Young children's developing sensitivity to discourse continuity as a cue to reference

Alexandra C. Horowitz
ahorowit@stanford.edu
Department of Psychology
Stanford University

Michael C. Frank
mcfrank@stanford.edu
Department of Psychology
Stanford University

Abstract

Children can learn new words in pedagogical contexts, but they may also infer reference using a variety of other information sources. Here we investigate children's sensitivity to the placement of novel labels within discourse structure as a possible mechanism for word learning. In Experiment 1, children ages 2–6 years participated in word learning trials featuring two novel items and one novel label. In critical trials, the labels were embedded *between* two sentences about the same item, whereas in a control condition, the label was introduced *after* two sentences about the item. Children of all ages were more likely to attribute the label to the toy whose descriptions bracketed the embedded label, and response strength increased with age. Children across all ages responded at chance in the control condition. In Experiment 2, adults showed the same patterns of responses as children in both critical and control conditions. Together, these results suggest that discourse continuity is a reliable cue to reference for both children and adults.

Keywords: word learning; discourse; social cues; language development.

Introduction

Children use a variety of strategies for learning new words. In overtly pedagogical situations, children can use cues such as pointing, joint attention, and labeling to establish a direct mapping between an object and name. However, many situations do not feature ostensive labeling events. In these cases, children must rely on other strategies to infer the referent of a novel word. One source of information may come from discourse structure (the order of utterances and how they relate to each other). Recognizing how speakers relate topic information may help children resolve reference that would otherwise be ambiguous in the absence of the broader context.

For example, a child may not have an idea of what *chinchilla* means from an utterance such as, "I love chinchillas!", but she may apply her knowledge of discourse structure to infer its meaning when the same utterance is related to topical information, such as "I got a new pet. I love chinchillas! They're so soft." Children are exposed to information about discourse structure whenever they hear speech, and their accumulation of experience may help them to update and refine their expectations about topic relationships. This developing expertise may allow children to infer meaning that is locally ambiguous, yet resolvable in the context of broader discourse.

Little work has explored whether children use discourse structure to scaffold reference disambiguation. Our aim is to address this question by investigating children's and adults' recognition of communicative structure when dissociated from other social and ostensive cues. Understanding the contribution of discourse knowledge in children's reference disambiguation may help identify opportunities for children to infer meaning from topic coherence.

A large body of research has been devoted to children's ability to map names to inferred referents through disambiguation of a single new item and label. In the presence of a known item and an unknown item, children map a novel label to the novel item rather than the already-named item (Markman & Wachtel, 1988; Merriman, Bowman, & MacWhinney, 1989; Mervis & Bertrand, 1994; Clark, 1990). Though the mechanisms at play are debated, this finding establishes that children can make inferences about a speaker's likely intended meaning in constrained contexts. While employing their repertoire of word-object mappings may allow children to disambiguate some novel referents, there may be other situations in which multiple items are unknown.

Children can also use social-pragmatic cues to infer novel word reference. By age 2, children map novel labels to novel objects that the speaker attends to rather than what they themselves may be attending to (Baldwin, 1991), and even after a time delay (Baldwin, 1993). Young children apply new terms to the target but not non-target items of a speaker's search (Tomasello & Barton, 1994), show evidence of considering not only their own novelty perspective, but also what is novel to a speaker in a discourse context (Akhtar, Carpenter, & Tomasello, 1996), and can recognize that speaker naming events may convey social-pragmatic implications about information that is expected to be shared (Diesendruck & Shemer, 2006).

Despite this body of evidence on general pragmatic cues, few studies have investigated children's sensitivity to discourse structure. Nevertheless, some work suggests that discourse continuity may provide important opportunities for learning. Frank, Fernald, and Tenenbaum (2013) examined a video corpus of caregivers interacting with their 6–18 month-old children. In these natural settings, they found discourse continuity (that utterances in close succession are likely to relate to the same topic) was as reliably linked to reference as were social cues such as pointing and gaze (Frank, Tenenbaum, & Fernald, 2013). Although this result suggests that discourse continuity may be an available cue to disambiguate reference in the presence of competitors, their work did not provide evidence that learners actually make use of this information. Therefore, our present studies provide the first test of whether discourse position can be used to determine reference in word learning.

