NORTHWESTERN UNIVERSITY

Children Use Contrast, Multiple Familiarization Scenes, and Multiple Object Categories to Learn Verbs

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

Thomas Bruce Piccin

EVANSTON, ILLINOIS

June 2007

ABSTRACT

Children Use Contrast, Multiple Familiarization Scenes,

and Multiple Object Categories to Learn Verbs

Thomas Bruce Piccin

Recent work exploring children's verb learning in the laboratory has generated some interesting contradictions. Some studies have found that children as old as 4 years old are unable to reliably map a novel verb to an action (e.g., Kersten & Smith, 2002; Imai, Haryu, & Okada, 2005), even though much younger children routinely acquire and produce verbs in a naturalistic setting. In contrast, Waxman, Lidz, Braun, and Lavin (under review) found that infants as young as 24 months could identify an action category across dynamic scenes and map a novel verb (but not a novel noun) to that action category. In the current research, I extend the work of Waxman et al. by independently manipulating three factors—multiple familiarization scenes, multiple object categories, and explicit contrast—in a forced choice verb-learning task with 3-year-olds. Children watched one or more familiarization scenes in which an actor performed a simple, durative action with a familiar object. Scenes were labeled with either a novel verb (e.g., "The man is larping a balloon!") or novel noun (e.g., "The man is waving a larp!"). A contrast phase showed the same actor perform a different action with a different object, labeled in the negative (e.g., "Uh, he's not larping that!"), which was followed by another familiarization scene. At test, children were shown two scenes simultaneously: one

scene showed the actor performing a new action with the familiar object, and the other showed the actor performing the familiar action with a new object. Participants were asked to point to the scene representing the novel word. Children learned verbs best when all three sources of information—multiple familiarization scenes, multiple object categories, and contrast—were available. When only one was available, contrast appeared to be the most useful. The use of multiple object categories across scenes facilitated verb learning more than a single object category. Children deprived of all three sources of information did not learn verbs at all, but did learn nouns. Ways in which the information required for learning a verb may differ from that required for learning a noun are discussed.

ACKNOWLEDGMENTS

I am deeply indebted to the many, many people who helped me to realize this goal. I must begin by acknowledging that I simply could not have achieved what I achieved in the timeframe that I achieved it without the support and guidance of my advisor, Sandy Waxman. I would like to say that she is my role model for how to think about developmental psychology, human cognition, and scientific research, but the truth is that I do not aspire to be like her. My goals are far more realistic. I am happy if I just get to sit next to her once in a while. She is brilliant, and I am a better scientist and a better person for having known her.

Many other people have proven to be sources of support and inspiration for me as I have traveled on this journey, and I will always be grateful to them. Scott Baker has always believed in me and was the first to encourage me to pursue this degree. Kathy Hirsh-Pasek, knowing nothing about me except that I had an intense interest in her work, graciously invited me into her home and into her profession and has been a source of encouragement ever since. She, along with Roberta Golinkoff, Liz Shipley, and the other members of the Mid-Atlantic Language Union, warmly welcomed me into their group and helped fuel my enthusiasm for learning about language development and developmental psychology. Pam Blewitt provided warmth, serenity, and wisdom. Rebecca Brand was a generous friend and partner at a time when I needed both. Bob Johnson, Scott Liddell, Ceil Lucas, and Carol Erting of Gallaudet University introduced me to the field of linguistics and encouraged me to view language from a new perspective.

At Northwestern, my committee members Amy Booth and Sue Hespos generously offered their time, expertise, and good humor. My officemate Heather Norbury was never

without a frank, thoughtful perspective and a pretty smile. I was very lucky to work with a wonderful group of very talented and fun people in the lab, including Irena Braun, Ann Bunger, Christy Call, Alissa Ferry, Elisa Sneed German, Ariel Grace, Patricia Herrmann, Erin Leddon, Kristen Syrett, Josh Viau, and Adriana Weisleder. I benefited from the wisdom and guidance of many excellent Northwestern professors, especially Eli Finkel, Dedre Gentner, Matt Goldrick, Jeff Lidz, Dan McAdams, Barbara O'Brien, and Steve Zecker. I also benefited from the friendship and counsel of many colleagues and fellow graduate students at Northwestern, especially Flo Anggoro, Jennifer Asmuth, Dan Bartels, Stella Christie, Julie Colhoun, Sam Day, Rumen Iliev, Jason Jameson, Anna Lane, Tracy Lavin, Andrea Proctor, David Shor, Andrzej Tarlowski, Sara Unsworth, and Jennie Woodring. I would also like to thank the staff of the Northwestern Psychology Department for their help and patience, especially Florence Sales, Ginger Gilmore, Kylah Eagan, and Tomeka Bolar.

While at Villanova University, I was very lucky to work with many wonderful professors, especially Mike Brown, Tom Toppino, and Bob Broderick, and to have the support of Nancy Hippert in the Psychology Department. I also enjoyed the company of my very gracious hosts Jeff and Gillian Szanto, and the companionship of Rocky, Clyde, and Rags.

I have been fortunate to have had the support of a number of close friends throughout this endeavor, especially Theresa and Greg Anderson and their daughter Rachel, Karen Desrosiers, Susan Miller, Laurie Lynn Drevlow, Celene de Miranda, and Connie Gartner. None ever once questioned my reasons for pursuing this goal or my ability to succeed.

Many people helped me conduct the research described in this dissertation. Adriana Weisleder and Christy Call, aside from being absolutely delightful to be around, kept things

running smoothly in the lab and helped tremendously with scheduling and logistics. Amara Stuehling and Bridget O'Brien worked tirelessly for an entire summer at recruiting subjects and collecting data. Jonathan Adler, Valerie Bernstein, Christy Call, Marina Chernov, Heather Norbury, Andrzej Tarlowski, and Josh Viau contributed their acting talents toward the preparation of stimuli. Rebecca Brand generously allowed me to utilize the staff and resources of the Cognitive Development Project at Villanova University. While I cannot list by name the more than 200 children who were involved in this research, I must acknowledge that this work could not have been completed without their enthusiastic participation. I am deeply indebted to them and their families.

Finally, my mother Barbara, my father Gino, my sister Nancy, and my brother Donald have stood behind me unfailingly, even during times when I may have given them little reason to do so. It seems unlikely that I could have accomplished this or anything else without their love and support. I dedicate this work to them.

CONTENTS

TABLES AND FIGURES

INTRODUCTION

The process of learning a word involves mapping a phonological form to an abstract concept (Waxman & Lidz, 2006). Except for noun phrases that identify individual objects (e.g., proper nouns), words refer to collections of things, not specific items. Thus, a crucial part of learning a word is deciding which items or events in the world are suitable referents for a particular word and which are not. For example, a child who learns that the family pet is called a *dog* must eventually decide that the neighbor's dog is also called a *dog*, but the neighbor's cat is not and neither is the coffee table. Making such distinctions may give a word learner pause temporarily but is certainly a manageable task. After all, careful inspection of the objects in question usually yields the clues necessary to correctly assign labels. However, the same child must eventually learn that *playing* can mean stacking colored blocks, but it can also refer to an activity on television in which many large men wearing helmets try to hurt each other while passing around an oddly shaped ball—something that does not look at all like stacking blocks. Even more puzzling, perhaps, is Mommy's assertion that the man who has come to build a brick wall in the back yard is certainly not *playing*, even though what he is doing looks a lot like stacking blocks.

The question of how children learn a new verb and how they determine how to extend newly-learned verbs is of great interest to researchers who study word learning. Verbs are commonly thought of as referring to "actions," but they actually do much more. A verb encodes an "event," which can be an action (e.g., *run*), but can also be a state (e.g., *suffer*), a process (e.g., *develop*), or a relation between two entities (e.g., *love*) (Frawley, 1992). Syntactically, a verb is the foundation of a sentence, and verbs are charged with conveying the

main idea of every utterance. As I will discuss, despite the important role that verbs play in language, it has only been within the last 10 years that the study of verb learning has really begun to move to the forefront of word learning research. In Part I, I take a broad perspective and review what is currently known about word learning generally and verb learning more specifically, following the field's historical trajectory from a fixation on object labels to a much more recent focus on verbs. In Part II, I describe a program of research that systematically examines the effects of three factors on children's verb learning.

PART I: BACKGROUND

The Noun Advantage

The acquisition of object labels has been the focus of word learning research for many years (Golinkoff & Hirsh-Pasek, 2006). One factor underlying this preoccupation with nouns to date could be the observation that across languages and cultures, children's earliest vocabularies appear to be dominated by words that refer to objects (Goldin-Meadow, Seligman, & Gelman, 1976; Fenson et al., 1994; Gentner & Boroditsky, 2001). For example, Gentner (1982) examined the first 60 words produced by Tad, an American English-learning boy, and broadly categorized the words as nominals (referring to objects), predicates (referring to actions, events, or changes of state), expressives (expressing feelings or performing social functions), and indeterminates (used in multiple ways). The child's first word was *dog*, produced at 11 months of age. By 16 months, Tad had a vocabulary of 13 words, 11 of which were nominals (e.g., *duck*, *cheese*, *Mama*). Only one word, *yuk*, could be considered a predicate term, and this only by the most generous of definitions. At 19 months, eight months after Tad had uttered his first word, he produced his first verb-like term, *down*, to indicate that he wanted to be taken out of his highchair. A handful of similar predicate terms appeared over the next two months, such as *up*, *off*, and *out*, along with dozens of nominals. By 21 months of age, Tad's vocabulary had reached 60 words and included 41 nominals and 9 predicates. Conspicuously absent from Tad's vocabulary, even as he approached his second birthday, were terms that described collections (e.g., *family*), abstract nouns (e.g., *joy*), adverbs (e.g., *slowly*), and verbs (e.g., *eat*).

Indeed, it is typically not until halfway through the third year of life that the object name dominance disappears from children's productive vocabularies and the proportion of nouns begins to approximate that which occurs in the speech of their adult caretakers (Huttenlocher, Smiley, & Charney, 1983; Fenson et al., 1994; Snedeker & Gleitman, 2004). This "noun advantage" cannot be attributed to properties of infant-directed speech that cause nouns to occur more frequently, because nouns are learned more easily than verbs even when input frequency is controlled for (e.g., Leonard, Schwartz, Morris, & Chapman, 1981; Rice & Woodsmall, 1988; Merriman, Marazita, & Jarvis, 1993). It is also not the case that the noun advantage occurs only in "noun-friendly" languages that possess linguistic properties (e.g., word order, word frequency, stress patterns) that might make nouns more salient than other forms (Fisher & Gleitman, 2002). A noun advantage in early vocabularies has been reported in "verb-friendly" languages as well, including Korean (Au, Dapretto, & Song, 1994), Italian (Caselli et al., 1995), and Navajo (Gentner & Boroditsky, 2001). Additional cross-linguistic evidence comes from Bornstein et al. (2004), who found a noun bias in the early vocabularies of children learning Spanish, Dutch, French, Hebrew, Italian, Korean, and American English.

Many researchers have taken these cross-linguistic data to be evidence of a "universal" noun advantage in early word learning, but some have challenged this idea. As this debate has intensified, attention has focused on languages such as Chinese, Japanese, and Korean. These languages—unlike English—allow arguments (i.e., subjects and objects) to be dropped, often making verbs more frequent than nouns in mothers' speech to children (Imai, Haryu, Okada, Lianjing, & Shigematsu, 2006). If the noun advantage were simply the result of children's lexical production mirroring that of their caretakers, then one would expect to find a verb

advantage—or at the very least, an attenuated noun advantage—in children learning these Asian languages. Results of studies exploring this question have been inconclusive, however. While some researchers have reported finding more verbs than nouns in early vocabularies of children learning Mandarin (Tardif, 1996) and Korean (Choi, 2000), others have reported a noun advantage in children learning Japanese (Ogura, 2001, as cited in Imai et al., 2006) and Korean (Au et al., 1994), while still others have found the proportions of nouns and verbs to be roughly equal in children learning Korean (Choi & Gopnik, 1995).

One factor that seems to have contributed to the controversy is that different methodologies have led to different results. Research based on parental reports of their infants' vocabulary consistently reports a noun advantage, while research based on language production in interactive play sessions sometimes has not (Lavin, Hall, & Waxman, 2006). Still, each of these methods is limited in terms of evaluating the relative ease with which nouns and verbs are learned. The fact that a child produces a word in a certain context says nothing about how fully the child understands the word's meaning, or how adept the child might be at using the same word in other contexts (Bowerman, 1980; Tomasello, 1995). This has led to attempts at direct comparison of noun learning and verb learning in an experimental setting.

Comparing Noun and Verb Learning

One study in which noun and verb learning were directly compared was reported by Childers and Tomasello (2002). English-speaking 2½-year-olds were taught novel nouns, verbs, or actions, presented according to different exposure schedules (e.g., massed vs. distributed exposures) over a period of one to four days. Overall, children learned actions

better than either nouns or verbs, and they learned both types of words best when exposures were spread out over four days. Learning was evaluated using both comprehension and production tests. Most relevant was the fact that on average, children produced three times as many nouns as verbs. However, as Childers and Tomasello point out, there are reasons to view this apparent noun advantage with caution. First, each verb that the children learned was associated with specific object. In other words, during training, the experimenter would select a novel object and perform a novel action with that object, labeling the action (e.g., "Look at this! It's dacking. See? It dacks."). Children never saw the action performed with any other object. To the extent that this type of learning may have differed from more naturalistic verb exposure, in which verbs often (but not always) apply to a wide range of objects, verb learning may have been impaired. Second, children's comprehension of verbs was tested by asking them to select the object that was associated with the novel action (e.g., "Show me the one that dacks. Which one was dacking?"). This procedure, while designed to parallel the noun comprehension test as closely as possible, may have served to bias the comprehension measure in favor of nouns.

Childers and Tomasello (2002) examined children's facility at establishing noun and verb mappings but did not test children's ability to apply those labels to novel objects or actions not seen during training. Kersten and Smith (2002) did look at children's willingness to generalize newly-learned nouns and verbs to novel instances. English-speaking adults and 3½- to 4-year-olds were shown 24 learning events and taught two novel nouns (noun condition) or two novel verbs (verb condition). Each learning event depicted a novel bug-like object in motion along a particular path. A novel word represented an object-path pair, such that each

novel word, regardless of its grammatical class, was equally associated with both an object and a path. The syntactic frame in which the novel word was presented varied according to its grammatical class. Nouns were always presented in the form "This is a zeebee," while verbs were presented in the present tense ("This one is morping."), past tense ("See that? He morped!"), and future tense ("This one's gonna morp."). During test, participants were shown four types of events and asked if the novel word applied (e.g., "Is this a zeebee?" or "Is this one morping?"). Test events included events that matched the learning event (a) in object and path, (b) in object only, (c) in path only, or (d) not in object or path.

In the noun condition, both adults and children correctly extended the novel noun to object matches but not to path matches. In the verb condition, however, adults correctly extended the novel verb exclusively to path matches, but children failed to systematically extend the verb to either path or object matches. One possibility for this failure is that children were attempting to map the verb to the manner of motion. Since this attribute varied randomly, attending to this feature would have been unsuccessful, and the children may have searched for some other aspect of the scene to associate with the verb (e.g., the object). To explore this possibility, a second experiment was conducted in which the manner of motion rather than the path was associated with the novel verb. In both the noun and verb conditions, children mapped the novel word to the object but not the motion. Thus, 3½- to 4-year-olds failed to systematically associate a novel verb with either a path or manner of motion, tending instead to link the verb to a novel object.

Both of the studies described above reported findings consistent with the proposal that nouns are learned more easily than verbs. However, both studies also tested children learning

English, a noun-friendly language. In a series of studies, Imai and her colleagues (Imai, Haryu, & Okada, 2005; Haryu et al., 2005; Meyer et al., 2003) directly compared noun and verb learning in monolingual children learning Japanese, Chinese, and English. Three-year-olds, 5-year-olds, and adults watched a standard event presented simultaneously on two computer screens for 30 seconds. The standard event depicted an actor performing a novel action with a novel object. The event was labeled by the experimenter three times with either a novel noun ("Look! There is an X!") or a novel verb ("Look! There is Xing!"). Two test events were then presented simultaneously and the participants were asked to generalize the novel noun ("Which one is X in?") or novel verb ("In which one is the girl Xing?"). One test event (AS) contained the same action as the standard, but a different object; the other test event (OS) contained the same object but a different action. Both test events contained the same actor as the standard. Six different standard events were used, each with two test events.

Comparable response patterns were seen across all three languages. Adults and 5-year-olds reliably generalized nouns to the OS event and verbs to the AS event. That is, they reliably mapped nouns to objects and verbs to actions. Three-year-olds, on the other hand, generalized nouns to the OS event but performed no different from chance when generalizing novel verbs. This finding is striking: Three-year-olds in three different languages failed to associate a novel verb with an action, despite the fact that children of this age routinely comprehend and produce dozens of verbs in a natural language environment (Clark, 1996, 2003).

In a follow-up study performed with Japanese children, Imai et al. (2005) explored several possible explanations for the failure of 3-year-olds to generalize verbs correctly. First,

the task demands may have been too great, because children were required to hold the standard event in memory while simultaneously examining two test events. The follow-up study used a yes-no paradigm to reduce processing load. Participants were shown the standard event alongside one test event at a time and asked whether the word used to label the standard event also applied to the test event. Second, the children in the original study may have mapped the novel verb to the object, to the object as well as the action, or to the object-action combination. To test for all of these possibilities, the second study included a still object (SO) test event which contained a still image of the object on a table. If children mapped the verb to the object as well as the action, they should generalize to all test events. If children mapped the verb to the object, they should generalize to OS and SO but not AS. Finally, if children mapped the verb to the object-action combination, they should not generalize to any of the test events.

In the follow-up study, participants watched the standard event on one screen for 20 seconds while the other screen was covered. The experimenter labeled the standard event twice with a novel verb. The second screen was then uncovered, and the children saw the test event while the standard event was still visible. The experimenter labeled the standard event once more, pointed to the test event, and asked, "Is the person also Xing here?" This was repeated for six standard events, with three different test events for each standard event.

The Japanese 3-year-olds reliably rejected the SO test event—meaning that they did not map the verb to the object—but performed at chance with the AS and OS events. In no case was the proportion of "yes" responses greater than chance. These results suggest that the children expected the verbs to refer to actions, because they did not generalize to the still objects. The results also suggest that the children were overly conservative in their

generalization of novel verbs, because they failed to generalize to an event in which the action was the same as the standard but the object was different. In fact, the children failed to generalize the verbs at all.

In a third study, Imai et al. (2005) found that children would generalize a verb to a novel scene with a different actor, but only if the action and the object were the same as that seen during training. The researchers concluded that Japanese 3-year-olds correctly generalize novel nouns based on similarity of the object, independent of the action involved, but apparently map novel verbs to "object-action interactions" (p. 346). This would explain why the children generalized a verb when the agent changed but not when either the object or the action changed. Imai et al. (2006) suggested that these findings point to the progressive learning of full verb meanings over a number of years, as opposed to a "fast-mapping" strategy of learning that children have demonstrated with object labels (Heibeck & Markman, 1987). They conclude that these cross-linguistic findings support the view that "noun learning is universally advantaged over verb learning" (Imai et al., 2006, p. 472).

Why Are Verbs Harder?

Why might it be harder to learn a verb than a noun? A number of researchers have noted some fundamental differences between the functions of nouns and verbs in human language and suggested that these differences could make verb learning especially challenging relative to noun learning. For example, Gentner (1978, 1982) argued that verb referents are less "accessible" than noun referents because verbs encode a number of semantic components (such as manner and direction) and can do so in a variety of ways, whereas nouns represent

concepts that are "preindividuated" by the perceptual world. To the extent that the boundaries of a noun referent are predefined by the natural world, part of the job of word learning is already done for the child. The child needs only to establish a mapping from the word to the predefined object. In contrast, learning a verb requires that the child first identify the semantic elements that are represented by the verb—in a sense, the boundaries of the verb referent—and then map the verb to the referent. As a result, verb acquisition is necessarily slower. This is the essence of the "natural partitions" hypothesis (Gentner, 1981, 1982; Gentner & Boroditsky, 2001).

Verbs also tend to refer to events or relationships that are ephemeral in nature, whereas noun referents tend to be more stable over time (Langacker, 1987; Slobin, 2001). A child who hears the word *bottle* used to refer to an object in her environment usually has ample time to inspect the object and encode its perceptual features. In contrast, verbs like *kick*, *hug*, and *fall* label events that are short-lived and may tend to look different every time the child sees them. Furthermore, verbs are more likely to be used in "non-ostensive" contexts (when the referent is not observable) than nouns, because parents often use verbs to request that a child perform some action (e.g., "Eat your peas") or in anticipation of an impending action (e.g., "Let's change your diaper") (Tomasello, 1992; Tomasello & Kruger, 1992). Other verbs like *love* and *wish* label activities that are not perceptible at all.

Object categories are readily separated into different hierarchical levels (e.g., animal, dog, beagle), and from an early age infants appear to be sensitive to this hierarchy with respect to their word learning. For example, infants assume that a novel word labels a basic level object category rather than a subordinate or superordinate one (Rosch, Mervis, Gray, Johnson,

& Boyes-Braem, 1976; Waxman & Senghas, 1992; Waxman & Markow, 1995). However, there is no basic level category for actions (Clark, 1996). Any number of verbs can be used to refer to the very same scene (Gleitman, 1990), depending on the perspective the speaker wishes to take, and the lack of hierarchical organization likely serves to complicate the child's task of determining the correct referent of a verb.

Gentner (2006) considered a number of possible "semantic-conceptual" bases for the noun advantage in early lexical acquisition. The first, maturation, suggests that verb learning requires a level of cognitive aptitude that early word learners do not yet possess. While it seems likely that the acquisition of some words which map to especially complex concepts (e.g., *justice*) must await cognitive development (Waxman & Lidz, 2006), there are at least two reasons to believe that maturational factors alone cannot explain the noun advantage in early vocabularies. The first is that adult second-language learners also show a noun advantage, learning verbs and prepositions more slowly and with a greater incidence of errors (Dietrich, 1985; Lennon, 1996; Källkvist, 1999). The second is evidence that has come from an innovative experimental paradigm known as the Human Simulation Paradigm (HSP), introduced by Gillette, Gleitman, Gleitman, and Lederer (1999) and adopted by several others since (Snedeker, Gleitman, & Brent, 1999; Snedeker, 2000; Snedeker & Gleitman, 2004; Kako, 2005; Lavin et al., 2006; Piccin & Waxman, in press). The goal of the HSP is to measure the contribution of linguistic information in word learning by holding conceptual factors constant. To accomplish this goal, researchers have used adult participants as "simulations" of early word learners. These participants are provided with access to the kind of information that is presumably available to learners at the onset of word learning. That is, adults are permitted to

observe the scenes in which a "mystery" word occurs, without the benefit of any additional supporting linguistic information, and asked to identify the word. Because the participants are adults, any difference in their ability to identify nouns versus verbs in this task cannot be attributed to a difference in their ability to represent the underlying concepts.

In the paradigm's original implementation, adult participants watched a series of short video clips of a mother interacting with her toddler. The mother uttered the same target word across several different scenes, but the audio track had been removed and each target word had been replaced by an audible beep. Thus, participants heard no speech—only the beeps. Their task was to guess the target words, some of which were nouns and some verbs, purely on the basis of their visual observation of the scenes. In fact, when adults were deprived of access to linguistic information, they were indeed more successful in identifying nouns than verbs (Gillette et al., 1999). This finding points to the value of syntactic and linguistic information in the mapping of words to meaning, and argues against the maturational view of the early noun bias by demonstrating that even fully developed adults show a noun advantage when their access to linguistic cues is restricted.

