UC Merced

UC Merced Previously Published Works

Title

Finding meaning in a noisy world: exploring the effects of referential ambiguity and competition on 2·5-year-olds' cross-situational word learning*

Permalink

Journal

ISSN

Authors

BUNCE, JOHN P
SCOTT, ROSE M

Publication Date

2017-05-01

DOI

Copyright Information

Peer reviewed

Finding meaning in a noisy world: Exploring the effects of referential ambiguity and competition

on 2.5-year-olds' cross-situational word learning

John P. Bunce

University of California Merced

Rose M. Scott

University of California Merced

Please address correspondence to:

Rose M. Scott
School of Social Sciences, Humanities, and Arts
University of California Merced
5200 North Lake Road
Merced, CA 95343
rscott@ucmerced.edu
(209) 228-4362

Acknowledgments

Abstract

While recent studies suggest children can use cross-situational information to learn words, these studies involved minimal referential ambiguity and the cross-situational evidence overwhelmingly favored a single referent for each word. Here we asked whether 2.5-year-olds could identify a noun's referent when the scene and cross-situational evidence were more ambiguous. Children saw four trials in which a novel word occurred with four novel objects; only one object consistently co-occurred with the word across trials. The frequency of distracter objects varied across conditions. When all distracter referents occurred only once (no-competition), children successfully identified the noun's referent. When a high-probability competitor referent occurred on three trials, children identified the target referent if the competitor was absent on the third trial (short-competition) but not if it was present until the fourth trial (long-competition). This suggests that although 2.5-year-olds' cross-situational learning scales up to more ambiguous scenes, it is disrupted by high-probability competitor referents.

Introduction

When children encounter a new word, the referential scene offers many potential interpretations. To illustrate, imagine a child eating dinner with her family. In front of each family member sits a plate of food, a cup, a napkin, a fork, and a spoon. As the child eats peas with her hands, her mother says, "Use your fork." In this scene, *fork* could refer to fork, spoon, plate, peas, rolling, green, eat, and so on. How might the child resolve this ambiguity and identify the correct referent for *fork*?

Studies have shown that children can use a variety of non-linguistic and linguistic cues to constrain their interpretation of a word (e.g., Baldwin, 1993; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Markman & Wachtel, 1988; Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Yuan, Fisher, & Snedeker, 2012). For instance, children can use distributional, phonological, and syntactic information to identify a word's likely grammatical category (e.g., Cauvet, Limissuri, Millotte, Skoruppa, Cabrol, & Christophe, 2014; Mintz, 2006; Zhang, Shi, & Li, 2015) and restrict reference accordingly (e.g., Bernal, Lidz, Millotte, & Christophe, 2007; Fisher, Klingler, & Song, 2006; Waxman & Booth, 2001; Waxman, Lidz, Braun, & Lavin, 2009). Thus, 24-month-olds who hear a novel word used as a noun (e.g., *The man is waving a larp.*) assume it refers to an object, whereas those who hear the same word used as a verb (e.g., *The man is larping a balloon.*) assume it refers to an action (e.g., Bernal et al., 2007; Waxman et al., 2009).

Returning to our hypothetical dining scenario, the child could likely use the sentence that *fork* occurs in to infer that the word is a noun and restrict her interpretation of the word to potential object referents. Even armed with this useful constraint, however, she would face a dilemma, as the referential scene offers a number of candidate object referents to choose from

(e.g., fork, spoon, plate, peas). Researchers have long assumed that one way learners cope with such referential ambiguity is by considering additional referential contexts in which the same word occurs (e.g., Fazly, Alishahi, & Stevenson, 2010; Fisher, Hall, Rakowitz, & Gleitman, 1994; Pinker, 1984; Siskind, 1996; Yu & Smith, 2007). Across situations, scene elements that are not relevant to the word's meaning should occur less consistently than those that are central to its meaning. If children could identify the elements that consistently co-occurred with a word across uses, then this would help them determine the word's likely referent.

Recent evidence suggests that under at least some circumstances, both adults and children can use cross-situational information to identify the referents of new words (e.g., Akhtar & Montague, 1999; Childers, 2011; Childers & Paik, 2009; Dautriche & Chemla, 2014; Gillette, Gleitman, Gleitman, & Lederer, 1999; Scott & Fisher, 2012; Smith, Smith, & Blythe, 2011; Smith & Yu, 2008; Yu & Smith, 2007, 2011; Yurovsky, Yu & Smith, 2013). For instance, Smith and Yu (2008) presented 12- and 14-month-olds with a series of training trials in which pairs of novel objects were accompanied by two novel labels. On each trial, it was ambiguous which label went with which object. However, across trials each word consistently co-occurred with only one object. Following training, infants saw test trials in which a label was presented with its referent and a distracter object. Infants looked longer at the referent object, suggesting they had used the cross-situational information to identify the words' referents. These findings indicate that the basic mechanism necessary for using cross-situational information in word learning is present in infancy: children can attach some amount of information about potential referents to a word's lexical entry under referential uncertainty and then retrieve and update this information based on later observations.

However, the extent to which children can exploit cross-situational information in

everyday word learning situations remains unclear. This is because prior experiments on children's ability to use cross-situational information have simplified the learning situation in two critical ways: the experimental setting involved minimal referential ambiguity and the cross-situational evidence overwhelmingly favored a single referent for each word. We begin by discussing each of these simplifications in more detail. We then turn to the present research, which explored whether children could use cross-situational information to identify a word's referent when both the referential scene and the cross-situational evidence were more ambiguous.

Referential ambiguity

Prior studies on cross-situational word learning in children have involved less referential ambiguity than is present in typical word-learning situations. In the dining room example described above, the child is confronted with dozens of potential referents for a new word. In contrast, in prior experiments, a novel word was accompanied by a maximum of two candidate object- or event-referents in each trial (e.g., Smith & Yu, 2008; Scott & Fisher, 2012; Vlach & Johnson, 2013). These simple test scenes may have been critical to children's success: recent evidence suggests that as the number of potential referents in a scene increases, even adults have difficulty using cross-situational information to identify the referents of words (e.g., Medina, Snedeker, Trueswell & Gleitman, 2011; Smith et al., 2011). These findings raise the possibility that if young children were confronted with a larger number of potential referents for each word, they might be unable to use cross-situational information to identify the words' correct referents.

To date, no studies have systematically explored how the number of potential referents in a scene affects children's cross-situational word learning. However, recent observational studies by Yu, Smith, and colleagues suggest that greater referential ambiguity might impede children's

cross-situational learning (e.g., Pereira, Smith, & Yu, 2014; Suanda, Foster, Smith, & Yu, 2013; Yu & Smith, 2012). In these studies, young toddlers wore head cameras while engaging in a play session with their parents. Dyads sat at a table and played with three sets of three novel objects, each of which had a novel label. Parents learned these labels prior to the session and were encouraged to use them throughout the play session. After the play session, children were tested on their knowledge of the nine labels. Comparison of parental labeling events for learned and unlearned words revealed that the words children learned had tended to occur when a single object (the referent) was dominant in the children's field of view, while words that they failed to learn had occurred when multiple objects were equally visible. These studies thus provide suggestive evidence that two-year-olds may have difficulty learning words from cross-situational information when referential scenes contain more than two potential referents.

