box") to denote ongoing action (Echols & Marti, 2004) and intention to label the manner rather than the result of the action (Behrend, et al., 1995). Research has shown that children are sensitive to the present progressive morpheme (*-ing*) by 18 months of age (e.g., Echols & Marti, 2004; Golinkoff, Hirsh-Pasek, & Schweisguth, 2001; Brown, 1973). In addition, the novel verb was followed by a prepositional phrase as a further cue that it was a manner verb. The combination of the present progressive verb and a prepositional phrase in the sentence "I'm febbing on the box," was intended to provide the strongest indication that the novel word *febbing* is a verb rather than another part of speech. Although linguistic cues alone did not provide sufficient information to determine which was the target action, linguistic and social information together unambiguously pointed to the target action.

*Testing phase.* To assess whether word learning had occurred, the Intermodal Preferential Looking Paradigm (IPLP; Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996) was used. The premise of the minimally demanding IPLP is that when presented with a split-screen display, children who learned the word should look longer at the side of the screen matching the language that they hear.

Following completion of the training phase, children and parents moved to a separate area of the room. Children were seated on their parent's lap approximately 3 feet away from a 32-inch television. Parents were instructed to close their eyes and refrain from speaking or directing their child's attention during the course of the video. Both novel actions from the box were presented simultaneously on the television screen; one side of the screen displayed the flipping action on the light switch, the other the pressing action on the Morse code key (see Figure 2). Crucially, the perceptually salient results of the actions (e.g., red light or tone) were not shown in the test videos, forcing the child to focus on the *action* rather than on the action's *results*. Children could not choose based on results alone.

The testing period had three phases (see Table 2). First, children saw a *salience* trial allowing them to examine both actions on the television accompanied by neutral audio. Because the salient results were removed from the action in the video presentation, we

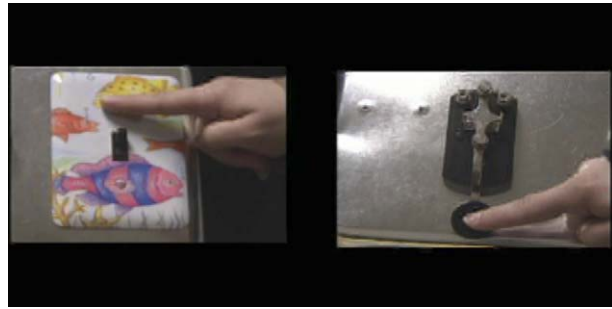


Figure 2. Video display during all test trials.

anticipated that children would look equally at each action during the *salience* phase. Next, children saw two *test* trials (each 6 seconds long) in which they were directed repeatedly to "Look at febbing on the box!" If children learned the name of the action, their attention to the referent action should be greater than that during the *salience* trial. To assess whether children were truly pairing a label with an action rather than simply attending to the previously interesting action, a third and fourth phase were required. These additional trials offered a more stringent test of word mapping. In the third, *new-label* test trial, children were asked to look at *glorping* rather than *febbing*. It was predicted that if children thought that the target action already had a name (*febbing*), the new name (*glorping*) would prompt them to look away from it, and, if they used the principle of mutual exclusivity, they might also gaze more at the novel, unnamed action. This method has been used successfully in prior research (e.g., Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996; Hollich et al., 2000). In the fourth, *recovery* trial, children were again asked to look at *febbing*. If children had indeed learned the original name for the action, they should renew their looking to the target action.

To assess verb learning, it was our a priori decision to look for a cubic pattern of results demonstrating that children increased their looking to the target from salience to the first pair of test trials, decreased their looking to the target during the third trial (the *new-label* trial), and then renewed their attention to the target action during the *recovery* trial. If children were unable to make use of speaker cues in the face of compelling perceptual cues, two possible outcomes could occur. Children might fail to map a label to the action without a result as evidenced by no significant preference for either action in any of the trials. Or, children in the conflict condition might mistakenly map the label to the action with a result, as evidenced by a pattern of looking identical to that of the children in the coincident condition (see Pruden et al., 2006).

Table 2  
Testing Phase of Experiments 1 and 2

Trial type	Audio	Presentation on left side of screen	Presentation in center of screen	Presentation on right side of screen
Centering trial (3 sec.)	Hey, look up here		Smiling baby	
Saliency trial (6 sec.)	What's going on up here? What's happening on the TV?	Hand flipping light switch up and down		Hand pressing Morse code key
Centering trial (3 sec.)	Hey, do you know how to feb on the box?		Smiling baby	
Test trial 1 (6 sec.)	Find febbing on the box. Do you see febbing on the box?	Hand flipping light switch up and down		Hand pressing Morse code key
Centering trial (3 sec.)	Do you know how to feb on the box?		Smiling baby	
Test trial 2 (6 sec.)	Look at febbing on the box. See febbing on the box?	Hand flipping light switch up and down		Hand pressing Morse code key
Centering trial (3 sec.)	Oh, do you know how to glorp on the box?		Smiling baby	
New-label trial (6 sec.)	Look at glorping on the box, not febbing, glorping.	Hand flipping light switch up and down		Hand pressing Morse code key
Centering trial (3 sec.)	Now do you know how to feb on the box?		Smiling baby	
Recovery trial (6 sec.)	See febbing on the box? Look at febbing on the box.	Hand flipping light switch up and down		Hand pressing Morse code key

Videotaped experimental sessions were viewed offline and in real-time by trained and reliable coders who were blind to the testing condition. In order to validate assumptions about differential saliency of the actions, during each free-play trial of the demonstration phase, coders indicated the amount of time children were engaged with the mechanism that created (or potentially created) a result (e.g., touching or flipping the light switch and touching or pressing the key pad). For the video test phase of the experiment, coders observed children's visual fixation during each trial and indicated the amount of visual fixation to the left, center, or right side of the screen by pressing the corresponding button on a timing device. Interrater reliability was calculated for 20% of the participants and averaged  $>.90$ .

### Results

#### *Were the Actions Differentially Salient to Young Children, as Predicted?*

An alpha level of .05 was used for all statistical tests. A two-tailed paired samples t-test revealed that indeed, prior to verb training, across experimental conditions, children spent a significantly greater amount of time performing the action that produced a result than the action that produced no result,

$t(30) = 3.48, p = .002$ . Specifically, children spent 9.72 ( $SD = 6.67$ ) seconds (73% of the time) enacting the action that produced an interesting result and 3.58 ( $SD = 4.97$ ) seconds (27% of the time) enacting the action that did not produce a result. There were no changes in action performance preferences following verb training: Children preferred to perform the action that produced an interesting result across all three free-play trials.

A one-way ANOVA with percentage of time spent performing the action that produced a result as the dependent variable and Condition (conflict/coincident) as the between subjects factor revealed no significant effect of Condition,  $F(1, 28) = 1.27, p = .270$ . Thus, children in both conditions preferred to perform the action that produced a result; it was indeed the more perceptually salient action.

#### *Did Children in the Coincident Condition Differ From Those in the Conflict Condition?*

For each child, a proportion of looking score was calculated by dividing the number of seconds spent looking at the action that previously produced a result by the number of seconds spent looking at both the action that previously produced a result and the action that previously did not produce a result. Thus, any value greater than .50 indicated a preference for



the action that previously produced a result and any value less than .50 indicated a preference for the action that previously did not produce a result. This proportion was calculated for all trials. Visual fixation during the two identical test trials was averaged resulting in four proportions: salience, test, new label, and recovery (see Figure 3).

