Research Report

Preschoolers' Perspective Taking in Word Learning

Do They Blindly Follow Eye Gaze?

Erika Nurmsoo and Paul Bloom

Yale University

ABSTRACT—*When learning new words, do children use a speaker's eye gaze because it reveals referential intent? We conducted two experiments that addressed this question. In Experiment 1, the experimenter left while two novel objects were placed where the child could see both, but the experimenter would be able to see only one. The experimenter returned, looked directly at the mutually visible object, and said either, "There's the [novel word]!" or "Where's the [novel word]?" Two- through 4-year-olds selected the target of the speaker's gaze more often on* there *trials than on* where *trials, although only the older children identified the referent correctly at above-chance levels on trials of both types. In Experiment 2, the experimenter placed a novel object where only the child could see it and left while the second object was similarly hidden. When she returned and asked, "Where's the [novel word]?" 2- through 4-year-olds chose the second object at above-chance levels. Preschoolers do not blindly follow gaze, but consider the linguistic and pragmatic context when learning a new word.*

When learning an object name, young children often assume that the name refers to the object the speaker is looking at, even if the children themselves are looking at a different object when the word is used (e.g., Baldwin, 1991). Eye gaze cannot be a necessary cue for word learning; blind children can learn object names, after all, and children can infer that a word refers to an object even if the speaker is not looking at it (e.g., through mutual exclusivity—Markman & Wachtel, 1988; when learning names for absent referents—Akhtar & Tomasello, 1996). Even so, eye gaze does seem to be *sufficient* for word learning.

Address correspondence to Erika Nurmsoo, University of Bristol, Department of Experimental Psychology, 12a Priory Rd., Bristol, BS8 1TU, Great Britain, e-mail: nurmsoo@aya.yale.edu.

What does this indicate about how children learn the meanings of words? Such results are often taken as demonstrating the importance of social cognition in word learning, under the assumption that children use a speaker's eye gaze as a cue to referential intent (e.g., Bloom, 2000). According to another account, however, children attend to a speaker's line of regard because of simpler, possibly unlearned, orienting responses (e.g., Moore & Corkum, 1994; but see Woodward, 2003); once they have done so, they come to associate the object they are attending to and the word they are hearing (Smith, Jones, & Landau, 1996; Plunkett, 1997). The precise details of such claims differ, but they share the view that, regardless of what role social cognition plays in other aspects of word learning, this primary and early-emerging sensitivity to eye gaze is unmediated by any inferences about the speaker's referential intent.

The two accounts can be distinguished by what they predict about how mandatory this sensitivity to eye gaze is. If children use gaze cues out of an understanding that eye gaze reflects the speaker's meaning, gaze information should be exploited when it reveals the speaker's intent, but ignored when it is irrelevant or uninformative. We tested this hypothesis in two studies.

EXPERIMENT 1: VISUAL PERSPECTIVE

In Experiment 1, the child and the speaker explored two novel unnamed objects together. Then, in the speaker's absence, the objects were placed such that the speaker would be able to see only one of them, although the child could see both. The speaker returned, looked at the mutually visible object, and said, "There's the [novel word]!" or "Where's the [novel word]?" If children mandatorily follow eye gaze to a target object when learning a new word, they would be expected to map the word to the mutually visible object in both conditions. Alternatively, if children are sensitive to the speaker's visual perspective and use gaze information in word learning only when it is informative, they would be expected to disregard eye-gaze cues on *where* trials and instead map the word to the object hidden from the speaker.

Method

Participants

Thirty-two older children (M = 49 months, range = 46–54 months; 16 boys and 16 girls) and 32 younger children (M = 31 months, range = 25–36 months; 13 boys and 19 girls) were tested in a university lab or at their day-care center.

Materials

We used an opaque cardboard screen (32 cm × 22 cm) containing two compartments, one with and the other without a window. This screen was placed between the child and the speaker so that the child could see into both compartments, but the speaker could see only through the window into one compartment.

We used four pairs of novel objects, and the novel labels *spoodle*, *nurmy*, *flurg*, and *gorp*. The target object, its location relative to the window, and the left/right position of the window in the screen were all counterbalanced.