In addition to presenting a limited number of potential referents, many prior cross-situational learning studies also involved exhaustive labeling: children encountered two potential referents on each trial and both of these referents were labeled with a novel word (e.g., Scott & Fisher, 2012; Smith & Yu, 2008; Suanda, Mugwanya, & Namy, 2014; Vlach & Johnson, 2013; Yu & Smith, 2011). Labeling both of the candidate referents present on each trial might have further reduced the referential ambiguity of the task in several ways. First, labeling both referents provides implicit contrast information: if a child sees two objects and hears *bosa* and *manu*, this implies that one of these objects is a *bosa* while the other is <u>not</u> a *bosa*. Contrast information has been shown to facilitate children's word learning and extension in a variety of tasks (e.g., Ankowski, Vlach, & Sandhofer, 2013; Childers, Hirshkowitz, & Benavides, 2014; Namy & Clepper, 2010; Waxman & Klibanoff, 2000). In prior cross-situational learning studies, implicit contrast could have encouraged children to form one-to-one mappings between words and

objects. If so, this would have reduced ambiguity by ruling out the possibility that a given label referred to both objects (i.e. it described a superordinate category like *toy*) or that both labels referred to a single object. Second, labeling both referents could have reduced referential ambiguity across trials via mutual exclusivity (e.g., Markman & Wachtel, 1988; Yurovsky, Yu, & Smith, 2013). Suppose a child hears *bosa* and *manu* while viewing object-A and object-B and then hears *bosa* and *kaki* while viewing object-A and object-C. Based on only two observations, the child can determine that *bosa* refers to the repeated object-A, *manu* refers to object-B, and *kaki* refers to object-C. If the child instead hears only *bosa* during the first trial and *kaki* in the next, then both trials would be more ambiguous: *bosa* could refer to either object-A or object-B, and *kaki* could refer to either object-B or object-C. The child would thus require additional observations to identify each word's referent.

No study has directly tested the role of exhaustive labeling in cross-situational paradigms and thus its impact on children's cross-situational word learning is currently speculative. However, it is noteworthy that the one prior study that tested young children in a Smith and Yu (2008) style task without exhaustive labeling produced negative results (e.g., Yurovsky, Hidaka, Yu, & Smith, 2010). In that experiment, 15-month-olds first viewed a series of training trials in which they saw two novel objects and heard only one novel label. Following training, children saw test trials in which a novel label was presented six times and accompanied by its target referent and a distracter. Infants looked equally at the two objects during the test trials, suggesting that they failed to use cross-situational information to identify the word's referents. Because the training phase differed from Smith and Yu (2008) in several ways in order to accommodate presenting a single label on each trial (i.e. the number of training trials was doubled and the duration of each was halved), it is impossible to attribute these results to the lack

of exhaustive labeling per se. Nevertheless, these negative results are consistent with the possibility that labeling each candidate referent on every learning trial played a key role in children's performance in prior cross-situational learning experiments. Given that children are unlikely to consistently encounter situations in which all potential referents are labeled, it is thus important to determine whether children can make use of cross-situational information in the absence of this supportive information.

Cross-situational evidence: the impact of competing referents

In prior experiments, children were presented with very straightforward cross-situational statistics: each novel word co-occurred with its referent 100% of the time and occurred with all distracters with equally low probability. Thus, the cross-situational statistics overwhelmingly favored a single referent for each novel word. In real life, however, the cross-situational evidence may not be so clear-cut. To illustrate, consider our initial example of a child attempting to learn the word *fork*. Quite often when children hear the word *fork*, both a fork and a spoon are likely to be present. Because words can and do occur in the absence of their referents (Gleitman, 1990; Harris, Jones, & Grant, 1983), children could also encounter the word *fork* when a spoon is present but a fork is not. This could result in children receiving similar levels of cross-situational evidence for both the word's target referent (fork) and a high-probability competitor (spoon), making it difficult for them to determine which object is the correct referent.

Suanda et al. (2014) recently demonstrated that the presence of a high-probability competitor affects school-aged children's ability to identify the correct referent for a novel word via cross-situational observation. Five- to seven-year-old children saw a series of 16 training trials in which two novel words were accompanied by two novel objects. Each word consistently co-occurred with a single target referent across trials. The frequency with which each word

occurred with non-target referents varied across conditions. In the high contextual diversity condition, each word occurred with its target referent four times (100% co-occurrence) and occurred once with four different distracter referents (each 25% co-occurrence). In the low contextual diversity condition, each word occurred with its target referent four times (100% co-occurrence), occurred with one of the distracters three times (75% co-occurrence) and occurred with another distracter only once (25% co-occurrence). In the low contextual diversity condition, the distracter that occurred with the word on three of the four trials in which it was presented served as a high-probability competitor for the target referent, much as in our fork/spoon example. Following training, children completed a series of test trials in which they heard a target word while viewing four objects: the target referent and three objects that had never co-occurred with the word before. Children were asked to pick the object that they thought went with the word. Although children in both conditions identified the correct referents at above chance levels, children in the high contextual diversity condition significantly outperformed those in the low contextual diversity condition. Thus, children who never encountered a repeated word-distracter pairing were significantly more likely to learn the words' referents than were children who repeatedly encountered words with a high-probability competitor.

There are several ways in which the high-probability competitor could have impacted learning in low contextual diversity condition. Given that words sometimes occur in the absence of their referents, an object that occurs with a word on three out of four occasions could be the correct referent for that word. Children might therefore have viewed both the target and the high-probability competitor as possible referents for the word. Recent evidence suggests that when adults encounter words with two referents, they learn those two referents less well than when the cross-situational evidence strongly favors one referent for each word (Yurovsky, Yu, & Smith,

2013). Thus, children in the low contextual diversity condition might have performed more poorly because they were attempting to learn two referents for each word rather than one. Alternatively, it could be that children in the low contextual diversity condition were attempting to identify a single correct referent for each word, but the high-probability competitor prolonged this process. Children in the high contextual diversity condition could identify each word's correct referent on the second observation, while children in the low contextual diversity condition had to consider multiple competing referents until the disambiguating trial in which the high-probability competitor was absent. Although the timing of this trial varied across children, for many children it occurred on the third or fourth observation (Suanda, personal communication). Thus, the high-probability competitor might have increased the memory burden of the task by increasing the length of time that children needed to consider competing referents before they could identify the target referent.

In either case, the fact that the presence of a high-probability competitor affected cross-situational learning in five- to seven-year-olds suggests that competitors might be quite problematic for younger children such as the infants and toddlers typically tested in cross-situational word learning studies. Given that younger learners' less-developed memory capacities constrain their ability to encode, retrieve, and aggregate information about potential referents across observations (e.g., Vlach & Johnson, 2013), it is likely their cross-situational learning would be more easily disrupted by the presence of a high-probability competitor. Moreover, the order of learning observations might be especially important for younger children: if they repeatedly encounter the word with both the target and high-probability competitor, then they might be unable to identify the correct referent when they eventually encounter it in the absence of the competitor.

Present research

In sum, existing studies demonstrate that children possess a basic mechanism for using cross-situational information to identify a novel word's referent. However, these studies have presented children with simplified learning situations involving minimal referential ambiguity and cross-situational evidence that overwhelmingly favored a single referent for each word. In the present research, we began to explore whether young children could use cross-situational information to learn words under more challenging conditions.

Specifically, we examined how greater referential ambiguity and the presence of a high-probability competitor referent affected 2.5-year-olds' ability to identify the referent of a novel noun. To test this question, we adapted Smith and Yu's (2008) paradigm: we increased the number of potential referents on each trial, labeled only a single object, and varied whether the target referent occurred with a high-probability competitor. Children viewed a series of trials in which they saw four novel objects accompanied by a single novel label and their looking time to each object was measured. Across trials, only the target object consistently co-occurred with the label. Children in the short- and long-competition conditions encountered a high-probability competitor referent, whereas those in the no-competition condition did not. Across the short- and long-competition conditions, we varied when the high-probability competitor occurred in order to investigate whether the order of observations affected children's cross-situational learning.

Several patterns of results were possible. If children were able to track which object consistently co-occurred with the label across trials despite greater referential ambiguity and the presence of a high-probability competitor, then children in all three conditions would successfully identify the novel word's referent. In contrast, if the increased referential ambiguity in our task interfered with children's ability to make use of cross-situational information, then

none of the conditions would show above chance performance. If children were able to cope with additional referential ambiguity but the competition created by the high-probability competitor impaired cross-situational learning, then performance would differ across the conditions: children in the no-competition condition would successfully identify the novel word's referent, while those in the short- and long-competition conditions would not. Finally, if children were affected by how long they needed to consider multiple competing referents before the cross-situational evidence identified the target referent, then performance should differ across the short- and long-competition conditions.