Preliminary analyses indicated that neither Gender nor Target Action (flipping the light switch or pressing the Morse code key) had a significant effect on looking behavior; therefore data were pooled across these factors. To determine whether children performed differently as a result of condition, proportions were entered into a repeated-measures ANOVA with Condition (conflict/coincident) as the between-subjects factor and Trial Type (salience/test/new label/recovery) as the within-subjects factor. The only significant result was a main effect of Condition,  $F(1, 29) = 5.312, p = .029, \eta^2 = .155, \text{Power} = .606$ . Across the four trial types, children in the coincident condition looked significantly longer at the action that had previously produced an interesting result than did the children in the conflict condition ( $M = .54$  and  $.42$ , respectively).

On the basis of these overall results alone, it is impossible to draw clear conclusions about whether

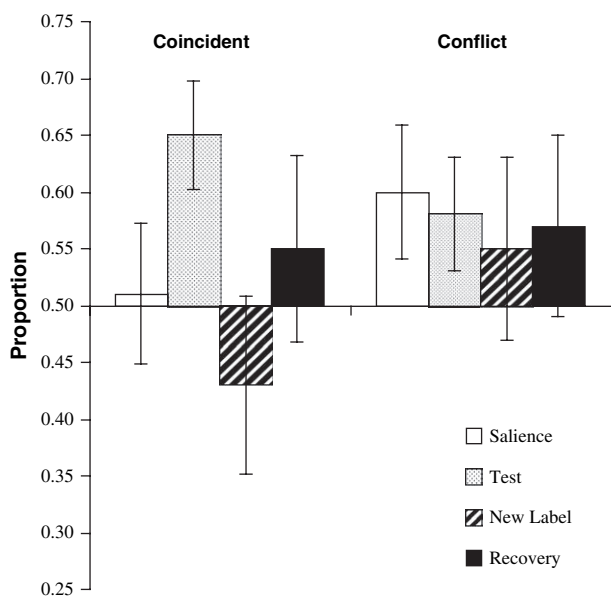


Figure 3. Experiment 1: Coincident and conflict conditions. Bars represent the mean proportion of looking (looking time to action with a result divided by looking time to both actions with and without a result) as a function of trial type. In the coincident condition, correct word-learning performance should be depicted by bars (for the test and recovery trials) above .50, whereas in the Conflict condition, the same performance should be depicted by bars below .50. Vertical lines depict standard errors of the means.

or not children in either condition learned the target verb. However, successful verb learning would be evidenced by a significant cubic pattern of results demonstrating that children increased their looking to the target between the salience trial and the first pair of test trials, decreased their looking to the target during the third trial (the new-label trial), and then renewed their attention to the target action during the recovery trial. Therefore, although there was no interaction between Trial Type and Condition, to determine whether children showed evidence of verb learning, we performed another repeated-measures ANOVA to test for the presence of the cubic pattern in each condition.

#### Did Children Learn the Names of the Actions?

*Coincident condition.* Proportions of looking time were entered into a repeated-measures ANOVA with Trial Type (salience/test/new label/recovery) as the within-subjects factor. Although there was no main effect of Trial Type,  $F(3, 42) = 1.693, p = .183, \eta^2 = .108, \text{Power} = .411$ , planned contrast analyses revealed a significant cubic pattern across trials,  $F(1, 14) = 4.715, p = .048, \eta^2 = .252, \text{Power} = .525$ , such that children increased their looking to the action that had previously produced a result during the test trial, decreased their looking to that action during the new-label trial, and increased their looking to that action once again during the recovery trial (see Figure 3).

A one sample t-test (compared against a chance value of .50) confirmed that children allocated significantly more attention to the action that had previously produced a result during the test trials,  $t(15) = 3.144, p = .007$ . During all other trials, children performed at chance level. Thirteen out of sixteen children increased their attention to the action that had previously produced a result upon hearing its name, and 11 out of 16 decreased their attention to that action upon hearing a new label. Thus, results suggest that children in the coincident condition were successful in mapping a label to the action that had produced an interesting result during training.

*Conflict condition.* Proportions were entered into a repeated-measures ANOVA, with the within-subjects factor of Trial Type. There was no main effect of Trial Type ( $F < 1$ ). Planned contrast analyses did not reveal a significant cubic pattern ( $F < 1$ ). One sample t-tests (compared with a chance value of .50) confirmed that children did not allocate significantly more attention to the action that previously produced a result on any trial. Only 7 out of 16 children increased their looking to the action that had previously not produced a result upon hearing its name and 6 out of 16 decreased their looking to that action

upon hearing a new label. In other words, children showed no evidence of word learning.

*Vocabulary and age comparisons.* To ensure that performance differences between children in the coincident and conflict conditions were not due to differences in age or vocabulary level, two sample *t*-tests were conducted. Results revealed no significant differences in the age or vocabulary level of children in the coincident and conflict conditions. To assess whether children's vocabulary level was related to performance on the verb-learning task, for each condition we performed a two-sample *t*-test comparing the vocabularies of children who failed against those of children who succeeded (as defined by an increase in attention to the labeled action between the salience and test trials and a decrease in attention to the labeled action between the test and new-label trials). In the coincident condition, there was no difference between the vocabularies of children who failed ( $n = 7$ ) and succeeded ( $n = 9$ ),  $t(14) = -.504$ ,  $p = .622$ . In the conflict condition, on the other hand, children who failed the verb-learning task ( $n = 13$ ) had a significantly lower vocabulary than did those who succeeded ( $n = 3$ ),  $t(13.36) = 2.92$ ,  $p = .012$  (Welch-corrected degrees of freedom for unequal variances). Specifically, the mean vocabulary score of children who were successful at verb learning was 62.00 ( $SD = 6.56$ ) words, whereas the mean vocabulary score of children who were not successful was only 36.38 ( $SD = 25.76$ ) words.

### Discussion

Results of Experiment 1 suggest that 2-year-olds successfully learned a novel verb when perceptual and speaker cues to verb meaning coincided. However, children failed to learn a novel verb when speaker and perceptual cues were put into conflict. Children in the conflict condition neither mapped the novel label to the action with no result nor mismapped the label to the action with a result (as in the noun-learning work of Pruden et al., 2006). This is surprising given that children witnessed only one action performed during training: the action being performed was the action being labeled. Although by 21- to 24-months, children are surely sensitive to social and linguistic cues to a word's meaning (e.g., Baldwin, 1993; Hollich et al., 2000; Naigles & Kako, 1993), when faced with compelling perceptual cues, they proved unable to rely on speaker information to solve the problem of verb mapping.

Children's success in the coincident condition cannot simply be construed as a preference to watch the outcome of the action with an interesting result for

two reasons. First, the salient results of the action were removed from the test videos. At test, children saw only the novel actions, not the actions' results. Second, as demonstrated by the cubic pattern, children switched their attention to the unnamed action during the new-label test trial. If children simply preferred to look at the action that had previously produced interesting results, then the introduction of the novel label should have made no difference in their visual fixation: children should have continued to watch the action that had previously produced a result. They did not. Thus, these findings present a compelling case for verb learning in the coincident condition of Experiment 1.