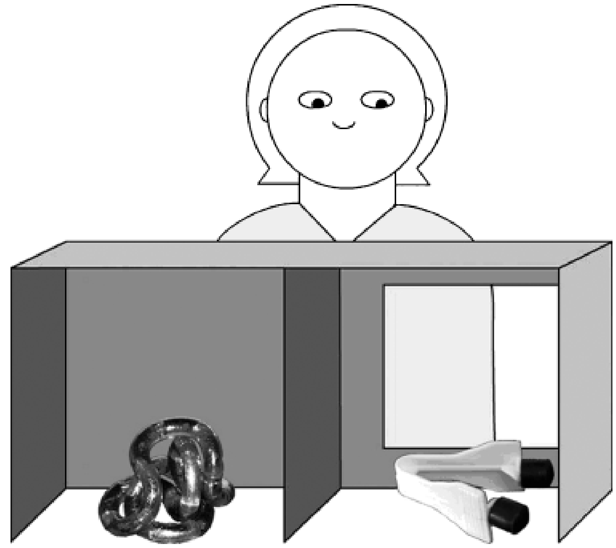

Fig. 1. The event shown during Experiment 1, with sample novel objects. On all trials, the experimenter fixed her gaze on the object in the window. On *there* trials, the experimenter said, ''There's the [spoodle/nurmy/flurg/gorp]! There it is!'' On *where* trials, she asked, ''Where's the [spoodle/nurmy/flurg/gorp]? Where is it?''

Procedure

In the familiarization phase, each participant was shown what the screen looked like from both sides. Two familiar toys were placed on the child's side of the screen, one in each compartment. The child was asked to identify which toy the speaker could see, and which she could not see. The screen was then rotated so that the child had the speaker's perspective and could see only the object in the window.

In the experimental phase, there were two *where* trials and two *there* trials, identical except for the test question. On each trial, the child explored a pair of novel objects with two adults (the speaker and the assistant). The speaker left the room or turned her back while the assistant placed each object in its compartment. From the child's perspective, both objects were visible, one in the compartment with the window and the other in the compartment without the window. When the speaker returned (or turned around), only the object in the window was visible from her perspective (see Fig. 1). She fixed her gaze on the object in the window and asked the test question: On *there* trials, the speaker said, "Oh! There's the [novel label]! There it is!" On *where* trials, she said, "Oh! Where's the [novel label]? Where is it?"

In both conditions, the speaker then looked up at the child, held out her hand, and asked, "Can I have the [novel label]?"

Results and Discussion

Older children correctly selected the mutually visible object on an average of 1.53 of the 2 *there* trials, whereas younger children averaged 1.38 correct *there* trials. Both levels of performance were above chance, $ts(31) = 4.48$ and 3.00, p_{rep}s = .996 and .97, respectively. On *there* trials, the speaker's eye gaze was directed at the correct target, and both older and younger children correctly used this cue to identify the referent.

Older children correctly selected the hidden object on an average of 1.41 of the 2 *where* trials, whereas younger children averaged 1.00 correct *where* trials. Only older children's performance was consistently above chance level, $t(31) = 3.23$, $p_{rep} = .97$. On *where* trials, the speaker's gaze information was uninformative, as it was directed at an incorrect object. Older children were able to disregard these gaze cues and identify the correct referent, although the younger children were not.

To determine whether children in each age group selected the visible object more often on *there* trials (when it was correct) than on *where* trials (when it was incorrect), we gave each participant a score of 1 if he or she selected the visible object more often on *there* trials and a score of −1 if he or she selected the visible object more often on *where* trials. Sign tests revealed that participants in both age groups showed the correct pattern of responses; both the older and the younger children selected the visible object more often on *there* trials than on *where* trials, ps < .05.

Children at both ages appear to have used the adult's eye-gaze cues flexibly, relying on gaze direction to map the novel word only when gaze direction was informative. The younger children, although uncertain of the correct answer in the *where* condition, nevertheless did not err by selecting the target of the speaker's eye gaze when it was uninformative. The older children appear to have determined the speaker's referential intent in the *where* condition.

An alternative analysis, however, is that the older children employed the strategy of selecting any object that was hidden from the speaker when asked a *where* question. That is, rather than making an inference about referential intent, they might have simply responded to this sort of question by looking for a hidden object. We explored this alternative in Experiment 2.

EXPERIMENT 2: KNOWLEDGE PERSPECTIVE

Method

Participants

Eighteen older children (M = 46 months, range = 41–56 months; 11 boys and 7 girls) and 18 younger children (M = 31 months, range = 29–33 months; 9 boys and 9 girls) were tested in a university lab or their day-care center. No child participated in both experiments.

