Available online at www.sciencedirect.com
ScienceDirect

Journal of Memory and Language 57 (2007) 596–615

Journal of Memory and Language

www.elsevier.com/locate/jml

Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation

Joy E. Hanna [a,*], Susan E. Brennan [b]

[a] Department of Psychology, Oberlin College, 120 W. Lorain Road, Oberlin, OH 44074, USA
[b] Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, USA

Received 2 July 2006; revision received 22 January 2007
Available online 16 April 2007

Abstract

In two experiments, we explored the time course and flexibility with which speakers' eye gaze can be used to disambiguate referring expressions in spontaneous dialog. Naive director/matcher pairs were separated by a barrier and saw each other's faces but not their displays. Displays held identical objects, with the matcher's arranged in a row and the director's mirroring the matcher's or else in a circle (Experiment 1) or in a reversed row (Experiment 2). Directors instructed matchers to move targets, which were unique or had a competitor nearby or far away. When mirrored displays held far competitors, matchers used directors' eye gaze to identify targets before the linguistic point of disambiguation. Reversed displays caused substantial competition, yet matchers still identified targets before the linguistic point of disambiguation, showing an ability to rapidly re-map directors' eye gaze. Our findings indicate eye gaze is a powerful and flexible disambiguating cue in referential communication.

Keywords: Discourse processing; Conversation; Eye gaze; Ambiguity resolution; Speech

Speech unfolds in time. This means that words, expressions, and utterances are temporarily ambiguous; much research has focused on how the language processing system copes with this ambiguity. Over the past decade, unobtrusive eyetracking technology has enabled the detection of very early effects of context upon the interpretation of ambiguous utterances, especially during interactive discourse (e.g., Henderson & Ferreira, 2004; Metzing & Brennan, 2003; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell & Tanenhaus, 2005). But the focus has largely been on addressees' eye gaze as a *measure* of interpretation. In this paper, we focus on a different aspect of eye gaze: speakers' eye gaze as a *constraint upon* interpretation. Where a speaker is looking is potentially a powerful cue about attention and intention in face-to-face communication. Yet no one to date has systematically examined the effects of speakers' visual orientation on addressees' processing of ambiguous spoken utterances.

The ability to use eye gaze cues develops very early in life. Infants not only are highly sensitive to another person's head position, but attend to the other's eyes by 2 months of age, to the direction of gaze by around 6 months, and, with accuracy, to the object that the other is fixating by around 12 months (Butterworth & Grover, 1990; Phillips, Wellman, & Spelke, 2002). It is unclear whether very young children truly interpret another's gaze as referential (as opposed to simply mimicking

* Corresponding author.
E-mail address: Joy.Hanna@oberlin.edu (J.E. Hanna).

0749-596X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2007.01.008

another's head turn in the same direction and happening to fixate upon the same object). Yet even 12 month olds will look more often and longer at an object when the eyes of another person turning toward it are open rather than closed (Brooks & Meltzoff, 2002). The ability to achieve shared visual attention facilitates word learning (Baldwin, 1995; Bloom, 2002; Repacholi, 1998; Tomasello, 1995) and heralds the developing awareness of other people's intentionality (Argyle & Cook, 1976; Brooks & Meltzoff, 2002; Phillips et al., 2002).

The ability to detect another animal's focus of attention most likely has evolutionary benefits. Many nonhuman animals are sensitive to one another's eye gaze; eye contact may signal threat, and gaze aversion may signal submission (Argyle & Cook, 1976). Sensitivity to eye gaze cuts across species; some animals are able to interpret the direction of human eye gaze, including dolphins (Tschudin, Call, Dunbar, Harris, & van der Elst, 2001) and dogs (Soproni, Miklósi, Topál, & Csányi, 2001).

Adult humans are quite accurate at detecting one another's face-directed gaze at normal conversational distances (Argyle & Cook, 1976; Pusch & Loomis, 2001). They use the orientation of both head and eyes, which often move together. People are most accurate in detecting the direction of gaze from someone who is directly facing them; when an interlocutor's head is turned 30° to the side, gaze is detected with an error of about 3 ° (Gibson & Pick, 1963). When head and eyes are oriented toward somewhat different directions, gaze is detected as aiming somewhere in between (Argyle & Cook, 1976). People tend to overestimate the amount of gaze that is directed at their own faces (e.g., interpreting over the shoulder gaze as face-directed, particularly at greater interpersonal distances), and they tend to underestimate the angle of gaze as the gazer turns away from their faces (Argyle & Cook, 1976).

In ordinary conversations, interlocutors attend closely to each other's patterns of eye gaze (as anyone knows who has ever tried to surreptitiously peek at their watch during a conversation). Interlocutors spend a great deal of time looking at either each other as well as at any objects under discussion (Argyle & Cook, 1976). The addressee gazes more at the speaker than the speaker gazes at the addressee (Kendon, 1967); in fact, the lack of an addressee's visual attention may lead a speaker to interrupt herself and restart an utterance in order to capture his attention (Goodwin, 1981).

On the production side, speakers' eye movements to objects show scanning patterns that reflect the incremental encoding of utterances on conceptual, syntactic, and phonological levels (Griffin, 2001; Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998; van der Meulen, Meyer, & Levelt, 2001). Under some circumstances, communication is more successful when a listener's eye movements in a scene correspond to those of a speaker's during a separate scene description (Richardson & Dale, 2005). Since eye movements are a naturally occurring accompaniment to reference production, they are informative, and it is possible that addressees can use them as visual cues during reference resolution. However, almost no work has investigated the time course with which gaze cues are available or used in a conversational setting. In fact, most referential communication experiments have prevented visual contact between participants (using barriers between interlocutors or by having confederate speakers wear sunglasses to conceal where they are looking; Hanna & Tanenhaus, 2004; Metzing & Brennan, 2003). One of the few studies to allow visual contact (Boyle, Anderson, & Newlands, 1994) showed that in a matching task, people who could see each other tended to look up from their displays, at each other, at points of difficulty; however, this result was mentioned without explanation.