We discuss two experiments indicating that both children and adults can use discourse position to resolve reference ambiguity. Additionally, our control condition shows that children and adults rely on the informativeness of discourse structure rather than simpler heuristics such as temporal proxim-

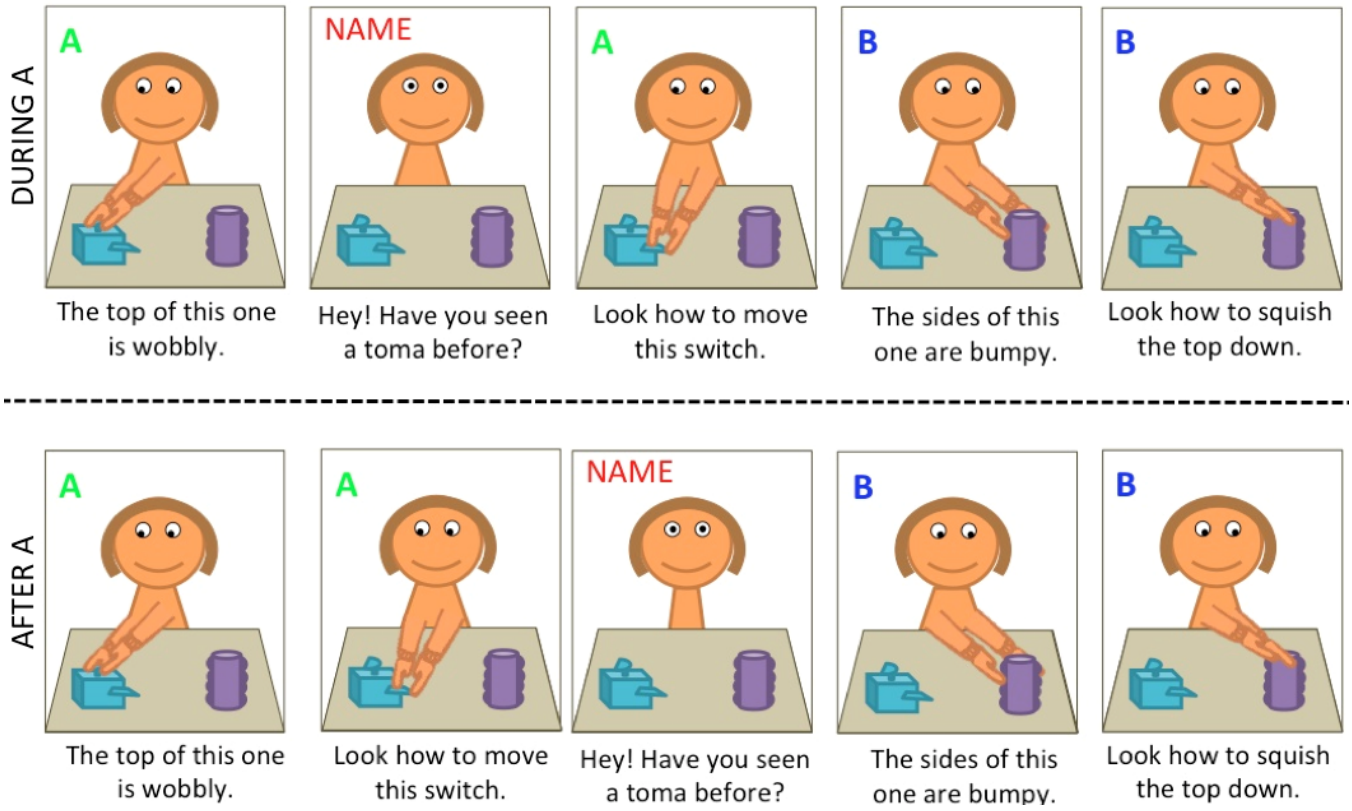

Figure 1: Schematic order of events for "A" trials across conditions. In the *During* condition, the experimenter makes eye contact without other gaze cues and introduces the naming event *between* two descriptions of a single toy. In the *After* control, events are identical except that the experimenter introduces the naming event *after* two descriptions of a toy.

ity. Our findings suggest that language users are able to apply their knowledge about how utterances relate within a discourse to make inferences about speakers' intended referents. Overall, this work suggests that children are not constrained to social and contextual cues from individual utterances, but can evaluate how information may relate across broader discourses.

