Can Adverbs Call Attention to Manner of Motion for 2-year-olds Learning Verbs?

Sudha Arunachalam, Kristen Syrett, and Sandra R. Waxman

1. Introduction

When a child acquires a new word, she must attend to multiple sources of information—including the linguistic context in which it appears and the extralinguistic environment in which it is uttered—in order to identify its meaning and grammatical properties. This is no small feat, since both information sources are plagued with ambiguity (e.g., Bloom, 2002; Gleitman, 1990; Pinker, 1984; Quine, 1960). It is widely recognized that verb acquisition is especially challenging (e.g., Gentner, 1982; Gentner & Boroditsky, 2001; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Gleitman & Gleitman, 1992; Golinkoff & Hirsh-Pasek, 2008; Tomasello, 1992). While toddlers produce their first words around their first birthday, verbs lag behind, entering the lexicon in significant numbers almost a year later (Fenson et al., 1994).

Experimental studies document that verb acquisition is challenging at least in part because acquiring the meaning of a novel verb requires more informational support than acquiring the meaning of a novel noun (Arunachalam & Waxman, 2011; Gillette et al., 1999; Imai, Haryu, & Okada, 2005; Imai et al., 2008; Kersten & Smith, 2002; Snedeker & Gleitman, 2004; Waxman, Lidz, Lavin, & Braun, 2009). One powerful source of informational support is the verb’s linguistic context (Gleitman, 1990; Landau & Gleitman, 1985). Children as young as 2 years of age can gather information about verb meaning from the number and position of its surrounding arguments (e.g., Fisher, 2002; Hirsh-Pasek & Golinkoff, 1996; Naigles, 1990), as well as the semantic content of the phrases occupying those argument positions (Fisher, Hall, Rakowitz, & Gleitman, 1994; Gillette et al., 1999; Gleitman et al., 2005; Piccin & Waxman, 2007; Snedeker & Gleitman, 2004).

Moreover, recent evidence reveals that in at least some verb learning

* Authors Sudha Arunachalam, Boston University (sarunach@bu.edu); Kristen Syrett, Rutgers University—New Brunswick (k-syrett@ruccs.rutgers.edu); Sandra R. Waxman, Northwestern University (s-waxman@northwestern.edu). This research was supported by grants from the National Institutes of Health to the second (HD-057699) and third (HD-030410) authors.
situations, the semantic content of the phrases that occupy the verb’s argument positions is crucial (Arunachalam & Waxman, 2010, 2011; Naigles, Reynolds, & Küntay, 2011). Arunachalam & Waxman (2010) (A&W) taught 27-month-olds novel verbs in one of two conditions. In both, toddlers viewed the same video of an agent performing an action with an inanimate object (e.g., a boy waving a balloon). In the Rich Content condition, the verb was flanked by Determiner Phrases with content nouns (“A boy is gonna pilk a balloon”). In the Sparse Content condition, the novel verb was instead flanked by pronouns, and the content nouns appeared in the previous sentence to establish reference (“Let’s see a boy, and a balloon. He’s gonna pilk it”). When toddlers were asked to identify the verb’s referent from among two new scenes (e.g., a boy tapping a balloon, a boy waving a rake), those in the Rich Content condition succeeded, but not those in the Sparse Content condition. Toddlers’ success when the verb’s arguments were expressed as content nouns, but not pronouns, led A&W to reason that the rich semantic information available in the content nouns helped toddlers to identify the event participants (the boy, the balloon), and in turn, to home in on the relation between them (waving). Nominal expressions thus serve as one entry point to identifying verb meaning, when they occupy the verb’s argument positions.

Notice however, that this outcome, interesting in and of itself, also presents a challenge for theories of acquisition. After all, pronouns are extremely frequent in speech to young children (Cameron-Faulkner, Lieven, & Tomasello, 2003; Mintz, 2003; Valian, 1991; Weisleder & Waxman, 2010). We would hope that toddlers are able to learn novel verb meanings when they are presented in these pronoun contexts. What other linguistic elements, then, might provide semantic support for verb learning in the absence of content nouns flanking the verb? In the current experiment, we focus on one linguistic element in particular, testing the hypothesis that manner-of-motion adverbs (e.g., slowly) (see, e.g., Shäfer, 2002; Tenny, 2000; Thomason & Stalnaker, 1973) can support verb learning by calling attention to the very aspect of the scene, i.e. the motion, to which the novel verb refers.

