Developmental Science

Developmental Science 19:1 (2016), pp 63–75 DOI: 10.1111/desc.12289

PAPER

Discourse bootstrapping: preschoolers use linguistic discourse to learn new words

Jessica Sullivan[1] and David Barner[2]

1. Department of Psychology, Skidmore College, USA
2. Department of Psychology, University of California, San Diego, USA

Abstract

When children acquire language, they often learn words in the absence of direct instruction (e.g. 'This is a ball!') or even social cues to reference (e.g. eye gaze, pointing). However, there are few accounts of how children do this, especially in cases where the referent of a new word is ambiguous. Across two experiments, we test whether preschoolers (2- to 4-year-olds; n = 239) can learn new words by inferring the referent of a new word from the surrounding linguistic discourse. Across two experiments, we show that children as young as 2 can learn a new word from the linguistic discourse in which it appears. This suggests that children use the linguistic discourse in which a word appears to learn new words.

Research highlights

- We tested the abilities of more than 200 preschoolers to relate new words to the content of the discourse in which they appear.
- We showed that children as young as 2 can successfully learn new words by relating them to the discourse context, and to the intentions of a speaker.
- We showed developmental changes – between the ages of 2 and 4 – in children's ability to learn new words from discourse context.
- This research provides compelling evidence that higher-level discourse and pragmatic processes may be fundamental to word learning.
- This research also suggests – and provides empirical evidence for – one way that children might learn new words when other attested word learning mechanisms fall short.

Introduction

While children sometimes learn words from explicitly instructive contexts (e.g. a parent pointing to an object and naming it), much of their language exposure is overheard (Snow & Ferguson, 1977), not explicitly instructive (Jaswal & Markman, 2001), or occurs in reference to events that have not yet happened (Tomasello & Kruger, 1992). This raises a challenge for word learning, because children must often attach content to new words on the basis of only indirect evidence (Akhtar, Jipson & Callanan, 2001; Tomasello, Strosberg & Akhtar, 1996). Despite this challenge, children learn words quickly in early acquisition (e.g. Goldfield & Reznick, 1990). Here, we explore one way that children might learn words on the basis of indirect evidence – by inferring the relevance of new words to the surrounding linguistic discourse and to the speaker's communicative intentions. To foreshadow our results, we find some of the first evidence that preschoolers can learn new words by inferring their relation to the content of the surrounding linguistic discourse as a whole, without other cues to reference. We also find that there are important developmental changes in this ability, and that these changes are best explained by children's developing world knowledge.

How do children learn new words in cases of referential ambiguity? Previous researchers have proposed that

Address for correspondence: Jessica Sullivan, Skidmore College Department of Psychology, Tisch Learning Center, #101, 815 N. Broadway, Saratoga Springs, NY 12866, USA; e-mail: jsulliv1@skidmore.edu

children rely on constraints (or simplifying assumptions) to reduce the number of possibilities for what a new word might refer to: Children assume that novel nouns label whole objects (Macnamara, 1972; Markman 1990; Soja, Carey & Spelke, 1991), that words contrast with one another and are often mutually exclusive (Clark, 1983, 1988, 1990; Markman, 1990; Markman & Wachtel, 1988), and that nouns encode taxonomic categories (e.g. Markman & Hutchinson, 1984). Social cues such as pointing and eye gaze, which socio-pragmatic theories of word learning claim are central to word learning, can also narrow the range of possible referents for a new word (Baldwin, 1993; Carpenter, Nagell & Tomasello, 1998; Caza & Knott, 2012; Grassman & Tomasello, 2010; Meyer & Baldwin, 2013; Tomasello & Barton, 1994). However, socio-pragmatic cues cannot always fully specify the referent of a novel word (Frank, Tenenbaum & Fernald, 2013; Meyer & Baldwin, 2013; Quine, 1960), and they are only useful in contexts where the speaker points to or looks at the object being discussed. In addition to socio-pragmatic cues, children can also use linguistic information – such as syntax – to make inferences about the referents of new words. For example, via a process known as syntactic bootstrapping (Gleitman, 1990), children make different inferences about the type of action that *glorp* refers to when the word is presented in a frame like, 'I'm going to *glorp* it to John' as opposed to 'I'm going to *glorp* John' (e.g. Almoammer, Sullivan, Donlan, Marusic, Zaucer *et al.*, 2013; Barner, 2012; Barner & Snedeker, 2006; Brown, 1957; Bloom & Wynn, 1997; Landau & Gleitman, 1985; Naigles, 1996; Sarnecka, Kamenskaya, Yamana, Ogura & Yudovina, 2007; Soja, 1992; Waxman & Booth, 2001).

Each of the attested word learning mechanisms described above fails to accurately determine reference in many contexts. To understand the limitations of these previous approaches, consider a scene where two equally salient (but previously unnamed) objects sit on a table, and a speaker, without looking up, asks for one of them using a novel label in count syntax: 'Give me the *blick*'. Although count syntax tells us that the speaker has requested a countable individual, the referent of the novel word cannot be inferred from syntax alone, since many words with identical syntactic profiles differ substantially in meaning and reference (e.g. see Pinker, 1989, for discussion). In such cases, other attested word learning mechanisms also fall short: mutual exclusivity, the whole object constraint, and the taxonomic constraint fail to distinguish between the two referents. And, in cases where pointing or eye gaze are unavailable, so too do socio-pragmatic cues. This raises the question: when attested word learning mechanisms fail, or are unavailable (e.g. in cases where the referent of a new word is not visible at the moment of naming), how do children make use of their linguistic input to learn new words?

In the present study, we explored the idea that children learn words by making inferences about their likely meanings on the basis of discourse context. Specifically, we tested whether children can identify the referent of new words by relating them to the content of the discourse contexts in which they appear (Frank *et al.*, 2013; Rohde & Frank, 2014). According to Grice (1975), language users generally assume that their interlocutors will be truthful, clear, relevant, and maximally informative in conversation. As part of this, they assume that utterances will be related to one another; that sentences combine to form a coherent discourse. This assumption, that speakers contribute relevant utterances to a conversation (Sperber & Wilson, 1986), is critical to the present study, since it implies that words within utterances – including novel words – will be meaningfully related to other words and ideas raised by the discourse. For example, if a speaker utters the sentence in (1), the listener can make the inference (a) that going out is related to hunger in some way, and therefore (b) that the speaker has likely proposed going out in order to eat.

(1) 'I'm hungry. Want to go out?'

This type of inference operates over the relation between words within the discourse, and also over the relation between the discourse and the speaker's mental states and communicative intentions. Critically, this inference is not dependent purely on the content of any individual word or clause in isolation: Utterances like those in (2) and (3) do not lead to the inference that the speaker has proposed going to a restaurant (at least not without substantial additional context). Thus, the *relations* between words (not the individual words themselves) determine the discourse coherence.