We expected our task would be more difficult than Smith and Yu's (2008) task, and thus we chose to test 2.5-year-old toddlers rather than one-year-old infants. Our selected age range was based on Scott and Fisher (2012), who found that 2.5-year-olds could use cross-situational information to identify the referents of novel intransitive verbs that labeled solo actions, but they experienced difficulty with transitive verbs that labeled two-participant actions. These findings indicate that although 2.5-year-olds can use cross-situational information to learn verbs, despite the additional difficulty these words impose (e.g., Gillette et al., 1999; Gleitman et al., 2005), their cross-situational learning ability is still fragile and can be undermined by small increases in task complexity. Thus, this is a particularly interesting age range for investigating children's ability to use cross-situational information to learn nouns under challenging conditions.

Method

Participants

42 2.5-year-olds participated in the experiment (M = 2;10, range 2;8-3;2, 21 male, 21 female). All children were native speakers of English. An additional 4 children were tested but eliminated because they were fussy (1), off-task (1), spent more than 40% of the experiment

looking to a single quadrant of the screen (1), or because of parental interference (1). Children's productive vocabularies were measured with the MacArthur-Bates Communicative Development Inventory, Level 3 (Fenson, Marchman, Thal, Dale, Reznick, & Bates, 2007). Vocabulary scores ranged from 3 to 95 out of a possible 100 with a median of 59 (this range corresponds to the 2nd to the 94th percentile, with a median at the 36th percentile). Equal numbers of children were randomly assigned to the no-competition, short-competition, and long-competition conditions.

To verify that it was possible to identify the noun's referent in all three conditions, we tested a separate group of adult participants. 42 adults ($M = 20$ years; range 18-25; 13 male, 29 female) completed the experiment for course credit. One additional adult was tested but eliminated due to confusion about the task. Equal numbers of adults were randomly assigned to each of the three conditions.

Apparatus

Children sat on their parent's lap in a dimly lit room, centered 91 cm in front of a 76 cm x 128 cm LCD television; the bottom of the screen was 96.5 cm above the floor. A camera centered below the screen recorded the children's eye movements. Parents were instructed to close their eyes or look down to avoid biasing their children's responses. Adult participants sat alone in a chair in the same position as the children. Adult participants were told that they would watch a video used in research with children, that the video would contain a made-up word, and that they should pay close attention because they could be asked questions about the video when it finished.

Stimuli

Stimuli consisted of high-resolution photos of 8 familiar objects and 16 unfamiliar objects (see Table 1); each image measured 19 cm x 37 cm. We selected familiar objects that the

majority of 2.5-year-olds know the labels for based on the MCDI Lexical Norms database (Dale & Fenson, 1996; Jørgensen, Dale, Bleses, & Fenson, 2010). The unfamiliar objects consisted of unusual household items that children would be unlikely to know the labels for. In order to verify that these objects were unfamiliar to 2.5-year-olds, we collected pilot data from 20 additional children, none of whom participated in the main experiment. Children were shown each of the objects and asked, "What is this called?" None of the children correctly identified any of the objects. Additionally, children did not provide consistent incorrect guesses for any of the objects (e.g., consistently mislabeling the bottle opener as a spoon). This suggests that the children did not have labels for any of the unfamiliar objects.

[Insert Table 1 here]

Procedure

Participants watched a video comprised of a series of eight 8-second trials. During each trial, participants saw 4 objects, one in each quadrant of the screen (see Figure 1). A soundtrack recorded by a native English speaker accompanied the objects. The procedure had two phases: familiar-word and novel-word.

[Insert Figure 1 here]

The familiar-word phase consisted of 4 trials in which participants saw a familiar target object (ball) and 3 familiar distracter objects presented on the screen. In each trial, the ball was labeled twice (e.g., *Look at the ball*); the label occurred 2.5s and 6s into the trial, respectively. The familiar-word phase was included for two reasons. First, most preferential-looking studies with 2.5-year-olds have presented only two objects or events per trial. We therefore wanted to verify children's ability to locate a target object when presented with four objects on the screen. Second, given children's natural desire to inspect all four of the objects presented on the screen,

it was unclear how long 2.5-year-olds would persist in looking at the target referent after hearing it labeled. Pilot data from the familiar-word phase was therefore used to identify an appropriate analysis window (see Coding and analyses section).

In the novel-word phase, participants saw 4 trials in which 4 unfamiliar objects were presented on the screen. In all conditions, the target object (oven mitt) appeared in all four trials; the 3 distracter objects varied across trials and conditions (see below). In each trial, children heard the novel noun *jop* twice (e.g., *Look at the jop*). The onset of the novel noun occurred 2.5s and 6s into the trial, respectively. While in any given trial it was unclear which object *jop* referred to, across trials only the target object consistently co-occurred with the novel word.

We chose to present the novel noun in a labeling phrase, rather than in isolation (e.g., *bosa, manu*) as Smith and Yu (2008) did, for two reasons. First, this better represents everyday learning contexts, where words often do not occur in isolation (e.g., Aslin, Woodward, LaMendola, & Bever, 1996). Second, children might have had difficulty determining that a single, isolated novel word was a noun rather than an exclamation or command (e.g., Fennell & Waxman, 2010). Presenting the word in a sentence context avoided this interpretive difficulty.

The three conditions differed in (1) the frequency with which the distracters occurred with the word and (2) the trial on which children could identify the correct target referent for *jop* based on cross-situational information. In the no-competition condition, each of the distracter objects appeared on the screen only once; only the target object was repeated across trials (see Appendix). In the short- and long-competition conditions, 3 of the distracter objects appeared once, 3 distracter objects appeared twice, and 1 distracter object appeared three times. The object that appeared three times served as the high-probability competitor for the target referent. The short- and long-competition conditions differed in when the high-probability competitor was

absent: in the short-competition condition it was absent in the third trial, while in the long-competition condition it was absent in the fourth trial. Thus, depending on condition, the cross-situational information clearly identified the target on either the second (no-competition), third (short-competition), or fourth (long-competition) novel-word trial. Henceforth, we refer to this trial as the disambiguation trial.

The selection and position of objects for each trial were generated randomly with the following constraints: the target appeared in all trials, the target appeared at least once in both the top/bottom and left/right positions, the distracters occurred with their assigned frequency, the high-probability competitor was absent on the appropriate trial (short- and long-competition conditions only), and no object appeared in same location for more than two consecutive trials. For each condition, two independent object layouts were generated to control for possible quadrant-specific effects.

Coding and analyses

Where participants looked (top-left, bottom-left, top-right, bottom-right, or away) was coded frame-by-frame from silent video by a trained, naïve coder. To assess reliability, 26% of the children's and 26% of the adults' videos were coded by a second naïve coder. The two coders agreed on the children's direction of gaze for 93% of coded video frames and adults' direction of gaze for 94% of the video frames.

Given children's and adults' well-documented tendency to look at objects and events that match what they hear (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), our analyses focused on whether participants looked at the correct object immediately after hearing the familiar or novel word. However, we anticipated that children would exhibit a natural desire to inspect all four objects on the screen and that this would impact how long they persisted in

looking at the target object after hearing it labeled. To determine an appropriate duration for our analysis window, we examined the first 8 children's performance in the familiar-word phase. These children looked at the target object (i.e. ball) for approximately 2s following the onset of the first familiar label. Therefore, for both the familiar-word and novel-word trials we examined where participants looked during a 2s test window that began 200ms after the onset of the first label (this 200ms offset was due to the time it takes to program an eye movement; Tanenhaus et al., 1995); this window was used for both children and adults.

Analysis windows in which the participants looked away from the screen for more than 67% of the window's duration were dropped from the analyses (Children = 4/336 windows; 1 familiar-word, 3 novel-word; Adults = 3/336 windows; 1 familiar-word, 2 novel-word). For the remaining analysis windows, we calculated the proportion of time spent looking to the target out of the total time spent looking at the four objects. Analyses conducted with arcsine-transformed proportions yielded the same pattern of significant results as analyses conducted on untransformed proportions. For ease of interpretation, we report untransformed data here.