The information available in eye gaze is itself ambiguous. A gazer may fixate an object while processing it for the first time, while searching for a particular target, while inspecting it in the service of another decision, while planning or interpreting a referring expression, or in the service of programming a motor action such as reaching toward the object. Looking back and forth between two objects may indicate that the gazer finds a referring expression to be truly ambiguous between those two referents, or that she has simply noticed an interesting coincidence. If gaze is being used communicatively, the gazer may fixate an object as a form of pointing, to (intentionally) draw an interlocutor's attention to it, or (less intentionally) while waiting for an interlocutor to provide evidence that he is attending to it. A few studies have examined computer mouse movements as signals about an interlocutor's visual attention during a collaborative task (e.g., Brennan, 1990, 2005); however, pointing with eye gaze has the potential to be an even earlier signal (if a more volatile and less intentional one) than pointing with a mouse. It is likely that people who observe another's eye gaze can process it, whether it was produced instrumentally or intentionally; in one study, computer programmers were faster to find and identify bugs in a screen of software code after they had viewed the (prerecorded) eye gaze cursors of other programmers finding the same bugs over an indistinct version of the same screen (Stein & Brennan, 2004).

In the current project, we raise three questions about the potential role of eye gaze cues in face-to-face communication:

- First, *can eye gaze be used in the resolution of temporary ambiguity?* Other sorts of information that is available when people are visually co-present (such as information about the state of a task, e.g., Brennan, 1990, 2005; Clark & Krych, 2004) have a profound effect on how they coordinate reference resolution and collaborative activity. Where a speaker is looking may certainly be informative about that speaker's utterance planning, but this does not necessarily mean that an addressee can use the information on-line. An interlocutor's eye gaze is a signal with its own dynamic characteristics that must be integrated with speech or action in order to be useful. Other studies that have used the visual worlds paradigm have looked at visual information as a static display against which linguistic expressions are unpacked. A few studies have looked at language processing with dynamic visual displays, where a partner's visual attention is represented as a moving cursor (e.g., Brennan, 1990, 2005; Brennan, Chen, Dickinson, Neider, & Zelinsky, in press). None of these have incorporated the speaker's face and triangulation of gaze into the information available to the addressee. The question is whether an interlocutor's eye gaze, produced dynamically and interactively, can be used as easily as other more stable cues in a visual display.

- Second, if eye gaze is used by addressees, *what is the time course by which it is integrated with linguistic processing?* Some have argued that partner-based or other contextual influences on language processing are achieved via inferences that are made *after* initial (and automatic) processing, useful only for downstream adjustments or repairs of initial egocentric interpretations (e.g., Horton & Keysar, 1996; Keysar, Barr, Balin, & Brauner, 2000). However, strong evidence exists that visual information, such as the nature of the objects present in a display, the actions it is possible to perform with them, and the identity of the speaker, can be integrated early during the resolution process (Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003; Metzing & Brennan, 2003; Nadig & Sedivy, 2002; Tanenhaus et al., 1995). Monitoring speakers' eye gaze may speed reference resolution, especially if addressees are drawn to look where speakers are looking, constraining the domain of interpretation.
- Third, *does eye gaze serve as an automatic orienting cue, or as a flexible cue?* If eye gaze is used automatically (an addressee is drawn to look where a speaker is looking), then it should be helpful in interpretation only when what a speaker sees corresponds to what an addressee sees. In this case, we would expect an early, automatic use of eye gaze that would be difficult or impossible to adjust to a context in which the interlocutors' views were not congruent. This result would be consistent with evidence that eye gaze operates as a reflexive attention-orienting mechanism (e.g., Langton & Bruce, 2000; Langton, Watt, & Bruce, 2000) even when it is uninformative (e.g., Friesen & Kingstone, 1998; Friesen, Ristic, & Kingstone, 2004). The other possibility is that a speaker's direction of gaze may be ignored or else flexibly remapped if the speaker's perspective is incompatible with the addressee's. On a constraint-based view, which proposes that multiple probabilistic sources of information are integrated simultaneously in order to constrain interpretation (Jurafsky, 1996; MacDonald, 1994; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Taraban & McClelland, 1988), eye gaze may act as just one of several possible constraints or cues that can be flexibly weighted. If this is the case, then people should be able to modify their use of eye gaze information depending on their goals, the communicative context, or their spatial perspectives.

In two referential communication experiments, we explored the time course and flexibility with which eye gaze can be used in spontaneous spoken dialog. Pairs of naive participants were separated by a low barrier so they could see each other's faces but not each other's displays of colored shapes. We balanced naturalness (having participants interact spontaneously) with control (evoking predictable linguistic expressions) by priming directors during the practice trials to produce referring expressions with the desired linguistic point of disambiguation. We predicted that while the director was producing a referring expression, her head and eye orientation might provide early cues to resolve ambiguity between the intended referent and a similar competitor.

Experiment 1

Method and design

Directors and matchers sat across from one another at a table separated by a low barrier. Their displays held identical copies of the same objects located either in a mirrored arrangement, so that a given object that was to the director's right was aligned with its duplicate object to the matcher's left (Congruent displays), or else in arrangements that were spatially uninformative (Non-congruent displays). In Experiment 1, the directors' Non-congruent displays had objects arranged in a circle in the center of the display, rather than left-to-right like matchers' displays. Both partners were aware of when their displays were spatially correspondent and when they were not. Display types were blocked, and all pairs experienced both types of displays, with the display order counterbalanced between participants.