Experiment 1

We designed a novel task to investigate children's recognition of discourse structure as a cue to reference. In a large sample of 2– 6 year–olds, we introduced children to two novel toys accompanied by only one novel label, and we manipulated where in discourse the name was introduced. We were interested in which toy children selected as the referent of the label. The experimenter made eye contact with the child but gave no gaze or gesture cues to the referent of the label during the naming event, so the only cue to reference was the location of the naming event within the broader discourse. In the critical *During* condition, children were introduced to a novel label within descriptions about either the first ("A") or second ("B") toy (see Figure 1). Because the label was embedded within descriptions of the same toy in critical trials, we were interested in whether children would infer topic continuity and link the label to that toy.

An alternative explanation for why children might choose the toy whose descriptions bracketed the naming events is that children are making a temporal association between the label and the toy descriptions rather than considering discourse structure per se. That is, children may be selecting the toy that is described closest to the naming event, which would always correspond with the toy surrounding the introduction of the label in *During* trials. Therefore, we also ran a control condition to dissociate temporal proximity from discourse coherence. In the control *After* condition, the naming event was introduced after descriptions of either Toy A or Toy B rather than between descriptions of that toy. Thus the naming event could occur next to descriptions of both toys in *After A* trials, or next to only Toy B in *After B* trials. Sample scripts for each trial type are listed in Table 1.

Our design allows us to make the following predictions: If children recognize discourse continuity as a cue to reference, they should infer that new information contained within a single topic is likely to also refer to that topic. Therefore, children should select Toy A in *During A* trials and Toy B in *During B* trials, but not have a clear strategy in *After* trials. If children rely on temporal association rather than topic coherence, they should select a referent according to what comes closest to the naming event, i.e. Toy A in *During A* and Toy B in *During B* trials, as well as Toy B in *After B* trials because

it is the only description proximal to the naming event. These predictions are outlined in Table 2.

Table 1: Sample scripts for each condition (*During* or *After*) and trial type (Toy A or Toy B). The green sentences are descriptions of Toy A, and the blue sentences are descriptions of Toy B.

During A	**After A**
The top of this one is wobbly.	The top of this one is wobbly.
Have you seen a toma before?	Look how to move this switch.
Look how to move this switch.	**Have you seen a toma before?**
The sides of this one are bumpy.	The sides of this one are bumpy.
Look how to squish the top down.	Look how to squish the top down.
During B	**After B**
The top of this one is wobbly.	The top of this one is wobbly.
Look how to move this switch.	Look how to move this switch.
The sides of this one are bumpy.	The sides of this one are bumpy.
Have you seen a toma before?	Look how to squish the top down.
Look how to squish the top down.	**Have you seen a toma before?**

Table 2: Predictions for reference selections (Toy A or Toy B) across each trial type if participants rely on discourse continuity or temporal association.

	Discourse Continuity	**Temporal Association**
During A	A	A
During B	B	B
After A	either	either
After B	either	B

Methods

Participants One hundred sixty-six children were recruited from the San Jose Children's Discovery Museum to complete a planned sample of 128 children. Children were given a sticker and certificate as compensation their participation. Parents were asked to fill out a short demographic form about their children's language background, and only children who were reported to hear English at least 75% of the time were included in the study. Twelve children were excluded due to insufficient English exposure, 12 children whose language information was not reported were excluded, and 14 children were excluded for not completing all four trials of the study.

Children were recruited in four age groups: 2–year–olds (n=32, 18 girls, mean age 2 years 6 months), 3–year–olds years (n=32, 10 girls, mean age 3 years 6 months), 4–year–olds (n=32, 14 girls, mean age 4 years 6 months), and 5–year–olds (n=32, 18 girls, mean age 5 years 4 months).

Stimuli Four pairs of unusual items (e.g. a faucet aerator and a spaghetti measure) served as the novel toys. An additional item was used for training.

Procedures Participants were seated across from the experimenter in a quiet room at the museum. Children participated in a training trial featuring ostensive labeling of a single toy ("This toy is called a blicket. Can you point to the blicket?") before seeing four discourse disambiguation trials. For the discourse disambiguation trials, the experimenter placed two toys on the table and described each in turn (see Figure 1). The toy pair remained in view of the child throughout the duration of the trial. All children heard the same scripts used to describe the toys; the only difference was the discourse location where the label was introduced. In a between-subjects manipulation, half the participants in each age group (n=16 per age) participated in four *During* trials; the other half participated in four *After* control trials. Order and toy pairs were counterbalanced across participants. At the end of each trial, the experimenter prompted the child to identify the named item by pointing. If children did not respond immediately, they were prompted again to make their best guess. The sessions were videotaped and coded offline. The entire task took about 5 minutes to complete.