(2) 'I feel like dancing. Want to go out?'
(3) 'I'm hungry. But I'm not done with my homework yet.'

The idea that children might infer the meanings of words by making assumptions about discourse coherence – an idea that we call discourse bootstrapping – stems from a larger theoretical literature which proposes that children leverage their understanding of the communicative and referential intentions of their interlocutors to infer meaning (for an early example, see Pinker, 1979, p. 243). However, although the idea seems intuitive, no previous studies have actually tested it.

The idea that inferences about the relation between novel words and the surrounding discourse context could underlie early word learning is plausible given recent studies that have shown that preschoolers can infer how elements within a discourse are relevant to one another and to the speaker's communicative intentions in some contexts (Akhtar, Carpenter & Tomasello, 1996; Stiller, Goodman & Frank, in press; Schulze, Grassman & Tomasello, 2013; but see Verbuk & Shultz, 2010). However, despite evidence that children may possess the requisite pragmatic and discursive skills to infer the relation between novel words and the surrounding discourse, this idea has not yet been empirically tested in the word learning literature. Those studies that come closest (e.g. Akhtar, 2002) still do not differentiate children's use of linguistic cues to salience (e.g. using the fact that the experimenter previously mentioned shapes to infer that a new word refers to a shape) from their use of discourse coherence when inferring the referents novel words. Only the latter involves making an inference about the relation between words within a discourse and the speaker's intentions, and only *it* predicts that children should correctly identify the referent of a new word in cases where the referent is *not* the item made most salient by the surrounding discourse. For example, in a sentence like 'I'm hungry, but sweet things hurt my teeth. Look at what I want, there's a *gazz!*', a child (or adult) who attends to the relation between the novel word *gazz* and the content of the surrounding discourse should arrive at the (correct) inference that *gazzes* are non-sweet foods. This involves (1) realizing that a speaker who mentions being hungry is likely to talk about food; (2) realizing that the speaker indicated an implicit dispreference for sweet things; and (3) realizing that the novel word *gazz* is importantly related to (1) and (2). In contrast, a child who attends to individual words (e.g. *sweet*) as cues to salience might arrive at the incorrect conclusion that a *gazz* could be anything sweet (e.g. a sweet drink or food). Unlike inferences based on social cues (like eye gaze and pointing) or linguistic cues to salience, discourse-based inferences about the referents of unfamiliar words emerge from understanding the relevance of an unfamiliar word to the communicative content of the discourse as a whole, and to the speaker's communicative intentions.

There are, of course, reasons to question whether discourse coherence could play a role in word learning, especially in early language acquisition. First, while there is some recent evidence that the ability to compute the relevance between utterances emerges early (Schulze *et al*., 2013), other work has shown that preschoolers have difficulty computing relevance, unless they have access to other socio-pragmatic cues (like eyegaze; Tribushininia, 2012). In addition, there is a large literature suggesting that children struggle with many pragmatic and social inferences, like theory of mind (see Wellman, Cross & Watson, 2001, for review). Finally, elementary school educators often explicitly *teach* children how to guess the meaning of a new word based on its relation to the surrounding discourse (termed 'using context clues'; see Fukkink & de Glopper, 1998, for review), and a sizeable literature details the situations under which elementary school-aged children, teenagers, and adults learn (or fail to learn) new words via their relevance to the surrounding text (Jenkins, Stein & Wysocki, 1984; McCullough, 1943; Nagy, Anderson & Herman, 1987). Given that even older children can have substantial difficulties learning new words via their relation to the surrounding discourse – it is unclear whether we should expect preschoolers to successfully leverage the content of a discourse to infer the referent of new words.

In the present study, we take the first step in exploring the role of discourse in bootstrapping the meanings of new words. Specifically, we explore the idea that children might use sentences containing familiar words to constrain their interpretation of novel words (e.g. Frank, Goodman & Tenenbaum, 2009; Pinker, 1979, 1989, 1994). In two experiments, we introduced children to a novel word (e.g. *pliff*) within a short discourse, and then asked whether they could use the discourse to identify the referent of the word (and, by extension, learn it). In Experiment 1, in order to learn the new word, children were required to infer the relation between a puppy's indirect statement of preference (e.g. 'I'm thirsty') and the novel word. In Experiment 2, we tested whether children successfully learned new words either (a) by attending to individual words within the discourse (e.g. *thirsty*) as cues to salience or (b) by modeling the relation between the novel word, the content of the discourse as a whole, and the communicative intentions of the speaker. We did this by presenting sentences in which the intended referent was *not* the item made most salient by the individual words within the utterance. For example, children saw sets of three items (e.g. a pretzel, a cookie, and a shoe) and heard a statement like, 'I'm hungry, but salty things makes me sick. Look at what I want! There's a *gazz* on the table.' Here, if children used individual words within the discourse as linguistic cues to salience, they should have (incorrectly) inferred that *gazz* referred to the pretzel, since the phrases 'hungry' and 'salty things' made pretzels salient. In contrast, if children related the novel word to the discourse as a whole and to the speaker's communicative intentions, they should have inferred that *gazz* referred to the cookie.

Experiment 1

Participants

One hundred and forty-four monolingual 2- to 4-year-olds were included in the final dataset. This included 32 2-year-olds (M = 2;7; Range = 2;0–2;11), 55 3-year-olds (M = 3;6; Range = 3;0–3;11), and 57 4-year-olds (M = 4;6; Range = 4;0–4;11).[1] Children were assigned to one of two experimental conditions: A Practice Condition (n = 75) in which the children received warm-up trials to practice the task format, and a No Practice Condition (n = 69), in which there was no practice with the task format. Twenty additional children were not included in the final dataset for failing to meet inclusion criteria: three children from the No Practice Condition (bilingual: n = 1; failure to follow directions, n = 2) and 17 children from the Practice condition (bilingual n = 4; failed more than half of the practice questions, n = 8; failure to follow directions, n = 4; experimenter error n = 1). Bilingualism was determined via parent report.

Materials

On each trial, children saw a training stimulus containing three pictures inscribed in a rectangle (e.g. soda, bird, and cookie), followed by a test stimulus containing *different* exemplars of two of the items that had been displayed during training (e.g. a different soda, a different bird) – full descriptions of stimuli are in the Supplementary Online Materials (SOM). Stimuli included food (apple; banana; grapes; cookie; sandwich); clothing (coat; hat; shoes; gloves); animals (bird; cat; bear); kitchen tools (plate; whisk; spoon); drinks (water; soda can); sleep-relevant items (pillow; bed); and other familiar items (pencil; bus; triangle; computer; soap). Children in the Practice condition also saw four practice trials prior to the start of the main experiment (including images of a book, orange, eyeglasses, carrot, ball, cheese, pineapple, plant, collie, cane, and mug; see SOM). Each stimulus was printed in black and white ink on an 8.5″ × 11″ piece of white paper. A toy stuffed animal served as the narrator.