During each trial, directors worked from a non-verbal schematic that indicated which objects they needed to tell matchers to select, and where they needed to place them. On critical trials, in addition to the target object (e.g., a blue circle with five dots), there was a highly similar competitor (e.g., a blue circle with six dots) located either in the space next to the target (Near competitor) or farther away (Far competitor), or else there was no competitor of the same color (No competitor). See Fig. 1 for a schematic of the Congruent No competitor, Near competitor, and Far competitor displays from the matcher's point of view. Matchers' eye movements were recorded with a head-mounted eyetracker, and directors' were captured via the scene camera mounted on the matchers' head. Therefore, we could link the matcher's moment-by-moment interpretation to visual points of disambiguation (where the director was looking) as well as to linguistic points of disambiguation (the point at which a referring expression could map to only one object in the display).

Predictions

We predicted that when displays were Congruent and there was No competitor, matchers would begin looking at the target more often than any other shape immediately after the onset of the disambiguating color word; this would replicate the standard finding that eye movements to an object in a relevant visual display usually rise after around 200 ms of the linguistic point of disambiguation (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus et al., 1995). In addition, we predicted that when a same-color competitor was Near, reference resolution would be delayed until the linguistic point of disambiguation, at the number of dots. When the same-color competitor was Far away on a Congruent display, however, we predicted that the director's eye gaze could be used by the matcher to disambiguate the referring expression prior to the linguistic point of disambiguation. Finally, we expected that for the Non-congruent displays, where the director's eye gaze does not map clearly onto the target location, reference resolution would be slower.

Materials

The barrier between the partners consisted of a three foot wide, low vertical display board. Each side of the board contained six two and a half inch wide shelves, equally spaced, that held six objects. The board was a foot high, which (in combination with adjustable chairs) allowed partners to see each other's faces, but prevented them from seeing each other's displays.

The objects were circles, triangles, and squares cut out of red, yellow, green, and blue poster board. Shapes were either blank or had randomly arranged clusters of black dots on them (e.g., a blue circle with five dots). Twelve shape arrays were constructed and paired with a two-instruction sequence to create nine experimental items and three fillers. Each array had six shapes associated with it: a target shape, two potential competitor shapes, and three filler shapes. On Near and Far competitor trials, the target and competitor shapes were

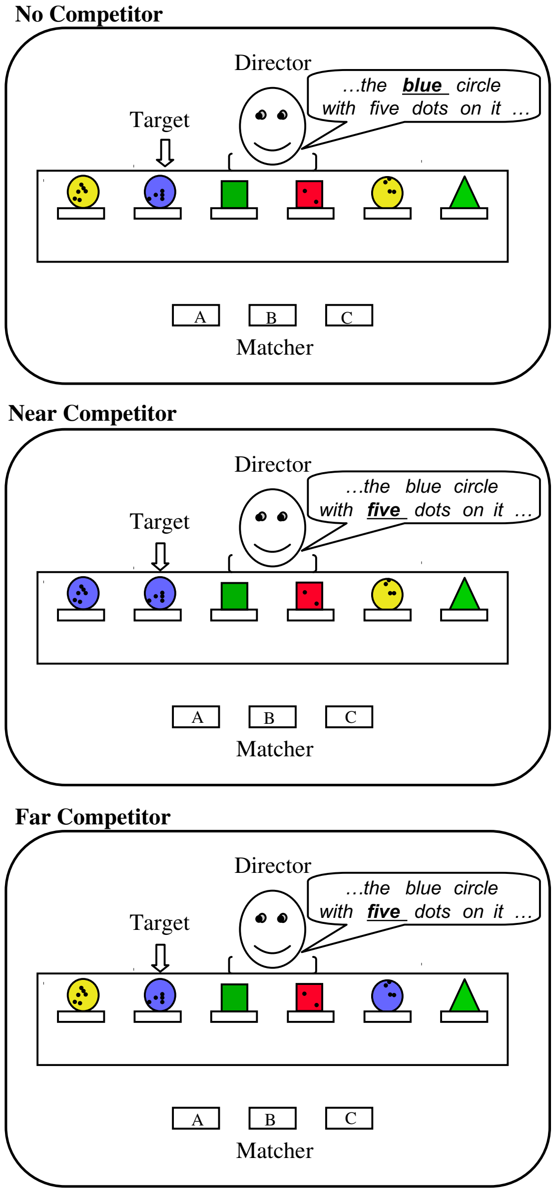

Fig. 1. Example displays for a single item rotated through the three competitor conditions (No competitor, Near competitor, and Far competitor) when the matchers' and directors' displays were spatially congruent (Congruent displays in Experiment 1, Mirror displays in Experiment 2). The displays are presented from the matcher's perspective. For Non-congruent displays in Experiment 1, the director's display (on the other side of the board) was arranged in a circle, and the director's eye gaze for this item would be central. For Reverse displays in Experiment 2, the director's display was the reversed order of the matcher's, and the director's eye gaze for this item would be to the other side of the board.

identical except for their number of dots and were the only two shapes of their color (e.g., the only two blue shapes). On No competitor trials, the target was the only shape of its color, and potential competitors were the same shape but a different, matching, color (e.g., the target was a blue circle with *five dots*, and the potential competitors were both yellow circles with different numbers of dots on them). The presence of potential competitors that were not referred to, in the No competitor experimental items and the three fillers, counteracted any possibility that matchers would expect the director to always refer to one of two nearly identical shapes in the display. The target and competitor shape locations were balanced across the experiment. The number of dots on the shapes varied and on critical trials had different initial consonants (e.g., blue circles with five dots versus six dots). Thirteen additional filler trials with arrays of abstract black and white geometric pictures known as Tangrams were also spread evenly throughout the experiment. The 12 critical shape arrays were repeated, with different spatial locations, in the Congruent and Non-congruent halves of the experiment.