Sample scripts for each condition and trial type are shown in Table 1. In *During* trials, the experimenter introduced the naming event between two sentences about the same toy (e.g. "You can push this button. **Hey [child's name]! Have you seen a toma before? Tomas are so neat!** What cool handles."). For two trials the label introduction was embedded during descriptions of the first toy (Toy A) and for two trial it was introduced during descriptions of the second toy (Toy B). When describing the toys, the experimenter directed her gaze to the toy and demonstrated a feature of the toy. There was a brief pause between each sentence. For the naming event, she disengaged from the toy and maintained a neutral position while drawing the children's attention using their names and establishing eye contact. The experimenter did not give any gaze cues or other indicators to the referent of the novel name. Thus, the naming event in itself carried no information to guide disambiguation; the only cue available was its location within discourse. In *During* trials, the naming event was always embedded between descriptions of a single toy. The *After* trials were identical except that the naming event appeared after the two descriptions about a toy (e.g. "You can push this button. What cool handles. **Hey [child's name]! Have you seen a toma before? Tomas are so neat!**").

Results and Discussion

Figure 2 illustrates the proportion of children selecting the second toy (Toy B) as the referent of the label across conditions (*During* and *After*) and trial types (whether the label was introduced with Toy A or Toy B). The figure also includes adult performance across conditions and trial types from Experiment 2.

Children showed increased sensitivity to discourse coherence over development. Overall, children in the *During* condition were more likely to select Toy B when the label was embedded during descriptions about Toy B, and were less likely to select Toy B (thus more likely to select Toy A) when the naming event was bracketed by descriptions about Toy A. This pattern became more pronounced as children got older.

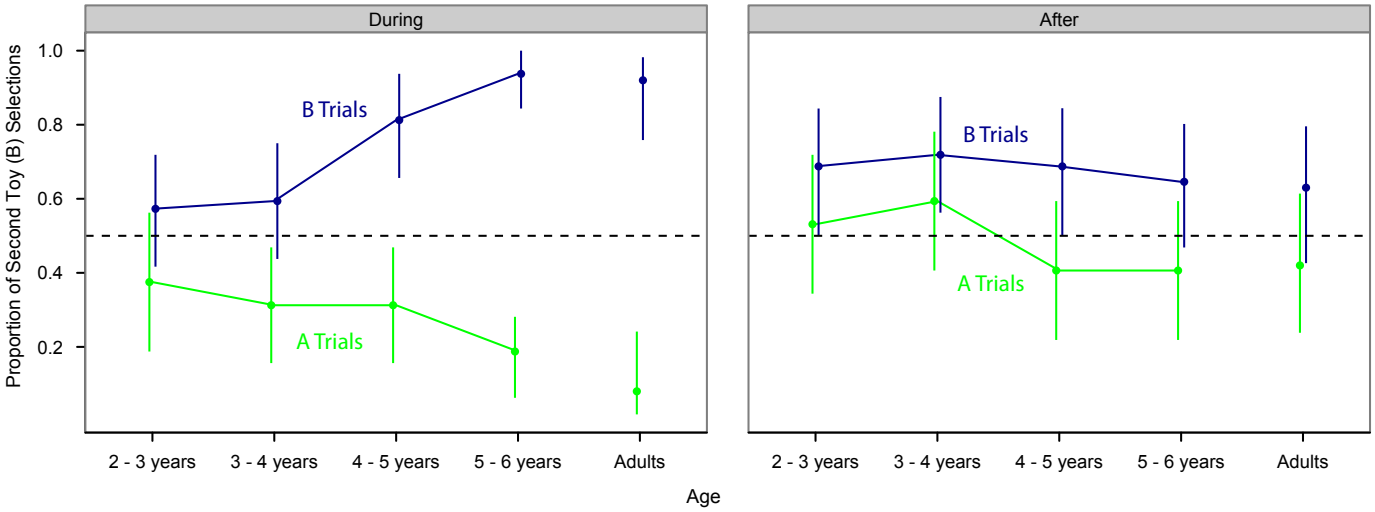

Figure 2: Combined data from Experiments 1 and 2. Mean proportion of selection of the second toy across condition (*During* or *After*) and trial type (label given with first toy: A, or second toy: B). Chance performance is at 0.5, error bars represent 95% confidence intervals.