Recent evidence (Syrett, 2007; Syrett & Lidz, 2010) demonstrates that toddlers do indeed attend to adverbs in the service of word learning. Specifically, adverbs support 30-month-olds’ acquisition of novel adjectives. In these studies, toddlers viewed, for example, a container that was both tall and clear, and heard a novel adjective describing it. Those that heard the object described as completely pelgy assigned the novel adjective pelgy a meaning like ‘clear’ rather than ‘tall’. The opposite pattern was found for toddlers who had heard the object described as very pelgy. Syrett and Lidz (2010) proposed that toddlers used semantic properties of the adverb to identify the meaning of the adjective it modified: while an object can be completely or maximally clear, it cannot be completely or maximally tall.

1 Corresponding A&W conditions: Rich Content = ‘+Nouns +Syntax’, Experiment 1; Sparse Content = ‘+Nouns +Syntax’, Experiment 2
In the current experiment, we propose that adverbs may likewise confer an advantage in verb learning. To our knowledge, no studies have examined the role of adverbs in verb learning in toddlers, although Wittek (2002) demonstrated that the adverb again can support verb learning in certain contexts with 5-year-olds. To test our hypothesis, we adapted A&W’s paradigm, identifying their Sparse Content condition (in which toddlers performed at chance, showing no evidence of verb learning) as an excellent testing ground for our hypothesis. We reasoned that a manner-of-motion adverb would highlight aspects of the event, and therefore aspects of the verb’s meaning.

2. Method
2.1. Participants

Forty typically-developing children (20 males; mean age: 27.3 months, age range: 24.9 to 29.4 months) participated. An additional twelve were excluded from the final sample for failure to meet selectional criteria (correct pointing on at least one of two training trials and clear pointing on at least one test trial), two due to parental interference, and one to fussiness.

We selected the age range based on evidence that children this age succeed in similar verb learning tasks (cf. Hirsh-Pasek & Golinkoff, 1996) and in word learning tasks with adverbs (Syrett & Lidz, 2010). Children were recruited from Evanston, IL, and surrounding areas. All children were acquiring English as a native language and were exposed to another language less than 25% of the time. Parents completed the MacArthur-Bates CDI Short Form: Words and Sentences (Fenson et al., 1994). Mean vocabulary production was 73 words (range: 22 to 99 words); there were no differences in vocabulary between conditions or genders.

2.2. Materials

Materials were adapted from Waxman et al. (2009) and Arunachalam & Waxman (2010).

2.2.1. Visual stimuli

Visual stimuli were digital video recordings of live actors performing child-friendly actions with inanimate objects, such as waving a balloon. Videos were edited using Final Cut Pro and presented to toddlers on a 20-inch television screen.

2.2.2. Auditory stimuli

Auditory stimuli were produced by a female native speaker of American English in a sound-attenuated recording booth. Utterances were edited, controlling for duration and intensity, using Praat software (Boersma &
Weenink, 2005), and were synchronized with the visual stimuli. Sound was presented through a speaker located directly below the center of the screen.

2.3. Apparatus and procedure

Toddlers and caregivers were welcomed into the waiting area. While the toddlers played with toys, their caregivers completed a consent form and the MacArthur-Bates CDI checklist. The toddler and caregiver were then brought into an adjoining room, where the toddler sat in a booster seat, approximately 14 inches from a 20-inch television screen. The caregiver sat behind the toddler and was asked not to interact with him/her during the session.

One experimenter controlled the video presentation from behind a curtain, while another sat next to the toddler to elicit responses. Toddlers were asked to indicate their choices by pointing. This method has been successfully employed in word learning tasks with participants of this age (Arunachalam & Waxman, 2010, 2011; Bernal, Lidz, Millotte, & Christophe, 2007; Dittmar, Abbot-Smith, Lieven, & Tomasello, 2011; Fernandes, Marcus, DiNubila, & Vouloumanos, 2006; Fisher, 2002; Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008; Noble, Rowland, & Pine, 2011; Syrett, Musolino, & Gelman, in press). Pointing responses were recorded with a video camera located directly above the screen.