In order to get directors to produce clear eye movements to the target, the director's shelves were numbered, and the schematic instruction card for each trial indicated only which number object needed to be moved. The instruction card for each trial was placed centrally, below the line of sight for the display objects. Therefore, directors had to first look down to read the instruction card and then look up at the appropriate shelf in order to describe the shape to the matcher. In order to encourage a period of gazing at the target by directors that might be perceptible by matchers, target shapes had to be described by their color, shape, and number of dots in order to be distinguished from the competitor. Directors' looks to a target shape were easily distinguishable from looks to a far away competitor, but were not easy to distinguish from looks to a nearby competitor. The first instruction with a given set of shapes was always the critical one. In order to encourage the director to give instructions of the form "the [color] [shape] with [number of] dots", the experimenter modeled this as a way of describing the objects during the practice trials.

Participants

Forty-eight Stony Brook University undergraduates (41 women, 7 men) volunteered for research credit or were paid for their participation. Participants were paired on the basis of availability; seven of the pairs knew each other. All were native speakers of English and were naive to the purpose of the experiment.

Later upon debriefing, none of the participants reported becoming aware of the experimental manipulations, although 26 of them (15 matchers, 11 directors) did notice that where the other person was looking could help them resolve which shape was being referred to. Of these, four matchers reported actually trying to ignore where the director was looking "because it might be cheating." Most participants reported thinking that the experiment had something to do with the Tangram (filler) trials.

Procedure

In order to introduce participants to the person they were paired with and encourage them to be comfortable working with and looking at one another, we had pairs first work on solving a puzzle together for 5 min. They were then randomly assigned the roles of director and matcher and were told that they would be working together to move some shapes and pictures. For each set of objects, the task was for the director to instruct the matcher to move two of the display objects to one of three other spaces (labeled A, B, or C) located on the matcher's side of the table. They completed two practice trials (one shape and one Tangram) during which the experimenter provided an example about how to refer to the shapes. Both partners were made aware of the corresponding or non-corresponding nature of their displays at the start of each of the two blocks of experimental trials. They were encouraged to communicate with each other if they had any difficulties. After the experiment, participants were fully debriefed.

Matchers' eye movements were monitored with an ISCAN ETL-500 head-mounted eyetracker, which has a temporal resolution of 60 Hz (or 30 frames per second) and a spatial resolution of better than 1 ° across a 20 ° range. Directors' eye gaze and verbal instructions were captured in the videotape from the eyetracker's head-mounted scene camera. Calibration of the eyetracker was performed after the puzzle task and before the practice trials.

Coding

Directors' spontaneous utterances were transcribed. To achieve control, studies of spoken utterance interpretation often have the same or similar utterance spoken by a confederate (Hanna, 2001; Keysar et al., 2000; Metzing & Brennan, 2003), although some studies have used naive speakers (Brown-Schmidt, Campana, & Tanenhaus, 2005; Brown-Schmidt, Gunlogson, Watson, Faden, & Tanenhaus, 2006; Kraljic & Brennan, 2005). Our speakers were all naive, and so their utterances varied somewhat, although the experimenter's example during the practice trials was successful in priming the desired form of referring expression (color, shape, modifier). Some utterances contained pauses or brief restarts. We eliminated the 25 cases (less than 1% of the data) in which speakers produced the wrong instruction or repaired utterances in such a way that the time course of interpretation was not comparable or could not be analyzed. We also removed one pair from the analysis completely because the director made an excessive number of speech errors.

The data from 23 pairs were coded from the videotape recordings using an editing VCR with frame-by-frame control and synchronized video and audio channels. Matcher looks were coded by noting which shape matchers were looking at frame-by-frame, with at least three frames (100 ms) of gazing required to constitute a look. Looks to the director were coded as well. Director looks were coded frame-by-frame from the videotape by noting whether the director was looking down (at the instruction card), at the matcher, at the target/same side of the display as the target, or at the other side of the display. Looks were coded from the onset of the director's verbal instruction until the matcher reached for the target object.

Data analysis

Fig. 2a and b shows the proportion of matchers' looks to the director's face and to objects in the display over 4 s, zeroed and aligned at the onset of the color word for all conditions. In the six panels in Fig. 2a, pairs had Congruent displays during the first half of the experiment and Non-congruent displays during the second half; in the six panels in Fig. 2b, pairs had Non-congruent displays first and then Congruent displays. To analyze the patterns of gazing in each condition, we computed the time matchers spent looking at the target, competitor, other objects, and director's face in four time windows, starting at the onset of the color word: 0–500, 501–1000, 1001–1500, and 1501–2000 ms.

Our statistical analyses were of two sorts. First, to look for effects of more-informative or less-informative directors' gaze cues in situations with and without temporary ambiguity, we computed two (by-subjects and by-items) $3 \times 2 \times 2 \times 4$ omnibus ANOVAs on matchers' looks to the target object: Competitor (No competitor, Near, vs. Far) × Display (Congruent vs. Non-congruent) × Order (Congruent first vs. Non-congruent first) × Window. Then, to examine how differences waxed and waned over time, we computed pairs of similar ANOVAs (Competitor × Display × Order) for each of the four windows, using planned contrasts to compare Far vs. No competitor and Near vs. No competitor. Statistical tests from these ANOVAS are reported in Table 1. Interactions with Order are discussed in the text.

Looking at a target object does not necessarily mean that an ambiguous expression has been resolved; a better measure is to compare the waxing and waning (respectively) of looks to targets and competitors. Our second set of analyses tested for reliable differences (or competition) between same-color target and competitor objects. For each of the four 500 ms windows, we computed two (by-subjects and by-items) $2 \times 2 \times 2 \times 2$ omnibus ANOVAs on matchers' looks to potentially ambiguous same-color objects: Object (Target vs. Competitor) × Competitor (Near vs. Far) × Display (Congruent vs. Non-congruent) × Order (Congruent display first vs. Non-congruent display first). This was followed by a complete set of contrasts for time spent gazing at targets vs. at competitors in each condition. Statistical tests and contrasts are given in Table 2.