A second possibility considered by Gentner (2006) is that verb learning is delayed because young children might be unable to perceive or comprehend the various semantic components that verbs encode. For example, a child cannot fully grasp the difference between the meanings of *give* and *sell* without some understanding of monetary transactions (Gentner, 1975). Again, while this factor is likely to play a role in the learning of some verbs (as well as some nouns), it cannot by itself explain the noun advantage. Pulverman and Golinkoff (2004) demonstrated that infants as young as 7 months are apparently able to identify and extract

critical motion verb semantic components such as path and manner. Infants were habituated to silent, computer-generated animated motion events that showed a starfish character moving in a particular manner (e.g., spinning) along a certain path (e.g., over) past a stationary ball. The infants were then shown test events in which (a) the path and manner were the same as that seen during habituation, (b) the path was the same but the manner was different, (c) the path was different but the manner was the same, and (d) both the path and manner were different. Infants dishabituated to all three test events in which either the path or manner (or both) had changed.

Additional supporting evidence comes from Casasola, Hohenstein, and Naigles (2003), who reported that 10-month-old infants habituated to a naturalistic motion event (e.g., a girl crawling toward a bush) noticed both path changes (e.g., a girl crawling away from a bush) and manner changes (e.g., a girl walking toward a bush). Furthermore, there is evidence that infants are in command of concepts that underlie many fundamental events and relations, such as *cause*, *containment*, and *support* (Baillargeon, 2000; Spelke, 2003; Hespos & Spelke, 2004), but acquisition of the corresponding verbs is delayed anyway. This suggests that the difficulty lies not in identifying the semantic components, but in packaging them and mapping to the verb (Gentner, 1982; Gleitman, 1990).

Finally, noting that children demonstrate a "relational shift" in learning across a range of domains—such that they initially focus on objects and only later on relations between objects—Gentner (2006) suggested that the early noun advantage might be one instance of this general learning tendency. For example, Gentner & Rattermann (1991) gave 3- and 5-year-olds a relational mapping task that required the children to disregard an object match and select

a relational match (e.g., same size and position). The 3-year-olds failed, tending to select the object match, but the 5-year-olds correctly selected the relational match. A similar relational shift was reported at a later age by Gentner and Flusberg (in preparation). Children were shown two different pictures that depicted the same relation between objects, but one object played a different role in each picture (e.g., a dog chasing a cat and a cat chasing a mouse). Five- and 7-year-olds tended to say that the cat in the first picture matched the cat in the second picture (object match) rather than the mouse (relational match). When the instructions were modified so that the relation in the first picture was emphasized (e.g., "Do you see this one that's chasing? What does it go with?"), 7-year-olds—but not 5-year-olds—selected the relational match. If young word learners tended to follow a similar developmental course, attending to objects first and relations later, one would predict early acquisition of object labels and later acquisition of relational terms such as verbs.

Gleitman and her colleagues (Landau & Gleitman, 1985; Gleitman, 1990; Gillette et al., 1999; Fisher & Gleitman, 2002) have taken a different tack, focusing on the linguistic rather than semantic requirements underlying word learning. According to this view, verbs are acquired relatively late not because the underlying concepts to which they refer are unavailable to young word learners, but because the linguistic information required to successfully learn verbs is not yet available to them. More specifically, these theorists argue that although the meaning of a concrete noun can often be inferred by observing the context in which it is uttered, the meaning of a verb depends more heavily on syntactic information and other linguistic cues. Because very young word learners have not yet established the ability to use linguistic information of this sort, they begin the task of lexical acquisition armed with

observation as their primary source of information. As a result, they are most likely to succeed in acquiring words whose underlying concepts can be identified from observation—primarily concrete nouns. These early-acquired nouns may then serve as a foundation for subsequent development, supporting the acquisition of additional nouns, making apparent critical aspects of linguistic structure, and facilitating the acquisition of predicates, including verbs and adjectives. Concrete nouns are therefore the stepping stones upon which subsequent word learning proceeds (Gleitman, 1990; Snedeker, 2000; Waxman & Lidz, 2006).

Looking Beyond Grammatical Class

While researchers such as Gleitman and Waxman have highlighted the differences between the ways nouns and verbs are learned, others have sought to construct a "unified" theory of word learning that attempts to look beyond grammatical class (Maguire, Hirsh-Pasek, & Golinkoff, 2006). This view acknowledges that some words are acquired before other words, but resists framing this difference in terms of grammatical class. That is, it is true that overall, nouns *tend* to be learned earlier and faster than verbs. But it is also true that some verbs appear very early in children's vocabularies (e.g., *sit*, *go*) and some nouns appear very late (e.g., *friend*, *justice*). The emergentist coalition model (ECM; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Hollich, Hirsh-Pasek, & Golinkoff, 2000) is a general model of word learning that attempts to describe the mechanisms underlying the acquisition of all words, regardless of grammatical class. The model acknowledges that children receive many different kinds of input in the course of their language development (e.g., attentional, social, and linguistic), and proposes that the weight that each type of input carries changes as a child develops from an

immature word learner to a mature one. Additionally, the model suggests that the word learning principles themselves are "emergent" in that they mature as the learner does.

A fundamental claim made by the ECM is that early word learners are primarily influenced initially by perceptual cues, and then later by social and linguistic cues (Maguire et al., 2006). This position is consistent with findings that emphasize the role of perceptual information such as shape in infants' word learning (Gentner, 1978; Landau, Smith, & Jones, 1988; Smith, Jones, & Landau, 1992), while taking into account older children's willingness to override perceptual cues in consideration of a speaker's apparent intent (Bloom & Markson, 1997; Gelman & Ebeling, 1998; Haryu & Imai, 1999). More importantly, however, the model offers a means of understanding why some words are learned earlier than others that is not tied to grammatical class, and therefore applicable to all kinds of words. If early word learning is driven primarily by perceptual information, and nouns tend to have referents that are more perceptually accessible than those of verbs (Gentner, 1982), then nouns will tend to appear earlier than verbs in children's vocabularies.

Additional support for the role of perceptual cues in word learning comes from researchers who examine word "imageability"—the degree to which a word brings to mind a mental picture, sound, or other sensory experience (Gillette et al., 1999). These researchers have noted that children's first words are typically highly imageable, and that the imageability of words is predictive of their order of acquisition (Bird, Franklin, & Howard, 2001; Bird, Howard, & Franklin, 2003; Lannon et al., in preparation). Furthermore, the imageability of words was a better predictor of learning success than was grammatical class for adults in the HSP (Gillette et al., 1999; Kako, 2005).

Maguire et al. (2006) have proposed that the concepts to which all words refer, regardless of grammatical class, can be thought of as existing along a continuum of abstractness which they label SICI—for shape, individuation, concreteness, and imageability, a collection of factors that are all believed to influence the difficulty of learning a particular word (Pinker, 1989; Gentner & Boroditsky, 2001; Snedeker & Gleitman, 2004). On one end of the continuum are referents that have stable, well-defined shapes, are easily individuated, and are high in concreteness and imageability. These would be referents of, most typically, concrete nouns and proper nouns. At the other end of the continuum are concepts that are encoded by more abstract terms such as abstract nouns and mental verbs. Importantly, placement along the continuum is not governed by word class but by the various contributions of the SICI components. Therefore, while most nouns may fall closer to the "concrete" end of the spectrum than most verbs, some verbs (e.g., *hammer*) may actually be more concrete by this measure than some nouns (e.g., *hope*). The SICI model thus purports to offer a solution to the "paradox" of verb learning—that some verbs are learned early but verbs as a class are learned later—and also to help chart the developmental course followed by children learning all types of words (Maguire et al., 2006).

The Challenges of Word Learning

What does it take to learn a word? Consider the following excerpt from an essay by British philosopher John Locke (1690, as cited in Gleitman, 1990, p. 3):

> If we will observe how children learn languages, we shall find that … people ordinarily show them the thing of which they would have them have the idea; and then repeat to them the name that stands for it, as "white," "sweet," "milk," "sugar," "cat," "dog."

As many psychologists, linguists, and philosophers have observed in the years since the lines above were written, the process of learning a word cannot be nearly as simple as Locke imagined (e.g., Quine, 1960; Macnamara, 1982; Goodman, 1983). To begin with, caretakers rarely explicitly label referents in this way for children, although there is cultural variation in this regard (Schieffelin & Ochs, 1987). Most often, a new word is introduced to a child implicitly, embedded in naturalistic speech (e.g., "Let's put on the jumper that Grandma bought for you!"). Even when adults do explicitly label things, the "things" they label are almost always objects, referred to by concrete nouns, only a subset of all the kinds of words that children eventually learn. Furthermore, even when presented with a concrete noun whose referent is explicitly labeled by an adult, the task of learning a word is not without challenges. When an adult points to something and says, "Look, that's called a cat," there are potentially an infinite number of objects, parts, properties, and actions that co-occur with the intended referent (e.g., head, ears, tail, fur, paws, white, fluffy, purring, sleeping) and it is still left to the child to determine which of these is the correct one (Quine, 1960). To make matters worse, naturalistic speech to infants and children contains any number of references to objects that are not present

(e.g., “Where is your bottle?”), activities that are not currently happening (e.g., “It’s time for your nap!”), or entities that cannot be perceived at all (e.g., “Make a wish!”).

Assuming that a child is able to identify the intended referent of a new word, there is still the problem of parsing the linguistic unit from the speech stream. Up to 90% of utterances directed at children are multiword (Brent & Siskind, 2001), and there are no obvious acoustic markers (such as pauses) that systematically define word boundaries (Klatt, 1979, 1989). Aslin (1993) found that even in situations where mothers were specifically asked to teach new words to their infants, the new word was spoken in isolation only 25% of the time. Thus, an essential and nontrivial step in the task of learning a new word is to identify and extract the word from a continuous stream of speech.

Finally, in order to learn a word a child must establish an association or “mapping” between the linguistic unit and the referent. But this is not simple either, because words do not map to tangible things, they map to concepts (Waxman & Lidz, 2006). That is, the word *cat* is not merely a pointer to the animal sleeping in the corner; the word refers to a collection of things that are all members of a conceptual category. This category includes the cat in the corner, but it also includes the neighbor’s cat and millions of other cats that the child has never seen and never will, as well as any number of toy cats, stuffed cats, drawings of cats, etc. Thus, in order to correctly establish a word-to-meaning mapping, the child must determine which things in the world (or properties or actions or relations) are appropriate referents for the word and which are not.

Despite the daunting challenges involved in learning a word, infants and children succeed with remarkable skill and swiftness. Beginning at around 7 months of age, infants are

able to recognize and learn new words (Jusczyk & Aslin, 1995), and they produce their first word between 10 and 15 months of age (Clark, 2003). By the age of 10 months, infants may have as many as 150 words in their comprehension vocabulary (Fenson et al., 1994). Sometime in their second year, infants begin to experience a "vocabulary explosion" (Benedict, 1979), during which they acquire an average of 5 to 10 new words every day from 18 months of age until they enter the first grade (Anglin, 1993). Some researchers have linked this dramatic increase in word learning to advances in cognitive or conceptual development that are believed to occur during this time period (e.g., Bloom, 1973; Gopnik & Meltzoff, 1986; Mervis & Bertrand, 1994).

How can this success be explained? Young word learners receive guidance from a variety of sources, internal and external, in the course of acquiring word meanings. One very influential source of information regarding word meanings is adults. In many cultures, when addressing infants, adults adopt a characteristic speech register known as "infant-directed speech" (IDS). Properties of IDS such as elevated emotional content, elongated pauses between utterances, repetition, and exaggerated prosody may serve to engage infants' attention and make word boundaries more salient (Fernald, 1992). Adults also provide a number of pragmatic cues to word meaning which infants exploit. For example, Baldwin (1991) showed that infants as young as 16 months old can use the direction of a speaker's eye gaze to determine the referent of a label. An experimenter showed an infant two novel toys, then gave one to the infant to play with (toy A) and put the other in a bucket (toy B) while she held the bucket in her lap. In the "follow-in labeling" condition, the experimenter waited until the infant was focused on toy A, then looked at toy A also and labeled it ("It's a toma"). In the

"discrepant labeling" condition, the experimenter waited until the infant was focused on toy A, then looked at toy B (in the bucket) and labeled it. In a subsequent comprehension test, infants associated the label with toy A in the follow-in labeling condition but not the discrepant labeling condition.

Children are also sensitive to a speaker's intention when learning a word. Tomasello and Barton (1994) taught 2-year-olds novel actions using dolls. The experimenter introduced a novel verb (e.g., "Let's plunk Big Bird") and then proceeded to perform an action on the doll. In half of the trials, upon completion of the action, the experimenter appeared satisfied and said "There!" In the other half, the experimenter performed the action clumsily and upon completion said "Whoops!" as if the action were accidental. Children tended to associate the verb with the action in the intentional but not the accidental condition, suggesting that they were not simply linking the word that they heard with the action that they saw—they were selectively mapping words only to what they perceived to be intentional acts on the part of the speaker.

The linguistic context surrounding a new word also holds cues to the word's meaning that children are able to utilize. In the first demonstration of children using the grammatical class of a word as a cue to its meaning, Brown (1957) showed 3- and 4-year-olds a drawing of a pair of hands manipulating an unknown material in a container. Three groups of children all saw the same picture, but each group heard a different description—one contained a novel count noun ("This is a sib"), one a novel mass noun ("This is sib"), and one a novel verb ("This is sibbing"). The children's interpretation of the novel word's meaning varied according to grammatical form. Those in the count noun condition interpreted "sib" as referring to the

container, those in the mass noun condition interpreted it as referring to the material, and those in the verb condition interpreted it as referring to the action.

More recently, Booth and Waxman (2003) showed 14-month-olds several exemplars of the same category (e.g., four animals) that shared the same property (e.g., purple) and described each item using a novel noun ("This is a blicket"), a novel adjective ("This is blickish"), or neutral language ("Look at this"). At test, infants were shown two objects that pitted the familiar category against the familiar property (e.g., a purple chair and a blue horse) and were asked to select one of the objects. Those in the noun condition heard "Can you give me the blicket," those in the adjective condition heard "Can you give me the blickish one," and those in the no-word condition heard "Can you give me one?" Only infants in the noun condition reliably selected the category match, suggesting that the infants were sensitive to the grammatical distinction between noun and adjective and mapped the noun (but not the adjective) to the category. Other work has revealed that by 21 months, infants can identify adjectives and map them to properties within basic level categories (Waxman & Markow, 1998).

Another way that children can use linguistic context to understand a word's meaning is through a process that has been called "syntactic bootstrapping" (Gleitman, 1990). Proponents of this theory suggest that word learners exploit correlations between the meanings of verbs and the syntactic frames in which they occur. According to this view, sensitivity to syntactic context allows word learners to induce the meaning of a verb in a way that would not be possible by simply observing the scene that is being described by the verb. For example, if a word learner watching a dog chasing a cat heard the scene described as *glorping*, there would

be no way for the learner to determine through visual inspection alone whether the verb meant *chasing* or *fleeing*, because both of those actions are occurring simultaneously. The learner would have to call upon rudimentary syntactic knowledge (e.g., the subject precedes the object in English) in order to determine who was glorping whom.

Many theorists believe that children's word learning is also guided by internal "constraints" that narrow the range of possible word meanings in a child's mind. For example, if a child saw a strange new animal and his mother said, "That's a rhinoceros," the child would naturally assume that the label referred to the entire animal, not just to the horn protruding from the animal's head. This assumption has been termed the "whole object constraint" (Markman, 1989). Not all words refer to whole objects, of course, but such an assumption is a reasonable starting point when identifying the meaning of a new word, especially if there is no other information available to suggest otherwise. Another assumption demonstrated by early word learners when they hear a new word is that the word applies to an object for which they do not already know a label. For example, shown a spoon and a whisk and asked to find the *blicket*, a 2-year-old will almost certainly point to the whisk (Merriman & Bowman, 1989). This response is thought by many to reflect a "mutual exclusivity constraint," or an assumption that each object has one and only one name (Markman, 1991). In this example, since the child already knows that a spoon is called a *spoon* but does not know what the whisk is called, it is natural to assume that *blicket* must refer to the whisk. As with the whole object constraint, children must be flexible when applying the mutual exclusivity constraint—otherwise, they would never learn that a cat is both a *cat* and an *animal*. For this reason, word learning

constraints are often referred to as “biases” or “default assumptions” to indicate that they guide, rather than dictate, children’s inferences of word meanings.

Waxman and Lidz (2006) suggest that infants begin the task of word learning already equipped with “powerful, albeit general, expectations about the links between linguistic and conceptual units.” These expectations would necessarily be specific enough to inform infants’ conceptual organization of the world and the mapping of concepts to words, but at the same time general enough to accommodate the fact that languages differ in terms of how they organize semantic categories and how they draw from these semantic categories to form grammatical categories. A dramatic demonstration of an early link between conceptual and linguistic units was given by Waxman and Markow (1995). Twelve-month-old infants were given, in succession, four different toys to play with, all from a particular category (e.g., four cars). One group of infants heard each toy labeled with a noun (e.g., “See the auto?”), another group heard an adjective (e.g., “See the autish one?”), and a third group heard no novel word (e.g., “See here?”). During the test phase, the infants were presented with two new toys, one from the same category (e.g., a car) and one from a different category (e.g., an airplane). The dependent measure was the time that the infants spent looking at or handling each of the two items during the test phase. Infants in the noun and adjective group showed a novelty preference at test, preferring to play with the toy from the new category. However, infants in the no word condition showed no novelty preference. These findings suggest that hearing a novel word prompted the infants to identify the commonalities across the four familiarization items and form a conceptual category in a way that infants in the no-word condition did not. This is a striking example of how linguistic and conceptual development are intertwined in

early word learners. Two additional points are of interest. First, infants in the noun and adjective conditions performed comparably, indicating that any words, regardless of grammatical class, facilitate object categorization at this early stage of word learning. Second, the categories formed by the infants in the two word conditions included not only the familiarization items but also the as-yet unseen items encountered during the test phase, indicative of the strength of the infants' expectations of the referential power of words.

Shifting the Focus to Verbs

While much of the word learning research to date has focused on object labels, it has become increasingly apparent that the way in which children learn nouns may not be representative of the way in which children learn words generally (Waxman & Lidz, 2006). Accordingly, many researchers have begun to turn their attention to the acquisition of other kinds of words, including **pronouns** (e.g., Girouard, Ricard, & Decarie, 1997; Leonard, Waters, & Caplan, 1997; Oshima-Takane, 1999; Song & Fisher, 2005), **proper nouns** (e.g., Hall, 1991; Imai & Haryu, 2001; Jaswal & Markman, 2001; Hall, Waxman, Bredart, & Nicolay, 2003), **adjectives** (e.g., Waxman & Klibanoff, 2000; Booth & Waxman, 2003; Waxman & Booth, 2003; Blackwell, 2005; Mintz, 2005), and **verbs** (e.g., Gentner, 1982; Huttenlocher et al., 1983; Naigles, 1990; Tomasello & Kruger, 1992; Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996; Imai et al., 2005).

Verbs have enjoyed particular attention recently, owing to the pivotal role that verbs play in language (Golinkoff & Hirsh-Pasek, 2006). A verb has been described as "the architectural centerpiece" of a sentence (Pulverman, Hirsh-Pasek, Golinkoff, Pruden, &

Salkind, 2006, p. 134). Clark (1996, p. 833) suggested that verbs "provide the core of most utterances" and "play a central role in the acquisition of syntax." Motion verbs, in particular, have served as a useful starting point for verb learning research. These verbs are often highly imageable and might therefore be more readily learned than more abstract verbs (Gentner & Boroditsky, 2001; Maguire et al., 2006), a position that is strengthened by the fact that verbs such as *jump* and *dance* are among the earliest verbs in children's lexicons (Fenson et al., 1994). Still, even the simplest of motion verbs comprises a number of semantic components, such as path, manner, and shape of motion (Talmy, 1975). In recent years, a number of studies have been designed to identify and isolate those semantic components that word learners find most relevant in determining a verb's meaning.

For example, Behrend (1990) showed participants a series of short videotapes, each depicting an event in which an adult performed a simple transitive action. Participants were asked to describe "what the people in the movies are doing." In constructing the events, Behrend considered three types of motion verbs: *action* verbs, which describe the physical action associated with the verb (e.g., *pound*); *result* verbs, which describe the end result (e.g., *flatten*); and *instrument* verbs, which describe the instrument used (e.g., *hammer*). The events were chosen such that each could be correctly labeled by at least two common verbs, one from each of these three types. One-third of the events had appropriate action and result verb labels (AR events), one-third had appropriate instrument and action verb labels (IA events), and one-third had appropriate instrument and result verb labels (IR events). For example, one of the IA events depicted a person shoveling sand in a sandbox and could have been labeled with an

instrument verb (*shovel*) or an action verb (*dig*). In another case, an AR event showed a person sweeping and could have been labeled with an action verb (*sweep*) or a result verb (*clean*).

After watching each event, participants were prompted for as many words as they could think of to describe the action. Behrend (1990) found that instrument verbs were most often given as a first response, despite the fact that instrument verbs were the least common type given overall. An age effect was also present, such that 3-year-olds used instrument verbs less frequently than 5-year-olds, 7-year-olds, or adults. Behrend speculated that instrument verbs were less common in English, which would explain the age effect, but were more common responses because they were more "efficient." That is, an instrument verb often incorporates an action or a result (e.g., to hammer can mean to pound with a hammer). Similarly, basic level category names for objects are efficient in that they reflect a maximal balance between similarities among members within a category and differences between members of different categories (Rosch et al., 1976). However, while children tend to interpret a novel noun as a basic level category (Bloom, 2000), they are unlikely to interpret a novel verb as an instrument verb, due to the infrequency with which instrument verbs occur in the language (Behrend, 1990).

In a second study, 3-year-olds, 5-year-olds, and adults were shown videotapes of an adult performing a novel action with a novel instrument and achieving a particular result. The event was labeled by the experimenter with a novel verb (e.g., "Watch this person, she is remming."). After three presentations of the training event, participants watched four different test events and were asked whether the novel verb applied to each (e.g., "Is she remming this time, or is she doing something else?"). One of the test events was identical to the training

event, and each of the other three test events differed from the training event in action, instrument, or result only. In contrast to the results from Study 1, (Behrend, 1990) found that participants were least likely to extend the novel verb when the test event differed in result, an effect which increased with the age of the participants. This result was an interesting contrast to that of Study 1, in which participants seemed to be suggesting through their labeling that instrument was the most important feature of an event reflected in a verb. In subsequent studies, Behrend explored and ruled out two possible explanations for the results of Study 2: that result changes in the test events were more salient than the other types of changes, and that participants were treating the novel verbs as synonyms for known result verbs.

In discussing these contrasting results, Behrend (1990) suggested that semantic differences in the ways that nouns and verbs are represented might lead to different naming strategies. Specifically, children might tend to label an event with the most specific verb available even while they generally avoid using specific subordinate labels to name objects. Since instrument verbs are often more specific than action or result verbs (because instrument verbs can include these other components), this would lead to a tendency to produce instrument verbs, as was seen in Study 1. However, Behrend argued, the infrequency of instrument verbs would make it unlikely that children would interpret a new verb as an instrument verb. An additional finding by Behrend (1990) was that 3-year-olds used significantly more action verbs and significantly fewer result verbs in describing familiar events than older children or adults, a finding that is consistent with results obtained by Gentner (1978).

Using a similar methodology, Behrend, Harris, and Cartwright (1995) asked whether the inflection used with a novel verb would influence whether a child would initially interpret

the verb as an action verb or as a result verb—a possibility that they termed "morphological bootstrapping" (p. 97). The logic was that if a child heard a new verb used only with a progressive *–ing* inflection, then she might be more inclined to infer that the verb described a "durative" event with a distinct action and would therefore be less likely to generalize the verb to an event with a different action. Similarly, if the child heard a new verb used only with a past tense *–ed* inflection, she might tend to assume that the verb referred to a "completive" event that described a particular result, and so would be less likely to generalize the verb to events with different results (see Antinucci & Miller, 1976).