What happened in the conflict condition? As Figure 3 demonstrates, children consistently devoted a nonsignificantly larger proportion of visual fixation to the action that produced a result during training. This pattern of results suggests that children may have been profoundly influenced by perceptual cues. Nonetheless, if children had relied *solely* on perceptual cues, they should have mismapped the label for the action that previously did not produce a result to the action that previously did (as in the noun-learning work of Pruden et al., 2006). However this was not the case. No mismapping occurred. One possible explanation for this finding is that, although children did not rely on perceptual salience alone to determine verb meaning, the presence of a perceptually salient action in training may have precluded them from detecting or fully processing speaker cues in service to verb learning. On this view, children failed to learn the verb during training and thus, at test, on all trial types, revealed a subtle yet nonsignificant preference to watch the action that had previously produced a result.

Results from the analyses comparing the vocabularies of children who successfully learned the verb with those of children who failed to learn the verb in the conflict condition also hint at the role of attention to speaker cues in word learning. Those children who successfully overcame the lure of perceptual salience and attended to the speaker's cues in the conflict condition had a larger expressive vocabulary than those who did not. Although a causal relation cannot be determined from these data, results suggest that the ability to attend to speaker cues in the face of compelling perceptual cues may predict word learning ability.

Because of the extreme perceptual asymmetry between the action with a result and the action without a result, it is unclear from the current data whether children in our task could use speaker information alone to guide verb learning. By removing

the lure of perceptual salience, it might be possible to evaluate the impact of speaker cues in a situation when these cues are not competing with other word learning information. Thus, in Experiment 2 we asked whether 21- to 24-month-olds can learn a novel verb when the compelling cue of perceptual salience is equalized through the use of actions that *both* produce interesting results.

### Experiment 2: Can Young Children Learn a Verb When Perceptual Salience is Equalized?

In this experiment we probe whether verb learning can occur when two actions offer an equally salient result. By offering equally salient actions, we may remove the asymmetry that may have prevented children in Experiment 1 from mapping the verb to the boring action. Alternatively, we may find that young 2-year-olds are only able to solve the verb mapping problem when coincident speaker information and perceptual cues spotlight a single action to the exclusion of others.

#### Method

##### Participants

Participants included 16 native English-speaking toddlers (8 male, 8 female) ranging in age from 21.10 to 24.03 months ( $M = 23.02$  months,  $SD = 1.06$  months). Participants were predominantly Caucasian and of middle- to upper-middle class. Data from an additional 15 participants were excluded for the following reasons: child ended the study ( $n = 9$ ), child watched less than 65% of the video ( $n = 1$ ), experimenter/technical error ( $n = 3$ ), and child could not perform both of the actions ( $n = 2$ ). Production vocabularies ranged from 5 to 80 words ( $M = 48.94$ ,  $SD = 22.77$ ), as measured by the MacArthur-Bates Communicative Development Inventory Toddler Short Form (Fenson et al., 2000).

##### Materials

Materials were identical to those used in Experiment 1.

##### Procedure

The procedure was identical to that in Experiment 1 with one important exception: the target (labeled) and nontarget (nonlabeled) actions *both* produced a salient result (i.e., flipping the light switch illuminated a red light *and* pressing the Morse code key

sounded a tone). No single action was highlighted perceptually. Thus, in order to learn the novel verb, children could not rely on perceptual cues alone; children must identify and attend to the speaker cues being offered for the action receiving a label.

#### Results

##### Were the Actions Equally Salient to Young Children?

A two-tailed paired samples t-test revealed that indeed, prior to verb training, the amount of time children spent performing the action that would later be labeled and the action that would not later be labeled did not differ significantly,  $t(15) = 1.67$ ,  $p = .117$ . Specifically, children spent 7.92 ( $SD = 4.88$ ) seconds (61% of the time) enacting the action that would later be labeled and 5.05 ( $SD = 5.36$ ) seconds (39% of the time) enacting the action that would not later be labeled. There were no changes in action performance preferences following verb training. However, a two-tailed t-test comparing the amount of time children flipped the light switch to the amount of time children pressed the Morse code key revealed a significant preference for the action performed on the light switch,  $t(15) = -2.581$ ,  $p = .021$ . Specifically, children spent 4.46 ( $SD = 5.53$ ) seconds (34% of the time) pressing the Morse code key and 8.51 ( $SD = 4.21$ ) seconds (66% of the time) flipping the light switch. Closer analysis revealed that this overall preference for the action performed on the light switch resulted from four children who performed this action exclusively. By eliminating these four children from the sample, the preference for the action on the light switch disappeared,  $t(11) = -1.276$ ,  $p = .228$ . Specifically, children spent 5.95 ( $SD = 5.66$ ) seconds (44% of the time) pressing the Morse code key and 7.71 ( $SD = 3.67$ ) seconds (56% of the time) flipping the light switch. To ensure that these four children were neither carrying nor masking our findings, in all subsequent analyses we considered both the full and the restricted samples.

##### Did Children Learn the Names of the Actions?

For each child, a proportion was again calculated. However, because both actions produced a result, the proportion was computed by dividing the number of seconds spent looking at the labeled action by the number of seconds spent looking at the labeled and the unlabeled action. Any value greater than .50 indicated a preference for the labeled action and any value less than .50 indicated a preference for the unlabeled action. This proportion was calculated for

all trials, again averaging the two identical test trials (see Figure 4).

Preliminary analyses indicated that gender had no influence on performance; therefore, the data were pooled across gender. To determine whether children preferred to watch the labeled or unlabeled action across the test trials, proportions were entered into a repeated-measures ANOVA, with Target Action (flipping the light switch/pressing the Morse code key) as the between-subjects factor and Trial Type as the within-subjects factor. Analyses both with and without the four children with exclusive preferences for the flipping action revealed no main effect of Trial Type. Crucially, despite the significant preference for the action performed with the light switch during free-play, there was neither a main effect of Target Action nor a Trial Type by Target Action interaction (all  $F_s < 1$ ). A priori contrasts also revealed no significant cubic pattern ( $F < 1$ ).

One sample  $t$ -tests (compared with a chance value of .50) confirmed that children did not allocate a significant amount of attention to the labeled action on any trial. Only 9 out of 16 children increased their attention to the labeled action upon hearing its name, and only 7 out of 16 decreased their attention to the labeled action upon hearing a new name. Therefore, no evidence of word learning was found.