Before the test session, toddlers engaged in a brief training procedure, designed to encourage pointing. They viewed two training trials, each introducing two dynamic scenes side-by-side on the screen. The experimenter asked the toddler to point to either a familiar character (e.g., Elmo) or an activity (e.g., dancing). No novel words were used during training. Toddlers who pointed incorrectly or who were reluctant to point were gently encouraged to point again. Those who failed to point or pointed incorrectly on both training trials were excluded from analysis.

2.4. Test session and trial structure

There were six different test trials, each featuring an agent engaged in an action with an object, described by a novel verb. Each trial lasted a little over half a minute, for a total time of just under five minutes. Participants were randomly assigned to one of two conditions. All toddlers viewed the same visual materials; the only difference was the adverb heard during the Linguistic Familiarization Phase (slowly or nicely). Each trial had the same structure (see Figure 1) and is described in more detail below.
Linguistic Familiarization Phase

During the Linguistic Familiarization Phase, toddlers saw a 5-second still frame of a live action scene and heard one of two versions of the auditory stimuli, as indicated in the Figure. Note that the stimuli were identical to A&W’s Sparse Content condition, with the sole exception of the utterance-final adverb (slowly or nicely) heard during the Linguistic Familiarization Phase.

Event Familiarization Phase

Next, during the Event Familiarization Phase (12 seconds), toddlers saw a dynamic scene in which the agent performed an action with the object (e.g., waving the balloon). The scene appeared on one side of the screen, and then on the other side of the screen (side counterbalanced across trials), with an exclamation accompanying each presentation. The novel verb occurred once during this phase in a bare syntactic frame (Look! Pilking!)

Test Phase

Finally, during the Test Phase, two different dynamic scenes appeared simultaneously on the screen side-by-side. In the Familiar Object scene, the agent performed a new action on the familiar object (e.g., tapping the balloon); in the Familiar Action scene, the agent performed the now-familiar action with a new object (e.g., a toy rake). Toddlers were first given an opportunity to inspect both scenes (6 seconds). The screen then went black for 2 seconds, and toddlers heard, Do you see pilking? The scenes then reappeared in their original locations for 12 seconds and toddlers heard, Find pilking! The experimenter encouraged the child to point.
2.5. Coding and Analysis

Pointing responses were recorded by the experimenter and independently verified by a condition-blind coder using the videos of the session. Toddlers’ first point after the test question served as the dependent measure. For each toddler, we calculated the number of trials on which she pointed to the Familiar Action, divided by the total number of trials on which she pointed.

2.6. Predictions

We predicted that if adverbs can call attention to aspects of the event to support verb learning, children in the adverb-supplemented Sparse Content condition should perform above chance level. Further, if the mere presence of a sentence-final adverb is sufficient, then children in both conditions should succeed. If, however, toddlers require the adverb to encode certain information about the event, such as specific manner of motion, then only toddlers in the slowly condition should succeed, while those in the nicely condition should perform at chance.

2.7. Results

The results are presented in Figure 2.

![Figure 1: Average proportion of points to the Familiar Action scene across trials and children for two adverb conditions](image)

As predicted, toddlers who heard *slowly* were more likely to point to the
Familiar Action scene at test (M = .65) than those who heard nicely (M = .49) (two-tailed t(1, 38) = -2.09, p < .05). Moreover, toddlers who heard slowly chose this scene significantly more often than predicted by chance (two-tailed t(1, 19) = 2.75, p = .01), while toddlers who heard nicely performed no differently than chance (t(1, 19) = -0.25, p = .80).

As a point of comparison, note that toddlers in A&W’s Rich Content condition, in which the content nouns occupied the verb’s argument positions, performed at this same above-chance level (M = .65), while toddlers who heard pronouns in the frame (the Sparse Content condition) were at chance (M = .50). Our findings thus demonstrate that even a pronominal frame can lead to success in verb learning, if properly supplemented with additional semantic information, either from content nouns framing the verb or an adverbial modifier calling attention to an aspect of verb meaning.

An analysis of individual toddlers’ performance offers strong converging evidence for this interpretation. We tallied the number of toddlers in each condition who favored the Familiar Action scene on a majority of their six trials. As Table 1 illustrates, those in the slowly condition were significantly more likely than those in the nicely condition to favor the Familiar Action scene (Pearson $\chi^2(1) = 4.15, p = .04$). Thus, the individual patterns reflect the overall trend captured by the parametric analysis of percentage pointing.