[1] We tested most of our children at a local science museum or in their preschool, where it is not always possible to ascertain the precise age of a child prior to testing. When recruiting children, we selected an age range of interest (2- to 4-year-olds), and calculated their precise age after testing. To avoid the biases introduced by recruiting children of different ages from different populations, we instead aimed to analyze a minimum n (in this case, n = 32) for the age with the fewest children.

Procedure

Children were tested in a quiet testing area within a lab, their preschool, or a local science museum. Each participant was introduced to a stuffed animal and told, ‘My friend Mr [animal name] speaks animal language! Wow! He needs your help. Are you ready to listen?’ Then, the experiment began. Each trial was structured so that there was first a word learning training (in which the participant heard a new word used within a discourse), followed by a referent identification measure (in which the participant selected the referent of the new word) and a word learning measure (in which the participant demonstrated knowledge of the new word in a new context).

As noted above, children in the Practice condition received a series of four trials prior to the start of the experiment to familiarize them with the task. This condition was introduced partway through data collection in order to address experimenter-reported concerns that some children were having difficulty with the task format, especially for early trials. As a result, the majority of data collection for the Practice condition was completed after the start of data collection for the No Practice condition. Practice trials followed the same format as those in the main experiment (described below), but did not require the child to infer the referent of a new word via its relation to the surrounding discourse. Children were shown images of objects and were taught the label for one of the objects via pointing and explicit labeling (e.g. ‘This is a *koojy*’, while pointing to the cheese). Thus, unlike in the main experiment, children were not required to use discourse cues to learn the new word, but instead could learn it via pointing. To test whether children had learned the word, they completed both a referent identification measure (‘Can you find the *koojy*?’) and a mutual exclusivity word learning measure (described below) in the same format presented during the main task. On each of the four practice trials, children had up to three chances to respond correctly. When a child got a practice trial wrong, the experimenter said, ‘Let’s try again!’ and then repeated the trial. In this way, children were only given indirect feedback, and were not trained on the correct way to respond. Only children who eventually got at least half of the practice trials correct were included in the final analyses. As noted above, 8/83 monolingual children who attempted the practice phase failed it and were excluded.

During the main experiment, there were eight trials. On each trial, the experimenter first presented the participant with a card containing three pictures and, using the stuffed animal, said a single statement of the form: ‘I’m very x. Look! There’s a y on the table’, where x was a mental state adjective and y was a novel word. On half of the trials – the Uninformative discourse trials – the mental state adjective (sad, happy, excited, angry) was uninformative and could not easily be used to infer the referent of the new word. For example, in a trial where soda, a cookie, and a bird were presented, the utterance ‘I’m very happy. Look! There’s a *pliff* on the table’ would be an Uninformative discourse trial. On the other half of trials – the Informative discourse trials – the mental state adjective (thirsty, hungry, cold, tired) could be used to infer the referent of the new word if the participant used the discourse bootstrapping. For example, in the same scenario described above, the utterance ‘I’m very thirsty. Look! There’s a *pliff* on the table’ would be an Informative discourse trial. Informative and Uninformative discourse trials alternated. Children were randomly assigned to one of two test versions. Test versions differed in one way: If a particular picture card was used for an Informative discourse trial in one version (e.g. an image of a bird, a cookie, and soda paired with the informative phrase ‘I’m very thirsty’), it was an Uninformative discourse trial in the other version (e.g. ‘I’m very happy’). Within each test version, there were two trial orders (forward or backward).

Children were given two tests of word learning. First, for the referent identification measure, the experimenter said, ‘Hmmm, [animal] says there’s a y on the table. Look! Do you see the y?’ and the child was given an opportunity to point to one of the three images on the training stimulus. If the child did not point to one of the images spontaneously, the experimenter said, ‘Can you show me the y? Point to the y!’ Regardless of the discourse context, the experimenter recorded whether the participant pointed to the item that would have been correct had the child heard the Informative discourse – for example, regardless of whether the child had been exposed to an Informative or Uninformative discourse, the experimenter recorded whether the participant, e.g. pointed to the soda (see Figure 1). Our logic was that if children represent discourse structure and use it to learn new words, then they should point to the correct (target) image (e.g. the soda) more often than chance only on the Informative trials, where they were exposed to an Informative discourse context (‘I’m very thirsty!’).

The second measure of children’s learning – the mutual exclusivity word learning measure – immediately followed the pointing measure. This measure ensured that children could understand the new word in a new context. Regardless of whether the child had pointed to the correct image during the referent identification task, the experimenter showed the child the test stimulus (containing new exemplars of two items from the

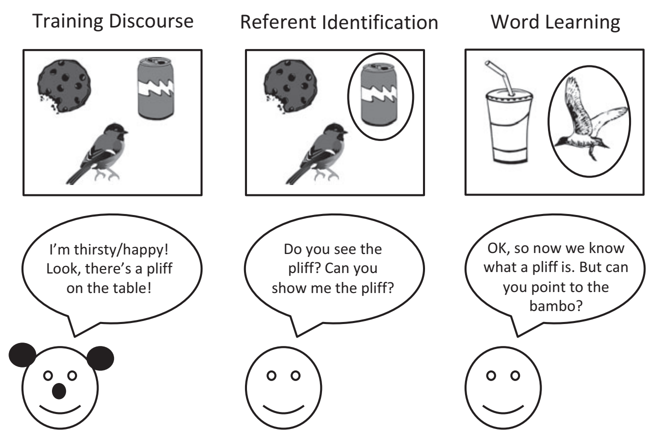

Figure 1 *Schematic of experimental setup for Experiment 1. Circles indicate correct response, and were not actually presented on the stimuli. Full description of stimuli are in the SOM.*

training stimulus – e.g. a new soda and a new bird) and said, ‘OK, so now we know what a y is. But now, can you point to the z?’, where y was the word that had been presented during the learning phase, and z was a novel word that had not previously been introduced. The experimenter then recorded which object the child pointed to. The logic behind this word learning measure was that if children mapped the word y onto a particular referent, they should be less willing to map the novel word z onto that same referent (Markman & Wachtel, 1988). This is because, by the mutual exclusivity principle, children assume that the meaning of a new word (z) differs from the meaning of a known word (y), and by the principle of contrast, children assume that differences in form (e.g. saying z instead of y) indicates differences in content (Clark, 1988, 1989).

Results

For both the referent identification and the follow-up word learning measures, we recorded whether the participant provided the ‘correct’ (target) response. As noted above, the correct response was to map the novel word onto the referent highlighted by the informative discourse context (regardless of whether the child actually heard the word in the informative or uninformative context). We conducted all analyses using the lme4 package in R (Bates & Sarkar, 2007). Responses were binary (correct vs. incorrect), and were therefore subjected to binomial logit analyses, and participant was always considered a random variable. For all analyses, we report parameter estimates (B), standard errors (SE), and p-values.