To assess whether children's vocabulary level was related to performance on the verb-learning task, we performed a two-sample  $t$ -test comparing the vocab-

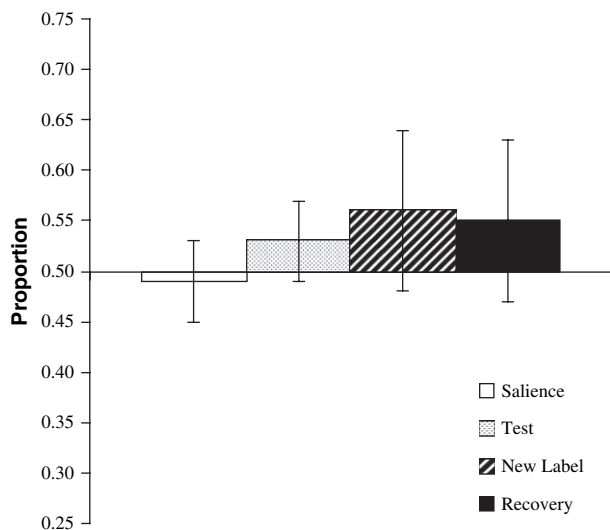


Figure 4. Experiment 2: Equal salience condition. Bars represent the mean proportion of looking (looking time to labeled action divided by looking time to both labeled and unlabeled actions) as a function of trial type. Correct word-learning performance should be depicted by bars (for the test and recovery trials) above .50. Vertical lines depict standard errors of the means.

ularies of children who failed ( $n = 11$ ) against those of children who succeeded ( $n = 5$ ; as defined by an increase in attention to the labeled action between the salience and test trials and a decrease in attention to the labeled action between the test and new-label trials). The vocabulary differences were nonsignificant,  $t(14) = -1.467$ ,  $p = .164$ . Performance on the verb-learning task was unrelated to the number of words in children's expressive vocabularies.

### Discussion

This experiment demonstrated that in a task involving perceptually compelling actions of equal salience, 21- to 24-month-olds were unable to attach a verb to its correct referent. Information from a speaker, in the absence of distinguishing perceptual cues, was not sufficient to promote word learning. Yet the cues to verb meaning provided by the speaker were *not* subtle: the action being performed by the speaker was the action the speaker labeled. These findings point to the pivotal role of perceptual information for early verb learning. Experiment 2 showed that when the cue of perceptual salience was equalized, young children were not able to override compelling perceptual information and attend to speaker cues during the task of verb learning. Although support does exist for young children's ability to use social and linguistic cues to determine verb meaning (e.g., Naigles & Kako, 1993), together, the experiments presented here demonstrate that when perceptually salient actions compete for children's attention, young children are apparently unable to muster the resources necessary to learn a verb.

The question remains, however, as to when children are able to override the lure of perceptual information to learn a verb. It is imperative that children master the use of social and linguistic cues to verb meaning because the meaning of a verb depends on more than its perceptual characteristics. The purpose of Experiment 3 was thus to determine when children can learn a novel verb in the most stringent of circumstances—when perceptual and speaker cues to the verb's meaning are put into conflict.

### Experiment 3: Can Children Eventually Learn a Verb When Perceptual and Speaker Cues Conflict?

#### Method

#### Participants

Participants included 16 native English-speaking children (8 male, 8 female) ranging in age from 32.77

to 36.30 months ( $M = 34.44$  months,  $SD = 1.05$  months). Participants were predominantly Caucasian and of middle- to upper-middle class. Data from an additional 11 participants were excluded for the following reasons: child refused to point ( $n=4$ ), experimenter error/technical problems ( $n=7$ ). Production vocabularies ranged from 65 to 100 words ( $M = 93.06$ ,  $SD = 10.78$  words) as measured by the MacArthur-Bates Communicative Development Inventory Toddler Short Form (Fenson, et al., 2000). Although the CDI is normed only for children between the ages of 16 and 30 months, to be consistent with the previous two experiments, the parents of these children were also asked to complete the CDI Toddler Short Form.

### Materials

Materials were identical to those used in Experiments 1 and 2.

### Procedure

The demonstration and training phases were identical to those in Experiment 1 except that all children were in the conflict condition. That is, the experimenter named the less interesting of the two actions (i.e., the action that did *not* produce a result), placing speaker and perceptual cues into conflict.

*Testing phase.* Because of the advanced age of these participants, pointing replaced visual fixation as the dependent measure. This substitution has been made in prior research with 3-year-olds, generating results comparable to looking time data (Imai et al., 2006; Maguire, Pruden, Hirsh-Pasek, Hansel, & Meyer, 2004). Prior to beginning the video, the experimenter explained to children that they would be playing a pointing game. Participants were told that pictures would appear on the screen, and they would be asked to point to them. Children were instructed to point in an exaggerated manner, using their whole arm. The experimenter helped children practice pointing by encouraging them to point to the television, the ceiling, etc., and praising them upon completion. Once children indicated that they understood the instructions and appeared comfortable with the procedure, the experimenter moved to a position behind the child and began the video. The videos were identical to those in Experiments 1 and 2. The audio however, was replaced by a script read by the experimenter asking children to point to specific actions demonstrated on the screen. Regardless of where children pointed, the experimenter responded neutrally (e.g., "Thanks!"). If children did not imme-

diately point, the experimenter paused the video and encouraged them to point by rephrasing the question without providing any additional labels (e.g., "Can you point for me, [child's name]? Which one is it?"). Left, right, or no-response answers were recorded during the course of the experiment by a second observer blind to the identity of the labeled action. Interrater reliability was calculated for 25% of the participants and averaged 100%.

### Results

#### *Were the Actions Differentially Salient to Young Children?*

A two-tailed paired samples t-test revealed that, in contrast to our predictions, prior to verb training, children exhibited only a marginal preference to perform the action that produced a result,  $t(15) = -2.01$ ,  $p = .06$ . Specifically, children spent 5.74 ( $SD = 4.05$ ) seconds (64% of the time) enacting the action with a result and 3.28 ( $SD = 4.31$ ) seconds (36% of the time) enacting the action without a result. Despite the marginal preference, 12 out of 16 children preferred to perform the action that produced a result. A sign test revealed this preference to be significant ( $p = .027$ ). Eliminating the four children preferring to perform the action without a result resulted in a significant preference for the action with a result,  $t(11) = -3.78$ ,  $p = .003$ . This subset of children spent 6.07 ( $SD = 4.31$ ) seconds (78% of the time) enacting the action with a result and 1.75 ( $SD = 2.39$ ) seconds (22% of the time) enacting the action without a result. To be sure that the preference of the four children preferring to perform the action without a result were neither carrying nor masking our results, subsequent analyses will consider both the full and the restricted samples.

#### *Did Children Learn the Names of the Actions?*

To assess whether word learning had occurred, we compared children's performance during each of four test trials against a theoretical chance performance of 50% (because children had the option to point to either the left or the right side of the TV screen on each trial). It was hypothesized that, if verb learning occurred, children would select the target action at a rate greater than 50%. A two-tailed binomial test was conducted to assess whether children would select the target action with greater than 50% accuracy on each of four test trials (Test 1, Test 2, New Label, Recovery). For all four trials, the observed proportion of correct responses differed significantly from the hypothesized value of .50 (all  $p$  values  $< .01$ ; all effect sizes were medium,  $d$  ranging from .37 to .43). Overall,

13 out of 16 children selected the correct action in both Test 1 and Test 2, and 12 out of 16 children selected the correct action with 100% accuracy, across all four test trials (see Figure 5).

To ensure that the subset of children who preferred the action without a result during free play did not carry these findings, we conducted a second analysis on only those children who *did* prefer the action with a result. Results from this subset of children did not differ from those from the entire sample. For all four trials, the observed proportion of correct responses differed significantly from the hypothesized value of .50 (all  $p$  values  $< .01$ ). Overall, 10 out of 12 children selected the correct action in both Test 1 and Test 2, and 10 out of 12 children selected the correct action with 100% accuracy, across all four test trials.