In one experiment, participants heard a single type of inflection. Three-year-olds, 5-year-olds, and adults watched a training event and heard a novel verb that had a progressive inflection (e.g., "Watch, she is remming"), a past inflection (e.g., "Watch, she remmed"), or no inflection ("e.g., "Watch, she will rem"). They then watched four different test events, one which was identical to the training event and three in which either the action, instrument, or result was different. After each test event, the experimenter asked if the novel verb applied. For example, the experimenter asked "Was she remming that time?" in the progressive condition, "Do you think she remmed that time?" in the past condition, and "Did she rem that time?" in the neutral condition.

The predicted effect was not seen in any of the age groups. That is, none of the participants reliably varied their verb generalization patterns as a function of the type of inflection they heard. However, responses for the 3-year-olds tended toward the predicted pattern, even though this tendency was not statistically significant (Behrend et al., 1995). This prompted the researchers to conduct a second experiment. Unlike Experiment 1, Experiment 2

had a within-subjects design. Each participant was taught three verbs with the progressive inflection and three with the past inflection. This time, 3-year-olds reliably responded as predicted. When taught a verb with the progressive inflection, the children were less likely to extend the verb to an event with a different action than to one with a different result. When taught a verb with the past inflection, however, the response pattern was just the opposite. Five-year-olds and adults did not demonstrate this morphological bootstrapping effect, possibly because by age 5 children are familiar enough with verb morphology that they do not rely on inflection to infer verb meaning (Behrend et al., 1995).

Forbes and Farrar (1993) considered four semantic components of a motion verb: continuity (whether the motion continued uninterrupted), direction (compass direction in which the motion proceeded), instrument (object used to perform the motion), and causative agent (whether a person performed the motion himself or was caused to perform it by another). The researchers systematically manipulated these four components in order to determine how the individual components might influence verb extension patterns. Three-year-olds, 7-year-olds, and adults watched a training videotape which showed two men performing a novel activity (e.g., one man using a pole to tow the other man, seated on a cart, in a circle). The activity was labeled by the experimenter with a novel verb (e.g., “Watch these people, they are romzing.”). Participants then watched four different generalization videotapes. Each generalization event was identical to the training event in all but one of the four semantic components. After each generalization event, participants were asked to generalize the novel verb (e.g., “Are they romzing this time, or are they doing something else?”).

One significant finding reported by Forbes and Farrar (1993) was that participants of all ages were less likely to generalize the novel verb to events in which the causative agent had been changed than they were to generalize to any of the other test events. This suggests that of these four action components, the one that was considered to be most critically linked to the verb's meaning was causative agent. For example, participants who initially saw one man towing another on a cart and heard the action labeled as *romzing* were unlikely to say that a man pushing himself around on a cart was also *romzing*. A second finding was that overall, children were less likely to generalize than adults, but children's patterns of extensions were identical to those of adults.

In later work, Forbes and Farrar (1995) used a similar procedure to manipulate agent, instrument, manner, and outcome while teaching children and adults novel verbs. In each of three conditions, participants watched three different training events. All three events were labeled by the experimenter with the same novel verb (e.g., "That was noffing."). In the "same" condition, all three training events depicted the same agent, instrument, manner, and outcome. In the "different" condition, one of these components was varied across the three training events while the other three were held constant. In the "mixed" condition, one component was held constant while the other three varied. Thus, the specificity of the verb definitions provided to participants varied across the three conditions, and the prediction was that participants' willingness to extend the verb to novel events would reflect this difference. Participants in the "same" condition saw the least amount of variability during training events, so they were expected to be least likely to extend the verb. Participants in the "different"

condition were expected to show the most number of extensions, and participants in the "mixed" condition were expected to fall somewhere in between.

During test, each participant watched four novel events. In each test event, one of the four components was different from that seen during the training phase. Participants were asked whether the novel verb applied to each test event (e.g., "Was that noffing, or was that something else?"). One of the principal findings was that different training conditions did not alter the relative importance that 10-year-olds and adults placed on the various action components when they were extending a verb to new situations, but this was not true for 3-year-olds.

In the "same" condition, all action components were given equal weight during training, since none of them varied. It was therefore up to the participants to decide which components were most integral to the meaning of the verb. However, in the "different" condition, participants were shown that various action components could change without changing the meaning of the verb. In both of these situations, 10-year-olds and adults consistently generalized less for manner and outcome, suggesting that these components were most strongly associated with the novel verb. In contrast, 3-year-olds' responses differed in these two situations. In the "same" condition, the children favored instrument and outcome; in the "different" condition, they favored manner. Still, across all three conditions, 3-year-olds' generalizations with regard to manner did not differ significantly. Forbes and Farrar (1995, p. 17) therefore concluded that "by 3 years of age, children are moderately biased to consider manner a prominent element of action verb meaning."

Fast Mapping

It is well established that children can acquire object labels after a small number of nonostensive exposures—sometimes after a single exposure (Bloom, 2000). To what extent is the same true of verbs? Golinkoff, Jacquet, Hirsh-Pasek, and Nandakumar (1996) demonstrated that 3-year-olds would "fast map" a novel verb to a novel action being performed by a Sesame Street character, and would then extend the verb to apply to the same action being performed by a different character. In one experiment, children were shown an array of four drawings, each of which portrayed a different character performing some action. Three of the actions were familiar (e.g., eating) and one was unfamiliar (e.g., bending over backwards). Children were prompted to point to one of the familiar actions (e.g., "Where's eating?") as well as the unfamiliar action, which was labeled with a novel verb (e.g., "Where's daxing?"). The purpose of this trial was to determine whether children would fast map the novel verb to the unfamiliar action.

In the second trial, children were shown two new familiar actions, a new unfamiliar action, and the unfamiliar action from trial 1 being performed by a different character. This trial tested whether the children would extend the novel verb learned in trial 1 to a new illustration that featured a different agent performing the same action. In the third trial, the stimulus set again contained two new familiar actions, as well as the original unfamiliar action seen in trial 1, and a new unfamiliar action. This trial was intended to test whether the children would associate a second novel verb with the new unfamiliar action. Finally, trial 4 tested whether the children would extend this second novel verb to a different agent. This trial

included a second exemplar of the unfamiliar action from trial 3, as well as two new familiar actions and the unlabeled unfamiliar action from trial 2.

As reported by Golinkoff et al. (1996), children did indeed map the novel verbs to the novel objects in trials 1 and 3, and they extended those verbs to new agents in trials 2 and 4. Golinkoff et al. interpreted these results as evidence in support of the notion that "verb extension, like noun extension, is based on the implicit recognition that a verb labels a category of actions, potentially performed by a variety of agents, in a variety of locales and orientations" (p. 3103). Building on the work of a number of researchers (e.g., Markman & Hutchinson, 1984; Merriman & Bowman, 1989; Waxman & Kosowski, 1990), Golinkoff and her colleagues proposed a hierarchical framework of "lexical principles" that emerge in infancy and serve to guide early word learning (Golinkoff et al., 1994; Golinkoff, Hirsh-Pasek, Mervis, Frawley, & Parillo, 1995). These principles form the basis of the emergentist coalition model (ECM), discussed earlier. The two-tiered model includes fundamental principles on the first tier that allow word learning to begin (e.g., "reference"—the idea that words refer to objects, actions, and attributes), and more advanced principles on the second tier that guide assumptions about new word meanings and promote rapid lexical acquisition. The model acknowledges that children receive many different kinds of input in the course of their language development (e.g., attentional, social, and linguistic), and proposes that the weight that each type of input carries changes as a child develops from an immature word learner to a mature one. The hierarchical framework offers a means of characterizing the ways in which word learning itself changes over the course of a child's development. According to the ECM, infants acquire the basic first-tier principles by the end of their first year. These principles lay the foundation for

the second-tier principles that drive infants to adopt conventional word forms (Clark, 1991) and enable the rapid vocabulary development seen in the second year.

Two of the principles posited by Golinkoff et al. (1995) are particularly relevant to the extension of verbs: "extendibility" on the first tier and "categorical scope" on the second tier. The principle of extendibility states that words refer not to individual things but to classes of things. This principle allows infants to understand that words do more than just label individual occurrences. Golinkoff et al. argued that perceptual similarity is initially the most common basis for extension, both for objects and actions. The question of what exactly defines the "shape" of an action, however, has yet to be answered. When infants advance to the second tier and acquire categorical scope, their patterns of word extension become more mature. While extendibility allows infants to generalize words beyond the original referent, categorical scope guides that generalization in meaningful ways. In the context of object labels, it is at this point that children begin to extend labels based on taxonomic category. With regard to verbs, Golinkoff et al. (1995, p. 208) argued that through categorical scope, "action labels are extended to other actions that appear to require the same semantic components as the original action."

Forbes and Poulin-Dubois (1997) put the principles of extendibility and categorical scope to the test in a preferential looking paradigm in which 20- and 26-month-olds were asked to generalize familiar verbs. Infants were familiarized to 8-second videotapes of two "standard" events: a *kick* event (a woman kicking three plastic balls into a box) and a *pick-up* event (a woman bending over and picking up three plastic balls). Participants were prescreened to ensure that they understood the verbs *kick* and *pick-up*. As the infants watched the

familiarization trials, they heard the events labeled with the familiar verbs (e.g., "Watch! This is kicking. See? Kicking."). Test events were identical to the standard events except that either the agent, manner, or outcome was changed. During test trials, infants saw two events simultaneously and heard a prompt directing them to look at a particular event. For example, during a *kick* agent test trial, infants simultaneously saw the *kick* event with a different agent and the *pick-up* event with a different agent and heard, "Look! Where is kicking? Find kicking." Looking times to the two different stimuli were measured.

Forbes and Poulin-Dubois (1997) argued that changing the manner or outcome of an event changes its shape, whereas changing the agent does not. Therefore, they reasoned, if participants extended the verbs *kick* and *pick-up* according to the principle of extendibility, then they should be less likely to generalize to events with a different manner or outcome than to events with a different agent. However, if participants were being guided by the principle of categorical scope, then shape of the event should be irrelevant and they should be equally likely to generalize to a different agent, manner, or outcome.

The researchers found that the 26-month-olds generalized the verbs to events with a different agent or manner, but not to events with a different outcome. This result is quite consistent with those discussed previously (e.g., Behrend, 1990). For the 20-month-olds, however, Forbes and Poulin-Dubois (1997) found an effect of expressive vocabulary size as measured by the MacArthur Communicative Development Inventory for toddlers (Fenson et al., 1991). Specifically, infants with high vocabulary production (those with scores above the median for the entire sample of $n = 22$) generalized to events with a different agent, but those with low vocabulary production did not. None of the 20-month-olds generalized based on

manner or outcome. Forbes and Poulin-Dubois interpreted these findings as evidence of the word learning principles posited by Golinkoff et al. (1995), and concluded, "the present study does suggest that the bases for extending [verb] concepts change significantly from 1;8 to 2;2: from strictly appearance to progressively more essential elements of semantic meaning" (p. 404).

Verbal Essence

Golinkoff et al. (2002) suggested that motion verbs are properly extended based on "an averaged representation of what that action looks like," what they termed the "verbal essence" (p. 605). They argued that for motion events, this verbal essence is fully characterized by two components: path and manner. If children could be shown to extend a verb based on these two features alone, then they could be said to have acquired the essence of the verb.

In order to test the hypothesis that path and manner are sufficient to characterize a motion verb, Golinkoff et al. (2002) sought to create visual displays of actions that included path and manner but masked all other semantic components, such as agent, instrument, and location. They could then show these displays to participants and test whether the participants could correctly identify the verbs that labeled the actions. To this end, Golinkoff et al. created "point-light" displays of adults performing familiar actions. Small white lights were attached to an adult's joints, and the adult was videotaped in the dark performing such actions as walking, dancing, and hopping. In the resulting display, only the lights were visible.

Three-year-olds were tested in a preferential looking paradigm. During salience trials, they saw two different actions simultaneously on different television screens, and heard speech

that identified the two actions but did not indicate which was which (e.g., "Hey, one is walking and one is dancing!"). It was expected that looking times to both screens would be equivalent. In test trials, children saw the same two actions but were prompted to look at a particular action (e.g., "Look at dancing! See dancing?"). Of course, it was predicted that during test trials children would look significantly longer at the screen that matched the auditory stimulus, and that, in fact, was the finding (Golinkoff et al., 2002). In a second experiment, children were shown point-light displays and asked to produce their own labels for what they saw. The majority of children readily recognized that the pattern of dots represented an action and produced the relevant verb.

While it is certainly impressive that young children can readily identify and label actions given nothing more than a pattern of moving lights, the extent to which these findings generalize to other types of verb learning is not clear. Golinkoff et al. (2002, p. 612) suggested that "verb extension proceeds as children abstract the invariant semantic components that describe an action." But their argument (and the entire point-light paradigm) seem applicable only to motion verbs, arguably a very narrow class of verbs. While the point-light paradigm is supposed to preserve verbal essence, in fact it preserves only physical essence. Many (perhaps most) of a child's earliest verbs—for example, *play*, *eat*, *fix*, *clean*, *cook*, *open*, *break*, *sleep*, *read*, and *look*—do not have consistent physical forms. It is not clear how word learning principles that are based on perceptual characteristics would explain the acquisition of these verbs.

Akhtar and Tomasello (1996) demonstrated rather dramatically that one thing that is *not* required for a child to learn a verb is temporal contiguity between the verb and the action to

which it refers. Two-year-olds were familiarized with four novel actions, each of which was associated with a different Sesame Street character (e.g., Oscar dropped through a chute, Grover spun on a plastic lid). Characters were removed individually from a canvas bag prior to each action, and then returned to the bag when the action was completed. After each action had been demonstrated 12 times with neutral language (e.g., "Watch what I can do to Oscar"), the language modeling phase began. In the referent condition, the experimenter removed one of the characters from the bag and labeled the action that was associated with that character (e.g., "Now let's meek Oscar. Let's meek him, okay? Let's meek him."). The experimenter proceeded to perform the appropriate action with the target character, and then performed the other three actions with their associated characters using neutral language. Thus, in the referent condition, one of the four actions was labeled with a novel verb, and the children saw the action performed immediately after hearing the novel label.

In the absent referent condition, children heard the action labeled in the language modeling phase but did not see the action performed. That is, after labeling the action, the experimenter pretended to be unable to find the target character in the bag. The experimenter then proceeded to perform the other three actions with the other characters using neutral language. In this condition, therefore, children never saw the action performed after they heard the novel verb. In order to correctly map the verb to the action, the children would have to remember the action that was associated with the target character and then map the new verb to that action. In a control condition, children heard only neutral language with all four actions during the familiarization and language modeling phases.

In a subsequent test phase, both comprehension and production of the novel verb were measured. First, the experimenter handed the child a new character (e.g., Cookie Monster) and asked the child to "meek" Cookie Monster. This was followed by the experimenter performing the target action with Cookie Monster and asking the child, "What am I doing to Cookie Monster?" Akhtar and Tomasello (1996) reported that on both measures, children in the absent referent condition performed comparably to those in the referent condition, and children in both experimental conditions performed significantly better than those in the control condition. That is, children who heard the novel verb and then saw the target action performed immediately afterwards were no more effective at learning the verb than were children who were never exposed to the label and the action at the same time. The researchers interpreted these results as strong evidence that word learners do not require perceptual pairing of word and action to learn a new verb, and in fact are quite adept at using social-pragmatic cues to infer a speaker's intended meaning even in the absence of the referent action. The researchers also argued that word learning principles—such as those proposed by Golinkoff et al. (1995)—would have been no help to children in this word learning situation.

Thus, while word learning principles such as those proposed by Golinkoff et al. (1995) may be quite useful in describing the acquisition of object labels, they may be much less helpful when discussing verbs. Indeed, as discussed previously, given the privileged role that object labels play in an infant's linguistic and conceptual development, it is likely that the process by which children learn nouns is not at all representative of word learning in general (Waxman & Lidz, 2006). The research that I describe in Part II takes a closer look at some specific factors that facilitate children's verb learning.

Summary

Young word learners face a number of challenges, from extracting a new word from a continuous stream of speech to identifying the intended referent to determining the manner in which the new word should be generalized to other objects, events, or contexts. Despite these challenges, infants and children succeed splendidly, amassing an impressive vocabulary in just a few years. To date, researchers exploring the question of how children learn words have focused on the acquisition of object labels, possibly influenced by the fact that early lexicons are dominated by concrete nouns. However, evidence suggests that nouns enjoy a privileged status for early word learners, owing to infants' inclination to seek out and establish links between conceptual and linguistic units. Investigators have begun to direct their attention at the processes by which other kinds of words—particularly verbs—are learned. Theoretical argument and empirical evidence converge to suggest that the key to verb learning is access to more and varied information—be it linguistic in nature (e.g., exposure to multiple syntactic frames containing the verb), conceptual (e.g., exposure to multiple contexts in which the verb is used), or, most likely, some combination of these. In Part II, I describe a series of experiments in which I examine some specific factors that influence young children's verb learning.

PART II: RESEARCH PROGRAM

Overview

In Part I, I referred to the "paradox" of verb learning—the fact that some verbs are learned early in a child's development, but verbs as a class are learned much later. Particularly puzzling is the fact that some researchers have reported that it is not until age 5 that children have been able to learn a verb in an experimental setting and successfully apply it to a new situation. For example, Imai et al. (2005) showed Japanese 3-year-olds, 5-year-olds, and adults a 30-second "standard" event in which a young woman performed a novel action with a novel object. Participants heard the event labeled three times with either a novel noun or a novel verb. Participants were then shown two test events simultaneously and asked to extend the novel word to one of the events. In one test event, the action was the same as in the standard event but the object was novel. In the other test event, the object was the same but the action was novel. In the noun condition, participants of all ages correctly extended the novel word to the object. However, in the verb condition, only adults and 5-year-olds correctly mapped the verb to the action. Three-year-olds performed at chance. This is a surprising finding, given that children of this age routinely comprehend and produce dozens or perhaps hundreds of verbs in a natural language environment (Clark, 1996).

Strikingly different results were obtained by Waxman and her colleagues, who developed a paradigm that incorporated two features that often have not been considered in previous verb learning studies: the use of multiple familiarization scenes and the use of explicit contrast (Waxman, Lidz, Braun, & Lavin, under review; see also Waxman & Booth, 2001). Twenty-four-month-old infants watched a series of familiarization scenes in which the same

actor (e.g., a woman) performed the same familiar action (e.g., washing) using four different exemplars of a particular familiar object category (e.g., cup). All infants saw the same video, but infants heard the scenes labeled with either a novel verb (e.g., "Look, the woman is larping a cup"), a novel noun (e.g., "Look, the woman is washing a larp"), or neutral language (e.g., "Look at this!"). The familiarization phase was followed by a contrast phase, which had two parts. In the first part, participants saw the actor from the familiarization phase perform a different action with a different object and heard the scene labeled with a negative tone (e.g., "Uh oh, she is *not* larping that!", "Uh oh, that's *not* a larp!", "Uh oh, look at that!"). In the second part of the contrast phase, participants saw a repeat of one of the familiarization scenes and heard the scene labeled with a positive tone (e.g., "Yay, she *is* larping that!", "Yay, that *is* a larp!", "Yay, look at this!").

At test, infants in all conditions saw two scenes simultaneously: a familiar scene (one of the scenes from the familiarization phase), and a novel scene, in which the now-familiar actor performed a different action (e.g., tapping) on the familiar object (e.g., cup). The two images were initially accompanied by neutral language ("Now look, they're different!") in order to assess the infants' baseline looking preference. In the second stage of the test phase, the images remained the same but infants heard a comprehension probe that varied depending on condition (e.g., "Where is she larping something?" in the verb condition, "Which one is a larp?" in the noun condition, or "What do you see now?" in the no-word condition). The prediction was that if infants mapped the novel verb to the action, then they should look toward the familiar scene (containing the familiar action) in response to the prompt. Otherwise, they should prefer the novel scene.

For infants in the verb and noun conditions, there was a clear practice effect. During the first block of test trials, infants in both conditions looked longer toward the familiar scene during the comprehension period relative to the baseline period. By the second block of test trials, this effect became more pronounced for those in the verb condition but disappeared for those in the noun condition. By the second block of test trials, therefore, infants in the verb condition showed a significant increase in time spent looking toward the familiar scene during the comprehension period relative to the baseline period, a difference not seen in the other two conditions.

Thus, 24-month-olds apparently detected the common action across familiarization scenes and mapped the novel verb (but not the novel noun) to that action. This mapping was independent of the particular object, since different exemplars were used in the familiarization scenes. This finding is striking, given the previously reported failures of much older children to learn novel verbs in the laboratory (e.g., Imai et al., 2005). With the current research, I extend the work of Waxman et al. (under review) by independently manipulating the factors of multiple familiarization scenes and explicit contrast and adding a third factor, the use of multiple object categories across familiarization scenes. The goal was to assess the contribution of each of these three factors to early verb learning. I use 3-year-old participants in a forced-choice pointing task, and I increase the difficulty of the test phase by pitting familiar action against familiar object. That is, I include one test scene in which the action seen during familiarization is performed with a new object, and a second scene in which the object seen during familiarization is used with a new action. Therefore, children who map the novel

word to the action seen during familiarization must extend the word to a new object category at test.

Three-year-olds were selected as participants for several reasons. Children of this age typically have vocabularies that are predominantly count nouns, and their ability to acquire and produce new verbs is still developing. At the same time, however, they have acquired enough syntax to allow them to parse the input properly and extract a novel noun or verb. Additionally, there are mixed reports of the ability of children at this age to fast map novel verbs. Some studies have found that they can (e.g., Golinkoff et al., 1996), but others have reported failures (e.g., Haryu et al., 2005). Apparently, the ability of children at this age to learn and extend novel verbs is present but fragile, making it an ideal age to explore the circumstances under which children can and cannot learn novel verbs in an experimental setting.

One notable difference between the experiments of Waxman et al. (under review) and those of Imai et al. (2005) is that Waxman et al. taught infants novel verbs for familiar actions, while Imai et al. used novel actions. It is possible that the use of familiar actions facilitated novel verb mapping the 24-month-olds; infants learn new words better in familiar, less-taxing contexts than in novel ones (Kersten & Smith, 2002; Casasola, Cohen, & Chiarello, 2003; Werker & Fennel, 2004). On the other hand, children may resist mapping new words to familiar referents due to mutual exclusivity constraints or resource conservation strategies (Markman & Wachtel, 1988; Woodward & Markman, 1998; Piccin & Blewitt, 2007). In the current work, as in Waxman et al., I use simple, familiar actions to ensure that the children's

word mapping is not hindered by "unnatural" novel actions that may be more complex or artificial than those encountered in a naturalistic environment.

The results reported by Waxman et al. (under review) provide empirical support for the belief that the three independent variables considered in the current work—multiple familiarization scenes, explicit contrast, and multiple object categories— may be particularly beneficial for verb learning. Researchers such as Gleitman (1990), Pinker (1989, 1994), and Gentner (1982) have long argued that cross-situational exposure is crucial to verb learning, because hearing a novel verb applied across a variety of contexts helps a child to identify common semantic aspects across those contexts and isolate those that are associated with the verb. Consistent with this idea is evidence that children's early verbs are context-specific. Children appear in some tasks to be quite conservative in the way that they generalize new verbs, and are often reluctant to extend a verb to scenarios where the actor or objects are different (Behrend, 1990; Behrend et al., 1995; Forbes & Farrar, 1993, 1995), or to produce a verb in a syntactic frame that is different from what they have been exposed to (Tomasello, 1992, 1995).

