Title
Preschoolers infer contrast from adjectives if they can access lexical alternatives

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

ISSN
1069-7977

Authors
Horowtiz, Alexandra
Frank, Michael

Publication Date
2014

Peer reviewed

Preschoolers infer contrast from adjectives if they can access lexical alternatives

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

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

Abstract

When speakers use modified noun phrases (e.g. "the long book"), they provide information not only about a salient feature of a single item (that this book is long), but also about implicit contrasts with possible alternatives (books can vary by length: some may be short). We investigate the development of preschoolers' ability to detect implicit contrasts from speakers' use of adjectives and make inferences about category structure. In Experiment 1, we found that adults and preschoolers can make contrast inferences from adjective use in a supportive frame, and this ability improves over the preschool years. In Experiment 2, we reduced the cues to contrast and found that adults still inferred implied contrast from adjective use alone, but preschoolers did not. Perhaps the issue for preschoolers was an inability to consider alternatives from explicit descriptions (e.g. bringing to mind "short" from hearing "long"). Experiment 3 tested this hypothesis by reading preschoolers a book containing relevant opposite pairs immediately prior to the task. After reading the book, older 4-year-olds were able to make contrast inferences reliably, suggesting that increasing children's access to lexical alternatives may boost their ability to make contrast inferences.

Keywords: Pragmatics; adjectives; language development.

Introduction

A challenge for children learning language is not only to learn the explicit meanings conveyed by semantic content, but also to pick up on available, but implicit, information. For example, if I ask "where's my left shoe?" a listener can learn not only that I am missing my left shoe (explicit), but also that I'm likely not missing its pair (implicit). Thus, speakers can provide cues about unstated information through their word choices. The goal of the current work is to investigate children's sensitivity to this kind of implicit information.

Although children have accumulated a substantial vocabulary by age 5, they still show surprising difficulties in making pragmatic inferences. Consider the scalar implicature that "some of the horses jumped over the fence." For adults this utterance typically implicates that SOME BUT NOT ALL of the horses jumped, or else the speaker would have used the word "all" (Grice, 1975). In contrast, 5- to 6-year-olds are happy to accept the interpretation that "some" means SOME AND POSSIBLY ALL (e.g. Papafragou & Musolino, 2003; Noveck, 2000). Adults infer implicatures y across a number of lexical scales (e.g. <OR/AND>, <MIGHT/MUST>), while children find these inferences challenging despite knowing the meanings of the individual lexical items.

Why do children fail to make scalar implicatures despite their familiarity with the component words? This failure is especially puzzling given children's proficiency in making other forms of pragmatic inference, such as in simple word learning tasks where they can disambiguate the referent of a novel term with respect to a set of possible objects (e.g. Clark, 1990; Diesendruck & Markson, 2001; Akhtar, Carpenter, & Tomasello, 1996). Although both reference disambiguation and implicature are inferences about speakers' lexical choices, a critical difference between them is the space of possible alternative interpretations. In reference disambiguation tasks, the possible alternatives are typically objects that are physically present and hence easy to reason about; in implicature tasks, the possible alternatives are typically *unsaid* lexical choices that may be hard to summon to mind. We refer to this idea (that implicatures are challenging because of the difficulty bringing to mind the relevant alternative possible utterances) as the *linguistic alternatives hypothesis* (Barner & Bachrach, 2010).

In support of this hypothesis, Barner, Brooks, and Bale (2011) found that preschoolers could make implicatures from familiar scales such as numbers and from explicit descriptions ("the cat and the cow are sleeping" when three animals were pictured sleeping), but not from quantifiers ("some of the animals are sleeping"). In addition, even when the quantifiers were strengthened by "only" ("only some of the animals are sleeping"), they still failed to make the implicature—indicating that the implicature per se was not the problem (because if it were, adding "only" would have helped children restrict their interpretation of "some"), but rather that children's difficulty in considering the alternative "all" may be responsible. In our current work, we further investigate the linguistic alternatives hypothesis as a potential explanation of children's performance in other types of pragmatic tasks.

In our previous work, we investigated preschoolers' inferences about implicit dimensions of contrast from adjective use (Horowitz & Frank, 2012). The intuition driving this paradigm was that if a novel item is described as "tall," it is likely that others of that kind may vary by height; if it's described as "red," others may vary by color. Older 4-year-olds were able to make this kind of contrast inference from both color and size terms, but younger 4-year-olds exhibited a bias to match by color regardless of the adjective used. This result may also be related to children's ability to access the appropriate linguistic alternatives. Although children are familiar with color names, they may not recognize color use as contrastive because there is not a particular implied contrast item per se ("red" implies that others may be "not red," but not that another item will necessarily be a specific color). So our task may have been difficult for younger children because they could not bring to mind the appropriate color alternative.