To assess whether children's vocabulary level was related to performance on the verb-learning task, we again performed a two-sample  $t$ -test comparing the vocabularies of children who failed against those of children who succeeded (as defined by pointing correctly in the test and new-label trials). The mean vocabulary score of the successful verb learners ( $n = 14$ ) was 96.83 ( $SD = 6.45$ ) words, whereas the mean vocabulary score of the unsuccessful verb learners ( $n = 2$ ) was 81.75 ( $SD = 14.17$ ) words. Although this difference was not significant, results revealed a trend toward significance,  $t(3.424) = -2.058, p = .120$  (Welch-corrected degrees of freedom for unequal variances).

### Discussion

Results showed that by 33 to 36 months of age, children successfully learned a verb when perceptual and speaker cues to verb meaning were put into

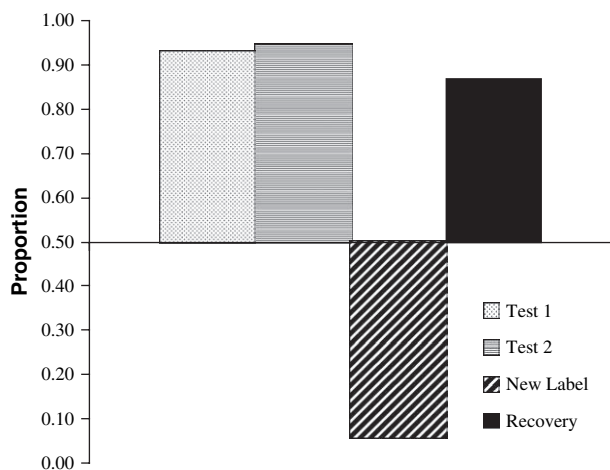


Figure 5. Experiment 3: Conflict condition. Bars represent the proportion of children pointing to the action without a result as a function of trial type.

conflict. Unlike their younger counterparts, by this advanced age, children were able to learn verbs even in a situation that was less than ideal.

The marginally significant results from the analyses comparing the vocabularies of children who successfully learned the verb with those of children who failed to learn the verb in the conflict condition again hint at the relationship between vocabulary and attention to speaker cues in word learning. As in Experiment 1, those children who successfully overcame the lure of perceptual salience and attended to the speaker's cues had a larger expressive vocabulary than those who did not. Although a causal relation cannot be determined from these data, results suggest that by 3 years of age, the ability to attend to speaker information in the face of compelling perceptual cues may predict word learning ability; sophisticated use of speaker information may result in a larger production vocabulary.

### General Discussion

Here we explored the problem of verb mapping: How do young children coordinate perceptual and speaker cues to verb meaning? Perceptual cues (whether or not an action produced a result) were manipulated while speaker information (including social and linguistic cues) was held constant. This design allowed us to reveal how children face the problem of referential ambiguity. We asked whether young children weight perceptual salience more heavily as a cue to verb meaning. And, if so, can they override their perceptual preferences in favor of speaker cues to verb meaning?

In Experiment 1, we asked whether 21- to 24-month-old children could learn a novel verb when perceptual salience and speaker cues either coincided or were put into conflict. Results demonstrated that children were successful in learning a novel verb only in the condition in which perceptual cues coincided with speaker cues to the verb's meaning. When presented with conflicting information, children failed. Young 2-year-olds could not override the perceptual salience of a competing action's results in order to attach a label to an action without a result.

Results of Experiment 2 provided further support for the notion that 21- to 24-month-olds cannot override action salience in order to learn a verb. When both actions offered perceptually salient results, children were unable to attach a verb label to its referent. That children failed to learn a verb in the conflict condition of Experiment 1 and the equal salience condition of Experiment 2 is striking given that in

these experiments children were not required to determine from an ambiguous situation which action was the verb referent. Instead, unlike in the complex labeling context of the real world, here the action being performed by the speaker was the action being labeled. Thus the verb-to-action mapping in these experiments was essentially provided to children for free. That children failed this transparent task suggests that the difficulty rests in children's attention to a speaker's cues in the presence of compelling actions and action results. The appealing result of an action (or actions) that the child could produce apparently disrupted attention to verb learning. Children were only able to muster the resources necessary to learn a verb when a single action with an interesting result was the focus of their attention. In the presence of a competing action with a salient result, 21- to 24-month-old children failed to learn a verb.

On the basis of these data, we do not claim that 2-year-olds never attend to or are unable to use social and linguistic cues to verb meaning. Rather, our data speak to the specific problem of learning a verb in the presence of a competing perceptually salient action. Young children may only be guided by perceptual preferences when perceptually salient actions are available. Note, however, that the complex labeling context of the real world is full of compelling actions and action results. Thus, the verb-learning situation presented in these experiments remains ecologically valid. Although we did not test a condition in which both actions were perceptually boring and neither action produced a salient result, based on the current findings and previous research demonstrating successful verb learning in young children, we predict that in such a condition, even 22-month-olds might be guided by speaker cues alone. Testing children's use of verb-learning cues in the absence of perceptual salience may help to determine the extent of children's reliance on perceptual cues and their ability to utilize social and linguistic information.

Despite the early conservatism demonstrated in Experiments 1 and 2, children must eventually come to rely on speaker cues to verb meaning—even in the presence of perceptual salience. Verb learning requires more than just attention to perceptual cues, as the meaning of a verb depends on more than its appearance. To learn verbs of perception (e.g., *see*, *perceive*), communication (e.g., *say*, *tell*), thinking (e.g., *believe*, *know*), or perspective (e.g., *chase*, *flee*) children must rely on speaker information from the syntax. To learn verbs that are distinguished by the actor's intent (e.g., *pour* versus *spill*, *topple* versus *knock over*), children must also be able to use the speaker's social cues. The purpose of Experiment 3 was therefore to

demonstrate that sometime after 2 years of age children are able to override salient perceptual cues to learn a verb from the perspective of the speaker. This experiment included only the conflict condition, arguably the most difficult test of verb learning. In this experiment 33- to 36-month-olds were able to successfully use a speaker's cues to learn a name for an action without a result. Older children made correct selections, not only on the reference test trials but also on the new-label and recovery trials, providing convincing evidence of their ability to learn a verb in a less than ideal situation. These data suggest that by 33- to 36-months, children were able to set aside the lure of action salience and attend to important cues to a verb's meaning. Because the linguistic and social cues to the verb's meaning were confounded in this series of experiments, we can only speculate on the extent to which children used the linguistic and social information independently; nonetheless, our findings demonstrate that the weighting children give to these more reliable and sophisticated cues changes sometime between 22 and 33 months of age.

The significant relationship between measures of production vocabulary and the ability to overcome perceptual salience to learn a verb in the conflict condition provides additional insight into the question of how children differentially weight cues to word learning across developmental time. Results demonstrate that children who were able to attend to speaker cues in the face of perceptual salience had a significantly higher vocabulary than children who failed to do so. One possible interpretation of this finding is that the ability to attend to speaker cues facilitates verb learning, resulting in a higher production vocabulary. In this case, children who assign greater weight to speaker cues are able to learn more words. Alternatively, it may also be that knowledge of a large number of words teaches children to attend to more sophisticated social and linguistic cues. In this case, word knowledge may lead children to shift their weighting of perceptual and speaker cues. On the basis of these data alone, we cannot determine the direction of this relationship between vocabulary and the use of speaker cues. Nonetheless, this effect helps to highlight the relationship between vocabulary knowledge and the emergent use of reliable and sophisticated word learning cues.