This finding suggests that children's sensitivity to discourse coherence as a cue to disambiguate reference increases across ages 2–6 years.

In *After* trials, children were at chance in selecting either toy for both *After A* and *After B* trials for all age groups. This result shows that children did not develop a consistent strategy to disambiguate reference when cues to topic coherence from discourse structure were not available.

To test the reliability of these patterns, we ran a generalized linear mixed model predicting toy selection as an interaction between condition (*During* or *After*), trial type (trial location A or B), and age with random effects of participant. There was a significant three-way interaction between condition, trial type, and age ($\beta = 0.89$, $p = 0.03$), indicating that, with increasing age, participants were more likely to select Toy A in A trials and Toy B in B trials only for the *During* condition. No other factors were significant.

We also ran a series of paired *t*-tests to examine response differences between trial types (naming event with Toy A or Toy B) within condition (*During* or *After*) for each age group (2–3s, 3–4s, 4–5s, and 5–6s) (Table 3). Significant differences in reference selection were found across trial types in the *During* condition for children ages 3–6 years; at these ages, children were more likely to select Toy A in *During A* trials and Toy B in *During B* trials. By age 3, children were able to consider the broader discourse structure to help disambiguate the target referent.

In the naming events, children shared eye contact with the experimenter and were introduced to a novel word, but there were no indicators of the referent of the label other than its location within the broader discourse. Children's systematic responding to selecting the toy whose descriptions bracketed the naming event thus suggests that they can recognize and refer to discourse coherence to infer reference in the absence of other social cues.

Could participants have assumed that the novel labels referred to both toys at once? The uniqueness of the toys makes this situation unlikely. Toys were distinct artifacts with different colors, shapes, and functions, and items in a pair were presented at opposite ends of the table, giving no visual signal that the toys were grouped. Additionally, if participants believed that both toys in a pair were examples of a novel category, we would have observed responding at chance across both *During* and *After* conditions. Thus we do not believe that a superordinate interpretation of the novel terms would explain the pattern of data we observed.

Finally, children did not appear to use temporal proximity to disambiguate reference in *After* trials. We found no significant differences between reference selections across *After A* and *After B* trials for any age group, suggesting that children did not have consistent strategies for disambiguating the intended meaning of a novel term when topic coherence information is not available. While the *After* trials lacked definitive cues to establish reference, we were unsure whether adult users might have strategies for interpreting these trials.

Experiment 2

In Experiment 2, we extended our design to adult participants. We wanted to confirm adults' sensitivity to discourse continuity as a cue to word learning in *During* trials, and assess strategies for referent disambiguation in *After* trials to compare with our developmental results.

Methods

Participants Twenty-five adult participants were recruited from the San Jose Children's Discovery Museum, and were

Table 3: Results from paired t-tests examining response differences across trial types (naming location with Toy A or Toy B) within condition (*During* or *After*) for each age group (2–3s, 3–4s, 4–5s, and 5–6s). Children ages 3–6 years show significant differences in their response selections between *During A* and *During B* trials. No other significant differences across trial types are found.

	During condition			*After* condition		
Age	t-value	df	p-value	t-value	df	p-value
2–3	-1.23	15	0.24	-1.58	15	0.14
3–4	-2.52	15	**0.02**	-1.23	15	0.24
4–5	-5.48	15	<**0.01**	-1.95	15	0.07
5–6	-8.22	15	<**0.01**	-1.58	15	0.14

offered a sticker and certificate for their participation. They were informed that the task was designed for children. Only participants who reported using English at least 75% of the time were included in the study. One participant was excluded for reporting English use under this threshold.

Stimuli and procedure The stimuli and procedure were identical to Experiment 1, with the exception that adults did not undergo a training trial to practice pointing. Otherwise, adults were randomly assigned to either the *During* or *After* condition, and trial order and toy pairs were counterbalanced across participants.

Results and Discussion

Results were coded for whether participants selected the second toy (Toy B) as the referent of the novel label (see Figure 2). Participants almost never selected Toy B in *During A* trials, but were near ceiling at selecting Toy B for *During B* trials. Responses to each trial type were significantly different from chance in exact binomial tests ($p < 0.01$ for both trial types), and significantly different from each other ($\beta = -2.06$, $p = 0.01$) in a generalized linear mixed model predicting toy selection by condition (*During* or *After*) and trial type (trial location A or B) with random effects of participant. These results illustrate that adults were sensitive to naming events within discourse structure as informative cues to referent disambiguation; like the results in Experiment 1, adults demonstrate referent selections that correspond with discourse coherence, selecting the toy whose descriptions bracketed the naming event.