Results

Remarkably, when matchers needed to distinguish targets from same-color competitors (bottom four panels in each of Fig. 2a and b), they were able to use directors' eye gaze well before the linguistic point of disambiguation (LPOD, marked on each graph by the second vertical line). Overall, these panels show a consistent pattern of rising looks to the target and falling looks to the competitor during the time when the referring expression was still linguistically ambiguous; in the competitor conditions, spoken disambiguating information was not heard until, on average, 1689 ms ($SD = 751$) after the onset of the color word.[1] Recall that the standard finding in psycholinguistic studies of spoken language processing is for eye movements to a referent object to begin rising at about 200 ms after the linguistic point of disambiguation.

Looks to target objects

When there was No competitor, matchers began looking at the target about 200 ms after the onset of the disambiguating color word in Non-congruent displays (top right panel, Fig. 2a) and even slightly *before* this point in Congruent displays (top left panel, Fig. 2a); a similar pattern of rising looks to targets appears to be slightly (although not significantly) delayed in Fig. 2b, in which participants had experienced Non-congruent displays before Congruent displays. In the first window, the order in which partners had

[1] While there was substantial variability in LPODs because directors' utterances were spontaneous, there were no reliable differences in mean LPODs across the competitor conditions (ANOVAS by-subjects and by-items for Near vs. Far competitor × Display showed no effects). To rule out the possibility that early looks were driven by the early LPODs in the distribution *within* a condition, we computed correlations of LPODs and target looks within each of the four conditions, for each of the four windows (the 16 contrasts on the lower half of Table 2). The only reliable contrast in Table 2 for which LPOD was correlated with gaze to target was for Incongruent displays/Far competitors in Window 4, $r = -.326$, $p = .01$, $N = 63$. That particular finding was not among the predictions; in fact, the correlation may explain in part why it emerged.

experienced the kinds of displays mattered; the rise in looks to the target began slightly earlier, yielding more looks to the target in the first window, $F_1(1,21) = 6.04$, $p = .023$; $F_2(1,08) = 6.29$, $p = .036$; Min $F'(1,24) = 3.08$, $p < .10$, when partners saw the Congruent displays before the Non-congruent ones rather than the other way around. This suggests that matchers may have made better use of any partner gaze cues when they experienced spatially correspondent cues (consistent with co-presence in real life) early in the experiment; with Non-correspondent displays first, they learned to inhibit or avoid using partner gaze as an orienting cue, having first experienced it as uninformative.

Table 1 shows that target looks increased over time, with a strong linear trend over the four windows in the omnibus ANOVA, $F_1(1,21) = 220.67, p < .001$; $F_2(1,8) = 157.05$, $p < .001$, Min $F'(1,20) = 91.75$, $p < .001$. Over all competitor conditions and windows matchers spent 53.8% of the time gazing at targets when displays were congruent with directors' displays and 49.9% when they were not. This difference was reliable when they saw the Congruent displays during the first half of the experiment (when they spent 58.8% of the time gazing at targets in Congruent displays and 48.2% gazing at targets in Non-congruent displays), but not for the reverse order (with Non-congruent displays first, they spent 45.8% of the time gazing at to targets in Congruent displays and 49.0% at targets in Non-congruent displays); this was a Display × Order interaction in the omnibus ANOVA, $F_1(1,21) = 9.70$, $p = .005$; $F_2(1,8) = 11.33$, $p = .01$; Min $F'(1,25) = 5.23$, $p < .05$.

Overall, having a competitor did result in ambiguity, in the form of fewer looks to the target in the presence of either Near or Far competitors compared to No competitor over the four 500-ms windows after the onset of the color word (which was equal to the LPOD when there was No competitor), $F_1(1,21) = 30.53$, $p < .001$; $F_2(1,8) = 16.68$, $p < .004$; Min $F'(1,17) = 10.79$, $p < .01$.

We expected that directors' eye gaze would be most informative early on for Congruent displays with competitors located far away from targets; this emerged as a Competitor × Display × Window interaction, $F_1(1,21) = 16.09$, $p = .001$; $F_2(1,8) = 10.65$, $p = .01$; Min $F'(1, 19) = 6.41$, $p < .03$. We explored this further in the contrasts summarized in Table 1. In the first 500 ms window (and only in this window), matchers gazed at targets (whether they were Near or Far) a higher proportion of the time when displays were Congruent than Non-congruent. In fact as far as very early looks to the target are concerned, there is some evidence that when a competitor was spatially distinct, it was like having No competitor at all. In this situation, matchers looked at the target in the first 500 ms after the color word as often (33.2% of the time) when the competitor was Far away as when there was No competitor at all (30.05% of the time). These levels of gazing were greater than those when the competitor was Near (22.5% of the time) (see Competitor contrasts for Window 1 of Table 1).

Competition

Additional evidence for early disambiguation comes from the fact that in conditions where directors' gaze was informative, matchers gazed at targets more than at competitors, beginning just after the onset of the color word. In the condition we expected to be most informative (Congruent displays with Far competitors, lowest left-hand panels of Fig. 2a and b), the first 500 ms contained 117 ms more gaze to targets than to competitors (see Table 2). This target advantage was reliable in the first window but not in the second and third windows, when matchers tended to look back and forth between the two similar objects (hence the noisy curves in these two windows); in the fourth window, looks to the target began to rise again and looks to the competitor fell. Although the difference between target and competitor in the fourth window was reliable by-subjects, it was not by-items. This may have been due to both matchers and directors tending to look back and forth at targets and competitors, perhaps to count and compare the numbers of dots distinguishing them (we address this issue with revised stimuli in Experiment 2).