An alternative explanation for the pattern of results observed here is that performance across experimental conditions may be explained by a misinterpretation of the linguistic cues on the part of 2-year-olds in Experiments 1 and 2. Young children may have interpreted the sentence "*febbing on the box*" to mean "*febbing on the box*" as in "*turning on*" or "*activating*

the box." If this were the case, for the 2-year-olds, mapping a label to the action that produced a result would be far easier than mapping a label to the action that did not produce a result. However, this interpretation is unlikely on the basis of the prosodic cues offered by the experimenter. Whereas, in the sentence "turning on the box," both the words "turning" and "on" are stressed, in our presentation of the sentence "febbing on the box," only "febbing" was stressed. These prosodic cues clearly indicate that "febbing" alone serves as the verb, whereas "on the box" serves as a separate prepositional phrase. Pretesting with adults confirmed the clarity of the prosodic cues resulting in their mapping of *febbing* to the manner of the action and not to the action's results. Moreover, results of Experiment 3 demonstrate that, like adults, 33-month-olds have no trouble interpreting the experimenter's prosodic cues. Although we do not have independent empirical evidence that children at 22 months of age can distinguish the specific prosodic cues used in our experiment, on the basis of the performance of the 3-year-olds and data demonstrating sensitivity to prosodic patterns within the first year of life (e.g., Jusczyk, Cutler, & Redanz, 1993), we argue that a misinterpretation of the prosodic cues offered by the experimenter does not fully account for the observed pattern of results.

What did children learn in our task? Although we argue that children were learning a novel manner verb, an alternative account for our findings may be that children were mapping a name to the *results* rather than to the manner of the actions (Behrend, 1989, 1990; Forbes & Farrar, 1993, 1995). This is unlikely in our experiment given that the results of the actions were not present in the test videos. If children had mapped the verb label to the action's results, in the absence of those results they should have consistently failed, regardless of their experimental condition. This was not the case. Another alternative is that in watching the test videos, children were calling to mind the action's former results. This would suggest that, instead of forming an abstract category of manner of action, children were mapping the novel label to a particular complex of (action + results). If this were the case, *febbing* would refer, for example, to "flipping this switch in order to activate this red light." This interpretation is possible given that our pressing and flipping actions are familiar, yet children had never seen them performed on this apparatus or with this label.

To determine whether children were truly mapping a label to the category of action rather than limiting the label to a particular context or to the result it generates, future research should include a further

extension trial during which children are asked to find *febbing* and *glorping* on an entirely new apparatus that they have never seen produce a result. Nonetheless, investigating verb reference, which is the focus of the present study, is a necessary and important precursor to investigating verb extension. Even if children in the current experiments had learned a more limited definition of the verb, our developmental trajectory still holds: When perceptually salient actions are available, children initially rely more heavily on perceptual cues and prefer to learn the names for more perceptually salient actions. Only later can they surmount that attraction to learn the name for a less perceptually salient action.

The developmental trajectory presented here is reminiscent of that offered by the Emergentist Coalition Model of word learning (ECM; Hirsh-Pasek et al., 2000; Hollich et al., 2000). This hybrid model suggests that children are sensitive to multiple cues in word learning (perceptual, social, and linguistic) and that they weight those cues differently as they develop from novice to expert word learners. Children have a range of cues at their disposal at all times; yet, not all cues are equally utilized in service to word learning. Initially, children rely most heavily on *perceptual* cues, including perceptual salience and novelty (Hollich et al., 2000). Children later begin to place greater weight on *social* (including speaker eye gaze, pointing, or handling of a referent) and *linguistic* (including syntactic structure and morphology) cues when deciphering a novel word's meaning (e.g., Echols & Marti, 2004; Gelman & Taylor, 1984; Waxman & Booth, 2001). Although social and linguistic cues are available from the beginning and evidence suggests that children are able to use this information from a very young age (e.g., Akhtar & Tomasello, 1996; Naigles & Kako, 1993), according to the ECM, early in the verb-learning process, young children will place greater weight on perceptual cues.

Hollich et al. (2000) and Pruden et al. (2006) validated the predictions of the ECM in the domain of noun learning by creating a situation, much like our own, where cues for word learning could be put into conflict. In these studies, perceptual cues were put into conflict with speaker cues by presenting children with two novel objects, one interesting and perceptually salient and one boring and less perceptually salient. The experimenter indicated via the use of social cues, such as eye gaze and body orientation, which of the two objects was being named. At test, children saw both objects and were asked to find the one that was labeled. Results suggest that, despite their preference to map to the perceptually salient object at 10 months (even when the boring object is



being labeled; Pruden et al., 2006), by around 19 months, children begin to exhibit a mature sense of reference. They overcome the perceptual salience of objects in their environment and weight more heavily the speaker's reliable social and linguistic cues to word meaning.

Although most work within the ECM has examined how children attach novel labels to object referents (but see Poulin-Dubois & Forbes, 2006), the current experiments extend these findings to the domain of actions. The combined results of the experiments presented here bolster the claim that verb learning is uniquely difficult for young children. It appears that because of the added perceptual appeal of action and the ambiguity of verb reference, children who have demonstrated mastery in the realm of noun learning experience considerable setbacks in the domain of verb learning. Although the pattern of results in these verb-learning experiments nearly parallels those of the noun-learning experiments (Hollich et al., 2000; Pruden et al., 2006), equivalent verb-learning abilities are not observed until over one year later. In this way, toddlers reprise with verbs what happens earlier with nouns: in the context of competing actions with varying perceptual cues children initially rely most heavily upon perceptual cues, and only later, with age and language experience, come to assign greater weight to the perspective of the speaker. These results lend further support to the Emergentist Coalition Model of word learning by extending it to include the learning of verbs.

In highlighting the significance of perceptual features as an early solution to the problem of verb mapping, the current experiments also help to explain what has previously been called the "paradox" of verb learning: that is, although some verbs are present in young children's early receptive and production vocabularies, the class of verbs is generally hard to master (Maguire, Hirsh-Pasek, & Golinkoff, 2006). According to the MacArthur-Bates CDI norms (Fenson et al., 2000), the earliest and most commonly used verbs are *eat* and *go*. Both actions can be considered highly salient. For example, *eating* involves food, taste, the cessation of hunger, and a specific set of actions with a strong perceptual basis, such as opening the mouth, biting, chewing, swallowing, etc. Likewise, *go*, although a "light" or "general-purpose" verb (Clark, 1978; Goldberg, 1995), is also perceptually based. According to Theakston, Lieven, Pine, and Rowland (2002), between 2 and 3 years of age, children do not operate with a single, integrated representation of the verb *go*. Instead, their earliest uses of *go* are to encode movement, a highly salient

component of action. Supporting the current findings, these perceptually salient and available actions (*eat*, *go*) prove easy to learn for the child who encounters difficulty attending to social and linguistic information in the face of compelling, perceptually salient actions. Although young children certainly do learn verbs that are not perceptually based, of the verbs produced by children under the age of 2, the majority (but not all) map to actions involving a strong perceptual component. This account is bolstered by recent research showing that the age of acquisition of the words reported on the MacArthur-Bates CDI correlates significantly with adult ratings of their imageability (McDonough, Lannon, Hirsh-Pasek, & Golinkoff, 2006; see also Gillette et al., 1999; Snedeker & Gleitman, 2004). It can be argued that imageability, or the ease with which a concept invokes a mental image, is a measure of perceptual availability.