We first considered the referent identification measure, where the child had to 'point to the *y*'. To begin, we constructed a model predicting performance on the referent identification task from all of our main predictors: Trial Type (Uninformative Discourse vs. Informative Discourse), Practice Condition (Practice vs. No Practice), Age, an Age * Practice Interaction, and an Age * Trial Type Interaction. We found an effect of Trial Type and an interaction of Trial Type and Age, suggesting that participants performed differently on the Informative vs. Uninformative discourse trials, and that this effect was mediated by age (Trial Type: $B = -1.487$, $SE = .5844$, $p = .01$; Age: $B = -.1591$, $SE = .1567$, $p = .310$; Practice: $B = -.4885$, $SE = .5958$, $p = .412$; Age * Practice: $B = .1495$, $SE = .1890$, $p = .429$; Age * Trial Type: $B = 1.1487$, $SE = .1872$, $p < .0001$). Because of the effect of Trial Type, we separated all subsequent analyses by Trial Type (Informative Discourse vs. Uninformative Discourse). Because of the absence of an effect of Practice (Practice vs. No Practice), we also combined data from across Practice conditions for all subsequent analyses.

In order to test the whether children learned new words by relating them to the content of the surrounding discourse, we first had to rule out the possibility that there was a baseline preference for selecting the target object in cases where the discourse was not informative. To do this, we compared children's rate of 'correct' responses for Uninformative Discourse trials to chance. Overall, children were slightly *less* likely (28%) than chance (33%) to choose the target object on Uninformative Discourse trials ($B = -.9175$, $SE = .0944$, $p = .018$). This suggests that there were no baseline preferences for selecting the target picture in the absence of an Informative discourse context.

We then asked whether children – as a group – selected the target object more often than chance after hearing the Informative sentences.[2] As a group, children were significantly more likely (74% correct) than would be expected by chance (33%) to select the target object when presented with an informative discourse context ($B = 1.406$, $SE = .149$, $p < .0001$; Figure 2). Because the development of the ability to use discourse bootstrapping was of special interest to us, and because our earlier analyses revealed an Age * Trial Type interaction, we next tested whether Age influenced performance on Informative Discourse Trials, and found that it did ($B = 1.1984$, $SE = .1758$, $p < .0001$), such that older children performed better than younger children. Thus, we analyzed the Informative trials for each age group separately. Children at all ages selected the target object more often than chance: 2-year-olds selected the target object 46% of the time ($B = -.1025$, $SE = .2154$, $p = .006$);[3] 3-year-olds selected the target object 76% of the time ($B = 1.2805$, $SE = .1938$, $p < .0001$); and 4-year-olds selected the target object 88% of the time ($B = 2.5795$, $SE = .2989$, $p < .0001$). These data suggest that by age 2, children can identify the referent of a new word via its relation to the surrounding discourse.

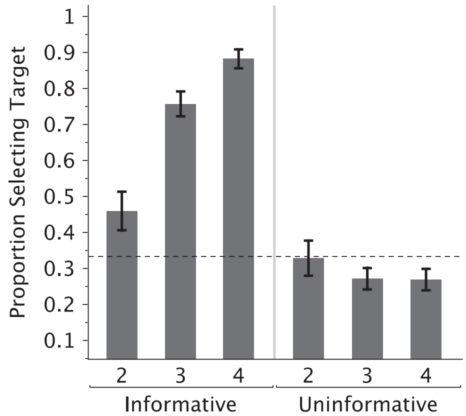

Figure 2 *Performance on the referent identification measure. Error bars are SEM, dashed line is chance. X-axis represents age in years and trial type. Y-axis is proportion 'correct', indexed by the proportion of trials on which the target picture was selected.*

To test whether children could use the newly learned word in a new context, we next analyzed results from the follow-up word learning measure. For this measure, participants saw new exemplars of two of the items that they had seen during training and were asked to find *z*, a new word that they had not previously heard. In this case, 'correct' performance on this task involves pointing to a *different* picture than the child pointed to on the referent identification task – they had to point to the ME-target item (the item that wasn't the referent of the most recently learned word; see Figure 1). As described above, we first predicted performance from Trial Type (Uninformative vs. Informative Discourse), Practice (Practice vs. No Practice), Age, a Practice * Age Interaction, and a Trial Type * Age Interaction. We found an effect of Age and an interaction of Trial Type and Age (Trial Type: $B = -.8184$, $SE = .5185$, $p = .114$; Age: $B = -.3601$, $SE = .1359$, $p = .008$; Practice: $B = -.3035$, $SE = .5191$, $p = .559$; Age *

[2] Note that because participants actually performed significantly below chance on the Uninformative trials, a direct comparison of Informative vs. Uninformative trials would be less conservative than the reported comparison to chance.

[3] Note that negative B values do not indicate below-chance performance here, because chance was .33 and a B of 0 indicates performance of .5; p-values are adjusted to represent a comparison to .33.

Practice: B = .1811, SE = .1604, p = .259; Trial Type * Age: B = .5114, SE = .1601, p = .001). As above, because there was no effect of Practice Condition on performance, we do not analyze Practice Conditions separately.

Again, we first considered performance on the Uninformative Discourse trials. As expected, children selected the ME-target object 47% of the time on Uninformative Discourse trials, which was not significantly different from the chance level of 50% (B = –.122, SE = .0837, p = .144). Next, we analyzed the Informative Discourse trials. Children were significantly more likely (66% correct) than chance (50%) to select the ME target object when they were presented with an informative discourse context (B = .8151, SE = .1252, p < .0001; Figure 3). We next tested whether there was an effect of Age on performance on this task, and found a marginal effect of age (B = .2851, SE = .1636, p = .081). While this effect was not significant, we were interested in whether children at each age were able to perform well on this task. For this reason, we conducted planned analyses on each age group separately. Two-year-olds selected the ME-target object more often (65%) than chance (B = .6388, SE = .1962, p = .001), as did 3-year-olds, who pointed to the ME-target object 59.6% of the time (B = .4767, SE = .1908, p = .013), and 4-year-olds, who selected the ME-target object 72.8% of the time (B = 1.351, SE = .251, p < .0001). This suggests that children were able to learn a new word by inferring its relation to the surrounding discourse.

Discussion

At all ages, and after exposure to a single utterance in which a novel word was presented alongside a mental state adjective, children inferred the referent of that novel word, and showed this by pointing to it more often than would be expected by chance alone. And, as revealed by our follow-up word learning (mutual exclusivity) measure, children used discourse to not only identify the referent of a new word, but also to learn that word well enough to apply it to new contexts. Our Uninformative Discourse trials revealed that this pattern of performance could not be attributed to baseline preferences for particular pictures in our study.