Other research has underscored the importance of contrast in lexical acquisition—as well as in learning generally (Quinn & Eimas, 1987; Krueger & Rothbart, 1990; Billman & Davila, 2001; Hampton, Estes, & Simmons, 2005). Clark (1990, 1992) has argued that contrast is a powerful tool that is recruited by language users from a very young age to infer intended meaning, and this view is supported by a large body of evidence (e.g., Au & Markman, 1987; Heibeck & Markman, 1987; Clark, 1988; Waxman & Kosowski, 1990; Waxman & Klibanoff, 2000; Diesendruck, 2005). This technique is especially helpful if a novel word is contrasted

with a known word from the same domain. For example, if children hear "Give me the beige one, not the red one," they are likely to infer that beige is a color because they know that red is a color. Thus, evidence of what a novel word does *not* refer to can often serve as additional information about what the word *does* refer to.

The third variable, the use of multiple object categories, was not used by Waxman et al. (under review) but was incorporated into the current research program with the hypothesis that the use of a variety of objects would help children identify the action as the referent of the novel verb. Evidence that supports this prediction comes from work reported by Forbes and Farrar (1995), who found that when 3-year-olds were taught a novel verb for an action using the same instrument across training trials, the children were unlikely to generalize the verb to an action using a novel instrument. However, if the children saw the action demonstrated using different instruments during training, then their tendency to generalize the verb to a novel instrument increased dramatically.

Findings by Maguire et al. (2002) may also be relevant to this issue. These researchers found that 19-month-old infants were able to map a novel verb to an action if the action was shown in a point-light display, but not when the action was shown performed by live actors. The interpretation was that the point light display served to strip away irrelevant information from the scene (such as the actor) and help the infants attend to only those perceptual features that were necessary to characterize the action. It is possible that for 3-year-olds, the presence of different object category exemplars across scenes may serve a similar purpose: a variety of objects across identically-labeled scenes may drive home the point that the object is irrelevant to the label.

In the current program of research, I systematically manipulate these three variables—use of multiple familiarization scenes, use of explicit contrast, and use of multiple object categories—in order to determine their effect on children's verb learning. The research program includes six studies. An overview of the design of each study in terms of the three variables is given in Table 1. My overarching goal with this research was to gain insight into the conditions that contribute to successful verb learning—while shedding some light on previously reported failures of toddlers to learn novel verbs—and to explore the ways in which the information required for learning a verb differs from that required for learning a noun.

	Multiple familiarization scenes	Contrast	Multiple object categories
Study 1	+	+	–
Study 2	–	+	–
Study 3	–	–	–
Study 4	+	–	–
Study 5	+	+	+
Study 6	+	–	+

Table 1. Overview of Study Designs

Study 1: A Paradigm for Successful Verb Learning

I began with the combination of the three independent variables that mirrored the work of Waxman et al. (under review): multiple familiarization scenes combined with explicit

contrast, with a single object category used across familiarization scenes. Given the success of the 24-month-olds reported by Waxman et al. using this paradigm, I expected that 3-year-olds would have no problem learning novel verbs. Children watched a series of four familiarization scenes on a computer monitor in which an actor performed a particular action using a particular object (e.g., a man waving a balloon). Each scene was accompanied by narration containing either a novel noun or a novel verb (Figure 1). While different objects were used in each scene, all objects were from a single object category (e.g., all balloons). Thus, there were a number of potential referents for the novel word, including the action as well as the object category.

The contrast phase included one scene in which both the action and object were different from those seen during familiarization, labeled in the negative. This scene was followed by a second scene that duplicated a familiarization scene, labeled in the affirmative. At test, children were shown two novel scenes simultaneously and asked to point to the scene that represented the newly learned noun or verb. One scene showed the familiar action being performed on an exemplar from a novel object category (i.e., an object category not seen during familiarization); the other showed a novel action (i.e., an action not seen during familiarization) being performed on an exemplar from the familiar object category. The prediction was that children in the verb condition would point to the familiar action, while children in the noun condition would point to the novel action (i.e., the familiar object).

Participants

Participants were 28 normally-developing 3-year-olds (16 girls, 12 boys; $M = 42.5$ months, $SD = 2.72$) whose first language was English. Participants were recruited from

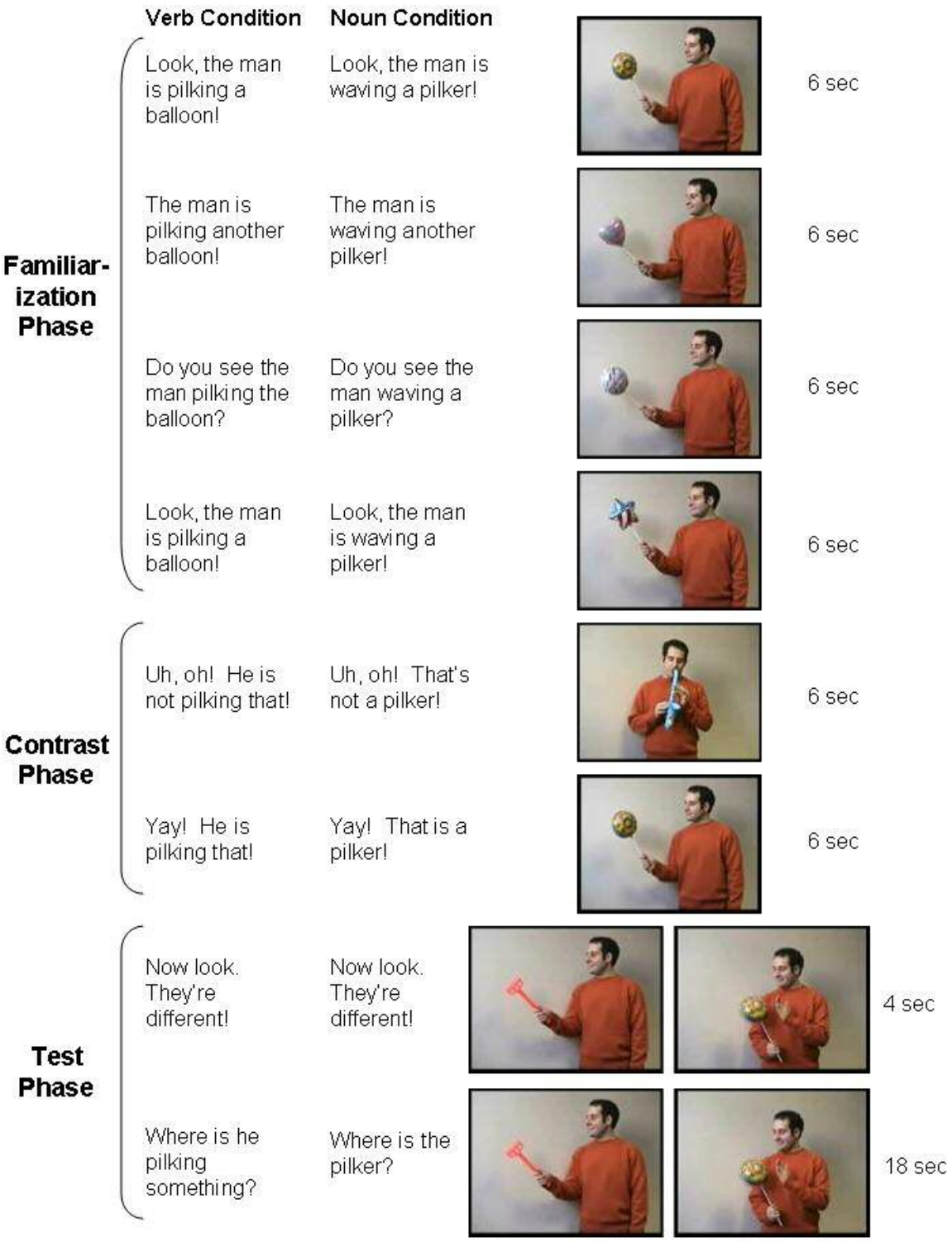

Figure 1. Study 1 Design

preschools and daycare centers or from the community. Half of the participants were recruited in a northern Chicago suburb and the other half were recruited in a western Philadelphia suburb. Children were randomly assigned to either the verb condition (n = 14; 7 girls, 7 boys; M = 42.5 months, SD = 2.77) or the noun condition (n = 14; 9 girls, 5 boys; M = 42.4, SD = 2.72).

Materials

Materials were the same as those used by Waxman et al. (under review), except that the test phase in each trial was modified to include two novel scenes, as described above. Scenes showed live actors performing simple, durative actions involving a variety of familiar objects (e.g., a man waving a balloon, a woman washing a cup) and were displayed on a computer monitor. Narration accompanying the audio contained the voice of an adult female, recorded in a sound-attenuated booth, using infant-directed speech to describe each scene using either a novel verb (verb condition) or a novel noun (noun condition). Children in both conditions watched the same video. A description of actions and objects used in all studies, and the corresponding novel verbs and nouns, is given in the Appendix.

Procedure

Children were seated in front of a computer screen in a quiet area of the laboratory or their school or daycare center. The experimenter sat on the child's right and controlled the video display using a keyboard or mouse placed out of the child's reach. When participants were brought to the laboratory accompanied by a parent, the parent sat behind the child, to the

left. The experimenter introduced the task by asking each child, "I have a pointing game that we can play. Are you a good pointer?" The experimenter then initiated the first of two training trials.

Training Phase. In each training trial, the child saw two video scenes simultaneously side-by-side on the computer screen and the experimenter asked the child to point to one. One training trial required the child to point to a familiar object (e.g., Big Bird or Elmo) and the other required the child to point to a familiar action (e.g., eating or dancing). The purpose of the training trials was to give the child the opportunity to become familiar with the task and to feel comfortable responding to the experimenter. Each trial was up to 30 seconds in duration, but ended immediately upon a correct response by the child. The experimenter began the video scene and prompted the child to point to one of the images, selected at random (e.g., "Can you point to Elmo? Where is Elmo? Can you point?"). To avoid ambiguous responses, the child was asked to touch the screen while pointing if he or she did not spontaneously do so. The experimenter continued to prompt the child until a correct response was given. If the child responded incorrectly, the experimenter asked the child to try again (e.g., "That's not Elmo! Try again. Can you point to Elmo?"). Once a correct response was given, the experimenter proceeded immediately to the second training trial. The process was repeated, with the experimenter selecting a target image that was on the opposite side of the screen as that selected in the first training trial. All children responded correctly in both training trials, and nearly every child responded correctly after a single prompt by the experimenter.

Familiarization Phase. Following training, each child participated in six experimental trials, each approximately one minute in duration. Each experimental trial began with a

familiarization phase, in which children watched four 6-second scenes in sequence, with a half-second pause between scenes. Each scene showed an actor performing a durative action on a different object, but all objects across familiarization scenes were from the same object category. Thus, in the first familiarization scene, the child might see a man waving a round orange balloon. In the second familiarization scene, the child might see the same man waving a silver star-shaped balloon, and so on. Each scene was accompanied by narration containing either a novel verb (verb condition; e.g., "Look, the man is pilking a balloon!") or a novel noun (noun condition; e.g., "Look, the man is waving a pilker!"). The same novel word was used in all four familiarization trials, but the sentences varied across scenes (e.g., "Look, the man is pilking a balloon! The man is pilking another balloon! Do you see the man pilking a balloon?").

Contrast Phase. The familiarization phase was immediately followed by the contrast phase, which consisted of two 6-second scenes separated by a half-second pause. The purpose of the contrast phase was to provide children with an opportunity to view a scene in which both the object and action were different from those seen during familiarization, and to hear the scene labeled in the negative using the novel word. Thus, in the first half of the contrast phase, children saw the familiar actor performing a new action with a new object and heard, for example, "Uh oh, he is *not* pilking that!" (verb condition) or "Uh, that is *not* a pilker!" (noun condition). The second scene was a repeat of one of the familiarization scenes, accompanied by "Yay, he *is* pilking that!" (verb condition) or "Yay, that *is* a pilker!" (noun condition).

Test Phase. In the test phase, children saw two scenes simultaneously, in which the action seen during familiarization was pitted against the object category. For example, if the

familiarization scenes showed a man waving a balloon, then one test scene showed the same man waving a rake (i.e., the familiar action with a novel object) and the other test scene showed the same man tapping a balloon (i.e., a novel action with the familiar object).

The test phase was divided into two stages. The purpose of the first stage, lasting four seconds, was simply to give the children a preview of the two test scenes. During this stage, children in both conditions heard, "Now look, they're different!" Following a half-second pause, the two test scenes were displayed again and the response period of the test phase began. In the verb condition, children heard, for example, "Where is he pilking something? Which one is he pilking?" In the noun condition, children heard "Where is the pilker? Which one is a pilker?" Following the recorded narration, the experimenter prompted the child, "Can you point?" This stage lasted up to 18 seconds with the child hearing the novel word up to six times. The experimenter continued to prompt until the child pointed to one of the scenes, at which time the experimenter said "Good job!" and proceeded immediately to the next trial. If the child failed to respond before the test phase ended, the experimenter simply proceeded to the next trial. However, most children responded after a single prompt by the experimenter. Because the test phase pitted the action seen during familiarization against the object category, any systematic response would demonstrate that the child had associated the novel word with either the familiar action or the familiar object category.

Results and Discussion

As predicted, children in the verb condition reliably mapped the novel word to the action, and children in the noun condition reliably mapped the novel word to the object

category. The dependent measure was the proportion of trials in which each child pointed to the test scene displaying the familiar action. A proportion of 1 would therefore indicate pointing to the familiar action on every test trial, a proportion of 0 would indicate pointing to the familiar object on every test trial, and a proportion of .5 would indicate chance performance. Consistent with prediction, children in the verb condition (M = .786, SD = .257) scored significantly above chance, $t(13) = 4.16$, $p = .0011$, two-tailed, and children in the noun condition (M = .119, SD = .178) scored significantly below chance, $t(13) = -8.00$, $p = .0000022$, two-tailed (Figure 2). The effect size, a relative index of treatment effect magnitude, was found to be $d_3' = 1.11$ for the former test and $d_3' = 2.14$ for the latter, both indicative of effects that are quite large (Cohen, 1988). An independent measures t test confirmed that the means of the two groups were significantly different from each other, $t(26) = 7.98$, $p = .000000019$, two-tailed. No significant effects of sex or location (i.e., Chicago vs. Philadelphia) were found.

Nonparametric tests of individual children's response patterns were consistent with the overall analysis. Children in the verb condition who pointed to the familiar action on four or more trials (out of six) were classified as learners, as were children in the noun condition who pointed to the familiar object on four or more trials. The probability of a participant exhibiting a learner response pattern through chance alone was .344. In the verb condition, 10 out of 14 children were learners, a proportion significantly different from chance, $p = .0080$, binomial test, two-tailed. In the noun condition, 13 out of 14 children were learners, a proportion significantly different from chance, $p = .0000089$, binomial test, two-tailed.

Children were classified as "inverse" learners if they pointed to the familiar object in the verb condition on four or more trials, or to the familiar action in the noun condition on four

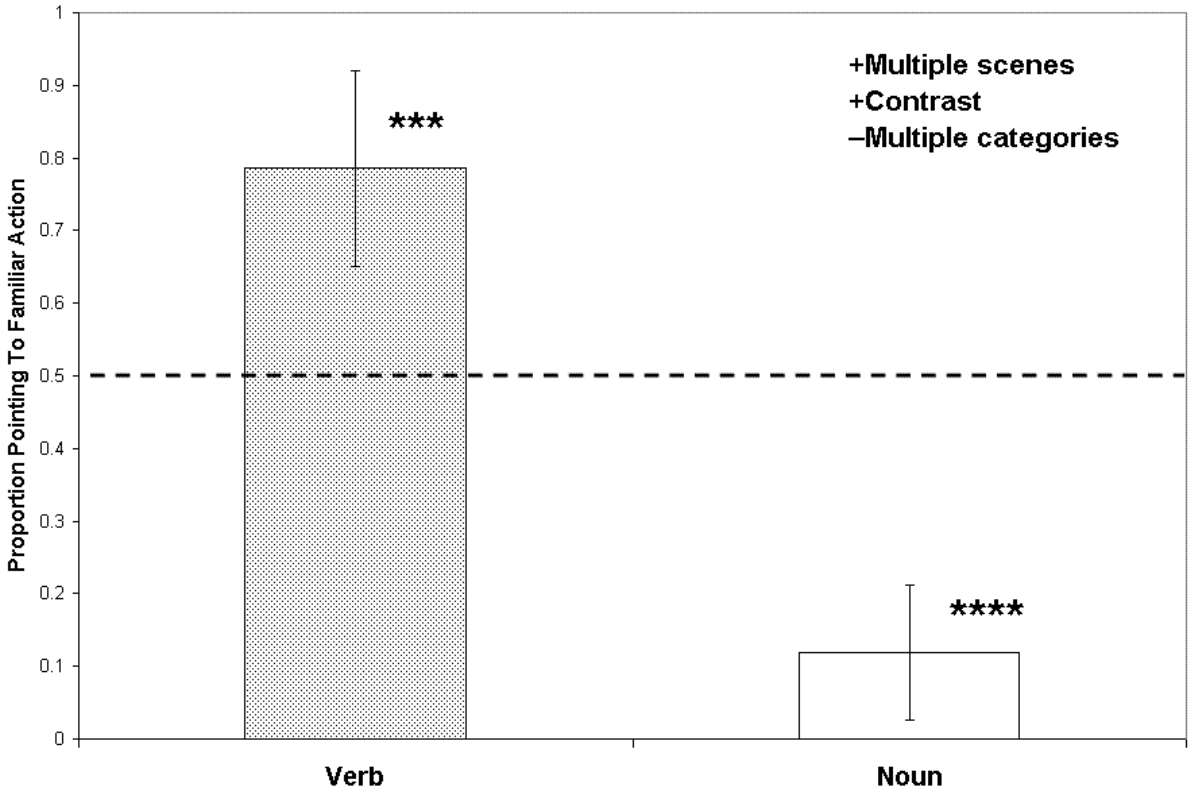

Figure 2. Study 1 Results (**** $p < .0001$; *** $p < .001$)

or more trials. There were two inverse learners in the verb group and none in the noun group; in neither case was this number greater than that expected by chance.

These results contribute to the body of research on early verb learning by demonstrating that 3-year-olds can "fast map" a novel verb to an action after watching a live action scene for 30 seconds and hearing the novel verb six times. Particularly noteworthy is the fact that these children succeeded when Imai et al. (2006) have reported that English-, Japanese-, and Chinese-speaking 3-year-olds all failed to learn verbs in a similar paradigm. This incongruity serves to underscore the importance of multiple familiarization scenes and explicit contrast in

verb learning, features of the current experimental design that are not present in many verb learning studies (including those of Imai et al.). In the next three experiments, I more closely examined the roles played by each of these factors in my verb learning task.

Study 2: The Value of Multiple Familiarization Scenes

The purpose of Study 2 was to determine whether exposure to multiple familiarization scenes is essential in order for children to learn a novel verb and successfully generalize it to a novel scene in this paradigm. The design was identical to that of Study 1, except that participants watched a single familiarization scene instead of four, and heard the novel word three times over a period of 18 seconds instead of six times over 36 seconds (Figure 3). If multiple familiarization scenes are instrumental for verb learning in this task, then the performance of children in the verb condition should suffer and their scores should be lower than those seen in Study 1. However, if multiple scenes are not necessary, then children should succeed as they did in Study 1. It was expected that children in the noun condition would be able to learn the novel noun.

Participants

Participants were 28 normally-developing 3-year-olds (16 girls, 12 boys; M = 42.3 months, SD = 2.78) whose first language was English. Participants were recruited in a northern Chicago suburb, from preschools and daycare centers or from the community. Children were randomly assigned to either the verb condition (n = 14; 9 girls, 5 boys; M = 42.3 months, SD = 2.89) or the noun condition (n = 14; 7 girls, 7 boys; M = 42.2 months, SD = 2.78).

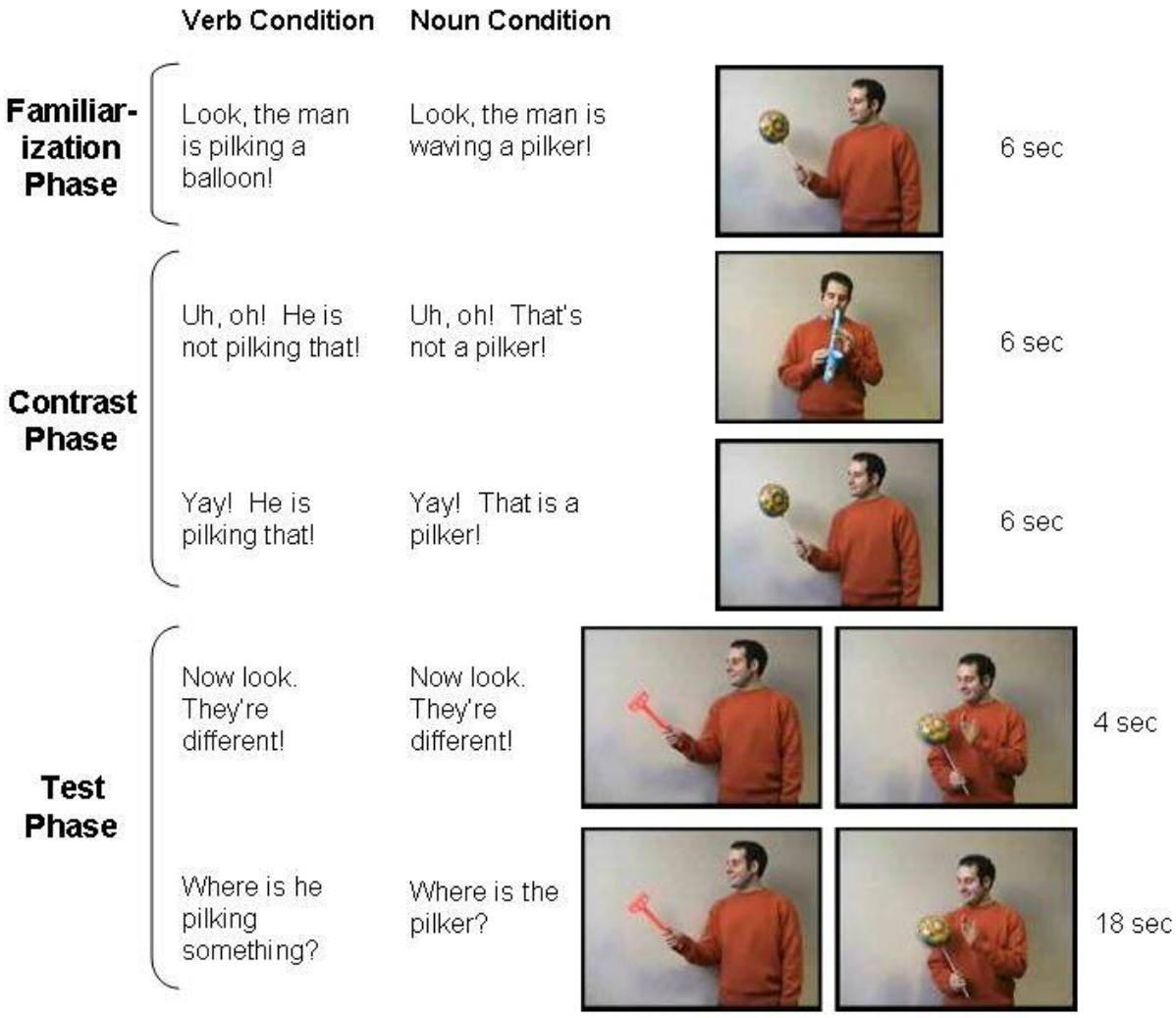

Figure 3. Study 2 Design

Materials

Materials were the same as those used in Study 1, modified as described below.

Procedure

The procedure was identical to that of Study 1, except that the familiarization phase comprised a single 6-second scene instead of four scenes, and included a single utterance of the novel word. The contrast phase and test phase were identical to those of Study 1.