In our current work, we modify our task to use scalar properties with familiar implicit alternatives to children (e.g.

"clean"/"dirty", "wet"/"dry") in order to determine whether their performance depends on the familiarity of linguistic alternatives. In Experiment 1, we found that adults (Experiment 1a) and preschoolers by age 3.5 (Experiment 1b) reliably made contrast inferences with these terms in a supportive context (contrasting with the younger children's failures with color terms in our previous work). In Experiment 2, adults' performance was sustained when framing cues supporting contrast were reduced (Experiment 2a), but preschoolers were at chance without these framing cues (Experiment 2b). In Experiment 3, before performing the same task as in Experiment 2b, we read a book highlighting the scalar opposites used in the test items (exposing children to the relevant linguistic alternatives for the later inferences). This exposure increased older children's contrast judgments. Overall, our findings support the hypothesis that children's performance in pragmatics tasks relates to their ability to consider lexical alternatives.

Experiment 1a: Adults

We first wanted to confirm adults' sensitivity to implicit contrast information conveyed though adjective use before investigating children's performance. We described a novel shape in a contrastive framing referencing either a feature or size adjective (e.g. "broken" or "small"), and asked adults to infer what other category members look like. If adults detect an implied dimension of contrast from adjective use, they should infer that other shapes are likely to vary along that property. This is precisely what we found.

Methods

Participants A planned sample of 128 adult participants was recruited from Amazon's Mechanical Turk online crowdsourcing service. Three subjects were excluded for failing to complete the task. All participants reported that they were native speakers of English and were informed that the task was designed for children.

Stimuli Participants were presented with cartoon images of items from outer space. They participated in a single trial featuring a set of three pictures: one referenced exemplar shape and two test shapes that each differed from the exemplar only by feature or only by size (Figure 1). Participants were randomly assigned to one of four sets of test shapes.

Procedures Participants read a story online in which a cartoon character, Allen the Alien, introduced them to a novel shape from outer space and said something about it, e.g. "This is a special kind of tibu. This is a [broken] tibu." Half of participants were presented with a feature adjective (e.g. "broken") and half were presented with a size adjective (e.g. "small"). They were then shown the two test shapes and asked, "What do you think [tibus] usually look like?", and prompted to select one of the two images. We measured the proportion of participants who selected the picture that contrasted with the named property.

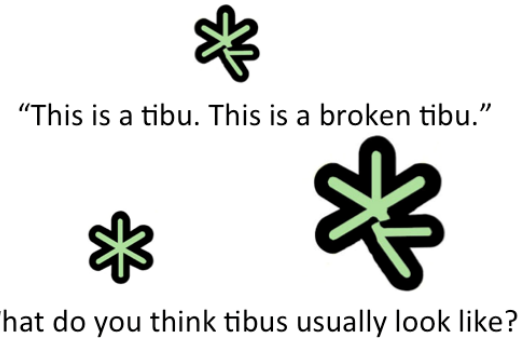

Figure 1: Example trial shape set. In the contrastive language conditions (Experiments 1a and 1b), the first sentence was modified from "This is a tibu" to "This is a special kind of tibu." All other expressions remained the same.

Results and Discussion

Responses were coded as correct if participants selected the shape that differed along the referenced dimension. In other words, we considered a response to be a correct contrast judgement if the participant selected the shape that differed by feature in feature adjective trials (e.g. heard "broken" and selected the shape that was unbroken), and differed by size in size adjective trials (e.g. heard "small" and selected the shape that was big).

Participants selected the contrasting dimension more often than chance and at nearly identical rates for both adjective types ($p < .001$ in exact binomial tests for feature and size terms; see Figure 2). Our results indicate that participants used the adjective referenced to make inferences about properties of novel category members, suggesting that adjectives are informative indicators of relevant property information to adults. They were able to consider the labeled property in order to infer that other novel category members are likely to differ along the referenced dimension. We next turned to investigate children's sensitivity to implicit information from adjective word choice.

Experiment 1b: Children

In Experiment 1b, we modified the task into a four trial storybook for children. Because young children are very familiar with feature and size opposites, we used a wide sample of 3- to 4-year-olds, broken into half-year age brackets. If children are sensitive to implicit contrasts conveyed through adjectives, they should select the image contrasting along the referenced dimension. If they do not recognize variable property cues from adjective use, then they should select at random or choose the image that matches (rather than contrasts with) the stated adjective.