Materials

An opaque cardboard screen (33 cm × 25 cm) was used. Objects placed behind the screen were visible to the child, but not to the speaker seated on the opposite side (see Fig. 2). We used three pairs of novel objects, and the novel labels *fendle*, *nurmy*, and *toma*. Each object served as target and as distractor an equal number of times.

Procedure

As in Experiment 1, in the familiarization phase of Experiment 2, participants were introduced to the screen from the speaker's and the child's perspective.

There were two *knowledge* trials and one *control* trial, identical except for the timing of the placement of the objects. On all trials, the child first explored a pair of novel objects with the speaker and the assistant. On knowledge trials, the speaker then placed one of the objects on the child's side of the screen, where the child could see it but the speaker could not. The other object remained in plain view as the speaker left or turned her back, at which point the assistant placed it behind the screen. Both objects were therefore hidden to the speaker, but she knew the location of one of them because she had placed it behind the screen herself. The speaker returned (or turned around), feigned surprise, and asked, "Where is the [novel label]? Where is it?" She held out her hand and asked, "Can I have the [novel label]?" The control trial was identical, except that both objects were placed behind the screen in the speaker's absence, and she could be seen as ignorant about the location of both.

The control trial appeared equally often as the first, last, or middle trial. Only one control trial was used because pilot testing revealed that children found the trial frustrating, as there was no clear correct answer.

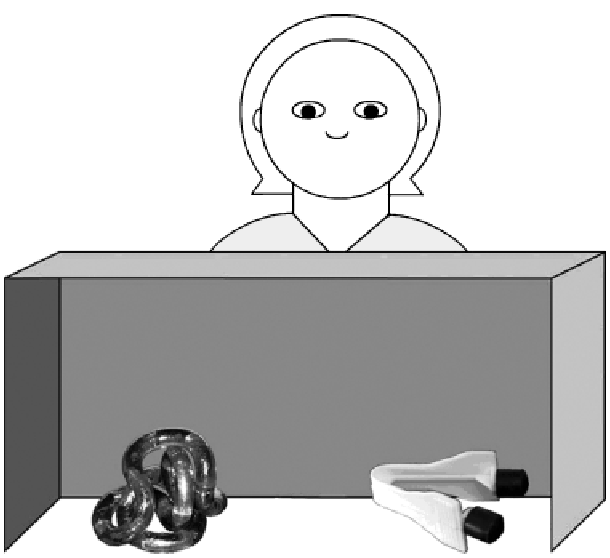

Fig. 2. The event shown during Experiment 2, with sample novel objects. On all trials, the experimenter asked, ''Where's the [fendle/nurmy/toma]? Where is it?''

Results and Discussion

On knowledge trials, the speaker was ignorant of the location of the object that was hidden second, in her absence. Children received a point each time they correctly selected this object, for a possible maximum score of 2. Children occasionally selected both objects (2 older children on a total of three trials, 1 younger child on one trial). This response was coded as incorrect, as the object hidden second was not chosen as the sole referent of the target word. The average scores were 1.72 for older children and 1.61 for younger children. Children in both age groups answered correctly significantly more often than predicted by chance, $ts(17) = 5.33$ and 5.17, respectively, $p_{rep}s = .996$.

On the control trial, the speaker was ignorant of the location of both objects. On this trial, many children responded by selecting both objects (8 older and 5 younger children). Of the children who selected only one object, 4 of 10 older children and 8 of 13 younger children chose the object hidden second, a level of performance no different from chance, $\chi^2(1, N = 14) = 0.4$, $p > .53$, and $\chi^2(1, N = 13) = 0.7$, $p > .40$, respectively.

GENERAL DISCUSSION

In our experiments, children ages 2 to 4 were able to take the perspective of a speaker (Experiments 1 and 2), and 3- to 4-year-olds were able to override eye-gaze cues when they were not relevant (Experiment 1). These findings might seem surprising, given that older children and adults have difficulty with similar tasks. For example, Epley, Morewedge, and Keysar (2004) explored perspective taking by having a speaker ask participants to move objects from one location to another in an upright array of boxes. They found that 6-year-olds and even adults often failed to take into account the speaker's visual perspective, responding as if the speaker could see an object that was visible only to the participant (but see Nadig & Sedivy, 2002). This task differs from our own in that it involves a much more complex array of objects and requires participants to determine the referents of existing words, not to make sense of new ones. Our study shows that 4-year-olds, and to some extent 2-year-olds, are able to consider the speaker's perspective in a referential task when given a simpler context.