Results and Discussion

As in Study 1, children learned both verbs and nouns. Children in the verb condition pointed to the familiar action more than 70% of the time (M = .712, SD = .177), a rate significantly higher than chance, $t(13) = 4.49$, $p = .00061$, two-tailed, $d_3' = 1.20$. Children in the noun condition pointed to the familiar action less than 20% of the time (M = .190, SD = .205), a rate significantly lower than chance, $t(13) = -5.64$, $p = .000080$, two-tailed, $d_3' = 1.51$ (Figure 4). Again, because the test phase pitted familiar action against familiar object, lower rates of pointing to the familiar action equated to higher rates of pointing to the familiar object. An independent measures t test confirmed that the means of the two groups were significantly different from each other, $t(26) = 7.20$, $p = .00000012$, two-tailed. No significant effect of sex was found.

Once again, nonparametric tests of individual children's response patterns were consistent with the overall analysis. In each of the verb and noun conditions, 12 out of 14 children were classified as learners, proportions significantly different from chance, $p = .00012$, binomial test, two-tailed. One child in each group was classified as an inverse learner, a number that was not greater than that expected by chance.

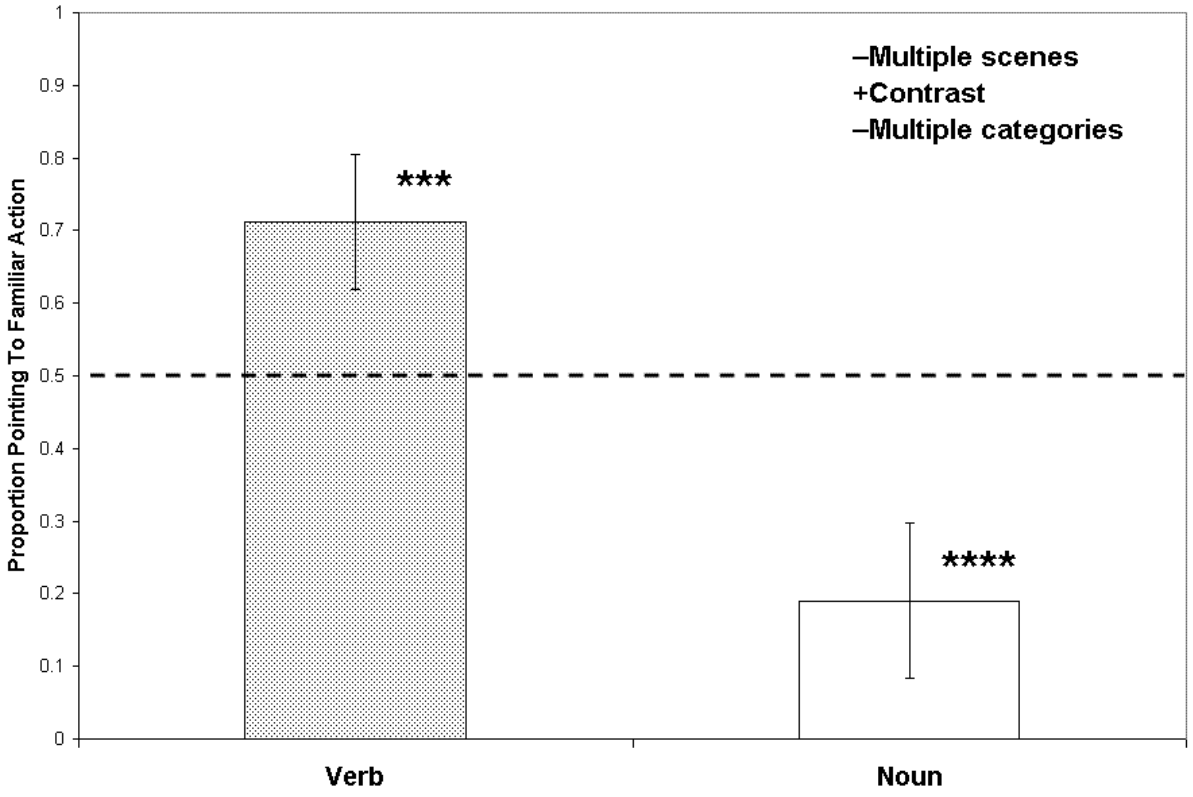

Figure 4. Study 2 Results (**** $p < .0001$; *** $p < .001$)

This is a remarkable finding, in light of other work that has found children of this age to be unable to learn novel verbs in a similar experimental setting (e.g., Imai et al., 2005). In the current study, children saw an action for a total of 12 seconds (six seconds during the familiarization phase and six seconds during the latter half of the contrast phase) and heard the novel verb just three times. Despite this limited exposure, at test the children were able to reliably select a novel instance of the action that incorporated a novel object. Was it the presence of the contrast phase that allowed the children to succeed? In Study 3, I eliminated

the contrast phase and retained a single familiarization scene, while holding constant the total trial duration and number of exposures to the novel word.

Study 3: The Value of Contrast

The purpose of Study 3 was to determine the effect of explicit contrast on children's ability to learn a novel verb for an action and generalize it to a new scene involving a new object category. The design was similar to that of Study 2, except that the contrast phase was eliminated and the single familiarization scene was extended to 18 seconds, with three exposures to the novel word (Figure 5). As with the previous two studies, there were three possibilities for children in each group: they could map the word to the action seen during familiarization (scoring above chance), they could map the word to the object (scoring below chance), or they could fail to make any systematic mapping at all (scoring at chance). The prediction was that without the benefit of either multiple familiarization scenes or contrast, participants would fail to map a novel verb to an action. However, it was expected that participants would still be able to map a novel noun to an object.

Participants

Participants were 28 normally-developing 3-year-olds (15 girls, 13 boys; M = 42.3 months, SD = 2.90) whose first language was English. Participants were recruited in a northern Chicago suburb, from preschools and daycare centers or from the community. Children were randomly assigned to either the verb condition (n = 14; 7 girls, 7 boys; M = 43.0 months, SD = 3.09) or the noun condition (n = 14; 8 girls, 6 boys; M = 41.6 months, SD = 2.62).

Figure 5. Study 3 Design

Materials

Materials were the same as those used in Study 1, modified as described below.

Procedure

The procedure was identical to that of Study 1, except that each trial comprised only a familiarization phase and a test phase. The familiarization phase included a single scene lasting

18 seconds accompanied by three utterances of the novel word (e.g., "Look, the man is pilking a balloon! Do you see the man pilking a balloon? Look, the man is pilking a balloon!"). The test phase was identical to that of Study 1.

Results and Discussion

In this experiment, children watched a single familiarization scene lasting 18 seconds, in which an actor performed a durative action with an object (e.g., a man waved a balloon). During the course of the familiarization scene, the children heard the scene labeled three times with a novel noun or novel verb. As predicted, and consistent with findings reported by other researchers, 3-year-olds failed to learn verbs under these conditions, pointing to the familiar action at test only 56% of the time (M = .560, SD = .311), a rate no different from chance, $t(13)$ = .717, p = .49, two-tailed (Figure 6). In contrast, children were successful at learning nouns, pointing to the familiar object 69% of the time—equivalent to pointing to the familiar action at a rate significantly less than chance (M = .310, SD = .284), $t(13)$ = -2.51, p = .026, two-tailed, $d_3' = .669$. An independent measures t test confirmed that the means of the two groups were significantly different from each other, $t(26)$ = 2.22, p = .035, two-tailed. No significant effect of sex was found.

Nonparametric tests of individual children's response patterns were consistent with the overall analysis. In the verb condition, only 7 out of 14 children were classified as learners, a proportion that was no different from chance, p = .26, binomial test, two-tailed. In the noun condition, however, 10 out of 14 children were learners, a proportion significantly greater than

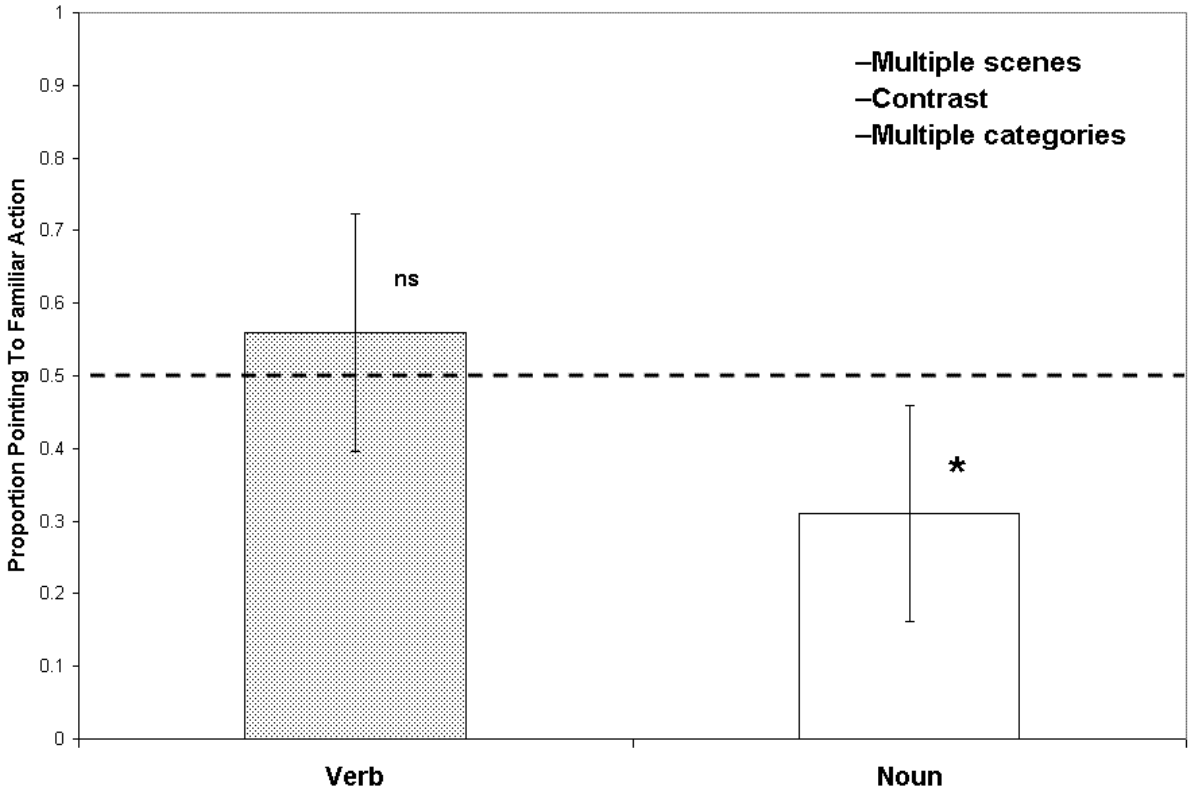

Figure 6. Study 3 Results (* $p < .05$)

chance, $p = .0080$, binomial test, two-tailed. There were 4 inverse learners in the verb group and 2 in the noun group; neither of these numbers was greater than that expected by chance.

Although the overall mean score for children in the verb condition was not different from chance, it may have been the case that the children's performance improved over the course of the six trials, as was the case in Waxman et al. (under review). In order to investigate this possibility, I next divided the six trials into two blocks of three trials and performed a separate analysis on each block. The results were consistent with the overall mean, however. Children in the verb condition performed comparably in the first block ($M = .548$, $SD = .310$)

and second block (M = .571, SD = .380), and in neither case was their score different from chance.

Thus, when deprived of both multiple familiarization scenes and contrast, children failed to learn verbs. Taken together, the results of Studies 1 through 3 may suggest that contrast alone, without multiple familiarization scenes, is sufficient for verb learning in this paradigm. However, it should be noted that the contrast phase itself includes multiple scenes. That is, there is no way within the current design to show both a valid referent and an invalid referent for a novel word without showing multiple scenes. Still, it is unlikely that the multiple scenes presented in the contrast phase of Study 2 provided information that was comparable to that provided by the multiple familiarization scenes of Study 1. This is because in Study 2, the contrast phase included the same scene as that seen during familiarization (Figure 3) and therefore offered no additional positive information as to the referent of the novel word.

Still unanswered is the question of whether multiple familiarization scenes without contrast would be sufficient to allow verb learning in this task. Study 4 was designed to answer this question.

Study 4: The Value Of Multiple Familiarization Scenes, Revisited

The purpose of Study 4 was to determine whether children will successfully learn verbs through exposure to multiple familiarization scenes without the benefit of contrast. The design was the same as that of Study 1, except that the contrast phase was eliminated. Thus, children saw four separate familiarization scenes and heard the novel word four times over a period of

24 seconds (Figure 7). As in Studies 1 through 3, it was expected that children in the noun condition would be able to learn the novel noun.

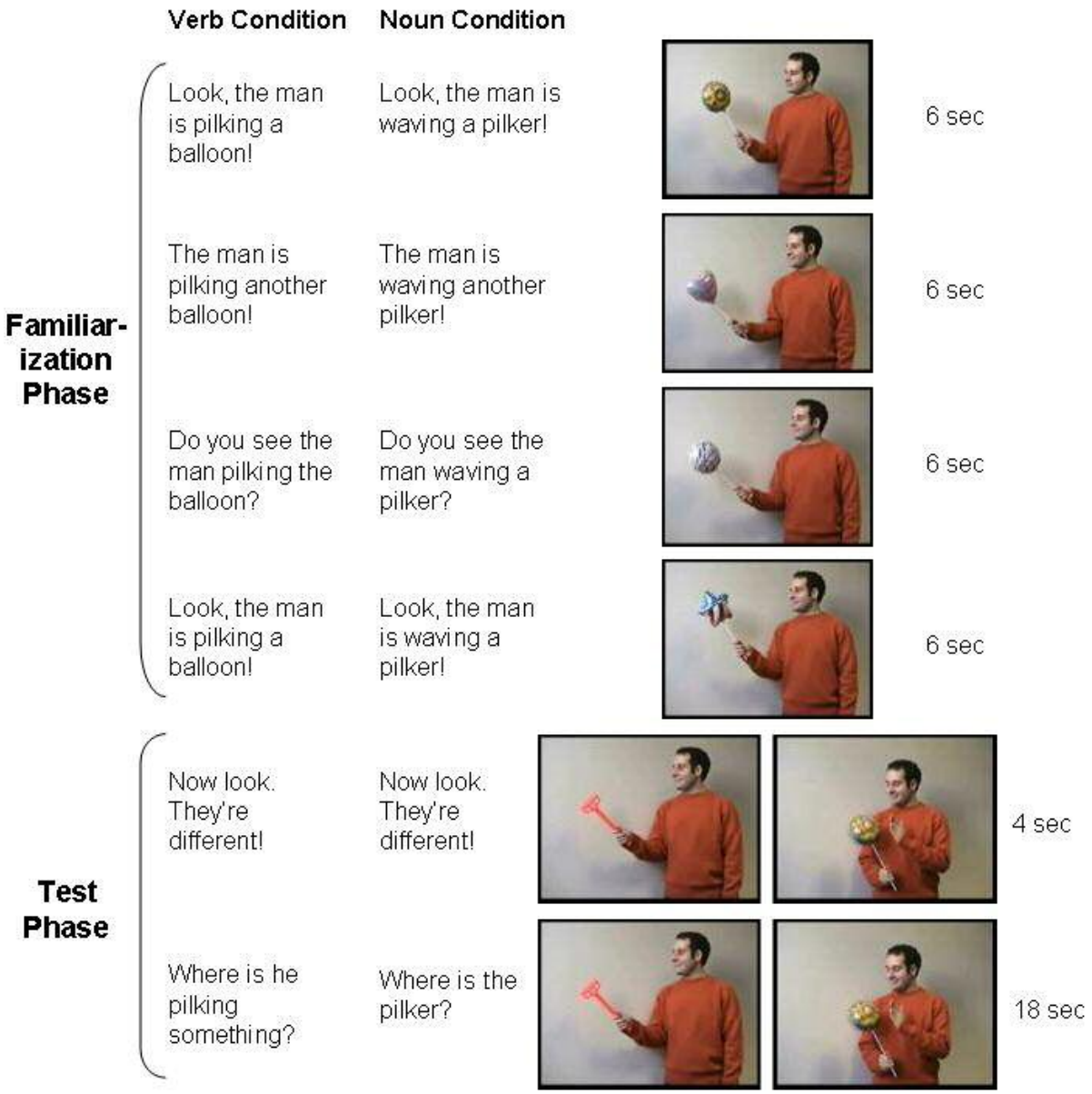

Figure 7. Study 4 Design

Participants

Participants were 28 normally-developing 3-year-olds (14 girls, 14 boys; $M = 41.1$ months, $SD = 2.70$) whose first language was English. Participants were recruited from preschools and daycare centers or from the community. Half of the participants were recruited in a northern Chicago suburb and the other half were recruited in a western Philadelphia suburb. Children were randomly assigned to either the verb condition ($n = 14$; 6 girls, 8 boys; $M = 41.2$ months, $SD = 2.67$) or the noun condition ($n = 14$; 8 girls, 6 boys; $M = 41.1$ months, $SD = 2.84$).

Materials

Materials were the same as those used in Study 1.

Procedure

The procedure was identical to that of Study 1, except that the contrast phase was eliminated.

Results and Discussion

In this experiment, children watched four different familiarization scenes and heard either a novel verb or a novel noun four times over a period of 24 seconds. No contrast phase was included. As in Study 3, the overall score of children in the verb condition ($M = .629$, $SD = .285$) was not different from chance, $t(13) = 1.69$, $p = .11$, two-tailed, but the overall score of children in the noun condition ($M = .214$, $SD = .240$) was significantly below chance, $t(13) =$

-4.46, $p = .00064$, two-tailed, $d_3' = 1.19$ (Figure 8). An independent measures t test revealed that the means of the two groups were significantly different from each other, $t(26) = 4.17$, $p = .00030$, two-tailed. No significant effects of sex or location (i.e., Chicago vs. Philadelphia) were found. Thus, on the basis of overall scores, children learned nouns but they did not learn verbs.

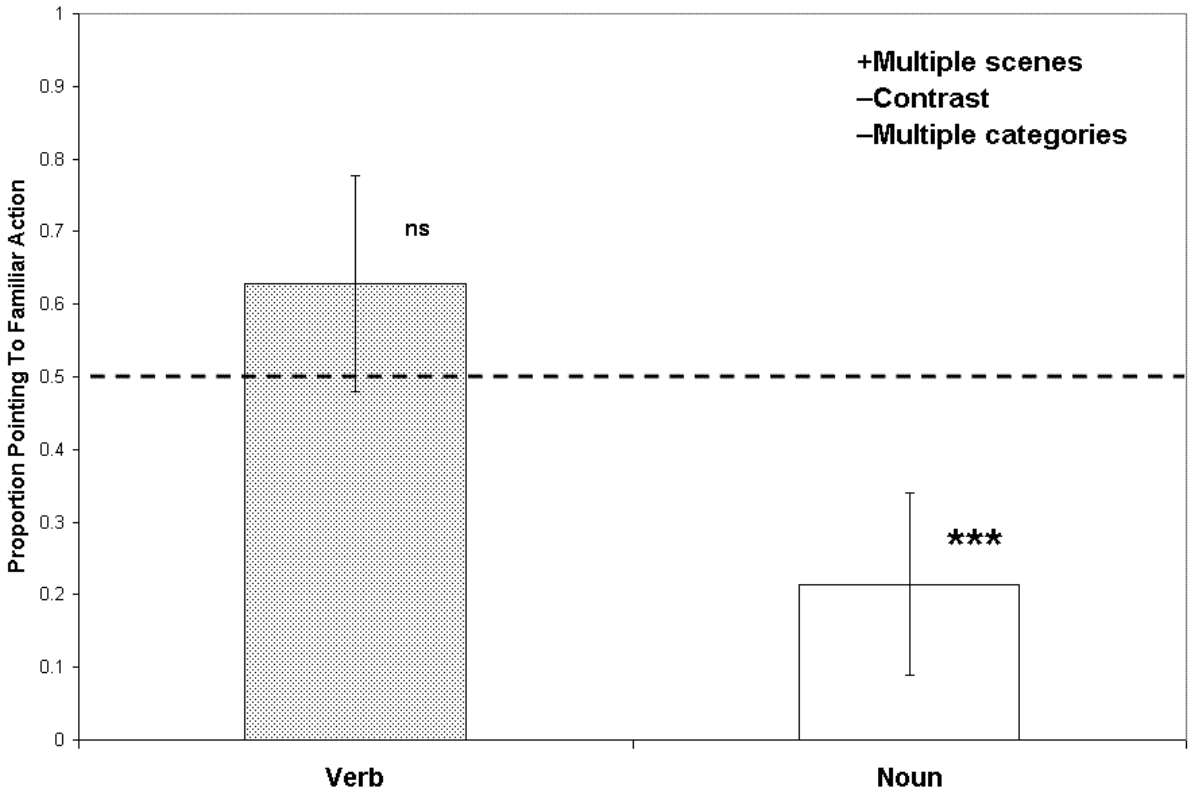

Figure 8. Study 4 Results—Overall (*** $p < .001$)

However, an analysis of children's performance in the first block of trials as compared to the second block revealed an interesting trend. In the verb condition, the children's mean

score *M* increased from .571 (*SD* = .356) in block 1 to .667 (*SD* = .226) in block 2. While the block 1 mean was not different from chance, $t(13) = .75$, $p = .47$, two-tailed, the block 2 mean was significantly greater than chance, $t(13) = 2.75$, $p = .016$, two-tailed, $d_3' = .74$ (Figure 9). This finding suggests that either the children were learning over the course of the experiment, or the procedure was sufficiently demanding that the children needed time to settle down and focus on the task. In either case, by the last three trials they were able to reliably associate the novel verb with the action seen during familiarization and not the object. In the noun condition, children's responses in block 1 (*M* = .238, *SD* = .305) and block 2 (*M* = .190, *SD* = .252) were comparable. In both blocks, the mean score was significantly below chance.

Once again, individual children's response patterns were consistent with the mean scores. In the verb condition, 8 out of 14 children were classified as learners, a proportion that differed from chance with marginal significance, $p = .091$, binomial test, two-tailed. This marginal result would be expected if children's performance started at chance and then improved over the course of six trials, as suggested by the previous analysis. In the noun condition, 11 out of 14 children were learners, a proportion significantly greater than chance, $p = .00093$, binomial test, two-tailed. There were 3 inverse learners in the verb group and 1 in the noun group; neither of these numbers was greater than that expected by chance.