Methods

Participants We recruited 96 3- to 5-year-old children from the Bing Nursery School at Stanford University and the San

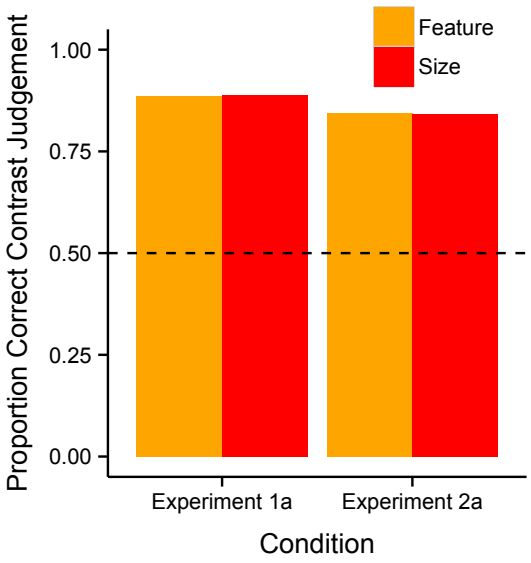

Figure 2: Adults' mean proportion correct performance in Experiments 1a and 2a. Feature trials are plotted in yellow and size trials in red. The dashed line represents chance (0.5).

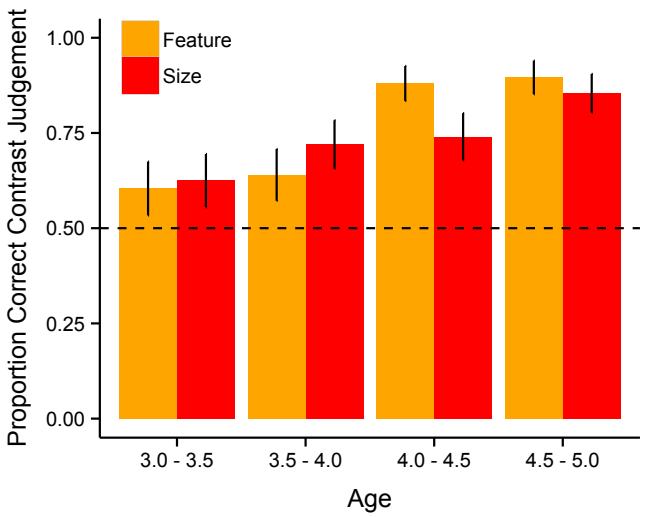

Figure 3: Preschoolers' mean proportion correct performance in Experiment 1b. Yellow bars depict feature adjective trials and red bars depict size trials. The dashed line represents chance (0.5). Error bars represent standard error.

Jose Children's Discovery Museum. Twenty-four children were recruited to each of four age groups: age 3.0 – 3.5 (M = 3;3), age 3.5 – 4.0 (M = 3;8), age 4.0 – 4.5 (M = 4;3), and age 4.5 – 5.0 (M = 4;8). About half of each age group was recruited from each location.

Stimuli We adapted the adult task into a physical book that the experimenter read with children. Each child participated in two training trials with pictures of familiar objects before undergoing four test trials comprised of the individual test sets used with adults. For each child, two of the test trials referenced a feature adjective, and two of the test trials referenced a size adjective. Test image sets were presented in one of two orders, and adjective type and image presentation were counterbalanced across participants.

Procedures We tested children individually in a quiet room at the nursery school or museum. They sat next to the experimenter at a table and were read a printed storybook with the images from Experiment 1a. Children were introduced to the character Allen the Alien and completed two training trials with common images to familiarize them with the task. If they did not select the correct image during a familiarization trial, they were prompted until the correct image was chosen.

The training trials were followed by four test trials. In each test trial, children were first shown an image of a single exemplar shape and heard Allen say something about it, e.g. "This is a special kind of tibu. This is a broken tibu." The experimenter then uncovered two test pictures, one that differed by the exemplar only by feature, and one that differed from the exemplar only by size. Children were asked, "What do you think tibus usually look like? What do most tibus look like?" and were prompted to point to one of the test images. The experimenter averted her eyes as children indicated their responses. Children also participated in a posttest following the test trials in order to demonstrate their knowledge of the adjectives used in the study. Test sessions took about 10 minutes to complete and were video-recorded.