When competitors were adjacent to targets in Congruent displays (middle left panels of Fig. 2a and b), there was substantial competition in the first two windows, and then targets attracted marginally more gaze than did competitors in the third window and reliably more in the fourth window. We had expected this Near competitor condition to be less informative than the corresponding Far competitor condition; however 1/3 of the Near trials had pairs of same-color objects located near the center of the display, right between the director and the matcher, so these cases appear to have been more informative than we expected. Certainly, people are better at detecting differences in a face-to-face partner's angle of gaze when objects to be distinguished are located directly in front of the gazer rather than obliquely (Gibson & Pick, 1963). In addition, any remaining competition could be dispelled rather rapidly; matchers could look back and forth between Near competitors more rapidly than between Far competitors.

As for the Non-congruent competitor conditions (lowest right hand panel in both Fig. 2a and b), these show a large amount of competition in the first three windows. Even so, there is evidence that matchers were able to use Non-congruent eye gaze cues to some extent. In the Far competitor conditions and by the fourth window, the higher proportions of looks to targets than to competitors (Table 2) indicated matchers had resolved the ambiguity. Even though we had intended for the director's gaze to the circular display in the

Non-congruent condition to be entirely uninformative when mapped to the matcher's display, it seems that matchers may have nevertheless been able to benefit from having directors' gaze cues (at the start of the session, they were shown the directors' displays in relation to their own and so were aware of the mapping).

Discussion

The results of Experiment 1 provide clear answers to two of our questions about the potential role of eye gaze cues in face-to-face communication. Eye gaze produced by a speaker can be used by an addressee to resolve a temporary ambiguity, and it can be used early. In this experiment, the directors' eye gaze helped matchers reliably begin fixating the target more often than the competitor well before the linguistic point of disambiguation; as predicted, the effect was strongest and earliest for the Congruent displays and when the competitor was spatially separated from the target. In fact when there were Far competitors, the proportion of target looks was as high immediately after the color word as

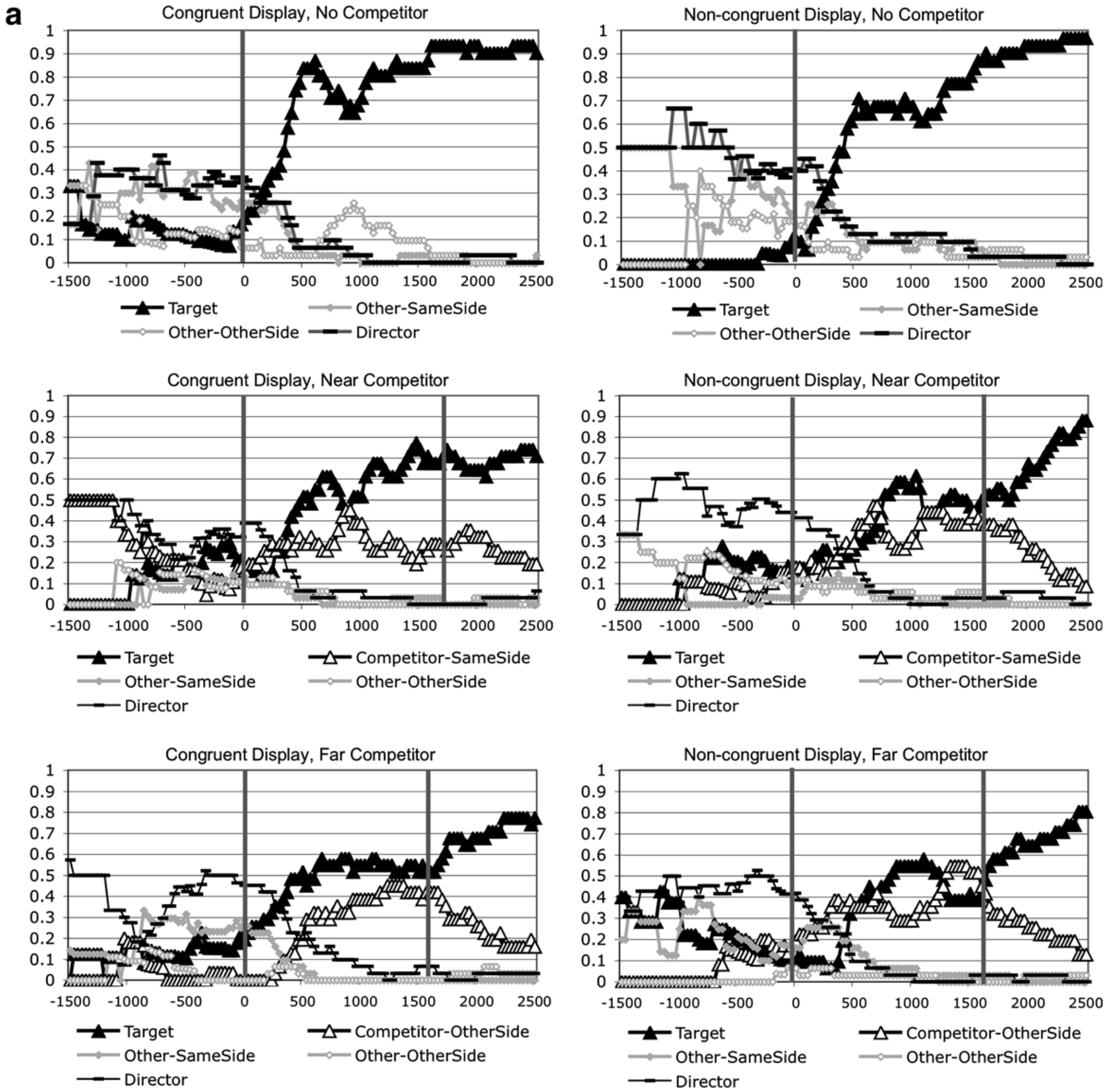

Fig. 2. Proportion of matcher's looks to the display objects and to the director over time in Experiment 1 for the No competitor (top panels), Near competitor (middle panels), and Far competitor (bottom panels) conditions, when the displays were Congruent (left panels) and Non-congruent (right panels). Objects other than the target are labeled as being on the same side of the display as the target, or the other side. The zero point is the onset of the color word and is marked with a vertical line; this is also the linguistic point of disambiguation (LPOD) for the No competitor conditions. The LPOD is marked with the second vertical line in the Near and Far competitor conditions. In (a) the Congruent displays were presented first; in (b) the Non-congruent displays were presented first.