Preliminary analyses of the familiar-word phase verified that children could locate the target when 4 objects were present: averaged across the four familiar-word trials, children looked significantly longer at the ball ($M = .39$, $SD = .17$) than expected by chance, $t(41) = 5.31$, $p < .001$, Cohen's $d = .82$. Adults also looked longer at the ball ($M = .71$, $SD = .21$) than expected by chance, $t(41) = 14.39$, $p < .001$, $d = 2.19$.

For the novel-word phase, we expected that children would not show a systematic preference for the target object until the available cross-situational information unambiguously identified the target referent. We therefore conducted our primary analyses on children's proportion of looking time to the target object during the disambiguation trial (no-competition:

novel-word trial 2; short-competition: novel-word trial 3; long-competition: novel-word trial 4). Preliminary analyses of children's performance in the disambiguation trial revealed no effects of sex, object layout, or whether the child's age, vocabulary, or performance in the familiar-word trials was above or below the median, all Fs < 2.07, all ps > .14. These factors were not examined further.

Results

Figure 2 shows the average proportion of looking time to the target object on the disambiguation trial separately by condition. Children in the no- and short-competition conditions looked longer at the target than did those in the long-competition condition, suggesting that children's cross-situational learning varied across conditions. This pattern was confirmed by a one-way analysis of variance (ANOVA) on children's proportion of looking time to the target during the disambiguation trial, which yielded a significant effect of condition, $F(2, 39) = 6.10$, $p = .005$, $\eta^2 = .24$. Planned comparisons indicated that children in the long-competition condition looked significantly less at the target than did children in the no-competition condition, $t(26) = -3.25$, $p = .003$, $d = 1.26$, or the short-competition condition, $t(26) = -3.04$, $p = .005$, $d = 1.17$. The no-competition and short-competition conditions did not differ, t < 1. To further explore these effects, we next examined each condition separately.

[Insert Figure 2 here]

We first asked whether children in the no-competition successfully identified the referent of the novel noun. Recall that in the no-competition condition, no high-probability competitor was present: only the target referent was repeated across trials. Performance in this condition thus provided an indication of children's ability to cope with the increased referential ambiguity in our task, independent of any effects of competition. Children in the no-competition condition

looked significantly longer at the target referent than expected by chance in the disambiguation trial ($M = .44$, $SD = .25$), $t(13) = 2.79$, $p = .015$, $d = .76$. A similar pattern emerged when examining children's proportion of looking time to the target averaged across the second through fourth novel-word trials (i.e. all the trials in which children in the no-competition condition had sufficient cross-situational information to identify the target referent): children looked at the target referent significantly longer than chance ($M = .38$, $SD = .19$), $t(13) = 2.47$, $p = .028$, $d = .68$. Thus, despite the additional referential ambiguity imposed by a larger number of potential referents and a lack of implicit contrast, children in the no-competition condition were able to use cross-situational information to identify the target referent for the novel noun.

To examine the impact of the high-probability competitor on children's cross-situational learning, we next analyzed the performance of children in the short- and long-competition conditions on the disambiguation trial (for these conditions, we did not analyze the proportion of looking time to the target averaged across the second through fourth novel-word trials because this measure would include trials in which the children in the short- and long-competition conditions did not have sufficient information to identify the target referent). Like children in the no-competition condition, children in the short-competition condition looked significantly longer at the target than expected by chance during the disambiguation trial ($M = .40$, $SD = .22$), $t(13) = 2.54$, $p = .025$, $d = .68$. Thus, the presence of a high-probability competitor did not prevent children in the short-competition condition from identifying the referent of the novel noun. In contrast, in the long-competition condition, children's proportion of looking time to the target during the disambiguation trial did not differ from chance ($M = .17$, $SD = .17$), $t(13) = -1.68$, $p = .12$. This suggests children in the long-competition condition were unable to cross-situational information to identify the noun's referent.

One possible explanation for the contrast between the short- and long-competition conditions is that children in these conditions exhibited different baseline preferences for the objects. The poor performance of children in the long-competition condition could have resulted from a lack of interest in the target object rather than an inability to track cross-situational information. Similarly, if children in the long-competition condition had a strong baseline preference for the high-probability competitor, then this might have undermined their ability to gather cross-situational information about the target referent. To address these possibilities, we examined the proportion of looking time to the target and high-probability competitor for the children in the short- and long-competition conditions during the first two novel-word trials. The information available to children in these two conditions was identical during the first two novel-word trials and thus looking patterns should not differ across conditions. A 2x2x2 mixed-model ANOVA with trial (1, 2) and object (target, distracter) as within-subject factors and condition (short-competition, long-competition) as a between-subjects factor revealed no significant effects, all *F*s < 1. This suggests that the difference in performance across the short- and long-competition conditions was not due to different baseline preferences for the target or high-probability competitor.

Finally, we analyzed the adult participants' performance in order to confirm that the novel noun was indeed learnable in all three conditions, and thus that the children's poor performance in the long-competition condition was not due to a problem with this condition. A one-way ANOVA on adults' proportion of looking time to the target on the disambiguation trial revealed no effect of condition, *F* < 1. Planned comparisons confirmed that participants looked significantly longer at the target than expected by chance in the no-competition condition (*M* = .47, *SD* = .32), *t*(13) = 2.59, *p* = .023, *d* = .69, short-competition condition (*M* = .46, *SD* = .30),

t(13) = 2.57, *p* = .023, *d* = .70 and the long-competition condition (*M* = .49, *SD* = .34), *t*(13) = 2.66, *p* = .02, *d* = .71. These results indicate that adult participants had no difficulty coping with the level of referential ambiguity in the task and that the presence of the high-probability competitor did not affect adults' ability to identify the target referent for the novel noun in either the short- or long-competition conditions. This suggests that the noun was in fact learnable in all three conditions, and thus the poor performance of children in the long-competition condition was not due to a task artifact.

Instead, our results suggest that the order of learning observations affected children's cross-situational learning. When the cross-situational information disambiguated the target referent on the third novel-word trial, children in the short-competition condition were able to use this information to identify the novel noun's likely referent. In the long-competition condition, children needed to cope with the competition imposed by the high-probability competitor for one additional trial. This prolonged competition interfered with their cross-situational learning: when the cross-situational information disambiguated the target on the fourth novel-word trial, children in the long-competition condition were not able to identify the noun's likely referent.

Although the difference in performance between the short- and long-competition conditions is striking, it is not unprecedented for small task differences to have pronounced effects on young children's behavior (see recent work on children's social cognition for abundant evidence of this; e.g., Rubio-Fernandez & Geurts, 2013; Scott & Roby, 2015; Yazdi et al., 2006). Such effects likely stem from the fact that children's limited attention and memory abilities are easily overwhelmed by minor increases in complexity. Relative to children in the short-competition condition, children in the long-competition condition needed to attend to multiple

objects for one additional trial, store an additional set of co-occurrence probabilities, and retrieve this information one additional time. This additional burden overwhelmed children in the long-competition condition, preventing them from identifying the word's likely referent.

Further results

Given concerns that have been raised regarding analyzing proportion data (transformed or otherwise) using ANOVA (e.g., Jaeger, 2008), we confirmed our main analyses using an empirical logit mixed effects model that included participant as a random effect. These analyses were conducted with R 3.1.2 (R Core Team, 2014) using the *nlme* package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2016) and planned comparisons were performed using the *multcomp* package (Hothorn, Bretz, & Westfall, 2008). The dependent measure was the empirical logit transform of the proportion of time spent looking at the target. Supporting our previous results, a linear mixed-effects model on children's proportion of looking to the target during the disambiguation trial yielded a significant effect of condition, $F(2, 39) = 4.05$, $p = .03$. Planned comparisons indicated that children in the long-competition condition looked significantly less at the target than did children in the no-competition condition, $z(26) = -2.52$, $p = .03$, or the short-competition condition, $z(26) = -2.34$, $p = .05$. The no-competition and short-competition conditions did not differ, $z < 1$. An identical model run on the adult's proportion of looking time to the target during the disambiguation trial revealed no significant effect of condition, $F < 1$.