Table 1: Number of toddlers displaying a proportion of points to the Familiar Action scene greater than or less than chance

<table>
<thead>
<tr>
<th>Mean proportion points to Familiar Action scene</th>
<th>x&gt;.50</th>
<th>x&lt;.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>nicely</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>slowly</td>
<td>14</td>
<td>4</td>
</tr>
</tbody>
</table>

3. General Discussion

These results provide insight into the kinds of linguistic information that toddlers take into account in their efforts to acquire a novel verb’s meaning. Unlike 27-month-olds in Arunachalam & Waxman (2010), who performed at chance levels when novel verbs were introduced in a pronominal frame, those in the current experiment succeeded, if an adverb describing a specific manner of motion (slowly) accompanied the utterance. These results, which provide the first evidence that adverbial modification supports verb learning in toddlers, supplement previous reports that toddlers use the semantic content of known words to acquire the meanings of novel words (e.g., Arunachalam & Waxman, 2011; Fennell & Waxman, 2010; Fernald, Zangl, Portillo, & Marchman, 2008; Fisher, et al., 1994; Gillette, et al., 1999; Gleitman et al., 2005; Hirsh-Pasek & Golinkoff, 1996; Piccin & Waxman, 2007; Portillo, Early, Quam, Zangl, & Fernald, 2005; Snedeker & Gleitman, 2004). However, in this case, the results go one step further to demonstrate this facilitative effect even for words that
may appear in the input but are not necessarily produced at the target age. A search of children’s speech in the Brown, Gleason, Sachs, and Suppes corpora in the CHILDES database (Brown, 1973; Gleason, 1988; MacWhinney, 2000; Sachs, 1983; Suppes, 1974) revealed that slowly appears to enter children’s speech somewhat productively only after three years of age.

The current results also reveal that not all adverbs promote learning in this context: toddlers succeeded with slowly, but performed at chance level with nicely. Why might one adverb, but not the other, facilitate verb learning? We envision two alternatives.

First, as we have argued, these results may reflect a core difference in the lexical semantics of these adverbs: slowly refers specifically to the manner of motion of the event and therefore highlights aspects of the event that support verb learning. Complementing the lexical semantics of the adverbs is an asymmetry in the lexical semantics of the adjectives at the root of these adverbs, which are often used in place of their adverbial counterparts. While nice describes a static property, slow inherently makes reference to events or motion, even when referring to an entity. Thus, there is a fundamental difference in the kind of information these words encode. It may be precisely this sort of information that is guiding the word learning process. We therefore predict that other modifiers beyond slowly, which refer to specific properties of an event, would also result in successful performance in this task.

The second hypothesis implicates sheer lexical frequency. A search of the Corpus of Contemporary American English (COCA) reveals that slowly is approximately four times as frequent as nicely (slowly: 2192 overall occurrences, or 24.32 per million; nicely: 536, or 5.95 per million). It is therefore possible that slowly’s success is a result of children being more familiar with that adverb than they are with nicely. Thus we predict that another adverb with comparable or greater lexical frequency would elicit the same response pattern. However, this story is complicated by the fact that nice (17,552 occurrences, or 194.88 per million) is much more frequent than slow (3809 occurrences, or 42.29 per million). An account appealing to raw lexical frequency would need to reconcile this pattern, while explaining away the lexical semantic differences, which we find particularly compelling.

The past decades have identified several components of the linguistic stream that young word learners take advantage of in verb learning. It is by now well established that the number and position of noun phrases accompanying a verb phrase provide important informational support for the acquisition of new verbs (cf. Fisher et al., 1994; Landau & Gleitman, 1985; Naigles, 1990; a.o.). We have demonstrated that when a verb’s syntactic frame is underinformative, and the presence of familiar content nouns in the previous sentence is not enough, the presence of a manner-of-motion adverb such as slowly can direct a toddler’s attention to the event in such a way that supports verb learning. Future research should probe the precise nature of this contribution and the ability of these findings to generalize over other modifiers.
References


Gentner, Dedre. (1982). Why nouns are learned before verbs: Linguistic relativity versus


Syrett, Kristen, Musolino, Julien, & Gelman, Rochelle. (in press). Can children use syntax to learn about number word meaning? *Language Learning and Development.