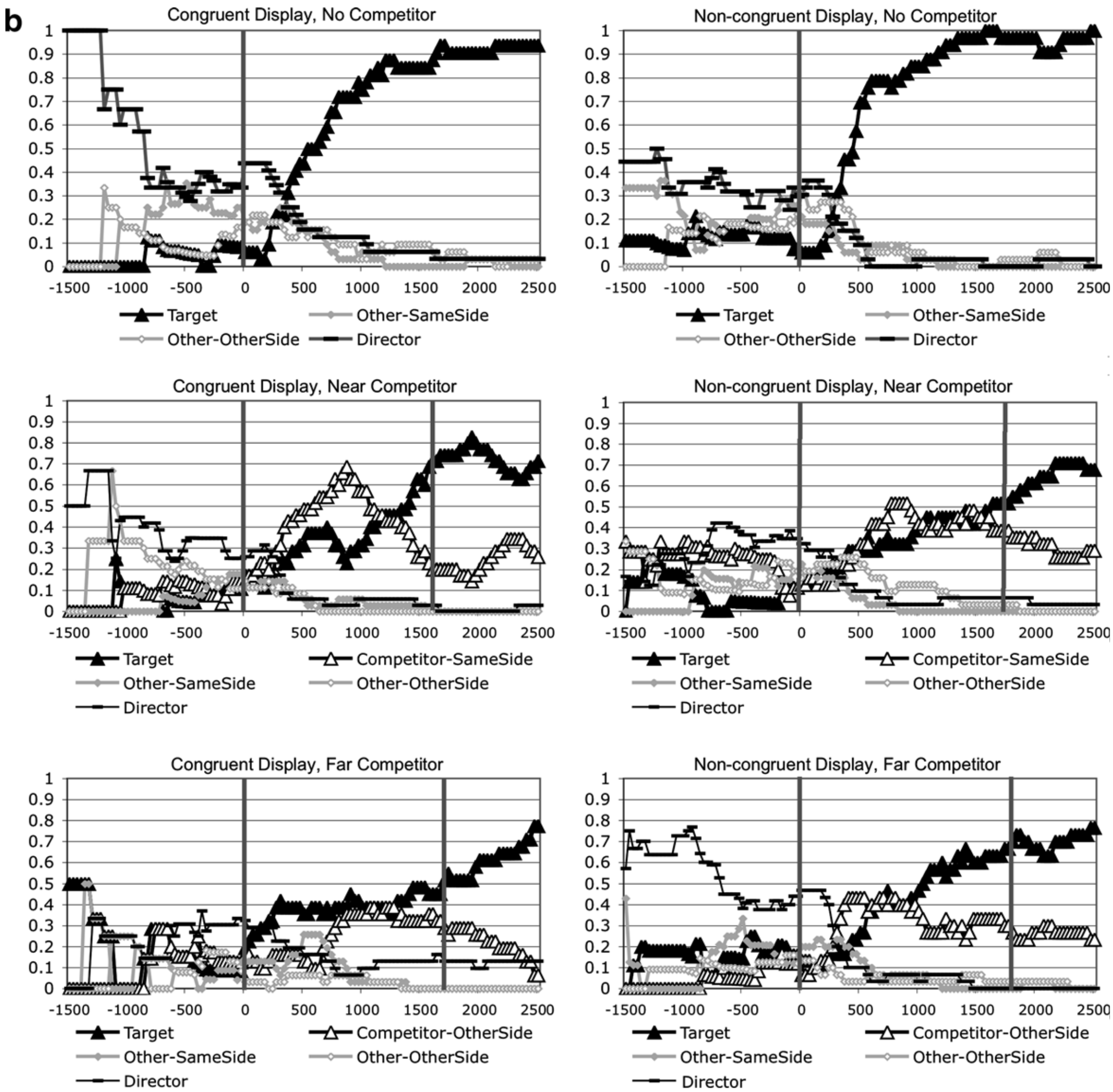

Fig. 2 (*continued*)

when there was No competitor. This evidence supports the view that eye gaze is another partner-based source of information that can constrain addressees' interpretations from the earliest moments of language processing.

Our results also hint at an answer to our third question. The immediacy with which eye gaze helped reference resolution is not inconsistent with the idea that eye gaze is an automatic orienting cue. However, our coding in this experiment began at the onset of the director's spoken utterance, which caused us to miss any fixations that took place after the instruction card was given to the director, but before she began speaking. We noticed that the director usually began looking reliably at the target before beginning to speak. In order to test the idea that eye gaze has an immediate attention-orienting effect, we needed to identify a relationship between the director's and matcher's eye movements in this period of time before there is any linguistic information available. Therefore, for Experiment 2, we began coding the trials from the point at which the experimenter placed the instruction card in front of the director in order to capture the directors' and matchers' earliest eye movements. Even more important, in order to more fully address the question of whether or not eye gaze is a cue that can be weighted differently and used flexibly, we needed to create a situation in which it might be possible for matchers to systematically remap the difference between the directors' and their own displays. We therefore changed the directors' spatially Non-congruent display so that it was a duplicate