General Discussion

Recent studies suggest that children are able to use cross-situational information to identify the referents of novel nouns under at least some circumstances (e.g., Smith & Yu, 2008). However, these prior studies have presented children with minimal referential ambiguity and

simplified cross-situational statistics. The present study examined whether 2.5-year-olds' cross-situational word learning would scale up to more challenging conditions in which both the referential scene and cross-situational information were more ambiguous. Children were presented with a series of trials in which a single novel word was accompanied by four novel objects. Only the target referent consistently co-occurred with the word across trials; the frequency with which the distracter objects occurred varied across conditions. When all of the distracter referents occurred with equally low probability (no-competition condition), children were able to successfully identify the target referent for the novel word. However, when a high-probability competitor was present on three of the four trials, children were able to identify the target referent only if the high-probability competitor was absent on the third trial (short-competition condition) but not if it was present until the fourth trial (long-competition condition). In contrast to the 2.5-year-olds, adults were able to identify the target referent in all three conditions.

These results expand our understanding of early word learning in several ways. First, the positive results in the no- and short-competition conditions replicate prior findings (e.g., Smith & Yu, 2008; Suanda et al., 2014; Yu & Smith, 2011) by demonstrating that when children encounter a novel noun in the presence of multiple potential object referents, they can use cross-situational information to determine to which object the word refers. Our results also extend prior findings by showing that 2.5-year-olds can engage in successful cross-situational noun learning even when faced with referential scenes that contain four candidate referents, and without the support of implicit contrast information. Thus, our results suggest that 2.5-year-olds' cross-situational word learning scales up to more ambiguous referential contexts than have been used in prior research.

Whether the referential contexts in the present experiment are more or less ambiguous than the average word learning situations that children in encounter in everyday life remains an open question, as there is currently considerable debate about how ambiguous typical learning situations are from the child's perspective (Medina et al., 2011; Yurovsky, Smith, & Yu, 2013). Our results simply demonstrate that by 2.5 years of age, children are capable of engaging in cross-situational word learning when confronted with a single noun accompanied by four or fewer potential referents. It is also unclear whether younger children would be able to make use of cross-situational information under similar circumstances. Recall that recent observational work by Smith, Yu, and colleagues suggests that 1.5-year-olds have difficulty using cross-situational information to learn nouns when faced with as few as three potential referents (Pereira et al., 2014; Suanda et al., 2013; Yu & Smith, 2012). Future work will need to examine whether the amount of referential ambiguity that children can cope with while tracking cross-situational information increases with age.

More generally, the positive results of the no- and short-competition conditions add to a growing body of research suggesting that when children encounter a word under referential uncertainty, they are capable of establishing a lexical entry for that word and attaching to it facts about the contexts in which the word occurs (e.g., Arunachalam & Waxman, 2010; Messenger, Yuan, & Fisher, 2015; Scott & Fisher, 2009, 2012; Smith & Yu, 2008; Suanda et al., 2014; Yuan & Fisher, 2009). This can include information about the accompanying referential scene, as demonstrated in the present experiment, as well as linguistic facts about the sentences in which the word is used (e.g., Arunachalam & Waxman, 2010; Messenger et al., 2015; Scott & Fisher, 2009; Yuan & Fisher, 2009). Children can subsequently retrieve these facts and use them to guide their interpretation of that word when they encounter it again in the future. This body of

literature thus paints a picture of a developing lexicon in which lexical entries gather up crumbs of information about words, allowing children to gradually refine their interpretation of words over time (e.g., extended mapping; Carey, 1978).

Our study also revealed clear limits on children's cross-situational learning. The negative results of the long-competition condition suggest that the consistent occurrence of a competitor referent across observations can interfere with 2.5-year-olds' cross-situational learning, and that the impact of this competitor depends in part on the order in which learning observations occur. Children in the short-competition and long-competition conditions encountered the same cross-situational information about the potential referents for the novel word. The only difference across conditions was the order in which this information was presented. Children in the short-competition condition only had to consider multiple competing referents for two consecutive trials, while children in the long-competition condition had to consider competing referents for three trials. This subtle difference proved crucial to children's performance: children in the short-competition condition successfully identified the noun's referent, but children in the long-competition failed to do so. Children in these conditions did not differ in their attention to the target or high-probability competitor during the first two learning trials. This suggests that children in these two conditions were equally willing to entertain these two objects as potential referents for the target word, and that the discrepancy between these conditions did not result from different baseline preferences for either object.

Instead, we speculate that children in the long-competition condition performed more poorly because the presence of the high-probability competitor prolonged the process of identifying the target referent for the novel noun. This need to consider multiple competing referents for an additional trial increased the memory burden of the task, thereby overwhelming

the children in the long-competition condition. It is noteworthy that this difficulty occurred despite the fact that our task, while more ambiguous than prior cross-situational learning tasks, still involved several simplifications that should have *eased* the memory burden imposed on children. Children encountered the novel noun in four consecutive trials that occurred only a few seconds apart and thus children did not need to retain the cross-situational information for long periods of time. In addition, these four observations of the novel noun occurred without any other novel words interleaved between them. Recent evidence suggests that both children and adults have greater difficulty aggregating cross-situational information across observations of a given word when those observations are interleaved with observations of other novel words (e.g., Smith et al., 2011; Vlach & Johnson, 2013), in part because of the greater memory burden imposed by encoding, retrieving, and comparing multiple sets of candidate referents. The fact that children in the long-competition condition failed even with consecutive observations that occurred close together in time indicates that the burden imposed by the high-probability competitor referents was substantial.

Our results thus suggest that when confronted with prolonged competition between referents, children may require support from other sources in order to make use of available cross-situational information in word learning. What might enable children to overcome the difficulties associated with competition? One possibility is that children's ability to cope with competition could depend on higher-order properties of the context in which that competition occurs. Recent cross-situational learning studies have shown that adults who encountered novel nouns in themed referential contexts where all of the objects were from the same category (e.g., animals, clothing, things in the kitchen) retained more information about the potential referents for each word across observations (Dautriche & Chemla, 2014) and demonstrated higher rates of

word learning (Chen & Yu, 2015; Dautriche & Chemla, 2014) than did adults who encountered the same nouns in non-themed contexts involving unrelated objects. Moreover, the advantage of themed over non-themed contexts held even when the former involved greater competition between referents because individual distracter referents co-occurred more frequently with the target (Chen & Yu, 2015). Together, these results suggest that encountering words in semantically coherent or themed contexts facilitates adults' encoding and retrieval of potential referents, and this offsets the challenges imposed by competition. If such contexts also facilitated children's encoding and retrieval of potential referents, then this might ease the memory burden they experience when tracking competing referents for a word. This in turn might allow children to make use of cross-situational information despite prolonged competition between referents.

This leads to a second question raised by our findings: to what extent, and in what types of referential contexts, are children confronted with competing referents in everyday word learning? Recent discussions about the nature of typical word learning situations have focused largely on the number of potential referents present in the referential scene (e.g., Medina et al., 2011; Yurovsky, Smith, & Yu, 2013) as well as the relative salience of target and distracter referents in the child's field of view (e.g., Pereira et al., 2014; Suanda et al., 2013). However, to our knowledge, no studies have addressed which referents are present in typical referential scenes and, more specifically, whether or not high-probability competitor referents are present. Given that the presence of a high-probability competitor undermines 2.5-year-olds' learning of a single novel noun in simplified laboratory conditions, it seems likely that high-probability competitors could have significant impact on real-world cross-situational learning. Quantifying the nature and extent of competition in everyday referential contexts therefore has important implications for the role of cross-situational information in early word learning.