In addition to explaining the earliest verbs in English-speaking children's receptive and production vocabularies, the developmental trajectory proposed here can also help to explain the finding that verbs appear earlier in the vocabularies of Mandarin-speaking children than in those of their English-speaking counterparts. Tardif (2006) argued that, whereas, in English, many frequent verbs are general in purpose and encode little specific meaning, in Mandarin, verbs encode much more and are used for very specific meanings. Ma, Golinkoff, Hirsh-Pasek, McDonough, & Tardif (2006) found that the relative verb advantage in early Mandarin vocabularies is related to the high imageability of early Mandarin verbs. Furthermore, early Mandarin verbs are significantly more imageable than early English verbs. Some early Mandarin verbs encode very specific, perceptually available actions. For example, among the earliest verbs appearing at 16 months (Tardif, Fletcher, Zhang, Liang, & Chen, in press) is *bao* ('to carry in the arms') and *na* ('to carry with the hands'). For these reasons, Mandarin verbs may be considered more salient and perceptually available than English verbs. This heightened perceptual availability may help to explain the relatively high proportion of verbs in Mandarin-speaking children's vocabularies.

Although our experimental findings suggest that children encounter difficulty attending to social and linguistic information in the face of perceptually salient actions, research has clearly demonstrated that even very young children can and do use social and linguistic information to learn verbs that do not map onto the most perceptually salient actions available. Consider, for example, Naigles and Kako's (1993) study showing that 2-year-olds are able to use

syntactic frames to override their extrasyntactic preference for perceptually salient novel actions. Why might children succeed in this study and not ours? There are at least two explanations. First, the actions in the Naigles and Kako (1993) study were different from those in the present study in terms of perceptual salience. Whereas, in the present study, children performed the actions and produced their results with their own hands, in the Naigles and Kako study, children simply observed actions taking place on a television screen. In addition, whereas the perceptually salient component of the current study involved an interesting result, the Naigles and Kako study concerned the size of the action as rated by adults. For these reasons, the effect of action salience in the two studies may not be comparable.

Second, the actions in the Naigles and Kako study were accompanied by syntactic frames that could be used to differentiate the actions from one another. For example, transitive frames (e.g., "The rabbit is gorp-ing the duck") correspond with causative, but not synchronous, actions. Because the syntactic frames in the present study were designed to describe a manner of action ("I'm gonna feb on the box"; "I'm febbing on the box"), these frames did not differentiate the perceptually salient and perceptually nonsalient actions. The syntactic frames used by Naigles and Kako served as an additional linguistic cue to differentiate the two actions, whereas the syntactic frames in the present study were held constant and did not serve to differentiate the target action from the nontarget action.

Our design was *not* intended to determine the age by which children could learn verbs, but rather to tax children's abilities and in so doing elucidate the *process* and the challenge of verb learning. We do not deny that even very young children can and do use social and linguistic information to learn names for actions. Rather, on the basis of our findings, we suggest that the challenge of verb learning may result in part from the difficulty young children encounter in attending to reliable social and linguistic word-learning information in a world full of compelling, perceptually salient actions.

Although these experiments help us to understand how children resolve the problem of referential ambiguity in verb learning, the present study has its limitations. For example, these experiments addressed only an early step in verb learning, namely verb reference. Second, our research was limited to arguably the easiest type of verb learning, the learning of self-performed action verbs with results. Although we deliberately chose this category of verbs based on prior verb-learning data (e.g., Huttenlocher et al.,

1983), the perceptual availability of these verbs, and on children's early preferences for these actions, in order to gain a more complete account of verb learning, research must include other categories of verbs, such as path, manner, intention, and mental verbs. Third, the linguistic cues used in Experiments 1, 2, and 3 may have been problematic in that they did not differentiate the target action from the nontarget action; they served only to label the manner of action and thus could have been applied to both the target and nontarget actions. For this reason the role of linguistic cues may have been minimized. Fourth, there is one condition that was not tested in the present experiment, namely, when both actions are perceptually boring and neither action produces a result. Young children may only be guided by perceptual preferences when perceptually salient actions are available. Testing children's use of word-learning cues in the absence of perceptual salience may help to determine whether children can be guided by social or linguistic information alone. Finally, the roles of action performance and action results were confounded within the cue of perceptual salience. Future research should disambiguate the roles of these factors to determine precisely which aspects of perceptual salience bias the young verb learner.

## Conclusions

This research helps to explain the complex process of verb learning by elucidating young children's solution to the problem of verb mapping. We have demonstrated that when perceptually salient actions are available, 21- to 24-month-old children are guided by perceptual preferences. As a result, young children learn verbs best when perceptual cues coincide with available social and linguistic cues to uniquely highlight a single action. We have also shown that, in the presence of perceptually salient actions, young children have difficulty overriding their perceptual preferences and attending to obvious social and linguistic information. Our findings suggest that, although young children are fully capable of noun learning and some verb learning, in the presence of perceptually salient actions, the acquisition of less perceptually available verbs is difficult until children are able to overcome the lure of perceptual salience and make use of relevant speaker information. Thus, for the young child learning a verb, in the words of Ralph Waldo Emerson, "What you [or I] do [sometimes] speaks so loudly that I cannot hear what you say."

## References

- Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology, 14*, 79–93.
- Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*(2), 395–418.
- Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. *Psychological Science, 9*, 40–45.
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms*. Cambridge, UK: Cambridge University Press.
- Behrend, D. A. (1989). Default values in verb frames: cognitive biases for learning verb meanings. *Proceedings of the 11th-annual meeting of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Behrend, D. A. (1990). The development of verb concepts: children's use of verbs to label familiar and novel events. *Child Development, 61*, 681–696.
- Behrend, D. A., Harris, L. L., & Cartwright, K. B. (1995). Morphological cues to verb meaning: verb inflections and the initial mapping of verb meanings. *Journal of Child Language, 22*, 89–106.
- Behrend, D. A., & Scofield, J. (2006). Verbs, actions, and intentions. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 286–305). New York: Oxford University Press.
- Bloom, L. (1978). The semantics of verbs in child language. Paper presented at the meeting of the Eastern Psychological Association, New York.
- Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression*. New York: Cambridge University Press.
- Bloom, L., Tinker, E., & Margulis, C. (1993). The words children learn: evidence against a noun bias in early vocabularies. *Cognitive Development, 8*, 431–450.
- Bornstein, M. H. et al. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75*, 1115–1139.
- Bowerman, M. (1974). Learning the structure of causative verbs: A study in the relationship of cognitive, semantic, and syntactic development. *Papers and Reports on Child Language Development, 8*, 142–178.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development, 21*, 315–330.
- Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*, 967–978.
- Childers, J. B., & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 311–335). New York: Oxford University Press.
- Clark, E. (1978). Discovering what words can do. In D. Farkas, W. Jacobsen, & K. Todrys (Eds.), *Papers from the parasession on the lexicon*. Chicago: Chicago Linguistic Society.
- DeLoache, J. S. (2004). Becoming symbol-minded. *TRENDS in Cognitive Science, 8*, 66–70.
- Echols, C. H., & Marti, C. N. (2004). The identification of words and their meanings: From perceptual biases to language-specific cues. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 41–78). Cambridge, MA: MIT Press.
- Fenson, L. et al. (2000). Short form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21*, 95–115.
- Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*, 55–64.
- Fisher, C., Hall, G. D., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. In L. Gleitman & B. Landau (Eds.), *The acquisition of the lexicon* (pp. 333–375). Cambridge, MA: MIT Press.
- Forbes, J. N., & Farrar, M. J. (1993). Children's initial assumptions about the meaning of novel motion verbs: Biased and conservative? *Cognitive Development, 8*, 272–290.
- Forbes, J. N., & Farrar, M. J. (1995). Learning to represent word meaning: Initial training events affect children's developing action verb concepts. *Cognitive Development, 10*, 1–20.
- Forbes, J. N., & Poulin-Dubois, D. (1997). Representational change in young children's understanding of familiar verb meaning. *Journal of Child Language, 24*, 389–406.
- Gelman, S. A., & Taylor, M. (1984). How two-year-old children interpret proper and common names for unfamiliar objects. *Child Development, 55*, 1535–1540.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, thought, and culture* (Vol. 2, pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gergely, G., Nadasdy, Z., & Csibra, G. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193.
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.
- Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. New York: Oxford University Press.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*, 135–176.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3–55.
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*, 23–64.