The results of Study 4 demonstrate that given multiple familiarization scenes, children can learn novel verbs and nouns without the benefit of explicit contrast. This finding is probably not surprising; the lack of availability of negative evidence is widely accepted as a given in the realm of language learning (Marcus, 1993), and it must be granted that children almost certainly learn words all the time without the benefit of explicit contrast. Still, learning

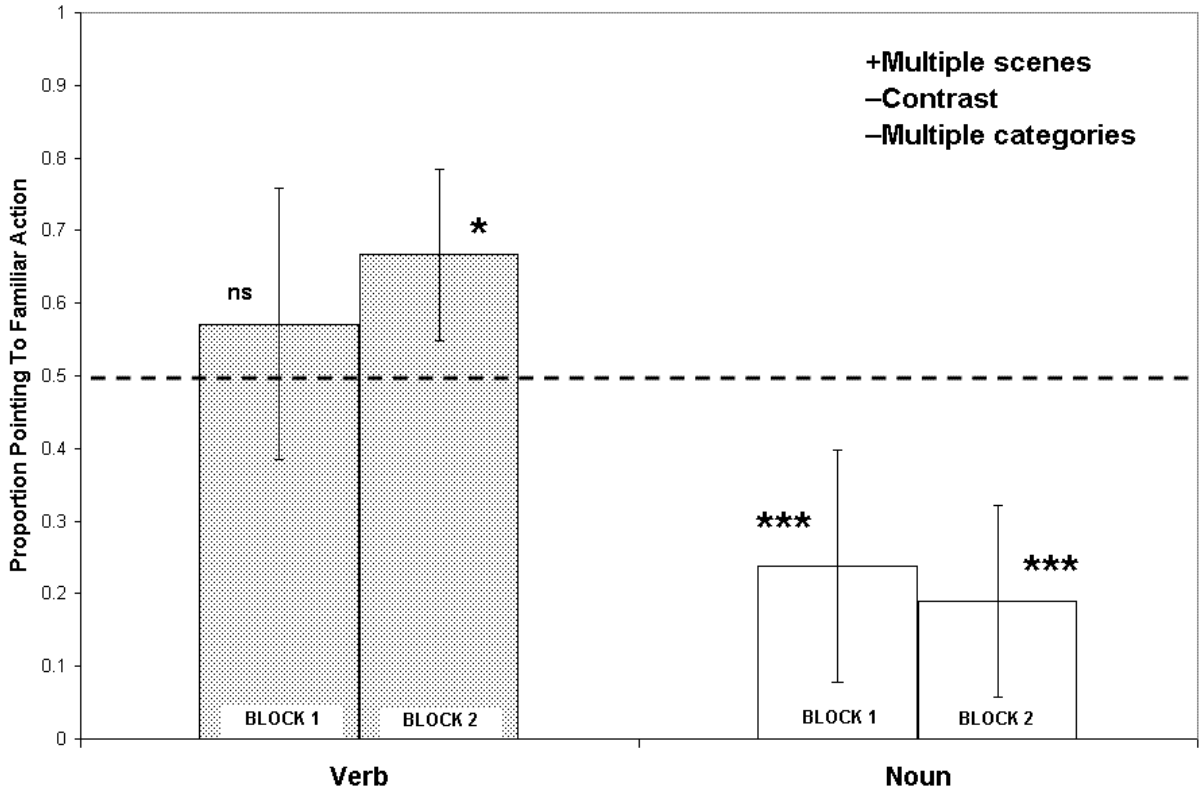

Figure 9. Study 4 Results—Blocks (*** $p < .001$; * $p < .05$)

verbs under these conditions was difficult for the children; it was not until the second block of trials that the children's responses became significantly different from chance. It seems likely that four familiarization scenes lasting a total of 24 seconds is much less exposure than children get when they learn a verb naturalistically.

Comparing Results Across Studies

The results of Studies 1 through 4 are summarized in Figure 10. As is evident from the figure, children learned both nouns and verbs most effectively when provided with multiple

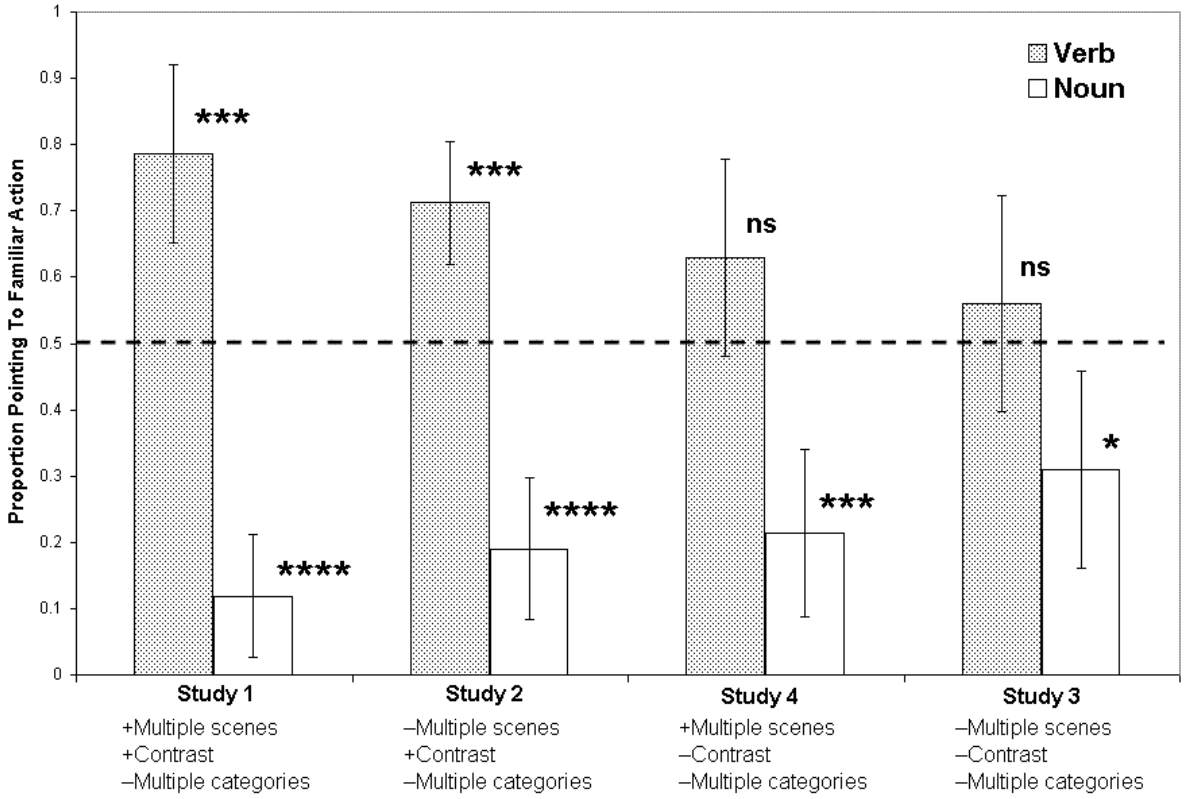

Figure 10. Results for Studies 1 through 4 (**** $p < .0001$; *** $p < .001$; * $p < .05$)

familiarization scenes supplemented with explicit contrast. The removal of either one of these sources of information impaired both types of word learning. Children found it particularly challenging to learn verbs without the benefit of contrast, even when multiple scenes were available. Given neither multiple scenes nor contrast, children could not learn verbs at all in this task. However, children reliably learned nouns even when neither multiple scenes nor contrast were available.

A (2) condition × (4) study independent measures analysis of variance (ANOVA) on the proportion of test trials in which children pointed to the familiar action revealed a main effect

of condition, $F(1, 104) = 98.7$, $p < .001$, as well as an interaction effect, $F(3, 104) = 3.55$, $p = .017$, but no main effect of study, $F(3, 104) = .1$, $p = .96$. The main effect of condition indicates that across all four studies, scores for children in the verb condition differed significantly from those for children in the noun condition. Pairwise comparisons confirmed that within each study, mean verb scores were greater than mean noun scores, $p < .001$ for Studies 1, 2, and 4, and $p < .01$ for Study 3.

In order to explore this condition × study interaction, we can compare results across these first four studies and examine the contributions of individual variables. For example, Study 1 differed from Study 4 only in terms of a contrast phase. With contrast (Study 1), children readily learned verbs, but without contrast (Study 4), children only began to learn verbs in the second block of trials. However, this comparison is slightly problematic because the contrast phase lasted 12 seconds and included two utterances of the novel verb. Thus, the argument could be made that it was this additional exposure to the novel word that facilitated learning, rather than the use of contrast.

A more meaningful way to examine the effect of contrast might be to compare Study 2 to Study 3. In this case, neither study included multiple familiarization scenes. However, between these two studies, the total trial duration and the number of exposures to the novel word were held constant, allowing for a direct comparison. With the benefit of contrast (Study 2), children learned verbs, but without contrast (Study 3), children performed at chance.

In order to examine the effect of multiple familiarization scenes, we can compare Study 3 (–multiple scenes) to Study 4 (+multiple scenes). Once again, however, this comparison is problematic, because Study 4 has a longer trial duration and more exposures to the novel word.

Additional control conditions were therefore created for Studies 3 and 4 in order to allow this type of comparison. First, the "verb-condensed" condition was created for Study 4. As with the verb condition of Study 4, the design of this condition was +multiple scenes, –contrast, and –multiple categories. However, the verb-condensed condition included only three exposures to the novel word over a period of 18 seconds, comparable to Study 3. Thus, a direct comparison between Study 3 (verb condition) and Study 4 (verb-condensed condition) was possible.

Study 4-VC: Verb-Condensed Condition

Participants

Participants were 14 normally-developing 3-year-olds (7 girls, 7 boys; $M = 41.4$ months, $SD = 3.06$) whose first language was English. Participants were recruited in a northern Chicago suburb, from preschools and daycare centers or from the community.

Materials

Materials were the same as those used in Study 4, except that the familiarization phase included three scenes rather than four (Figure 11).

Procedure

The procedure was identical to that of Study 4, except that there was only one condition, the verb-condensed condition.

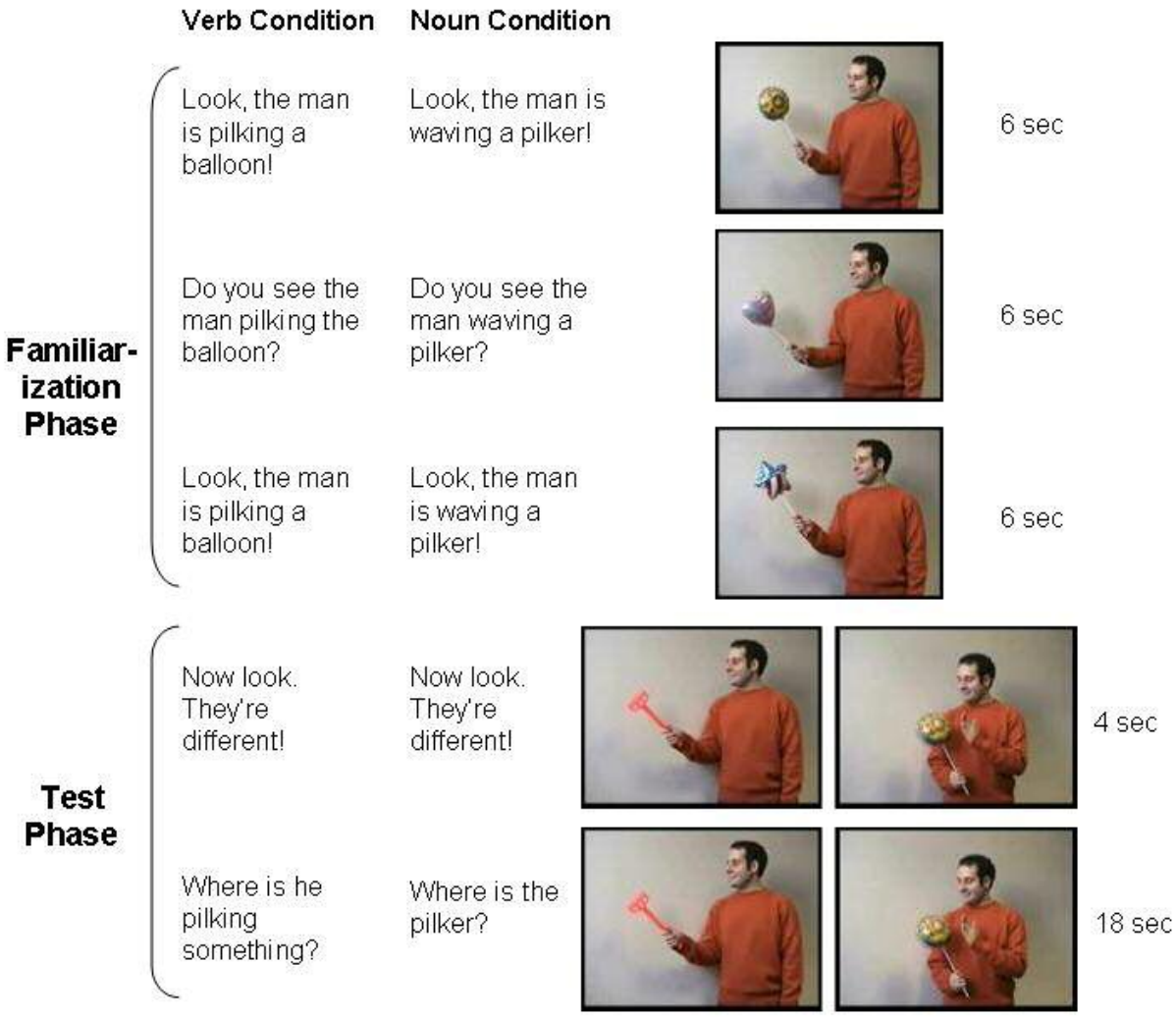

Figure 11. Study 4-VC Design

Results and Discussion

In this study, children saw three familiarization scenes and heard the novel verb three times over a period of 18 seconds. Three-year-olds failed to learn verbs under these conditions, pointing to the familiar action at test only 49% of the time (M = .488, SD = .265), a rate no

different from chance, $t(13) = -.168$, $p = .87$, two-tailed. A comparison of children's performance during the first block of trials to that during the second block showed no change over time. Children performed comparably in the first block ($M = .500$, $SD = .285$) and second block ($M = .476$, $SD = .386$), and in neither case was their score different from chance. No significant effect of sex was found.

Nonparametric tests of individual children's response patterns were consistent with the overall analysis. Eight out of 14 children were classified as learners and 5 were classified as inverse learners. Neither proportion was significantly different from chance, $p = .091$ and $p = .18$, respectively, binomial test, two-tailed.

The results of Study 4-VC are shown in Figure 12, along with the results of Study 3 for comparison. The designs of the two studies were identical, except that Study 3 included a single familiarization scene and Study 4-VC contained multiple scenes. The mean verb-learning scores of the two studies did not differ from each other, $t(26) = .654$, $p = .52$, two-tailed, and in neither case was the children's performance different from chance. This finding indicates that without the benefit of contrast, and given an extremely brief exposure to a novel verb—three utterances over a period of 18 seconds—children's learning was not facilitated by exposure to multiple familiarization scenes. This is not to say that exposure to multiple scenes is not important for verb learning, but rather that by itself such exposure is not sufficient, given the extremely taxing learning conditions with which the children were faced in this task.

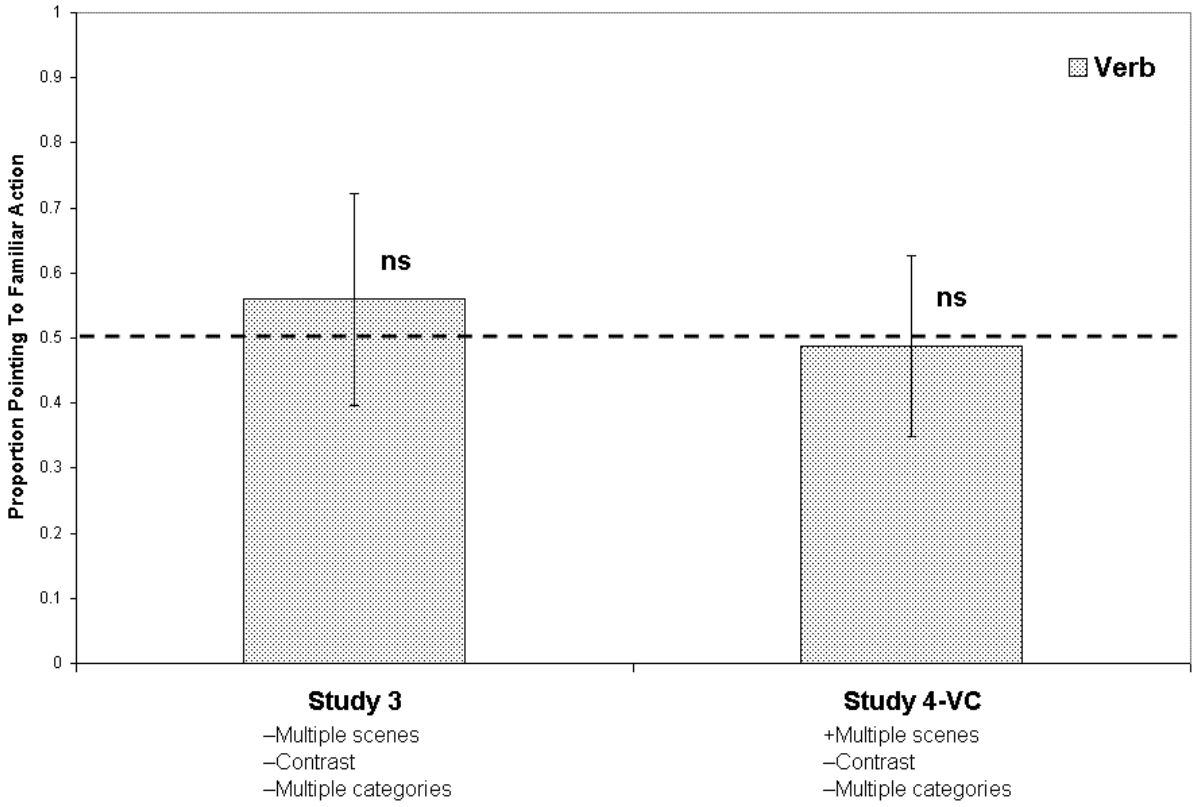

Figure 12. Study 4-VC Results

Study 3-VE: Verb-Extended Condition

The second control condition created was the "verb-extended" condition for Study 3. This condition had the same design as the verb condition but included four exposures to the novel verb over a period of 24 seconds. This allowed for another direct comparison of Study 3 (verb extended condition) to Study 4 (verb condition), in order to examine the effect of multiple familiarization scenes on verb learning in this task.

Participants

Participants were 14 normally-developing 3-year-olds (6 girls, 8 boys; $M = 42.4$ months, $SD = 2.34$) whose first language was English. Participants were recruited in a northern Chicago suburb, from preschools and daycare centers or from the community.

Materials

Materials were the same as those used in Study 3, except that the single familiarization scene was extended to include four exposures to the novel verb over a period of 24 seconds (Figure 13).

Procedure

The procedure was identical to that of Study 3, except that there was only one condition, the verb-extended condition.

Results and Discussion

Like those in Study 3, children in this study were not exposed to either multiple familiarization scenes or contrast. Nevertheless, although children failed to learn verbs in Study 3, children succeeded in the current study, apparently due to the increased duration of the single familiarization scene. Children pointed to the familiar action in response to the novel verb 72% of the time ($M = .720$, $SD = .211$), a proportion significantly greater than that predicted by chance, $t(13) = 3.91$, $p = .0018$, two-tailed, $d_3' = 1.04$. No significant effect of sex was found. A comparison of children's performance over time revealed that although their

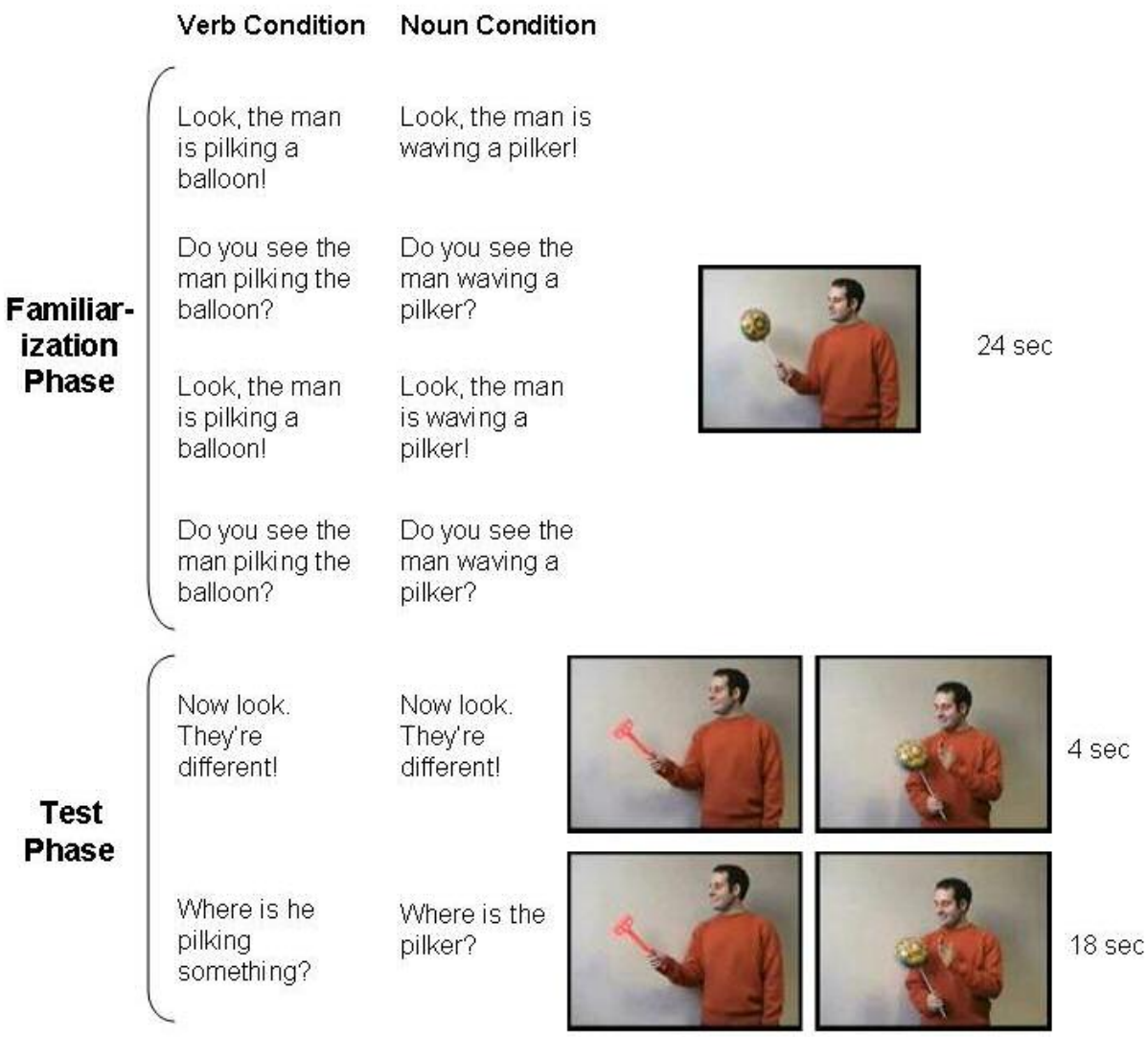

Figure 13. Study 3-VE Design

mean score was only marginally significant in block 1 (M = .643, SD = .276), $t(26) = 1.93$, p = .064, two-tailed, by block 2 their mean score had increased markedly (M = .762, SD = .242), $t(26) = 4.05$, p = .00041, two-tailed. Thus, in the absence of multiple familiarization scenes and

contrast, children were able to learn a novel verb—albeit with some difficulty—after hearing four utterances of the word over 24 seconds (Study 3-VE), but not after hearing three utterances over 18 seconds (Study 3). This finding highlights the fact that the minimal amount of information offered to children in this paradigm taxed them to the extreme limits of their word learning capabilities.

Nonparametric tests of individual children's response patterns were once again consistent with the overall analysis. Eleven out of 14 children were classified as learners, p = .00093, binomial test, two-tailed.

The results of Study 3-VE are shown in Figure 14, along with the results of Study 4 for comparison. The designs of the two studies were identical, except that Study 3-VE included a single familiarization scene and Study 4 contained multiple scenes. The mean verb-learning scores of the two studies did not differ from each other, $t(26) = .968$, $p = .34$, two-tailed. Somewhat paradoxically, the overall mean score for Study 3-VE (–multiple scenes) was greater than chance, but the overall mean score for Study 4 (+multiple scenes) was not (although note that this score did reach significance by the second block of trials).