Cross-situational learning mechanisms

Our results add to a growing body of evidence that under at least some circumstances, children and adults can use cross-situational information to learn words. These findings have raised many questions regarding the nature of the mechanism that underlies this cross-situational word learning ability. At present, two broad possibilities have been outlined in the literature.

One possibility, ACCUMULATIVE LEARNING, is that learners simultaneously accrue information about an entire set of potential referents for a given word (Fazly et al., 2010; Smith & Yu, 2008; Yu, 2008; Yurovsky, Fricker, Yu, & Smith, 2014). When learners first encounter a new word, they encode whatever referents co-occurred with that word. This co-occurrence information could be in the form of low-level associations between words and scene elements (e.g., Smith & Yu, 2008) or candidate interpretations that learners generate using the linguistic and non-linguistic cues described in the Introduction (e.g., Frank, Goodman, & Tenenbaum, 2009; Siskind, 1996). The next time they encounter the word, learners compare the current set of potential referents to the set previously stored in memory, adding new possibilities and updating the co-occurrence probabilities for previously encountered referents.

An alternative possibility is that learners, especially young children, are unable to track all of the candidate referents that co-occur with a word. Instead, learners might engage in CONJECTURE-BASED LEARNING (e.g., Medina et al., 2011; Trueswell, Medina, Hafri, & Gleitman, 2013). On this view, when learners encounter a new word, they make a guess or conjecture about what the word refers to using whatever linguistic and non-linguistic cues are available. Learners retain this single hypothesis, discarding information about alternative referents. The next time learners encounter the word, they retrieve and evaluate their conjecture. If the hypothesized referent is present, then they strengthen and retain the hypothesis. If the hypothesized referent is

absent, then the hypothesis is discarded and learners generate a new guess based on the current referential scene.

On both of these accounts, learners should eventually converge on the word's correct referent. However, these accounts disagree on how much information learners retain about the potential referents for a word as well as how the learning process should unfold. These accounts also generate different predictions regarding children's performance in our experiment. Although our task was not designed to test these competing accounts, a comparison of these predictions with our observed findings reveals several implications for these accounts and the mechanisms that support cross-situational word learning.

From an accumulative-learning perspective, the children in our task should have gathered information about the set of potential referents that occurred with the word on each observation. This should have led children in all conditions to converge on the target referent as the most probable referent for *jop* because it co-occurred with the word more frequently than any other object. The fact that children in the long-competition condition failed to identify the target referent is inconsistent with this prediction, suggesting that 2.5-year-olds are not always capable of tracking the set of potential referents for a word (see also Scott & Fisher, 2012).

Did the children in our task instead track only a single hypothesized meaning? According to the conjecture-based account, on the first novel-word trial children would have made an initial guess at the word's meaning. If they happened to select the target, then their guess would be confirmed on all subsequent trials. This predicts that children who guessed correctly in the first trial should have succeeded on the disambiguation trial, regardless of condition. In contrast, if children guessed incorrectly in the first novel-word trial, then at some point their guess would be disconfirmed. In the no-competition condition, children who initially selected any object other

than the target would have their guess disconfirmed on the subsequent trial, as no distracter objects were repeated across trials. In the short- and long-competition conditions, children who initially guessed the high-probability competitor would have their guess confirmed until the disambiguation trial, when the high-probability competitor was absent. These three groups of children should perform at chance on their respective disambiguation trials because they would need to select a new guess at random from the available referents.

If children in the short- and long-competition conditions initially selected an object other than the target or high-probability competitor, their guess would be disconfirmed on the second trial and they would need to guess again at random. Because this results in many possible patterns of guesses across the first several trials, for these conditions we focus on children who initially selected either the target or high-probability distracter, for whom predictions are more straightforward.

In order to evaluate the predictions of the conjecture-based account, we took the object that children looked at longest on the first novel-word trial as their initial hypothesized referent for the novel word. We then examined individual children's performance on the disambiguation trial as a function of their initial guess. As can be seen in Table 2, the resulting patterns of performance are inconsistent with the predictions of the conjecture-based account in two ways. First, of the four children who initially guessed correctly in the long-competition condition, only one succeeded in the disambiguation trial. This poor performance, which is worse than predicted by the conjecture-based account, suggests that the children were hampered by their attempt to track multiple competing referents for a prolonged period of time. Second, children in the no-competition condition who guessed incorrectly on the first trial succeeded in the disambiguation trial, as did children in the short-competition condition who initially selected the high-probability

competitor. If these children retained only their single hypothesized referent across trials, then they should have performed at chance when their hypothesis was disconfirmed. The fact that these two groups succeeded in the disambiguation trial suggests they were tracking more than one potential referent, enabling them to recover when their preferred referent was no longer present.

[Insert Table 2 here]

These results suggest that children retained more than a single conjecture across learning observations. However, as noted above, children in the long-competition condition performed worse than would be expected if they tracked the co-occurrence probabilities for the entire set of candidate referents. Our results are thus most consistent with recent models that offer a middle ground between accumulative and conjecture-based approaches (e.g., Smith et al., 2011; Yurovsky & Frank, 2015). Although adults and children are sometimes capable of accumulating information about multiple potential referents for a word (e.g., Vouloumanos & Werker, 2009; Yurovsky et al., 2014), their ability to do so depends on the difficulty of the learning situation and the demands imposed on attention and memory (e.g., Smith et al., 2011; Yu & Smith, 2011; Yurovsky & Frank, 2015). Recent studies suggest that increasing the number of potential referents in a scene or the length of time between observations can impair learners' ability to track multiple candidate referents (Smith et al., 2011; Vlach & Johnson, 2013; Yurovsky & Frank, 2015). Our study indicates that children's cross-situational learning is also affected by the presence of high-probability competitor referents. Thus, the effects of competition should be incorporated into future models of children's cross-situational word learning.

References

Akhtar, N. & Montague, L. (1999). Early lexical acquisition: the role of cross-situational learning. *First Language* **19***,* 347-358.

Ankowski, A. A., Vlach, H. A. & Sandhofer, C. M. (2013). Comparison versus contrast: task specifics affect category acquisition. *Infant and Child Development* **22**, 1-23.

Arunachalam, S. & Waxman, S. R. (2010). Meaning from syntax: evidence from 2-year-olds. *Cognition* **114***,* 442-446.

Aslin, R. N., Woodward, J. Z., LaMendola, N. P. & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (eds.), *Signal to syntax*, 117-134. Mahwah, NJ: Erlbaum.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language* **20**, 395-418.

Bernal, S., Lidz, J., Millotte, S. & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development* **3**, 325-341.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan & G. A. Miller (eds.), *Linguistic theory and psychological reality*, 264–293. Cambridge, MA: MIT Press.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D. & Christophe, A. (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development* **10**, 1-18.

Chen, C. & Yu, C. (2015). The effects of learning and retrieval contexts on cross-situational word learning. In *Proceedings of the Joint IEEE International Conference on Development and Learning and Epigenetic Robotics* (*ICDL-Epirob*).

Childers, J. B. (2011). Attention to multiple events helps two-and-a-half-year-olds extend new

verbs. *First Language* **31***,* 3-22.

Childers, J. B., Hirshkowitz, A. & Benavides, K. (2014). Attention to explicit and implicit

contrast in verb learning. *Journal of Cognition and Development* **15**, 213-237.

Childers, J. B. & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational

information to learn novel predicate terms. *Journal of Child Language* **36**, 201-224.

Dale, P. S. & Fenson, L. (1996). Lexical development norms for young children. *Behavior*

Research Methods, Instruments, & Computers **28**, 125-127.

Dautriche, I. & Chemla, E. (2014). Cross-situational word learning in the right situations.

Journal of Experimental Psychology: Learning, Memory, and Cognition **40***,* 892-903.

Fazly, A., Alishahi, A. & Stevenson, S. (2010). A probabilistic computational model of cross-

situational word learning. *Cognitive Science* **34**, 1017-1063.