Our findings are consistent with those of Moll and Tomasello (2006), who found that when an adult asked for help searching for an object, 2-year-olds did not offer the object in plain view, but instead fetched the object visible to the child but hidden from the adult. In this task, however, the adult's eye gaze alternated between the visible object and the location of the hidden object, whereas in the parallel condition in our Experiment 1, the adult's eye gaze was directed at the mutually visible nontarget object, making the mapping problem more difficult. Although 2-year-olds were not able to correctly select the hidden object on these trials, as a group they did not mistakenly map the novel word to the target of the adult's gaze.

These findings provide support for a pragmatic approach to word learning, in which children do not simply follow surface cues to the speaker's referential intent, but instead seek to understand the situation as a whole. Studies consistent with this perspective have shown that somewhat older children do not learn a new label despite clear referential cues toward a novel object when the reliability of the speaker is called into question. For example, Koenig and Harris (2005) showed that 4-year-olds, and in some cases 3-year-olds, preferred to learn a novel label from a previously reliable speaker rather than from one who claimed ignorance of the names of familiar objects or who labeled them inaccurately (see also Birch, Vauthier, & Bloom, in press; Jaswal & Neely, 2006; Nurmsoo & Robinson, in press). And Sabbagh and Baldwin (2001) found that when a speaker explicitly claimed ignorance about the correct referent of a novel word (e.g., "I don't know what a blicket is. Maybe it's this one."), 3- and 4-year-olds did not learn the word. The present set of studies adds to this body of research, by showing that even when faced with a speaker who shows every sign of competent naming, and no uncertainty about the meaning of a novel label, children do not rely solely on eye-gaze cues to determine the referent of the new word, but use additional pragmatic or linguistic information to infer the correct target.

Experiment 2 demonstrated that when two objects were hidden from the speaker, who could be seen as ignorant about the location of only one of them, 2- to 4-year-olds correctly identified the speaker's intended referent when she asked, "Where is the [novel word]?" This finding converges with research suggesting that children use other cues to a speaker's referential intent when eye-gaze information is not sufficient. For example, when preschoolers are told, "Let's find the toma," they will map *toma* to the object that satisfies the searcher even when multiple novel objects are examined (and rejected; Tomasello & Barton, 1994) or when the unseen novel object is hidden in a locked toy barn (Akhtar & Tomasello, 1996). Somewhat older children are sensitive to a speaker's false beliefs when interpreting the meaning of a novel word used to refer to the contents of a container, mapping the word to the object that the speaker falsely believes is hidden within (Carpenter, Call, & Tomasello, 2002; Happe & Loth, 2002).

There is considerable evidence, then, that children will track other cues to a speaker's referential intent when there are no eye-gaze cues that can be used to identify the correct target. We have shown here that preschoolers are sensitive to a speaker's referential intent even when eye-gaze cues are present and directed at an incorrect target. Preschoolers do not blindly follow the direction of gaze.

REFERENCES

Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology*, *14*, 79–93.

Baldwin, D.A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*, 875–890.

Birch, S.A.J., Vauthier, S., & Bloom, P. (in press). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Carpenter, M., Call, J., & Tomasello, M. (2002). A new false belief test for 36-month-olds. *British Journal of Developmental Psychology*, *20*, 393–420.

Epley, N., Morewedge, C.K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*, 760–768.

Happe, F., & Loth, E. (2002). 'Theory of mind' and tracking speakers' intentions. *Mind & Language*, *17*, 24–36.

Jaswal, V.K., & Neely, L.A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, *17*, 757–758.

Koenig, M.A., & Harris, P.L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*, 1261–1277.

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

Moll, H., & Tomasello, M. (2006). Level I perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*, 603–613.

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, *14*, 349–372.

Nadig, A.S., & Sedivy, J.C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*, 329–336.

Nurmsoo, E., & Robinson, E.J. (in press). Identifying unreliable informants: Do children excuse past inaccuracy? *Developmental Science*.

Plunkett, K. (1997). Theories of early language acquisition. *Trends in Cognitive Sciences*, *1*, 146–153.

Sabbagh, M.A., & Baldwin, D.A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070.

Smith, L.B., Jones, S.S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, *60*, 143–171.

Tomasello, M., & Barton, M. (1994). Learning words in non-ostensive contexts. *Developmental Psychology*, *30*, 639–650.

Woodward, A.L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, *6*, 297–311.

(Received 5/30/07; Revision accepted 8/27/07)