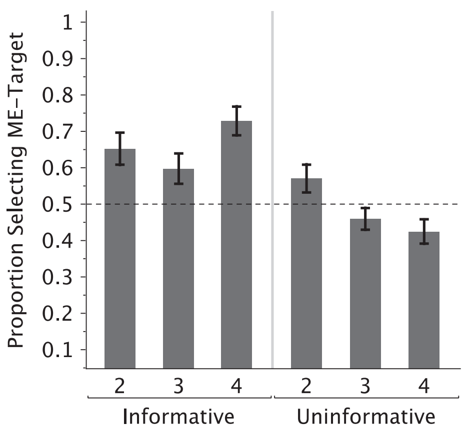

Figure 3 *Performance on the follow-up word learning (Mutual Exclusivity) measure. Error bars are SEM, dashed line is chance. X-axis represents age in years and trial type. Y-axis is proportion correct, as indexed by the proportion of trials on which participants selected the ME-target object.*

While children as young as 2 learned new words by inferring their relation to the surrounding discourse, Experiment 1 did not test the nature of this inference. One possibility is that children did not create a model of the discourse as a whole during our task, but instead used individual words within the discourse as linguistic cues to salience. For example, the training sentence 'I'm thirsty' may have made the drinkable item most salient, causing children to map the new word onto the drink. If this is true, then children learned a new word from the surrounding discourse, but did not necessarily form a model of the discourse, or of the speaker's communicative intentions. A second possibility is that children attended both to the communicative intentions of the speaker and to how words within the discourse were related to one another. Thus, children might have inferred that a novel word referred to a drink *not* because the word 'thirsty' makes the drinkable item most salient, but instead because they inferred something very specific about the relation between items within the discourse, and between the discourse and the speaker's communicative intentions (e.g. that a speaker is thirsty, and therefore that the speaker is likely to talk about drinks). We differentiated these possibilities in Experiment 2 by presenting children with sentences such as, 'I'm very thirsty, but hot things hurt my teeth! Look at what I want, there's a *dram* on the table!' Here, if children use the discourse as a whole to identify the referent of *dram* they should choose a cold drink (over a hot drink) at above-chance levels.

We made two additional changes to Experiment 2. First, we created a new follow-up word learning task to test whether children could use newly learned words in different contexts. Numerically, children in Experiment 1 did worse on the (mutual exclusivity) word learning measure than on the referent identification measure: Children selected the correct item 74% of the time in the referent identification measure (compared to a chance level of 33%), and 66% of the time on the mutual exclusivity measure (compared to a chance level of 50%). Because the mutual exclusivity measure required children to switch tasks (go from pointing to y to pointing to not-y), we couldn't evaluate whether this difference in performance across tasks emerged because (a) some children couldn't use the newly learned word in a new context or (b) the task-switching demands lowered some children's performance for reasons unrelated to their ability to learn or use new words. Thus, in Experiment 2, we modified the follow-up word learning task.

Second, we introduced a post-test to assess children's world knowledge of the ideas (e.g. 'hot drink') presented in our sentences. We made this change to better understand the age-related differences in performance in Experiment 1, and to understand why performance wasn't at ceiling for any of our measures at any of the tested ages. One possibility is that older children have greater inferential abilities, and this allows them to more efficiently relate new words to the surrounding discourse – but that inferential abilities are still not adult-like by age 4. Another possibility (not mutually exclusive with the first), however, is that our age differences stemmed from differences in the world knowledge required to understand the relation between our sentences and the referents used in our study (e.g. whether the child knows which of the three items is a drink). By accounting for children's world knowledge, we could test these possibilities.

Experiment 2

Participants

Ninety-six monolingual 2- to 4-year-olds were analyzed. This included 30 2-year-olds (M = 2;6; Range = 2;0–2;11), 31 3-year-olds (M = 3;5; Range = 3;0–3;11), and 35 4-year-olds (M = 4;6; Range = 4;0–4;11). An additional 18 children were not included in the final dataset for failing to meet inclusion criteria (bilingual n = 11; parent failed to report birth date n = 3; already participated in a related study n = 2; parent intervention n = 1; failure to respond to test trials n = 1).

Materials

Each participant saw two trials. For each trial, children saw a training stimulus of three images (e.g. pretzel, cookie, shoe), followed by a test stimulus of three images (e.g. a different pretzel, different cookie, and different shoe). Each set of images was printed in color on an 8.5″ × 11″ piece of white paper. Possible images included a pretzel, a cookie, a shoe, a steaming cup of coffee, a glass of ice water, and a truck (see SOM). A post-test stimulus contained new exemplars of a pretzel, a cookie, a steaming coffee, and an ice water, arranged in a 2 × 2 grid. A toy stuffed animal served as the narrator. Each child participated in two trials.

Procedure

The basic structure of the task was identical to Experiment 1. For each of the two trials, there was first a word learning training (in which the participant heard a new word used within a discourse), followed by two word comprehension tests. However, there were three important differences between the methods in Experiments 1 and 2.

The first difference was to the training discourse, which took the form, e.g. 'I'm very [hungry/thirsty], but [salty/sweet/hot/cold] things [make me sick/hurt my teeth]. Look at what I want! There's a *gazz* on the table.' This allowed us to test whether children integrated information about the discourse when learning the new word, instead of simply attending to individual words within the discourse. To minimize item effects, children were randomly assigned to one of two task versions. In one version, 'hungry' was paired with 'salty' (in the other it was paired with 'sweet'), and 'thirsty' was paired with 'cold' (in the other it was paired with 'hot'). Also, in one version the 'hungry' trial preceded the 'thirsty' trial, while in the other the opposite order was used (see SOM for full list of stimuli). In this way, trial order and word pairing were fully crossed across task version.

The second difference, alluded to above, was that the follow-up test of word learning was another pointing measure: children saw the test stimuli (which contained new exemplars of all three of the training stimuli) and were asked, 'Can you find another *y*?' This new measure allowed us to test whether children could interpret the newly learned word in a different context. However, it did not require children to switch tasks (as our mutual exclusivity measure in Experiment 1 did), making the task format easier for our children to learn.

Finally, we introduced a post-test to check whether children had sufficient world knowledge to learn from the discourse: children saw the four post-test items and were asked to find the sweet food, hot drink, salty food, and cold drink.

Results

We measured whether the participant pointed to the target object when asked to find *gazz* after the initial exposure (e.g. a pretzel after hearing 'I'm hungry, but sweet things make me sick. Look, there's a *gazz* on the table'), and also whether they pointed to a new exemplar of the target object during the follow-up word learning test. Participants performed similarly on these two measures, providing the same response for the initial referent identification task and the follow-up word learning task 81% of the time.

We first asked whether children selected the correct object more often than expected by chance. For this and all analyses, chance was 33%, since there were three possible items to choose from. As a group, children performed well on both measures of learning (49.7% correct for the Referent Identification (RI) task and 44.3% correct for the follow-up Word Learning (WL) measure), and selected the item picked out by the surrounding discourse significantly more often than chance (RI: $B = -.0313$, $SE = .1636$, $p < .0001$; WL: $B = -.2462$, $SE = .1585$, $p = .005$; Figure 4).