Fennell, C. T. & Waxman, S. R. (2010). What paradox? referential cues allow for infant use of

phonetic detail in word learning. *Child Development, 81,* 1376-1383.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S. & Bates, E. (2007).

MacArthur-Bates communicative development inventories. Baltimore, MD: Paul Brookes

Publishing.

Fisher, C., Hall, D. G., Rakowitz, S. & Gleitman, L. R. (1994). When it is better to receive than

to give: syntactic and conceptual constraints on vocabulary growth. *Lingua* **92***,* 333-375.

Fisher, C., Klingler, S. L. & Song, H. J. (2006). What does syntax say about space? 2-year-olds

use sentence structure to learn new prepositions. *Cognition* **101**, B19-B29.

Frank, M. C., Goodman, N. & Tenenbaum, J. (2009). Using speakers' referential intentions to

model early cross-situational word learning. *Psychological Science* **20***,* 578–585.

Gillette, J., Gleitman, H., Gleitman, L. & Lederer, A. (1999). Human simulations of vocabulary

learning. *Cognition* **73**, 135-176.

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

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

Harris, M., Jones, D. & Grant, J. (1983). The nonverbal context of mothers' speech to infants. *First Language* **4**, 21-30.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal* **50**, 346-363.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**, 434-446.

Jørgensen, R. N., Dale, P. S., Bleses, D. & Fenson, L. (2010). CLEX: A cross-linguistic lexical norms database. *Journal of Child Language* **37**, 419-428.

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

Medina, T. N., Trueswell, J., Snedeker, J. & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *PNAS* **108**, 9014-9019.

Messenger, K., Yuan, S. & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development* **11**, 356-368.

Mintz, T. H. (2006). Finding the verbs: distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (eds.), *Action meets word: how children learn verbs*, 31-63. New York: Oxford University Press.

Namy, L. L. & Clepper, L. E. (2010). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology* **107**, 291-305.

Nappa, R., Wessel, A., McEldoon, K. L., Gleitman, L. R. & Trueswell, J. C. (2009). Use of speaker's gaze and syntax in verb learning. *Language Learning and Development* **5**, 203–234.

Pereira, A. F., Smith, L. B. & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review* **21**, 178-185.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2016). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-124, http://CRAN.R-project.org/package=nlme.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rubio-Fernández, P. & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science* **24**, 27-33.

Scott, R. M. & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes* **24**, 777-803.

Scott, R. M. & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition* **122**, 163-180.

Scott, R. M. & Roby, E. (2015). Processing demands impact 3-year-olds' performance in a spontaneous-response task: New evidence for the processing-load account of early false-belief understanding. *PLoS ONE* **10**, e0142405.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition* **61**, 39-91.

Smith, K., Smith, A. D. M. & Blythe, R. A. (2011). Cross-situational learning: an experimental study of word-learning mechanisms. *Cognitive Science* **35***,* 480-498.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L. & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science* **13**, 13-19.

Smith, L. B. & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* **106***,* 1558-1568.

Suanda, S. H., Foster, S. B., Smith, L. B. & Yu, C. (2013). Attentional constraints and statistics in toddlers' word learning. In *Proceedings of the Joint IEEE International Conference on Development and Learning and Epigenetic Robotics* (*ICDL-Epirob*).

Suanda, S. H., Mugwanya, N. & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology* **126**, 395-411.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* **268***,* 1632-1634.

Trueswell, J. C., Medina, T. N., Hafri, A. & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology* **66***,* 126-156.

Vlach, H. A. & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition* **127**, 375-382.

Vouloumanos, A. & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology* **45***,* 1611-1617.

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. & Klibanoff, R. S. (2000). The role of comparison in the extension of novel

adjectives. *Developmental Psychology* **36***,* 571-581.

Waxman, S. R., Lidz, J. L., Braun, I. E. & Lavin, T. (2009). Twenty four-month-old infants'

interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology* **59**, 67-

95.

Yazdi, A. A., German, T. P., Defeyter, M. A., & Siegal, M. (2006). Competence and

performance in belief-desire reasoning across two cultures: The truth, the whole truth,

and nothing but the true about false belief? *Cognition* **100**, 343-368.

Yuan, S. & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn

combinatorial facts about verbs by listening. *Psychological Science* **20***,* 619–626.

Yuan, S., Fisher, C. & Snedeker, J. (2012). Counting the nouns: simple structural cues to verb

meaning. *Child Development* **83**, 1382-1399.

Yu, C. (2008). A statistical associative account of vocabulary growth in early word learning.

Language Learning and Development **4***,* 32-62.

Yu, C. & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational

statistics. *Psychological Science* **18***,* 414-420.

Yu, C. & Smith, L. B. (2011). What you learn is what you see: using eye movements to study

infant cross-situational word learning. *Developmental Science*, *14*, 165-180.

Yu, C. & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition* **125***,*

244-262.

Yurovsky, D. & Frank, M. C. (2015). An integrative account of constraints on cross-situational

learning. *Cognition* **145***,* 53-62.

Yurovsky, D., Fricker, D. C., Yu, C. & Smith, L. (2014). The role of partial knowledge in

statistical word learning. *Psychonomic Bulletin & Review* **21***,* 1-22.

Yurovsky, D., Hidaka, S., Yu, C. & Smith, L. B. (2010). Linking learning to looking: habituation and association in infant statistical language learning. In S. Ohlsson & R. Catrambone (eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 1589-1594. Austin, TX: Cognitive Science Society.

Yurovsky, D., Smith, L. B. & Yu, C. (2013). Statistical word learning at scale: the babies view is better. *Developmental Science* **16***,* 959-966.

Yurovsky, D., Yu, C. & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive Science* **37***,* 891-921.

Zhang, Z., Shi, R. & Li, A. (2015). Grammatical categorization in Mandarin-Chinese-learning infants. *Language Acquisition* **22***,* 105-114.

Appendix

Matrix of the Word-Object Co-occurrence Frequencies Across Conditions

	No-competition	Short-competition	Long-competition
Object 1	4	4	4
Object 2	1	3	3
Object 3	1	2	2
Object 4	1	2	2
Object 5	1	2	2
Object 6	1	1	1
Object 7	1	1	1
Object 8	1	1	1
Object 9	1		
Object 10	1		
Object 11	1		
Object 12	1		
Object 13	1		

Note: The above table illustrates the frequency with which each object co-occurred with the novel word *jop* in each condition. In all conditions, object 1 is the target referent. For the short- and long-competition conditions object 2 is the high-probability competitor.

Table 1

Objects Used in the Familiar-word and Novel-word Phases of the Experiment

Familiar	Novel
Ball, Banana, Book, Bottle, Car, Cup, Keys, Socks	Apple slicer, Bottle opener, Bottle stopper, Box cutter, Cherry pitter, Egg scrambler, Grip exerciser, Juicer, Lint brush, Light saber, Massager, Oil wrench, **Oven mitt,** Peeler, Potato masher, Scoop

Note: Bold indicates the target object during the respective phase of the experiment.