Results and Discussion

Preschoolers in our task did show sensitivity to implied contrast dimensions conveyed through adjective use. Only the youngest children in our sample (ages 3 – 3.5 years) did not reliably select the adjective contrast above chance. By age 3.5, children were able to infer category membership from word choice cues, choosing the image contrasting by feature when a feature term was referenced and the image contrasting by size when a size term was referenced, and performance increased with age (Figure 3).

We analyzed our results using a logistic mixed model, predicting correct responses as an interaction between age and contrast type with random effects of participant and shape. Children increasingly made more correct contrast judgments with age ($\beta = 1.51$, $p < .0001$). There was no significant effect of contrast type (feature vs. size adjectives), and there was no interaction between age and contrast type, suggesting that participants across ages did not differ in their responses to different property types. Overall, these analyses show that children demonstrate an increasing sensitivity to implicit contrast information from adjectives.

The posttest data reveal that children overwhelmingly were knowledgeable about the terms we used in our task. Correct identifications averaged 90% ages 3.0 – 3.5, 95% ages 3.5 –

4.0, 96% ages 4.0 – 4.5, and 98% ages 4.5 – 5.0. These scores indicate that children's comprehension is improving with age, but that even the youngest children knew the terms we used.

Our results from Experiment 1b indicate that preschoolers by age 3.5 were sensitive to the adjective provided as an indicator of implicit contrast. Although each of the two test images were equally similar to the exemplar because each differed by only a single property, children *avoided* selecting the property match (i.e. selecting the picture that was the *same* property as the one referenced, e.g. hearing "broken" and selection the other broken image) and instead selected the image *differing* along the referenced dimension (e.g. hearing "broken" and selecting the picture that was unbroken). This suggests that preschoolers are able to consider the pragmatic implications of word choice in our task to infer that the adjective selection conveys information about a relevant property dimension of interest: remarking on a novel shape's size implies that size is salient and may vary across category members, while reference to a feature highlights that the feature property may vary by individual.

We next turned to investigate the robustness of these inferences. If listeners are sensitive to adjective choice generally, then they should be able to maintain inferences about implicit contrasts from the minimal cue of a modified noun phrase. If their recognition of the informativeness of adjective choice is more fragile, they may rely on a supportive linguistic framing (e.g. highlighting property salience with "This is a special kind of tibu...") to guide a contrastive interpretation.

Experiment 2a: Adults

In order to test the extent of adults' sensitivity to implicit contrasts, we reran Experiment 1a with the framing cues to contrast removed. We found that adults were just as likely to form contrast inferences from adjective use alone as they were with the supportive framing.

Methods

Participants A new planned sample of 128 adult participants were recruited from Amazon's Mechanical Turk online crowd-sourcing service. Two subjects were excluded for failing to complete the task. All participants reported that they were native speakers of English and were informed that the task was designed for children.

Stimuli Stimuli were identical to Experiment 1a.

Procedures Procedures were identical to Experiment 1a with the exception that the referential statement was reduced by removing the phrase "special kind of" so that listeners heard only "This is a [tibu]. This is a [broken tibu]." This subtle change allowed us to examine listeners' inferences from adjective use without drawing attention to it via the framing.

Results and Discussion

As above, we measured the proportion of correct contrast judgments for which participants selected the test picture that differed along the referenced property dimension. Adults performance was significantly about chance ($p < .001$ in exact binomial tests for feature and size terms) and did not differ by adjective type. They showed only a slight decrease in performance in this adjective only framing from the contrastive language framing in Experiment 1a (see Figure 2). These results indicate that adjective use in our task is a strong indicator of relevant property information of novel category members for adults. Their nearly equal performance across Experiment 1a and 2a suggests that adjectives provided salient cues to implicit contrast dimensions on their own without the necessity of additional semantic support.

Experiment 2b: Children

We reran Experiment 1b with the contrastive framing removed in order to examine children's sensitivity to adjective use alone. Although 3.5-year-olds reliably inferred contrasts from properties referenced in Experiment 1b, children a full year older still had difficulty succeeding in the present experiment without the support of contrastive framing.

Methods

Participants A new sample of 41 children was recruited from Bing Nursery School. Because of the presumed increased difficulty of this task, we recruited children from the old age groups: 4.0- to 4.5-year-olds (M = 4;3) and 4.5- to 5.0-year-olds (M = 4;8).

Stimuli Stimuli were identical to Experiment 1b.