Next, we asked whether (a) passing the post-test and (b) age predicted performance on this task. Thus, we classified each child according to whether they passed the post-test (by getting all of the items right) or failed the post-test (by getting at least one item wrong).[4] We then predicted responses from Age and Post-Test performance. For the RI measure, age was not a predictor of task performance ($B = .0586$, $SE = .2164$, $p = .787$), performance on the post-test marginally predicted task performance ($B = .6669$, $SE = .3560$, $p = .061$), and an ANOVA comparing the model containing post-test performance and Age to one just containing Age revealed a marginal benefit to including post-test performance ($p = .069$). For the WL measure, Age was not a predictor of task performance ($B = -.2030$, $SE = .2132$, $p = .341$), but post-test performance was ($B = 1.044$, $SE = .3536$, $p = .003$), and an ANOVA comparing a model containing only Age to one that also contained post-test performance showed that post-test performance significantly contributed to model fit ($p = .003$).

Because post-test performance appeared to predict performance on the RI and WL measures – and because we had theoretical reasons to expect differences in performance depending on world knowledge – we next separated children by whether they passed or failed the post-test. Children who failed the post-test did not perform differently (RI: 41.17% correct; WL: 32.94% correct) from chance (RI: $B = -.4031$, $SE = .2478$, $p = .242$; WL: $B = -.7939$, $SE = .2575$, $p = .696$), while those who passed the post-test performed significantly above (RI: 57.14% correct; WL: 54% correct) chance (RI: $B = .3067$, $SE = .2177$, $p < .0001$; WL: $B = .1610$, $SE = .2017$, $p < .0001$). Thus, children who knew something about the words within the discourse (e.g. that cookies are sweet) were able to use this information to identify the referent of a new word from the surrounding discourse context. This is very strong evidence that children can use the content of a surrounding discourse (and not just individual words within that discourse) to identify the referents of new words. Unsurprisingly, children who failed our post-test tended to be younger (M = 2;8) than those who passed (M = 3;4), and while only 26.7% of 2-year-olds passed our post-test, 74.29% of 4-year-olds passed it. These data suggest that the age-related differences that we saw in Experiment 1 may have resulted from differences in world knowledge, and not age-related differences in inferential abilities.

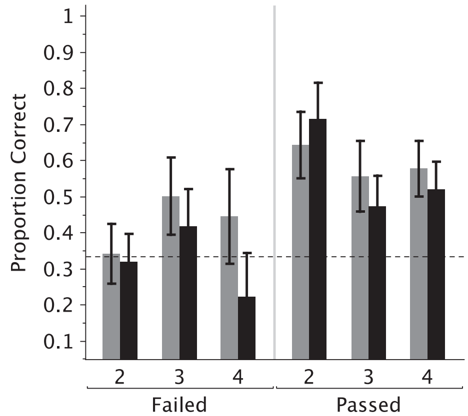

Figure 4 *Proportion correct in Experiment 2. Gray bars indicate the Referent Identification measure and black bars indicate the Word Learning follow-up measure. Error bars are SEM; horizontal dashed line indicates chance performance. The x-axis indicates age in years, and whether the child passed or failed the post-test.*

One possible explanation of these data is that children attended to an individual word within the discourse (e.g. *hungry*) and used this to restrict the range of possible referents to a set of category-congruent objects (e.g. to edible things). They then might have chosen randomly between these category-congruent objects. Such a strategy would not require children to represent the discourse as a whole when learning new words. To test this possibility, we first asked how frequently children chose category-congruent objects (edible items on *hungry* trials and drinkable items on *thirsty* trials). Children who passed our post-test selected the category-congruent objects 76% of the time (which is different from a chance level of 66%, $B = 1.2886$, $SE = .2547$, $p = .019$), and those who failed the post-test selected the category-congruent objects 71% of the time ($B = 1.1349$, $SE = .3106$, $p = .155$).

[4] While it would be ideal to analyze performance on the post-test at a finer grain, the types of responses that children gave typically precluded this. Children never refused to respond, so any child who gave an incorrect answer did so by choosing at least one item twice (e.g. selecting the pretzel as both the 'sweet food' and the 'hot drink'). Thus, getting one item wrong on this task requires incomplete knowledge of at least two of our four items, and often (as in the example above) indicated incomplete knowledge of more than two of the items.

To test whether children actually modeled the content of the discourse as a whole, we next asked whether – given that a child selected a category-congruent object – they also selected the discourse-consistent item more often than chance. Here, chance would be calculated as 50%, since each trial had two category-congruent items, and we restricted our analyses only to participants who chose one of those two category-congruent items. As a group, on trials where participants chose category-congruent objects (e.g. edible things when they heard the word *hungry*), they also chose the discourse-consistent object more often than would be expected by chance (RI = 62.2%; B = .5887, SE = .2083, p = .005; WL = 59.9%; B = .4675, SE = .1987, p = .019). We next split this group by post-test performance. Those who failed the post-test but who still selected a category-congruent object selected the discourse-consistent object at chance levels (RI = 53.3% of the time, B = .1332, SE = .2614, p = .610; WL = 46.67% of the time, B = –.1388, SE = .2639, p = .599). However, those who passed the post-test and selected a category-congruent object also selected the discourse-consistent object significantly more often than would be expected by chance (RI: 69.33% of the time, B = 1.5457, SE = .4522, p = .001; WL: 70.13% of the time, B = 1.3082, SE = .3593, p = .0003). These analyses suggest that our main finding can only be explained by discourse bootstrapping, and not by other, less sophisticated mechanisms.

Finally, because we had an *a priori* interest in the developmental trajectory of discourse bootstrapping, we tested the performance of each age group separately. At no age did children who failed our post-test successfully learn a new word on the basis of the surrounding discourse (all ps > .05). However, children who passed the post-test (and therefore had sufficient knowledge of the words under consideration to benefit from the discourse context) did very well at all ages. Two-year-olds (RI: B = .5878, SE = .5578, p = .022; WL: B = .9163, SE = .5916, p = .007), 3-year-olds (RI: B = .3451, SE = .4297, p = .016; WL: B = –.0572, SE = .3382, p = .060), and 4-year-olds (RI: B = .2318, SE = .3196, p = .004; WL: B = .1236, SE = .3005, p = .007) performed above chance on nearly all measures (3-year-olds performed only marginally better than chance on the WL measure). This provides evidence that children are able to process the entirety of a discourse when picking out the referent of a new word as soon as they know the other words within the discourse.