For *After* trials, adult participants were at chance in selecting Toy B in both *After A* and *After B* trials ($p > 0.3$ in exact binomial test for each). Performance was not significantly different between *After A* and *After B* trials ($\beta = 0.85$, $p = 0.15$). These findings indicate that adults did not exhibit a strategy for disambiguating reference in *After* trials; they were at chance in determining a referent when the naming event followed the descriptions of either toy. This pattern of results also parallels the developmental results. Adults, like children, did not show systematic response patterns when discourse information was not available.

Together, these results suggest that language users are sensitive to how information relates within discourse structure. Adults and children systematically disambiguated reference when they could infer topic coherence in *During* trials. In contrast, we found that listeners did not develop consistent response strategies for information that is isolated from social and discourse context in *After* trials. Neither children nor adults followed heuristics such as resolving reference by temporal association.

General Discussion

We investigated whether adults and children could use position in discourse as a cue to resolve reference. In our experiments, adults made use of discourse position effectively and children showed increasing sensitivity to discourse position across childhood. All groups except the youngest in our study successfully used discourse position to infer the mapping of a label. Taken together, our findings suggest that language users learn to make inferences about reference not only from pragmatic or social cues, but also from information about the general discourse in which a novel label is embedded.

Our experimental design ruled out two alternative explanations. The first is that children were simply selecting the referent most proximal to the naming event. In the *During* condition, this temporal proximity account would make the same predictions as a discourse-based account. Our *After* condition allowed us to rule out this possibility. While temporal proximity remains ambiguous in *After A* trials, it is unambiguous in *After B* trials because Toy B is the only toy described proximate to the naming event. However, children at all ages responded around chance for both *After A* and *After B* trials with no difference between naming location, suggesting that children did not use temporal proximity alone to make their judgments.

Our *After* condition rules out a second possible interpretation as well: that children's mappings are driven purely by novelty. By age 2, children state novel rather than given information in their productions (Baker & Greenfield, 1988) and apply similar expectations to other speakers by mapping new labels to items novel to the speaker's perspective (Akhtar et al., 1996). Children's behavior in the *During* condition is consistent with a novelty account: they chose the toy that was most recently attended to by the experimenter (the one newest to the discourse). But on this account, children should assign Toy B as the referent in *After A* trials, because the naming event directly precedes the introduction of this toy. Instead, responses from the *After A* condition were at chance between the two toys, suggesting that novelty alone also did not account for our findings.

Children can learn from ostensive naming events when they are available, but many situations they encounter are not overtly pedagogical, and the ability to extract information that is embedded in discourse may help children deduce the meanings of words in these cases. As our initial chinchilla example illustrated, discourse position is a powerful informa-

tion source for understanding language and for learning new words. Children who can infer how new information relates to the current topic may be able to accumulate knowledge more accurately and more efficiently. Because of its accessibility to young children in our study (and the possibility that even younger children might use discourse position in a simpler task), the use of discourse structure to help disambiguate reference might be one of the array of learning mechanisms that helps explain children's rapid vocabulary growth. Using topic coherence to make inferences about novel terms and information may allow children to access learning opportunities that would otherwise be unavailable.

Acknowledgments

Special thanks to the staff and families at the San Jose Children's Discovery Museum.

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, *67*(2), 635–645.

Baker, N., & Greenfield, P. (1988). The development of new and old information in young children's early language. *Language Sciences*, *10*(1), 3–34.

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 874–890.

Baldwin, D. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*(5), 832.

Clark, E. (1990). On the pragmatics of contrast. *Journal of Child Language*, *17*(02), 417–431.

Diesendruck, G., & Shemer, G. (2006). Young children's expectation of competence in word learning. *Journal of Child Language*, *33*(2), 321.

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*.

Markman, E., & Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive psychology*, *20*(2), 121–157.

Merriman, W., Bowman, L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the society for research in child development*.

Mervis, C., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child development*, *65*(6), 1646–1662.

Tomasello, M., & Barton, M. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, *30*(5), 639.