Table 2

Proportion of Children Who Succeeded in the Disambiguation Trial, Separately by Condition and Initial Guess

Initial guess	No-competition	Short-competition	Long-competition
Target	5/5	4/5	**1/4**
High-probability competitor	--	**3/4**	1/4
Other	**7/9**	3/4	3/6

Note: Initial guess was defined as the object the child looked longest at during the first novel-word trial. Success was defined as looking at the target numerically longer than chance (.25). One child in the short-competition condition was excluded for looking away during the first novel-word analysis window. Bold cells are inconsistent with predictions from the conjecture-based account.
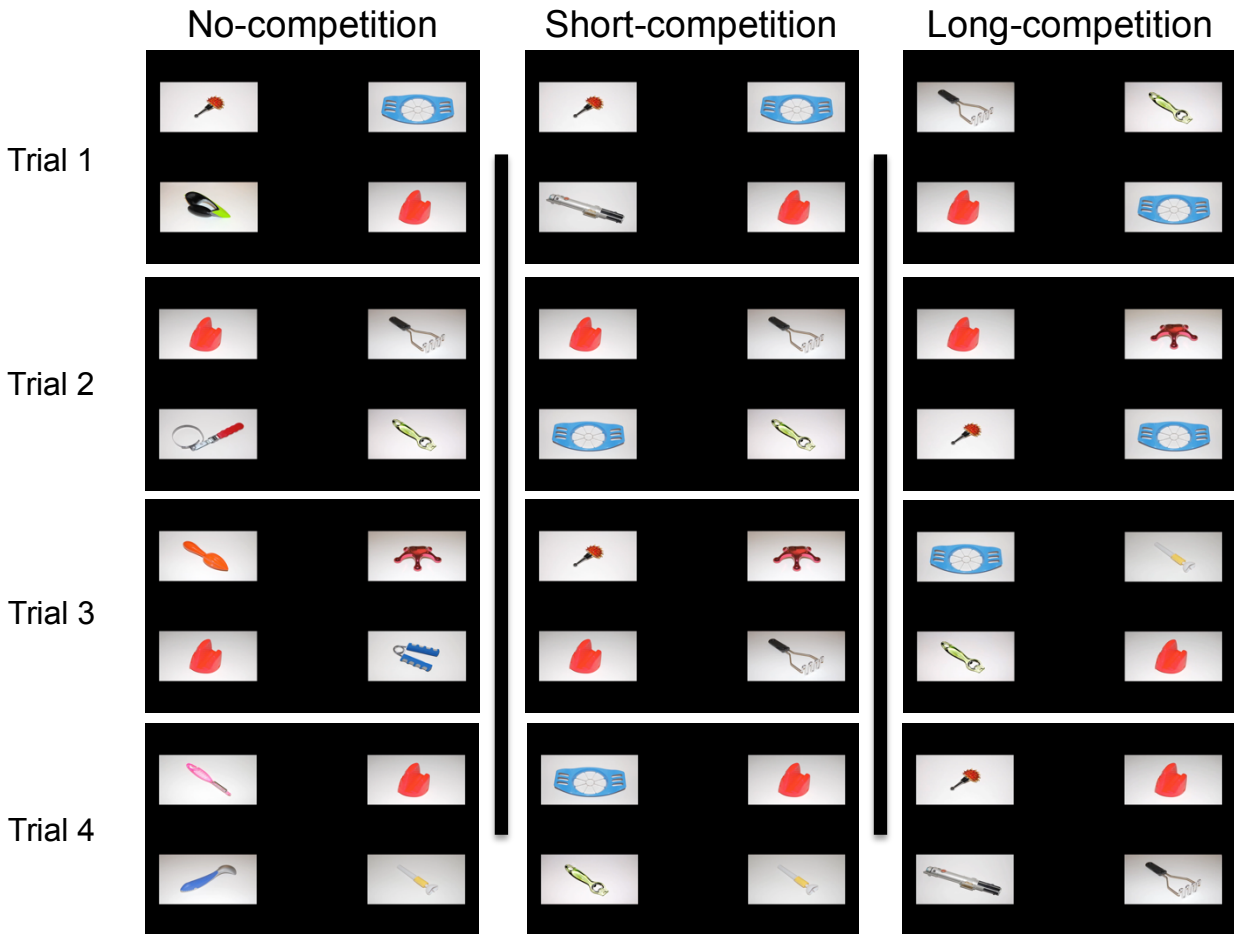
Figure 1. **Video sequences from the novel-word phase in the no-competition (left), short-competition (center), and long-competition (right) conditions.** On each trial, children heard the same novel word (e.g., "Look at the jop!"). Across trials and conditions, only the target object (red oven mitt; bottom-right object in Trial 1 of no-competition condition) consistently occurred with the novel word. In the no-competition condition, each distracter appeared on only one trial. For the short- and long-competition conditions, the high-probability distracter (blue apple slicer; top-right object in Trial 1 of short-competition) occurred on three of the four trials.
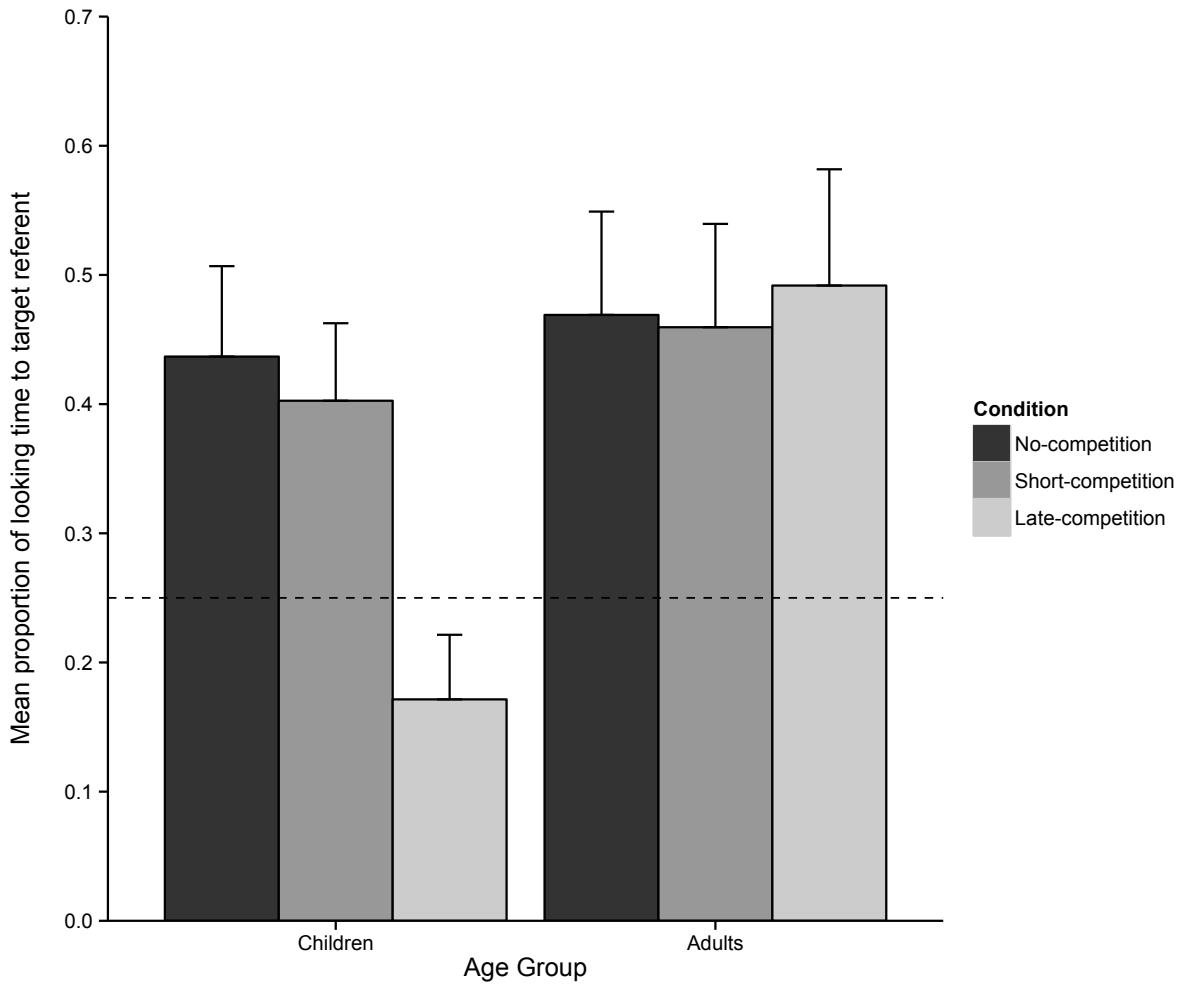
Figure 2. **Results for children and adults.** Mean proportion of looking time for children and adults during the disambiguation trial of the novel-word phase, separately by condition. Error bars represent one standard error of the mean. The dashed line indicates chance performance.