Discussion

Experiment 2 found that children as young as 2 successfully identified the referent of a novel word after hearing it presented only one time in a discourse. However, children only identified the correct referent if they had enough world knowledge to understand the ideas in the discourse (e.g. that a cookie is a ‘sweet food’). Critically, in Experiment 2, we ruled out the possibility that children used only low-level salience cues to connect novel words to the surrounding discourse (e.g. after hearing a word like *sweet*, assuming a new word refers to something sweet; Akhtar, 2002). Instead, children successfully integrated information from across the discourse in order to isolate the appropriate referent of a new word. This is consistent with the Discourse Bootstrapping hypothesis, in that it suggests that children can learn new words via an inference about their relation to the content of the surrounding discourse.

Second, we also found that world knowledge limits children’s ability to deploy discourse bootstrapping. Knowledge of the items presented within the sentences (e.g. that cookies are a sweet food) predicted children’s performance on our task, and, among those children who passed our world knowledge post-test, there were no age-related differences in performance. This suggests that developmental changes in performance on our task may be better explained by developmental changes to world knowledge than by developmental changes to inferential or other abilities.

General discussion

We tested whether preschoolers could use the content of a discourse to infer the referents of new words. We found evidence that preschoolers can learn new words by inferring their relation to the surrounding discourse and to the speaker’s communicative intentions. These data support what we’ve called the ‘discourse bootstrapping hypothesis’ – that children learn new words by making inferences about how these words are related to the surrounding discourse and to the speaker’s communicative intentions. In Experiment 1, children were shown images of objects (e.g. a bird, a soda, and a cookie) accompanied by either an informative discourse context (‘I’m thirsty, there’s a *pliff* on the table!’) or an uninformative discourse context (‘I’m happy, there’s a *pliff* on the table!’). Children as young as 2 used discourse context to infer that *pliff* referred to an edible item, but only when the new word was presented in an informative discourse context. In Experiment 2, we showed that word learning was driven by an inference about the relation between the new word and the surrounding discourse, and not merely by inferences based on individual words within the discourse. Together, these experiments provide evidence that children learn new words by relating them to the communicative content conveyed by the discourse as a whole. This is the first study to demonstrate that preschoolers recruit such a mechanism when learning new words. In doing so, this study provides an account of how children might learn words from conversation when other attested mechanisms are too general to differentiate candidate referents (e.g. syntactic bootstrapping), and when pointing or other ostensive acts of labeling are not available.

Although we showed that children can use discourse bootstrapping to learn new words by age 2, we also found that children's ability to use this mechanism changes with age. In Experiment 1, for example, we found that 2-year-olds were much worse than older children at learning new words from discourse. Experiment 2 suggested one reason for this difference: Success on our word learning task depended on knowledge of the other words and ideas presented in the discourse (e.g. knowledge that a pretzel is a salty food), which many 2-year-olds lacked. Thus, discourse bootstrapping is only useful in contexts that contain familiar words. The observation that discourse bootstrapping is most effective in contexts where many of the surrounding words are familiar is consistent with the finding that word learning quickly accelerates in early development, with each new word ratcheting up the child's capacity to learn subsequent words (Dapretto & Bjork, 2000; Ganger & Brent, 2004; Goldfield & Reznick, 1990). In the spirit of Quine (1960), the words that children acquire act as bricks in a lexical chimney, each new word supporting the learner as they scramble up the chimney, making inferences about the meanings of new words and utterances as they go.

Other factors not related specifically to vocabulary size likely also affect children's ability to use discourse in early learning, and may explain why younger children had greater difficulty than older children on our tasks. First, hearing words used across a variety of sentences and contexts is likely important for learning, especially because each sentence will possibly contain many words that very young children do not yet comprehend (Frank *et al.*, 2009; Yu & Smith, 2007). While children in our study were exposed to a single sentence as context for learning (and only heard the novel word once prior to test), previous studies suggest that they may receive much richer evidence in real conversation, making it easier for the youngest learners to acquire word meanings from discourse (Frank *et al.*, 2013). Second, identifying the relation between ideas within a discourse – a requirement of discourse bootstrapping – often involves understanding the beliefs, desires, and background assumptions of the speaker. This may require both a basic theory of mind and a relatively sophisticated understanding of how the world works – both of which change substantially during the preschool years. Finally, the successful use of discourse bootstrapping may also hinge on an understanding of subtle linguistic connectives like *because* and *but* that may be difficult to master early in acquisition (Bloom, Lahey, Hood, Lifter & Fiess, 1980; Peterson & McCabe, 1991). In support of this, recent work has shown that children's comprehension of these connectives is critical to forming appropriate models of a discourse and to using the discourse to disambiguate pronouns (Kehler, Sullivan, Hayes & Barner, in preparation), and learn new words (Sullivan, Jung & Barner, in preparation).

A challenge raised by our finding – that children use discourse bootstrapping to learn new words – is that relatively little is known about *how* children (or adults!) figure out the relations between words and ideas within a discourse. More generally, despite decades of research on the topic, we nevertheless lack generalized theories of discourse coherence or relevance that generate clear predictions across case studies. The present paper suggests that complex computations about the relevance of individual words to the content of a discourse are important to word learning. If this is the case, then the path forward for researchers interested in word learning, pragmatics or the development of communication will require the development of a strong and testable theory of discourse, and methods for testing such theories in children. Recent work from the pronoun-interpretation literature has taken important first steps in developing and testing a subset of highly constrained theories of discourse coherence and pragmatic reasoning (e.g. Kehler, Kertz, Rohde & Elman, 2007; Rohde, Levy & Kehler, 2011), as has recent computational work on pragmatic and social reasoning (e.g. Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013).

We have argued that even when preschoolers can't use cues such as syntax, socio-pragmatic signals, or ostensive instruction, they can still learn new words by making an inference about the relation between the word and the surrounding discourse context. This suggests that high-level processes – like discourse comprehension – that are most frequently discussed with respect to older children's and adults' language skill may actually be operative in children's early language acquisition. While future research will be required to better understand the emergence of children's reliance on discourse bootstrapping (and to develop a constrained and testable theory of discourse), this study provides a critical first step in documenting the importance of discourse to early word learning, and in describing how children can learn new words in cases where other attested word learning cues are not available.

Acknowledgements

Thank you to members of the LADLab for helpful comments on this manuscript. Thank you to Karen Bejar, Eleanor Chestnut, Jinsol Jung, Katie Kimura, Ieva Razhas, Nestor Tulagan, and Shirlene Wade for help gathering data, to Kevin Smith for statistical advice, and to the staff and families at the Reuben H. Fleet Science Center and the many preschools where we tested for their collaboration. Images in Figure 1 were used under Creative Commons Deed CC0. Data were collected while JS was affiliated with the University of California, San Diego. This work was supported by a NSF Graduate Research Fellowship to JS, and by a James S. McDonnell grant to DB.

References

Akhtar, N. (2002). Relevance and early word learning. *Journal of Child Language*, **29**, 677–686.