Procedures Procedures were identical to Experiment 1b with the exception that the referential phrase was minimized by removing the phrase "special kind of" to reduce contrast cues other than the adjective. Instead, they heard only "This is a [tibu]. This is a [broken tibu]," isolating the adjective as the only available indicator of category membership.

Results and Discussion

Although preschoolers showed increasing contrast selections from adjective use with age in Experiment 1b, they were essentially at chance when the contrastive language framing was removed. We analyzed our results using a logistic mixed model, predicting correct responses as an interaction between age and contrast type with random effects of participant and shape, and we found no significant effects and no significant interaction. In post-hoc followup tests, older 4s showed a significant feature contrast bias ($p = .001$, exact binomial test), but this contrast was not reliable in the full model when controlling for participant and item effects, and may have been driven primarily by the "broken" and "clean" items. Although adults remained attentive to implicit contrast information in both the contrastive language and adjective only framings, children performed substantially worse without the additional linguistic cues to guide their contrast judgements.

One explanation for the discrepancy in children's performance from Experiment 1b to Experiment 2b is that they require more information than an adjective alone to cue a con-

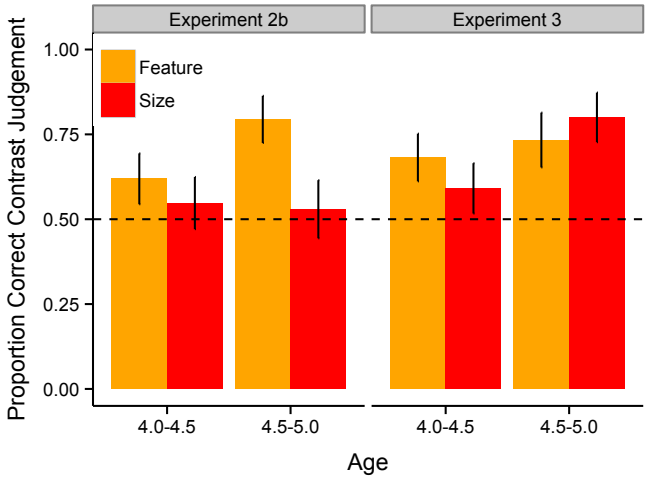

Figure 4: Preschoolers' mean proportion correct performance in Experiments 2b and 3. Feature adjective trials are plotted in yellow and size trials in red. The dashed line represents chance (0.5). Error bars show standard error.

trast inference. Although adults can conjure implicit contrast information from individual word choices, children may rely on the combination of informative lexical selections with the addition of supportive linguistic framing.

Another reason for children's shift in performance across experiments is not that they necessarily need contrast intentions to be explicitly conveyed, but rather that they require an awareness of what lexical alternatives *could* have been used in place of the ones chosen. Extending the linguistic alternatives hypothesis to our current task, children's contrast inferences may not relate to framing per se, but rather their access to scalar alternatives. We investigated this idea in Experiment 3 by increasing the availability of property contrasts.

Experiment 3

In Experiment 3, we provide a test of the linguistic alternatives hypothesis by increasing preschoolers' access to the relevant lexical alternatives. Before the experimental procedure, the experimenter read a seemingly unrelated book featuring the opposites referenced in the test trials (see Figure 5). This exposure to linguistic alternatives boosted older 4-year-olds' contrast selections.

Methods

Participants A new sample of 38 children from Bing Nursery School. Participants were grouped into two age groups: 4.0- to 4.5-year-olds (M = 4;3) and 4.5- to 5.0-year-olds (M = 4;8).

Stimuli Stimuli were identical to Experiment 1b and 2b with the addition of a separate training book read prior to the testing procedure. The training book consisted of clip art pairs of familiar images depicting the size and feature scalar contrasts portrayed in the test book (e.g. *small/big, broken/fixed*). Opposite pairs were labeled consecutively to maximize the salience of a given contrast dimension.

Procedures Children were told that they would be reading two books for the session. The procedure was identical to that of Experiment 2b with the addition of the opposites training book immediately preceding the test book. The experimenter read the training book with children, labeling the picture in a neutral way on each page (e.g. "This is a small teddybear. This is a big teddybear."). Although the properties used in both books were the same, no child explicitly noted any connection between the books.

Results and Discussion

Increasing preschoolers' access to relevant linguistic alternatives helped older 4s select property contrasts for both feature and size adjectives. These results suggest that supporting children's abilities to bring relevant alternatives to mind plays a strong role in their pragmatic inferences. Beyond relying on rich semantic framing cues to intended meaning, which are not always available in natural speech, reminding children of different types of modifiers increases their likelihood of forming contrast inferences from adjectives alone.