- Gleitman, L., & Gillette, J. (1995). The role of syntax in verb learning. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 413–427). London: Blackwell.
- Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M. & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23–45.
- Golinkoff, R. M., Hirsh-Pasek, K., & Schweisguth, M. A. (2001). A reappraisal of young children's knowledge of grammatical morphemes. In J. Weissenborn & B. Hoehle (Eds.), *Approaches to bootstrapping: Phonological, syntactic and neuro-physiological aspects of early language acquisition* (pp. 176–188). Amsterdam-Philadelphia: John Benjamin.
- Golinkoff, R. M., Jacquet, R. C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development*, 67, 3101–3119.
- Gopnik, A. & Meltzoff, A. N. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development*, 57, 1040–1053.
- Hampson, P. J. (1989). Aspects of attention and cognitive science. *Irish Journal of Psychology*, 10, 261–275.
- Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar*. Cambridge, MA: MIT Press.
- Hirsh-Pasek, K., & Golinkoff, R. M. (Eds.). (2006). *Action meets word: How children learn verbs*. New York: Oxford University Press.
- Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An emergentist coalition model for word learning. In R. M. Golinkoff et al. (Eds.), *Becoming a word learner: A debate on lexical acquisition*. New York: Oxford University Press.
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model of the origins of word learning. *Monographs of the Society for Research in Child Development*, 65 (3, Serial No. 262).
- Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*, 90, 72–93.
- Imai, M., Haryu, E., & Okada, H. (2002). Is verb learning easier than noun learning for Japanese children? 3-year-old Japanese children's knowledge about object names and action names. In B. Skarabela, S. Fish, & A. H. J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (Vol. 1, pp. 324–335). Somerville, MA: Cascadilla Press.
- Imai, M., Haryu, E., Okada, H., Lianjing, L., & Shigematsu, J. (2006). Revisiting the noun-verb debate: a crosslinguistic comparison of novel noun and verb learning in English-, Japanese- and Chinese-speaking children. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 450–476). New York: Oxford University Press.
- Jusczyk, P. W., Cutler, A., & Redanz, N. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675–687.
- Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, 96, 1–33.
- Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, 25, 265–288.
- Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual differences in early vocabulary development: Redefining the referential-expressive distinction. *Journal of Child Language*, 19, 287–310.
- Ma, W., Golinkoff, R., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2006). Imagine that! Imageability predicts verb learning in Chinese children. Unpublished manuscript, University of Delaware.
- Maguire, M. J. (2004). *Children's use of universal and language-specific cues in verb learning*. Unpublished dissertation. Temple University, Philadelphia, PA.
- Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A unified theory of word learning: Putting verb acquisition into context In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 364–391). New York: Oxford University Press.
- Maguire, M. J., Pruden, S. M., Hirsh-Pasek, K., Hansell, N., & Meyer, M. (2004, April). More isn't always better: Three-year-olds master intransitive verbs. Poster presented at the biennial meeting of the International Society on Infant Studies, Chicago, IL.
- McCune-Nicolich, L. (1981). Toward symbolic functioning: Structure of early pretend games and potential parallels with language. *Child Development*, 52, 785–595.
- McDonough, C., Lannon, B., Hirsh-Pasek, K., Golinkoff, R. M. (2006). The effect of imageability on early word learning. Unpublished manuscript, Temple University, Philadelphia, PA.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts of by18-month-old children. *Developmental Psychology*, 31, 838–850.
- Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, 64, 1665–1687.
- Naigles, L. G. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374.
- Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, 58, 221–251.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38 (Serial No. 149).
- Nelson, K., Hampson, J., & Shaw, L. (1993). Nouns in early lexicons: Evidence, explanations, and implications. *Journal of Child Language*, 20, 61–84.
- Pine, J. M. (1992). How referential are "referential" children? Relationships between maternal-report and observational measures of vocabulary composition and usage. *Journal of Child Language*, 19, 75–86.

- Poulin-Dubois, D., & Forbes, J. N. (2006). Words, intentions, and action: A two-tiered model of action word learning. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 262–285). New York: Oxford University Press.
- Poulin-Dubois, D., & Forbes, J. N. (2002). Toddlers' attention to intention-in-action in learning novel action words. *Developmental Psychology, 38*, 104–114.
- Pruden, S., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development, 77*, 266–280.
- Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press.
- Regier, T. (1997). Constraints on the learning of spatial terms: A computational investigation. In R. Goldstone, P. Schyns, & D. Medin, *Psychology of learning and motivation. Mechanisms of perceptual learning* (Vol. 36, pp. 171–217). San Diego, CA: Academic Press.
- Smiley, P., & Huttenlocher, J. (1985). Conceptual development and the child's early words for events, objects and persons. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 21–61). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts? In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 603–636). Cambridge, MA: MIT Press.
- Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and the lexicon, Vol. III: Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge University Press.
- Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*, 492–504.
- Tardif, T. (2006). But are they really verbs? Chinese words for action. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 477–498). New York: Oxford University Press.
- Tardif, T., Fletcher, P., Zhang, Z. X., Liang, W. L., & Chen, P. (in press). *The Chinese Communicative Development Inventory (Putonghua and Cantonese versions): Manual, forms, and norms*. Beijing: Peking University Medical Press.
- Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2002). Going, going, gone: The acquisition of the verb 'go.' *Journal of Child Language, 29*, 783–811.
- Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*, 311–333.
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language, 4*, 197–212.
- Tomasello, M. (1992). *First verbs*. Cambridge, UK: Cambridge University Press.
- Tomasello, M. (1995). Pragmatic contexts for early verb learning. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 115–146). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language, 23*, 157–176.
- Wagner, L., & Carey, S. (2005). 12-month old infants represent probably endings of motion events. *Infancy, 7*(1), 73–85.
- Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*, 217–242.
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.