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, **8**, 199–209.

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, **67**, 635–345.

Akhtar, N., Jipson, J., & Callanan, M. (2001). Learning words through overhearing. *Child Development*, **78**, 416–430.

Almoammer, A., Sullivan, J., Donlan, C., Marusic, F., Zaucer, R., et al. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, **110**, 18448–18453. doi:10.1073/pnas.1313652110

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

Barner, D. (2012). Bootstrapping numeral meanings and the origins of exactness. *Language Learning and Development*, **8**, 177–185.

Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development*, **2**, 163–194.

Bates, D., & Sarkar, D. (2007). Lme4: Linear mixed-effects models using S4 classes (R package version .9975-12). Available at http://CRAN.R-project.org

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, **7**, 235–261.

Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, **24**, 511–533.

Brown, R. (1957). *Words and things*. Glencoe, IL: Free Press.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, **63** (4, Serial No. 255).

Caza, G., & Knott, A. (2012). Pragmatic bootstrapping: a neural network model of vocabulary acquisition. *Language Learning and Development*, **8**, 113–135.

Clark, E.V. (1983). Meanings and concepts. In J.H. Flavell & E.M. Markman (Eds.), *Handbook of child psychology, Vol. 3. Cognitive development* (pp. 787–840). New York: Wiley.

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

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

Dapretto, M., & Bjork, E. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development*, **71**, 635–648. doi:10.1111/1467-8624.00172

Frank, M., & Goodman, N. (2012). Predicting pragmatic reasoning in language games. *Science*, **336**, 998.

Frank, M., Goodman, N., & Tenenbaum, J. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, **20**, 579–585.

Frank, M., Tenenbaum, J., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, **9**, 1–24.

Fukkink, R., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: a meta-analysis. *Review of Educational Research*, **68**, 450–469.

Ganger, J., & Brent, W. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, **40**, 621–632. doi:10.1037/0012-1649.40.4.621

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

Goldfield, B., & Reznick, S. (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. *Journal of Child Language*, **17**, 171–183. doi:10.1017/S0305000900013167

Goodman, N., & Stuhlmüller, A. (2013). Knowledge and implicature: modeling language understanding as social cognition. *Topics in Cognitive Science*, **5**, 173–174. doi:10.1111/tops.12007

Grassman, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, **13**, 252–263.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics III: Speech acts* (pp. 41–58). New York: Academic Press.

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

Jenkins, J., Stein, M., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, **21**, 767–787.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. (2007). Coherence and coreference revisited. *Semantics*, **25**, 1–44.

Kehler, A., Sullivan, J., Hayes, E., & Barner, D. (in preparation). Grammatical and pragmatic constraints on children's pronoun interpretation.

Landau, B., & Gleitman, L. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.
McCullough, C. (1943). Learning to use context clues. *Elementary English Review*, **20**, 140–143.
Macnamara, J. (1972). The cognitive basis of language learning in children. *Psychological Review*, **79**, 1–13.
Markman, E. (1990). Constraints children place on word meanings. *Cognitive Science*, **14**, 57–77.
Markman, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cognitive Psychology*, **16**, 1–27.
Markman, E., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, **20**, 121–157.
Meyer, M., & Baldwin, D. (2013). Pointing as a socio-pragmatic cue to particular vs. generic reference. *Language Learning and Development*, **9**, 245–265. doi:10.1080/15475441.2013.753802
Naigles, L. (1996). The use of multiple frames in verb leaning via syntactic bootstrapping. *Cognition*, **58**, 222–251.
Nagy, W., Anderson, R., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, **24**, 237–270.
Peterson, C., & McCabe, A. (1991). Linking children's connective use and narrative macrostructure. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure* (pp. 29–53). Hillsdale, NJ: Lawrence Erlbaum Associates.
Pinker, S. (1979). Formal models of language learning. *Cognition*, **7**, 217–283.
Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
Pinker, S. (1994). How could a child use verb syntax to learn verb semantics? *Lingua*, **92**, 377–410.
Quine, W. (1960). *Word and object*. Cambridge, MA: MIT Press.
Rohde, H., & Frank, M. (2014). Markers of topical discourse in child-directed speech. *Cognitive Science*, **38** (8), 1634–1661.
Rohde, H., Levy, R., & Kehler, A. (2011). Anticipating explanations in relative clause processing. *Cognition*, **118**, 339–358. doi:10.1016/j.cognition.2010.10.016
Sarnecka, B., Kamenskaya, V., Yamana, Y., Ogura, T., & Yudovina, Y. (2007). From grammatical number to exact numbers: early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognitive Psychology*, **55**, 136–168.
Schulze, C., Grassman, S., & Tomasello, M. (2013). 3-year-old children make relevance inferences in indirect verbal communication. *Child Development*, **84** (6), 2079–2093. doi:10.1111/cdev.12093
Snow, C., & Ferguson, C. (1977). *Talking to children*. Cambridge: Cambridge University Press.
Soja, N. (1992). Inferences about the meanings of nouns: the relationship between perception and syntax. *Cognitive Development*, **7**, 29–45.
Soja, N., Carey, S., & Spelke, E. (1991). Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition*, **38**, 179–211.
Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.
Stiller, A., Goodman, N., & Frank, M. (in press). Ad-hoc implicature in preschool children. *Language Learning and Development*. doi: 10.1080/15475441.2014.927328
Sullivan, J., Jung, J., & Barner, D. (in preparation). Discourse coherence as a cue to word reference.
Tomasello, M., & Barton, M. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, **30**, 639–650.
Tomasello, M., & Kruger, A. (1992). Joint attention on actions: acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, **19**, 311–333. doi:10.1017S0305000900011430
Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, **23**, 157–176. doi:10.1017/S0305000900010138
Tribushininia, E. (2012). Comprehension of relevance implicatures by pre-schoolers: the case of adjectives. *Journal of Pragmatics*, **44**, 2035–2044.
Verbuk, A., & Shultz, T. (2010). Acquisition of relevance implicatures: a case against a rationality-based account of conversational implicatures. *Journal of Pragmatics*, **42**, 2297–2313.
Waxman, S., & Booth, A. (2001). Seeing pink elephants: fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, **43**, 217–242.
Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Cognitive Development*, **72**, 655–684.
Yu, C., & Smith, L. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, **18**, 414–420. doi:10.1111/j.1467-9280.2007.01915.x

Received: 5 March 2014
Accepted: 17 December 2014

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Stimuli for the practice portion for participants in the Practice condition. All labeling was done via pointing. All ME Images were *new* exemplars of the same kind as in the Training Image.

Table S2. Stimuli for referent identification task in Exp. 1.

Table S3. Stimuli for Word Learning Measure in Exp. 1 All images were *new* exemplars of the images presented during the Referent Identification task in Exp. 1 (see Table 2).

Table S4. Stimuli for Experiment 2 .