Is it reasonable to conclude that contrary to prediction, exposure to multiple familiarization scenes actually hindered verb learning in this task, as these findings suggest? I believe so. Consider that in this task, children were being asked to view a series of scenes in which an action was performed with an object, and then later to generalize the verb that they heard to the same action performed with a different object. Seeing several different scenes all involving the same kind of object might have led the children to conclude that the meaning of the verb was linked to the object, a conclusion that may have made them less likely to

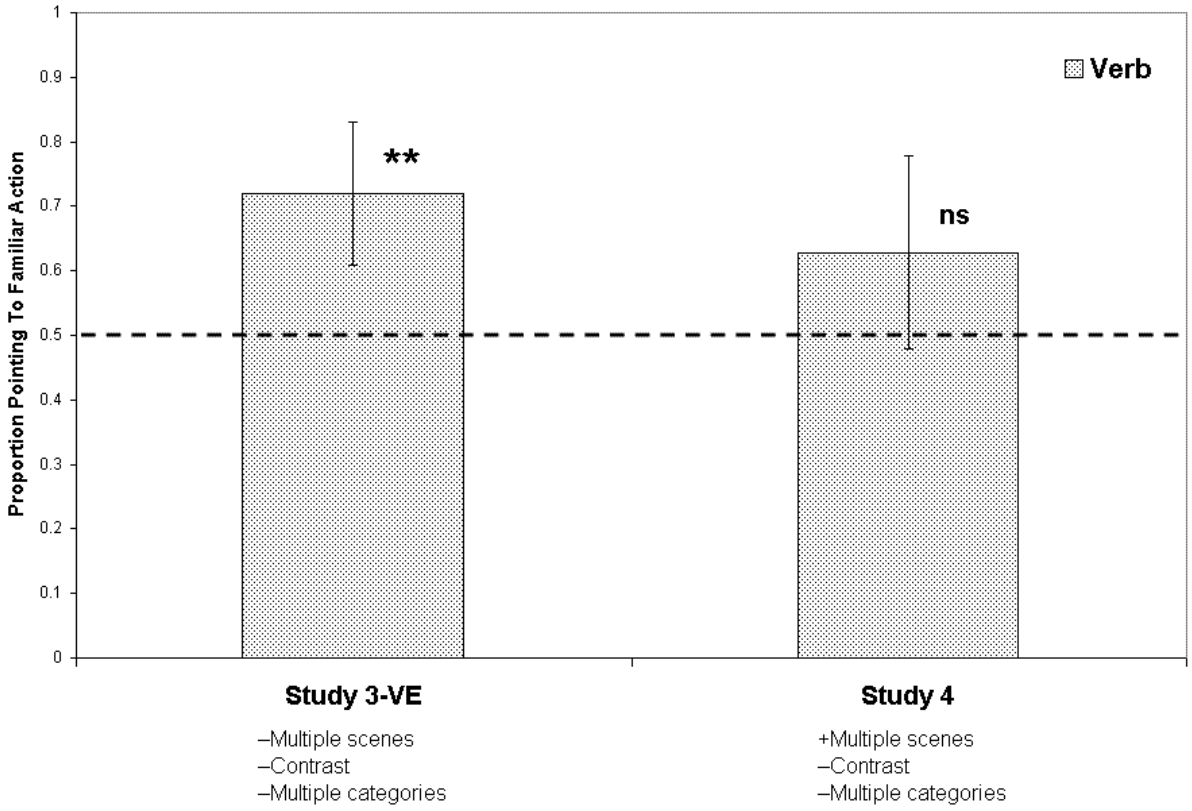

Figure 14. Study 3-VE Results (** $p < .01$)

generalize the verb successfully to a new object at test. On the other hand, a single scene (involving a single object) may have offered the children less evidence that the meaning of the verb was associated with the object, allowing the children to be more liberal with their generalizations. If this is true, then children should be more successful in this task if objects from different categories are used across familiarization scenes. This prediction was tested in Studies 5 and 6.

Study 5: The Value of Multiple Object Categories

All of the studies discussed thus far have incorporated objects from a single object category into the familiarization phase. With the final two studies, I investigated the influence of using objects from different categories. There is reason to suspect that with younger infants, the use of a single object category might facilitate verb learning in that it would reduce the cognitive load on the learner; less attention paid to the objects would mean that more attention could be paid to the action (Kersten, Smith, & Yoshida, 2006). However, given the greater cognitive capacities of 3-year-olds, it is possible that a greater variety of objects across familiarization scenes might help these older children to rule out the objects as potential referents of the novel verb. That is, to the extent that a learner seeks commonalities across scenes to "zero in" on a correct word meaning, greater variability of objects would serve to highlight the fact that the action—but not the object—remains constant across scenes. Thus, my hypothesis was that the use of a variety of objects would facilitate verb learning in this task.

The design of Study 5 was the same as that of Study 1, except that objects from different categories were used in the familiarization scenes (Figure 15). In the verb condition, children heard the novel word presented in the same sentence structure across scenes (e.g., "The man is pilking a hammer," "The man is pilking a carrot," etc.). Because the use of multiple object categories precluded the use of the same novel noun across scenes, Study 5 included only a verb condition.

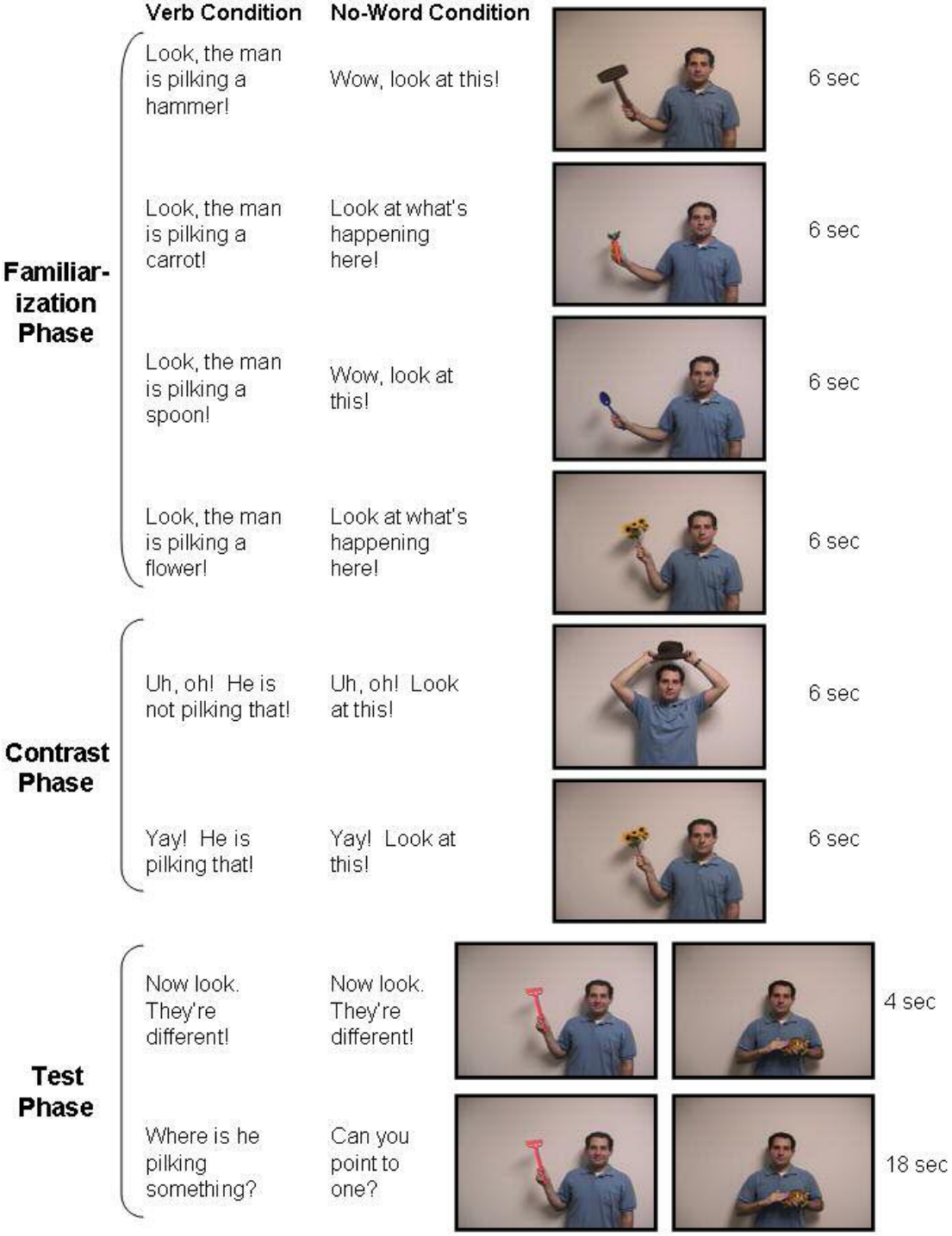

Figure 15. Study 5 Design

Participants

Participants were 14 normally-developing 3-year-olds (10 girls, 4 boys; $M = 41.3$ months, $SD = 2.49$) whose first language was English. Participants were recruited from preschools and daycare centers or from the community. Half of the participants were recruited in a northern Chicago suburb and the other half were recruited in a western Philadelphia suburb.

Materials

Materials were similar to those used in Study 1, except that four different objects were used across the four familiarization scenes. During the test phase, both scenes showed the familiar actor manipulating a different novel object. One test scene showed the action seen during familiarization, and the other showed a novel action.

Procedure

The procedure was identical to that of Study 1, except that there was only a verb condition.

Results and Discussion

As predicted, children reliably mapped the novel verb to the familiar action at a rate significantly higher than chance ($M = .774$, $SD = .192$), $t(13) = 5.34$, $p = .00013$, two-tailed, $d_3' = 1.43$. No significant effects of sex or location (i.e., Chicago vs. Philadelphia) were found.

Individual children's response patterns were consistent with the aggregate analysis. Twelve out of 14 children were classified as learners, a proportion significantly different from chance, $p = .00012$, binomial test, two-tailed. No children were classified as inverse learners.

Thus, given the benefit of multiple familiarization scenes, contrast, and multiple object categories—arguably the combination of factors that would be most likely to facilitate verb learning in this task—children did, in fact, readily learn novel verbs. But what was the relative contribution of the addition of multiple object categories, as compared to the other two variables? We saw in Study 4 that children could learn verbs without the benefit of contrast, but with some difficulty, as evidenced by the fact that they succeeded only in later trials. Might the use of multiple object categories rather than a single category compensate for the lack of contrast and make verb learning easier for the children in this scenario? Study 6 addressed this question.

Study 6: The Value of Contrast, Revisited

The design of Study 6 was identical to that of Study 5, except that the contrast phase was eliminated, and identical to that of Study 4, except that multiple object categories were used during familiarization (Figure 16). Study 6 therefore offers opportunities for comparison with previous studies along two separate dimensions. As in Study 5, children saw two novel objects at test; one appeared in conjunction with the action seen during familiarization, and the other with a novel action. Study 6 contained only a verb condition.

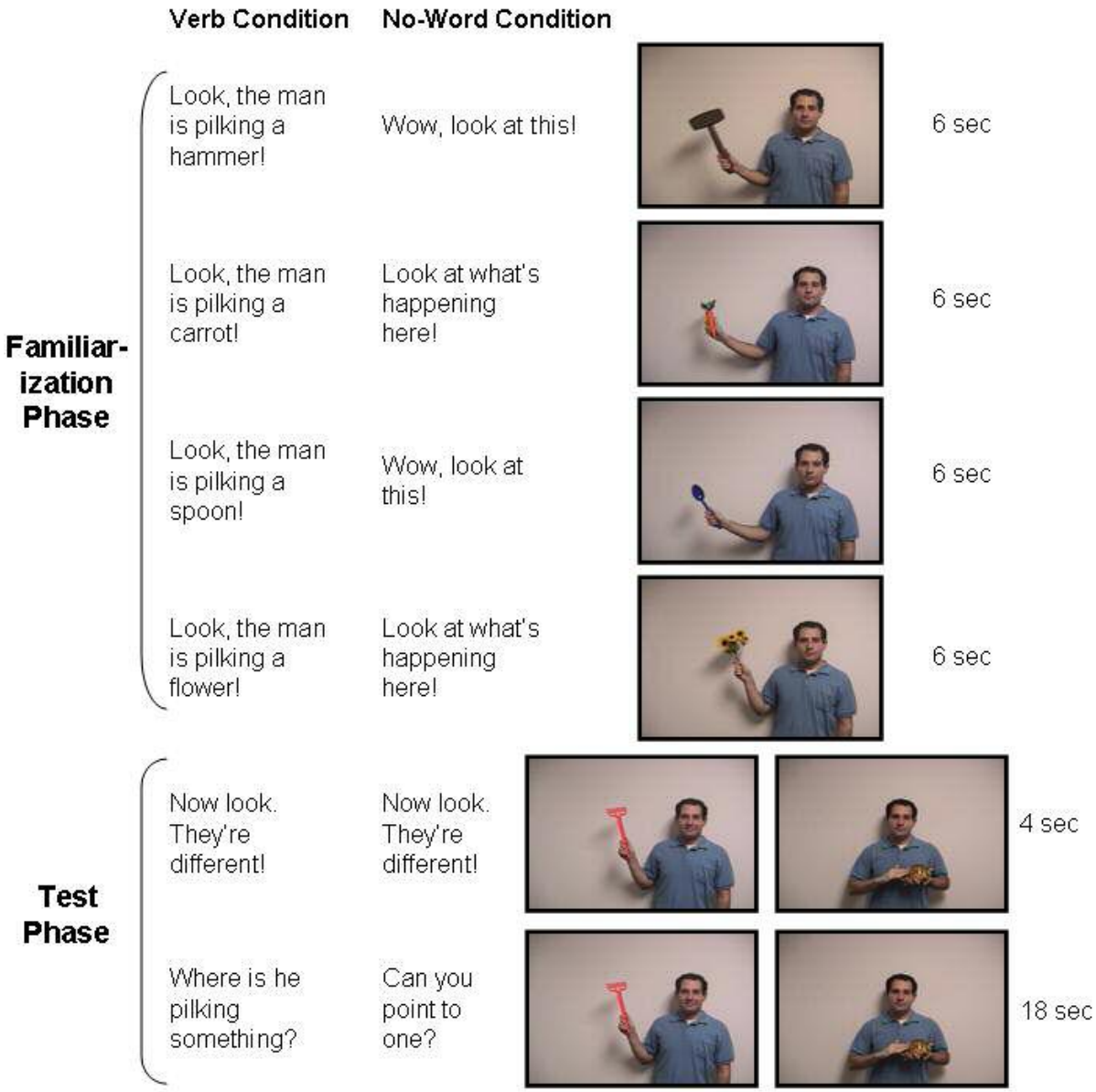

Figure 16. Study 6 Design

Participants

Participants were 14 normally-developing 3-year-olds (7 girls, 7 boys; M = 40.8 months, SD = 2.67) whose first language was English. Participants were recruited in a northern Chicago suburb, from preschools and daycare centers or from the community.

Materials

Materials were the same as those used in Study 5, modified as described below.

Procedure

The procedure was identical to that of Study 5, except that the contrast phase was eliminated.

Results and Discussion

In this final study, I investigated the ability of children to learn a novel action verb given four different familiarization scenes, wherein the referent action was performed using objects from four different object categories, but without the benefit of contrast. Children heard the action labeled with the novel word four times, once per familiarization scene. Children in this study reliably mapped the novel verb to the familiar action at a rate significantly higher than chance (M = .750, SD = .193), $t(13) = 4.84$, p = .00032, two-tailed, $d_3' = 1.30$. No significant effect of sex was found.

Individual children's response patterns were consistent with the aggregate analysis. Eleven out of 14 children were classified as learners, a proportion significantly different from chance, $p = .00093$, binomial test, two-tailed. No children were classified as inverse learners.

Mean verb-learning scores for children in Studies 5 and 6 are shown in Figure 17. Children's performance did not differ between the two studies, $t(26) = .327$, $p = .75$, two-tailed.

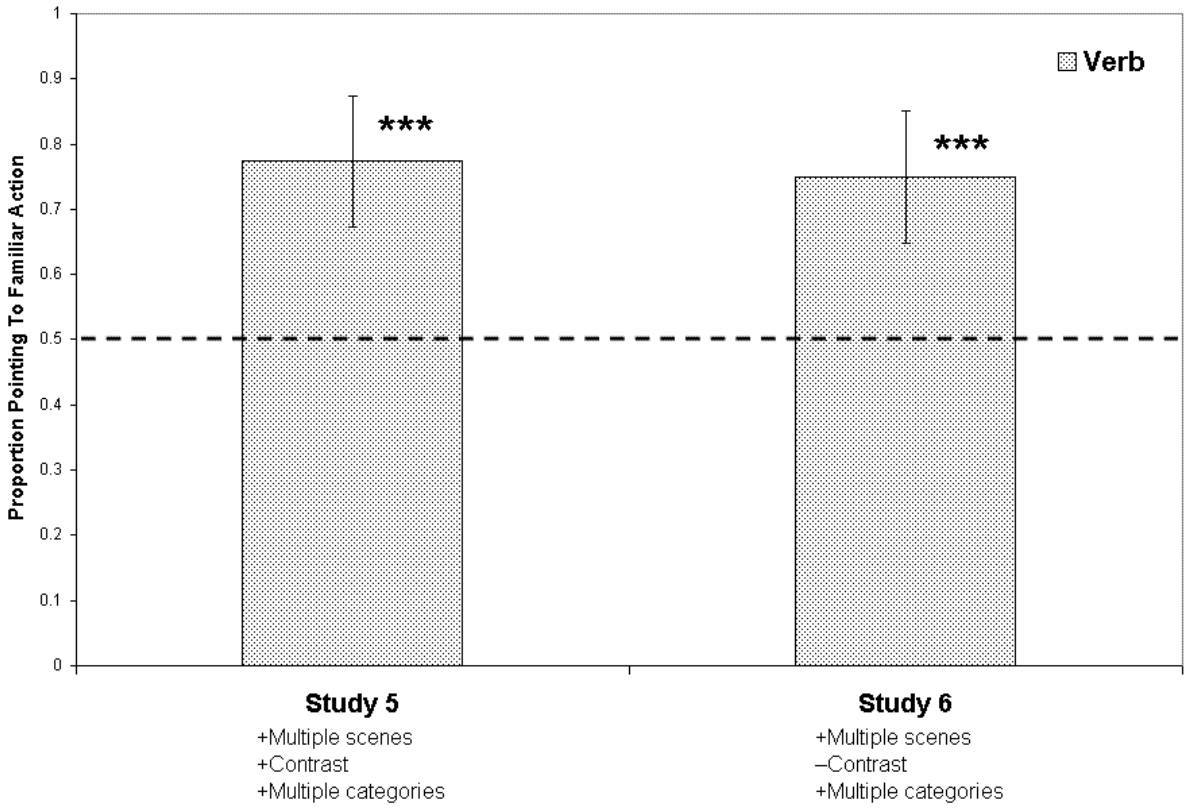

Figure 17. Results for Studies 5 and 6

As previously discussed, the results of Studies 2 and 3 point to the importance of contrast in verb learning. However, the results of Study 6 suggest that given sufficient variety

in the input, explicit contrast is not necessary for learning. That is, the use of multiple object categories across familiarization scenes may have bolstered learning by providing additional information as to the word's meaning, just as the presence of contrast apparently plays a similar role.

Conclusion

The current series of experiments explored the effects of multiple familiarization scenes, explicit contrast, and multiple object categories on 3-year-olds' learning of novel verbs. In the first four studies, only multiple scenes and contrast were manipulated, while a single object category was used for each verb. In Study 1, when children had the benefit of both multiple familiarization scenes and explicit contrast, they readily mapped a novel verb to an action and then generalized the verb to a new scene in which the action was performed with a new object. In Study 2, children saw only a single familiarization scene in conjunction with contrast and were still able to successfully learn the novel verb. In Study 3, children saw a single familiarization scene without contrast and were unable to learn verbs. In Study 4, children saw multiple familiarization scenes and no contrast. In this case, children did successfully learn verbs, but only by the second block of trials. In all four studies, children reliably mapped a novel noun to an object.

In the last two studies, multiple object categories were used across multiple familiarization trials. Study 5 included contrast; Study 6 did not. In both cases, children readily learned novel verbs.

The results of these experiments help to shed some light on the surprising reports of toddlers' inability to learn novel verbs in the laboratory. While young children have shown an impressive ability to "fast map" object labels as early as age 13 months (Woodward, Markman, & Fitzsimmons, 1994), some researchers have reported that children as old as 4 years have been unable to learn a novel action verb in a laboratory setting (e.g., Childers & Tomasello, 2002; Imai et al., 2005). These failures are puzzling in light of evidence that (a) children much younger than 4 years old routinely acquire and produce action verbs in a naturalistic setting, and (b) even young infants have demonstrated an ability to represent the semantic components that underlie many common action verbs (Pulverman et al., 2006).

How can we explain toddlers' difficulties at learning novel verbs in the laboratory? One approach is to examine the differences in informational requirements between nouns and verbs. Verbs tend to be more variable than nouns in terms of the kinds of semantic components that they encode and the way that these components are combined (Talmy, 1975; Gentner & Boroditsky, 2001), and their meanings are therefore more dependent on exposures across varying contexts. Simply put, learning a verb requires more variety in the input than does learning a noun. As is illustrated by the current work, a single familiarization scene—such as that offered in some previous verb learning studies—may simply be insufficient information for a child to learn a novel verb in some cases, regardless of the duration of the scene or the number of times that the word is uttered. In fact, one of the messages that is clear from the current work is that several short scenes may potentially provide much more information than one long scene. This empirical evidence is consistent with theoretical arguments that posit both

conceptual- and linguistic-based reasons that verb acquisition is delayed relative to noun learning.

The usefulness of contrast in verb learning can be seen clearly by comparing the results of Studies 2 and 3. In Study 2, children saw one familiarization scene plus a contrast phase. All told, the trial duration leading up to the test phase was 18 seconds (although the referent action was visible for only 12 seconds). Children heard the novel verb three times over the course of these 18 seconds. Despite this limited exposure, children reliably associated the verb with the action and not the object. In Study 3, children saw a single familiarization scene without contrast. The familiarization scene lasted 18 seconds, and once again, the children heard the novel verb three times. Although the children heard the novel verb the same number of times as in Study 2, and saw the referent action for a full 18 seconds, they nevertheless failed to learn verbs in Study 3. This is a dramatic illustration of the importance of contextual variety in verb learning.

Additional evidence for the value of contrast in verb learning comes from an examination of the results of Studies 1 and 4. Both studies incorporated multiple familiarization scenes, but Study 1 included a contrast phase and Study 4 did not. Children learned verbs easily in Study 1, but had much more difficulty in Study 4, as indicated by the fact that their performance did not increase significantly above chance levels until the second block of trials.

Although explicit contrast was clearly helpful in this verb learning task, it was not necessary for success. As mentioned, children did eventually learn verbs in Study 4 aided by multiple familiarization scenes, without exposure to contrast. Furthermore, children easily

learned verbs in Study 6, without contrast but with multiple scenes and multiple object categories. Contrast therefore appears to be one way—but not the only way—of providing the necessary contextual variety to facilitate children's verb learning in this task.

The contribution of multiple familiarization scenes to verb learning in this task was less clear. It was originally intended that a comparison of the results of Study 3 (–multiple scenes) to those of Study 4 (+multiple scenes) would offer an indication of the importance of cross-situational exposure to verb learning. Neither study included contrast. Children were unable to learn verbs in Study 3, but succeeded by the second block of trials in Study 4, suggesting that multiple scenes may have facilitated learning. As discussed previously, however, this comparison is somewhat problematic, because the total trial duration and number of exposures to the novel word were not controlled for. That is, children saw the referent action for 18 seconds and heard the verb three times in Study 3, but saw the action for 24 seconds and heard the verb four times in Study 4. This was the motivation for the additional control conditions included in Studies 4-VC and 3-VE.

It is clear from a comparison of Study 3 to Study 4-VC that exposure to multiple scenes is not by itself sufficient for verb learning. In both studies, children heard the novel word three times over a period of 18 seconds, with no contrast phase. Study 3 included multiple scenes and Study 4-VC did not, but children's performance was equivalent to chance in both cases. Thus, given this extremely limited duration of exposure, and with no contrast phase to provide additional cues to the word's meaning, children were unable to learn verbs, even when presented with multiple familiarization scenes.