Older children selected the contrast property for both feature and size terms more often than chance ($p < .001$ and $p < 0.01$ respectively in exact binomial tests) and younger children for feature terms ($p = .01$ in exact binomial test), though younger children's performance did not differ across feature and size trials. A logistic mixed model predicting correct responses as an interaction between age and contrast type with random effects of participant and shape revealed no significant effects or interaction, however.

When we combine results with those of Experiment 2b, we find a three-way interaction between experiment, adjective type, and age, such that older children show improved contrast inferences for size terms only after the opposites book ($\beta = 2.60$, $p = 0.04$). Increased access to lexical alternatives seemed to help older children reliably select the dimension contrast according to the property.

Our results from Experiment 3 suggest that exposing children to a book of unrelated pictures with the scalar alternatives used in our test trials helped older children to select opposites more consistently, without any framing cues. An alternative hypothesis, that the initial book served to train children to always select named opposites, is not supported because we did not see a change in performance for the youngest children. In addition, anecdotally none of the children remarked on any relationship between the books, even though they conveyed the same adjective properties. Instead, we believe that the opposites book served to make the lexical scales more accessible to children so that, at least for the oldest children in our task, they could spontaneously infer implicit contrast information from an adjective produced alone.

Figure 5: Examples of training pairs used in Experiment 3. The teddybears depict a size contrast (small – big) and the piggybanks depict a feature contrast (broken – fixed).

General Discussion

We set out to investigate whether the linguistic alternatives hypothesis might be applied to pragmatic tasks beyond computing scalar implicatures. Our results suggest that children's performance in our tasks is related to their ability to consider relevant lexical alternatives. In Experiment 1, participants robustly selected the image contrasting with the property referenced when supported by contrastive language framing. In Experiment 2, adults maintained performance but preschoolers were at chance when framing was reduced. We suggest that children were unable to recognize implicit property variability from adjective use alone. In Experiment 3, we helped increase 4-year-olds' access to lexical alternatives by previewing the task with a book of opposites, and found that this boosted older childrens contrast inferences.

The inferences measured in our task are fairly counterintuitive: a correct contrast judgment requires selecting the property *contrast* instead of the property *match*, although both choices are available in each set. Our findings that young children are able to make contrast inferences demonstrates that they are sensitive to the pragmatic implications of speakers' word choices.

Their diminished success in the adjective only condition but recovered contrast inferences after the opposites exposure further suggests that the linguistic alternatives hypothesis may explain this pattern of performance; children can form pragmatic inferences when they recognize word choice as conveying implicit contrasts with relevant implicit alternatives, but fail when they cannot access implied lexical alternatives spontaneously from adjective use alone. In other words, children can recognize adjective use as conveying information about implicit contrast dimensions, but unlike adults, they may need additional cues supporting these inferences until they gain enough experience to form these inferences on their own.

The opposites exposure may help children in a few ways. Visually, images of familiar objects may make contrasts more salient. Aurally, hearing named opposites may help children better retrieve these terms, attend to the relevant property categories, or select matches for these specific lexical items. Attentionally, portraying contrast pairs may support children's ability to make comparison judgements (Christie & Gentner, 2010). In future work, we hope to identify how each of these contributions may impact children's inferences about implied contrast from word choice.

The ability to infer implicit information can allow children to learn about the world more efficiently. When children can recognize implied contrasts conveyed through word choices, they can learn not only about a particular instance (e.g. "This is a small tibu"), but can also form inferences about additional information conveyed about the speaker's knowledge or perspective of the world from their word choices (e.g. tibus are likely to vary by size). In order to form these inferences, children may need enough experience with language and particular property comparisons to recognize pragmatic opportunities. This hypothesis could explain both why children fail to compute scalar implicatures for weak quantifiers (Barner et al., 2011) and why they failed to form contrast inferences in our previous work with color terms (Horowitz & Frank, 2012). Informative underlying cues to meaning are constantly conveyed through speech, and children sensitive to these cues will be able to learn more effectively and efficiently from their interactions.

Acknowledgments

Special thanks to the staff and families at the Bing Nursery School and 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*, 635–645.

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, *60*, 40–62.

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*, 84–93.

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, *11*, 356–373.

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

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630.

Grice, H. P. (1975). Logic and conversation. *1975*, 41–58.

Horowitz, A., & Frank, M. C. (2012). Learning from speaker word choice by assuming adjectives are informative. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society.*

Noveck, I. A. (2000). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*, 165–188.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, *86*, 253–282.