The results of Study 3-VE suggest that even without multiple scenes, children can learn a verb if they hear the verb enough times over a long enough duration. In this study, children heard a verb four times over a period of 24 seconds and were able to learn it successfully. This finding was contrary to prediction, given that children failed in Study 3, after hearing the verb three times over a period of 18 seconds. It may be surprising that such a subtle difference would lead to such a definitively different result. However, one cannot conclude from this result that in general, children will learn a new word after hearing it four times but not after hearing it only three times. The goal of the task, after all, was not to simulate real-world conditions, but to examine the effects of a limited number of factors in a highly controlled task with an extremely restricted set of potential verb referents. In naturalistic settings, the meanings of some words may be acquired slowly, over the course of many months or even years (Gropen, Pinker, Hollander, & Goldberg, 1991; Bloom, 2004).

The fact that children learned verbs in Study 3-VE (without multiple scenes) but had much more difficulty in Study 4 (with multiple scenes) indicates that the value of multiple scenes depends on the specific nature of the scenes. For example, scenes that utilized objects from different object categories across scenes were more useful to the children in this task than those utilizing objects from a single category. The fact that the multiple scenes in this task were artificially restricted in scope and contextual variety may have made them less effective than would be the case in a natural learning environment. It is also important to note that the factor that was most effective in facilitating verb learning—contrast—itself relied on multiple scenes. In other words, the contrast phase by definition incorporated multiple scenes. It is

therefore difficult to examine the influence of multiple scenes in this task independently of any other factors.

Studies 5 and 6 demonstrate the value of multiple object categories in verb learning. When a variety of object categories was used across familiarization scenes, children learned verbs just as easily without contrast (Study 6) as they did with contrast (Study 5). In the absence of contrast, children had difficulty learning verbs when shown a single object category (Study 4), but were more successful when shown multiple object categories (Study 6).

Taken together, the results of the current studies suggest that multiple familiarization scenes, contrast, and multiple object categories all contributed, to varying degrees, to verb learning in this paradigm. Children learned verbs best when all sources of information were available. When only one was available, contrast appeared to be the most useful. Children deprived of all three sources of information did not learn verbs at all, but did learn nouns.

It is important to note that in the current series of experiments, children were faced with the task of learning a new word and then extending it to a new situation. This learning measure is in some ways more informative than a standard comprehension or production test. In a typical comprehension test, for example, children are shown an array of referents and asked to select the one that corresponds to the novel word. In a typical production test, children are shown the referent and asked to produce the novel word. Rather than asking children to select or label a referent that they have already seen, the extension task asks children to decide which of two novel scenes could correctly be described by the novel word. This goes to the core of the word's meaning by requiring children to make an inference about which semantic components of the scene are relevant to the meaning of the word.

It should also be noted that while previous verb learning studies have often used "novel" actions and objects, I chose to use familiar (but not necessarily easily nameable) actions in conjunction with familiar objects. I did this partly to ensure that the children's word mapping was not hindered by unnatural novel actions that may have been more complex or artificial than those encountered in a naturalistic environment. Also, in light of evidence that children learn novel verbs more easily in the context of familiar objects than in the context of unfamiliar objects (Kersten & Smith, 2002; Kersten et al., 2006), I felt that the use of familiar objects might facilitate verb learning in that it would reduce the cognitive load on the learner; less attention paid to the objects would mean that more attention could be paid to the action (Kersten et al., 2006). The fact that the objects I used were easily nameable but the actions were not may have created an imbalance between the ease of verb learning and noun learning in these studies. However, in such a case, mutual exclusivity would be expected to impair noun learning but not verb learning.

To summarize, the current research offers valuable new evidence that the information requirements for successful verb learning differ markedly from those for noun learning. Specifically, exposure to a new verb across a variety of contexts is essential in order to allow a child to identify those aspects of a scene which are relevant to the verb's meaning, to rule out those which are not, and to establish the necessary verb-to-world mapping. While these findings extend verb learning research by tapping into learning mechanisms (i.e., cross-situational observation, contrast, and multiple object categories) that have not typically been addressed in other verb learning studies, these studies are just the beginning. Pathways for future research built on these studies are already clear. For example, to what extent does

children's conceptual knowledge of the objects involved facilitate verb learning? Under what conditions will children extend a new verb to a novel actor? Most importantly, perhaps, what roles do the variables studied here play in the acquisition of less-concrete verbs such as *love*, *play*, and *want*? It has become increasingly clear that the informational requirements for verb learning and noun learning are dramatically different, and the goal for future verb-learning research must be to more precisely characterize the nature of these requirements at different developmental points and to examine the ways in which children's conceptual development equips them to extract and utilize this information in the service of verb acquisition.

REFERENCES

Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology, 14*, 79-93.

Anglin, J.M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58* (10, Serial No. 238).

Antinucci, F., & Miller, R. (1976). How children talk about what happened. *Journal of Child Language, 3*, 167-189.

Aslin, R.N. (1993). Segmentation of fluent speech into words: Learning models and the role of maternal input. In B. de Boysson-Bardies, S. de Schonen, P.W. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 305-315). New York: Kluwer Academic.

Au, T.K., Dapretto, M., & Song, Y.K. (1994). Input versus constraints: Early word acquisition in Korean and English. *Journal of Memory and Language, 33*, 567-582.

Au, T.K., & Markman, E.M. (1987). Acquiring word meanings via linguistic contrast. *Cognitive Development, 2*, 217-236.

Baillargeon, R. (2000). How do infants learn about the physical world? In D. Muir & A. Slater (Eds.), *Infant development: The essential readings* (pp. 195-212). Malden, MA: Blackwell Publishing.

Baldwin, D.A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*, 875-890.

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.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6*, 183-200.

Billman, D., & Davila, D. (2001). Consistent contrast aids concept learning. *Memory & Cognition, 29*, 1022-1035.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, and Computers, 33*, 73-79.

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: The importance of being imageable. *Journal of Neurolinguistics, 16*, 113-149.

Blackwell, A.A. (2005). Acquiring the English adjective lexicon: Relationships with input properties and adjectival semantic typology. *Journal of Child Language, 32*, 535-562.

Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague: Mouton.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bloom, P. (2004). Myths of word learning. In D.G. Hall & S.R. Waxman (Eds.), *Weaving a lexicon* (pp. 205-224). Cambridge, MA: MIT Press.

Bloom, P., & Markson, L. (1997, April). *Children's naming of representations*. Poster session presented at the biennial meeting of the Society for Research in Child Development, Washington, DC.

Booth, A.E., & Waxman, S.R. (2003). Mapping words to the world in infancy: Infants' expectations for count nouns and adjectives. *Journal of Cognition and Development, 4*, 357-381.

Bornstein, M.H., Cote, L.R., Maital, S., Painter, K., Park, S., Pascual, L., Pecheux, M.G., Ruel, J., Venuti, P., & Vyt, A. (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. (1980). The structure and origin of semantic categories in the language learning child. In M.L. Foster & S.H. Brandes (Eds.), *Symbol as a sense: New approaches to the analysis of meaning* (pp. 277-299). New York: Academic Press.

Brent, M.R., & Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81*, B33-B44.

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

Casasola, M., Cohen, L.B., & Chiarello, E. (2003). Six-month-old infants' categorization of containment spatial relations. *Child Development, 74*, 679-693.

Casasola, M., Hohenstein, J., & Naigles, L. (2003, April). Ten-month-olds infants' discrimination of manner and path in motion events. In M. Casasola (Chair), *From infancy to adulthood: Exploring the effect of linguistic input on the discrimination of manner and path in motion events*. Symposium conducted at the biennial meeting of the Society for Research on Child Development, Tampa, FL.

Caselli, M.C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development, 10*, 159-199.

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.

Choi, S. (2000). Caregiver input in English and Korean: Use of nouns and verbs in book-reading and toy-play contexts. *Journal of Child Language, 27*, 69-96.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22*, 497-529.

Clark, E.V. (1988). On the logic of contrast. *Journal of Child Language, 15*, 317-335.

Clark, E.V. (1990). On the pragmatics of contrast. *Journal of Child Language, 17*, 417-431.

Clark, E.V. (1991). Acquisitional principles in lexical development. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 31-71). New York: Cambridge University Press.

Clark, E.V. (1992). Conventionality and contrast: Pragmatic principles with lexical consequences. In A. Lehrer & E.F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 171-188). Hillsdale, NJ: Lawrence Erlbaum.

Clark, E.V. (1996). Learning to do things with verbs. *Contemporary Psychology: A Journal of Reviews, 41*, 833-835.

Clark, E.V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology, 41*, 451-463.

Dietrich, R. (1985). *Nouns and verbs in the learner's lexicon*. Paper presented at the Workshop on European Second Language Acquisition Research, Linguistic Institute, Georgetown University, Washington, DC.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (5, Serial No. 242).

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D.J., Bates, E., Hartung, J.P., Pethick, S.J., & Reilly, J.S. (1991). *Technical manual for the MacArthur communicative development inventories*. San Diego, CA: San Diego State University.

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens, & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 262-282). New York: Cambridge University Press.

Fisher, C., & Gleitman, L.R. (2002). Language acquisition. In R. Gallistel (Ed.), *Stevens' Handbook of Experimental Psychology, Vol. 3: Learning, motivation, and emotion* (pp. 445-496). New York: John Wiley & Sons.

Forbes, J.N., & Farrar, M.J. (1993). Children's initial assumptions about the meaning of novel motion verbs: Biased and conservative? *Cognitive Development, 8*, 273-290.

Forbes, J.N., & Farrar, M.J. (1995). Learning to represent word meaning: What initial training events reveal about 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.

Frawley, W. (1992). *Linguistic semantics*. Hillsdale, NJ: Lawrence Erlbaum.

Gelman, S.A., & Ebeling, K.S. (1998). Shape and representational status in children's early naming. *Cognition, 66*, 835-847.

Gentner, D. (1975). Evidence for the psychological reality of semantic components: The verbs of possession. In D.A. Norman, D.E. Rumelhart, & The LNR Research Group (Eds.), *Explorations in cognition* (pp. 211-246). San Francisco: Freeman.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development, 49*, 988-998.

Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory, 4*, 161-178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development, Vol. 2: Language, thought, and culture* (pp. 301-334). Hillsdale, NJ: Lawrence Erlbaum.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544-564). New York: Oxford University Press.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S.C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215-256). Cambridge: Cambridge University Press.

Gentner, D., & Flusberg, S. (in preparation). The role of relational language in the development of analogical reasoning.

Gentner, D., & Rattermann, M.J. (1991). Language and the career of similarity. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225-277). New York: Cambridge University Press.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*, 135-176.

Girouard, P.C., Ricard, M., & Decarie, T.G. (1997). The acquisition of personal pronouns in French-speaking and English-speaking children. *Journal of Child Language, 24*, 311-326.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3-35.

Gleitman, L.R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J.C. (2005). Hard words. *Language Learning and Development, 1*, 23-64.

Goldin-Meadow, S., Seligman, M.E., & Gelman, R. (1976). Language in the two-year-old. *Cognition, 4*, 189-202.

Golinkoff, R.M., Chung, H.L., Hirsh-Pasek, K., Liu, J., Bertenthal, B.I., Brand, R.J., Maguire, M.J., & Hennon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology, 38*, 604-614.

Golinkoff, R.M., & Hirsh-Pasek, K. (2006). Introduction: Progress on the verb learning front. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 3-28). New York: Oxford University Press.

Golinkoff, R.M., Hirsh-Pasek, K., Mervis, C.B., Frawley, W.B., & Parillo, M. (1995). Lexical principles can be extended to the acquisition of verbs. In M. Tomasello & W.E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 185-221). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Golinkoff, R.M., Mervis, C.B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21*, 125-155.

Goodman, N. (1983). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.

Gopnik, A., & Meltzoff, A.N. (1986). Words, plans, things, and locations: Interactions between semantic and cognitive development in the one-word stage. In S. Kuczaj & M. Barrett (Eds.), *The development of word meaning* (pp. 199-223). New York: Springer-Verlag.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language, 18*, 115-151.

Hall, D.G. (1991). Acquiring proper nouns for familiar and unfamiliar animate objects: Two-year-olds' word-learning biases. *Child Development, 62*, 1142-1154.

Hall, D.G., Waxman, S.R., Bredart, S., & Nicolay, A. (2003). Preschoolers' use of form class cues to learn descriptive proper names. *Child Development, 74*, 1547-1560.

Hampton, J.A, Estes, Z., & Simmons, C.L. (2005). Comparison and contrast in perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1459-1476.

Haryu, E., & Imai, M. (1999). Controlling the application of the mutual exclusivity assumption. *Japanese Psychological Research, 41*, 21-34.

Haryu, E., Imai, M., Okada, H., Li, L., Meyer, M., Hirsh-Pasek, K., & Golinkoff, R.M. (2005). Noun bias in Chinese children: Novel noun and verb learning in Chinese, Japanese, and English preschoolers. In A. Brugos, M. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the*

29th Annual Boston University Conference on Language Development (pp. 272-283). Somerville, MA: Cascadilla Press.

Heibeck, T.H., & Markman, E.M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*, 1021-1034.

Hespos, S.J., & Spelke, E.S. (2004). Conceptual precursors to language. *Nature, 430*, 453-456.

Hollich, G.J., Hirsh-Pasek, K., & Golinkoff, R.M. (2000). Breaking the language barrier: An emergentist coalition model for 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. (2001). Learning proper nouns and common nouns without clues from syntax. *Child Development, 72*, 787-802.

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development, 76*, 340-355.

Imai, M., Haryu, E., Okada, H., Lianjing, L., & Shigematsu, J. (2006). Revisiting the noun-verb debate: A cross-linguistic 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.

Jaswal, V.K., & Markman, E.M. (2001). Learning proper and common names in inferential versus ostensive contexts. *Child Development, 72*, 768-786.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1-23.

Kako, E. (2005). Information sources for noun learning. *Cognitive Science, 29*, 223-260.

Källkvist, M. (1999). *Form-class and task-type effects in learner English: A study of advanced Swedish learners*. Lund, Sweden: Lund University Press.

Kersten, A.W., & Smith, L.B. (2002). Attention to novel objects during verb learning. *Child Development, 73*, 93-109.

Kersten, A.W., Smith, L.B., & Yoshida, H. (2006). Influences of object knowledge on the acquisition of verbs in English and Japanese. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 499-524). New York: Oxford University Press.

Klatt, D.H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical analysis. *Journal of Phonetics, 7*, 279-312.

Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169-226). Cambridge, MA: MIT Press.

Krueger, J., & Rothbart, M. (1990). Contrast and accentuation effects in category learning. *Journal of Personality and Social Psychology, 59*, 651-663.

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

Landau, B., Smith, L.B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*, 299-321.

Langacker, R.W. (1987). *Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Lannon, B., McDonough, C., Ma, W., Hirsh-Pasek, K., Golinkoff, R.M., & Tardif, T. (in preparation). The role of imageability in early word learning: Evidence from English and Chinese.

Lavin, T.A., Hall, D.G., & Waxman, S.R. (2006). East and west: A role for culture in the acquisition of nouns and verbs. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 525-543). New York: Oxford University Press.

Lennon, P. (1996). Getting "easy" verbs wrong at the advanced level. *International Review of Applied Linguistics in Language Teaching, 34*, 23-36.

Leonard, C.L., Waters, G.S., & Caplan, D. (1997). The influence of contextual information on the resolution of ambiguous pronouns by younger and older adults. *Applied Psycholinguistics, 18*, 293-317.

Leonard, L.B., Schwartz, R.G., Morris, B., & Chapman, K. (1981). Factors influencing early lexical acquisition: Lexical orientation and phonological composition. *Child Development, 52*, 882-887.

Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: MIT Press.

Maguire, M.J., Hennon, E.A., Hirsh-Pasek, K., Golinkoff, R.M., Slutzky, C.B., & Sootsman, J. (2002). Mapping words to actions and events: How do 18-month-olds learn a verb? In B. Skarabela, S. Fish, & A.H.J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 371-382). Somerville, MA: Cascadilla Press.

Maguire, M.J., Hirsh-Pasek, K., & Golinkoff, R.M. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 364-391). New York: Oxford University Press.

Marcus, G.F. (1993). Negative evidence in language acquisition. *Cognition, 46*, 53-85.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Markman, E.M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 72-106). New York: Cambridge University Press.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16*, 1-27.

Markman, E.M., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*, 121-157.

Merriman, W.E., & Bowman, L.L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54* (3-4, Serial No. 220).

Merriman, W.E., Marazita, J., & Jarvis, L.H. (1993). Four-year-olds' disambiguation of action and object word reference. *Journal of Experimental Child Psychology, 56*, 412-430.

Mervis, C.B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child Development, 65*, 1646-1662.

Meyer, M., Leonard, S., Hirsh-Pasek, K., Imai, M., Haryu, E., Pulverman, R., & Addy, D. (2003, November). *Making a convincing argument: A crosslinguistic comparison of noun and verb learning in Japanese and English*. Paper presented at the 28th annual Boston University Conference on Language Development, Boston, MA.

Mintz, T.H. (2005). Linguistic and conceptual influences on adjective acquisition in 24- and 36-month-olds. *Developmental Psychology, 41*, 17-29.

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

Oshima-Takane, Y. (1999). The learning of first and second person pronouns in English. In R. Jackendoff, P. Bloom, & K. Wynn (Eds.), *Language, logic, and concepts: Essays in memory of John Macnamara* (pp. 373-409). Cambridge, MA: MIT Press.

Piccin, T.B., & Blewitt, P. (2007). Resource conservation as a basis for the mutual exclusivity effect in children's word learning. *First Language, 27*, 5-28.

Piccin, T.B., & Waxman, S.R. (in press). Why nouns trump verbs in word learning: New evidence from children and adults in the Human Simulation Paradigm. *Language Learning and Development*.

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

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

Pulverman, R., & Golinkoff, R.M. (2004). Seven-month-olds' attention to potential verb referents in nonlinguistic events. In A. Brugos, L. Micciulla, & C.E. Smith (Eds.),

Proceedings of the 28th Annual Boston University Conference on Language Development (pp. 473-480). Somerville, MA: Cascadilla Press.

Pulverman, R., Hirsh-Pasek, K., Golinkoff, R.M., Pruden, S., & Salkind, S.J. (2006). Conceptual foundations for verb learning: Celebrating the event. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 134-159). New York: Oxford University Press.

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

Quinn, P.C., & Eimas, P.D. (1987). On categorization in early infancy. In J. Oates & S. Sheldon (Eds.), *Cognitive development in infancy* (pp. 131-161). Hillsdale, NJ: Lawrence Erlbaum.

Rice, M.L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development, 59*, 420-424.

Rosch, E. Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Schieffelin, B.B., & Ochs, E. (1987). *Language socialization across cultures*. New York: Cambridge University Press.

Slobin, D.I. (2001). Form/function relations: How do children find out what they are? In M. Tomasello & E. Bates, (Eds.), *Language development: The essential readings* (pp. 267-289). Malden, MA: Blackwell Publishing.

Smith, L.B., Jones, S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Developmental Psychology, 28*, 273-286.

Snedeker, J. (2000). Cross-situational observation and the semantic bootstrapping hypothesis. In E. Clark (Ed.), *Proceedings of the Thirtieth Annual Child Language Research Forum*. Stanford, CA: Center for the Study of Language and Information.

Snedeker, J., & Gleitman, L.R. (2004). Why it is hard to label our concepts. In D.G. Hall & S.R. Waxman (Eds.), *Weaving a lexicon* (pp. 257-293). Cambridge, MA: MIT Press.

Snedeker, J., Gleitman, L., & Brent, M. (1999). The successes and failures of word-to-world mapping. In A. Greenhill, M. Hughs, H. Littlefield, & H. Walsh (Eds.), *Proceedings of the 23rd Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Song, H., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language, 52*, 29-57.

Spelke, E. (2003). Core knowledge. In N. Kanwisher & J. Duncan (Eds.), *Attentions and performance, volume 20: Functional neuroimaging of visual cognition* (pp. 1233-1243). Cambridge: Oxford University Press.

Talmy, L. (1975). Semantics and syntax of motion. In J.P. Kimball (Ed.), *Syntax and semantics* (vol. 4, pp. 181-238). New York: Academic Press.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*, 492-504.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. New York: 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., & Barton, M.E. (1994). Learning words in nonostensive contexts. *Developmental Psychology, 30*, 639-650.

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.

Waxman, S.R., & Booth, A.E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*, 217-242.

Waxman, S.R., & Booth, A.E. (2003). The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds. *Developmental Science, 6*, 128-135.

Waxman, S.R., & Klibanoff, R.S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*, 571-581.

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

Waxman, S.R., & Lidz, J. (2006). Early word learning. In D. Kuhn & R.S. Siegler (Eds.), *Handbook of Child Psychology, 6th Edition, Vol. 2*. Hoboken, NJ: John Wiley & Sons.

Waxman, S.R., Lidz, J.L., Braun, I.E., & Lavin, T. (under review). 24-month-olds' interpretations of novel verbs and nouns in dynamic scenes.

Waxman, S.R., & Markow, D.B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*, 257-302.

Waxman, S.R., & Markow, D.B. (1998). Object properties and object kind: 21-month-old infants' extension of novel adjectives. *Child Development, 69*, 1313-1329.

Waxman, S.R., & Senghas, A. (1992). Relations among word meanings in early lexical development. *Developmental Psychology, 28*, 862-873.

Werker, J.F., & Fennell, C. (2004). Listening to sounds versus listening to words: Early steps in word learning. In D.G. Hall & S.R. Waxman (Eds.), *Weaving a lexicon* (pp. 79-109). Cambridge, MA: MIT Press.

Woodward, A.L., & Markman, E.M. (1998). Early word learning. In D. Kuhn & R.S. Siegler (Eds.), *Handbook of child psychology, Volume 2: Cognition, perception, and language* (pp. 371-420). New York: John Wiley & Sons.

Woodward, A.L., Markman, E.M., & Fitzsimmons, C.M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30*, 553-566.

APPENDIX

Description of Stimuli

Studies 1 through 4

Familiarization scene[a]	Novel verb/noun	Contrast scene[b]	Test scenes
Man waving balloon	pilking/pilker	Playing saxophone	Waving rake vs. tapping balloon
Woman washing cup	semming/sem	Playing guitar	Drinking cup vs. washing plate
Man pushing chair	dacking/dacket	Bouncing ball	Pushing box vs. lifting chair
Woman twirling umbrella	wugging/wugget	Lifting hat	Spinning umbrella vs. twirling pillow
Boy pulling bunny	toping/topin	Sweeping broom	Pulling drum vs. tossing bunny
Girl petting dog	larping/larp	Drinking mug	Kissing dog vs. petting Frisbee

[a]Studies 1 and 4 contained four familiarization scenes; Studies 2 and 3 contained one familiarization scene. [b]Studies 1 and 2 only.

Studies 5 and 6

Familiarization scenes	Novel verb	Contrast scene[a]	Test scenes
Man waving hammer/ spoon/carrot/flower	pilking	Lifting hat	Waving rake vs. petting turtle
Woman washing cup/ boat/duck/ball	semming	Dancing doll	Playing slinky vs. washing plate
Man pushing bottle/ truck/book/bowl	dacking	Bouncing die	Pushing apple vs. drinking cup
Woman twirling basket/ telephone/teapot/monkey	wugging	Playing saxophone	Brushing hair vs. twirling plate
Man tossing ball/ shoe/bucket/bat	toping	Tapping tambourine	Tossing bear vs. playing guitar
Girl tapping pig/lamp/ drum/hat	larping	Dancing cat	Sweeping broom vs. tapping watering can

[a]Study 5 only.