Table 1
Matcher gaze at target object after onset of color term in Experiment 1

Factors	Window 1: 0–500 (ms)	Window 2: 501–1000 (ms)	Window 3: 1001–1500 (ms)	Window 4: 1501–2000 (ms)
Display	$F_1(1,21) = 11.51$, $p = .003$ $F_2(1,8) = 7.07$, $p < .03$ Min $F'(1,18) = 4.38$, $p = .05$	$F_1(1,21) = .47$, $p = .50$ $F_2(1,8) = .09$, $p = .77$ Min F' is n.s.	$F_1(1,21) = .24$, $p = .63$ $F_2(1,8) = .05$, $p = .84$ Min F' is n.s.	$F_1(1,21) = 1.163$, $p = .29$ $F_2(1,8) = .44$, $p = .53$ Min F' is n.s.
Competitor				
Near vs. No	$F_1(1,21) = 6.87$, $p < .02$ $F_2(1,8) = 3.12$, $p = .11$ Min $F'(1,16)=2.15$, $p = .16$	$F_1(1,21) = 42.67$, $p < .001$ $F_2(1,8) = 15.32$, $p = .004$ Min $F'(1,14)=11.27$, $p < .005$	$F_1(1,21) = 34.20$, $p < .001$ $F_2(1,8) = 17.15$, $p = .003$ Min $F'(1,16)=11.42$, $p < .005$	$F_1(1,21) = 38.89$, $p < 001$ $F_2(1,8) = 22.22$, $p = .002$ Min $F'(1,18)=14.14$, $p = .001$
Far vs. No	$F_1(1,21) = 1.78$, $p = .20$ $F_2(1,8) = .66$, $p = .44$ Min F' is n.s.	$F_1(1,21) = 47.03$, $p < .001$ $F_2(1,8) = 24.42$, $p = .001$ Min $F'(1,17)=16.74$, $p < .001$	$F_1(1,21) = 40.04$, $p < .001$ $F_2(1,8) = 30.00$, $p < .001$ Min $F'(1,20)=17.15$, $p < .001$	$F_1(1,21) = 58.62$, $p < .001$ $F_2(1,8) = 13.15$, $p = .007$ Min $F'(1,12) = 10.74$, $p < .01$
Display × Competitor				
Near vs. No	$F_1(1,21) = .11$, $p = .75$ $F_2(1,8) = .03$, $p = .88$ Min F' is n.s.	$F_1(1,21) = .88$, $p = .36$ $F_2(1,8) = .44$, $p = .53$ Min F' is n.s.	$F_1(1,21) = 1.09$, $p = .31$ $F_2(1,8) = .43$, $p = .53$ Min F' is n.s.	$F_1(1,21) = 6.72$, $p < .02$ $F_2(1,8) = 10.67$, $p = .01$ Min $F'(1,27) = 4.12$, $p = .05$
Far vs. No	$F_1(1,21) = 1.41$, $p = .25$ $F_2(1,8) = 2.17$, $p = .18$ Min F' is n.s.	$F_1(1,21) = .38$, $p = .54$ $F_2(1,8) = .55$, $p = .48$ Min F' is n.s.	$F_1(1,21) = .41$, $p = .53$ $F_2(1,8) = .28$, $p = .61$ Min F' is n.s.	$F_1(1,21) = .10$, $p = .76$ $F_2(1,8) = .08$, $p = .78$ Min F' is n.s.

ANOVAs of Display × Competitor for each of four 500-ms windows, with planned contrasts for Near competitor vs. No competitor and Far competitor vs. No competitor; significant or marginal effects highlighted.

Table 2
Matcher gaze at target (T) versus same-color competitor (C) after onset of color term in Experiment 1

	Window 1: 0–500 (ms)	Window 2: 501–1000 (ms)	Window 3: 1001–1500 (ms)	Window 4: 1501–2000 (ms)
Factors				
Object gaze (Target vs. Competitor)	$F_1(1,21) = 1.0$, $p = .33$ $F_2(1,8) = .15$, $p = .71$ Min F' is n.s.	$F_1(1,21) = .31$, $p = .59$ $F_2(1,8) = .76$, $p = .41$ Min F' is n.s.	$F_1(1,21) = 3.78$, $p = .06$ $F_2(1,8) = 5.7$, $p = .04$ Min $F'(1,27)=2.27$, $p < .14$	$F_1(1,21) = 27.36$, $p < .001$ $F_2(1,8) = 7.61$, $p = .03$ Min $F'(1,19) = 5.95$, $p < .03$
Object × Display	$F_1(1,21) = 4.76$, $p = .04$ $F_2(1,8) = 4.6$, $p = .06$ Min $F'(1,23) = 2.34$, $p < .14$	$F_1(1,21) = .83$, $p = .37$ $F_2(1,8) = .25$, $p = .63$ Min F' is n.s.	$F_1(1,21) = .17$, $p = .68$ $F_2(1,8) = .10$, $p = .76$ Min F' is n.s.	$F_1(1,21) = 2.33$, $p = .14$ $F_2(1,8) = 1.88$, $p = .21$ Min F' is n.s.
Object × (Near/Far) Competitor	$F_1(1,21) = 1.18$, $p = .29$ $F_2(1,8) = .47$, $p = .51$ Min F' is n.s.	$F_1(1,21) = 3.40$, $p < .08$ $F_2(1,8) = 2.62$, $p = .14$ Min $F'(1,19) = 1.67$	$F_1(1,21) = .12$, $p = .74$ $F_2(1,8) = .12$, $p = .74$ Min F' is n.s.	$F_1(1,21) = .08$, $p = .77$ $F_2(1,8) = .35$, $p = .57$ Min F' is n.s.
Interaction: Object × Display × (Near/Far) Competitor	$F_1(1,21) = 13.16$, $p = .002$ $F_2(1,8) = 10.09$, $p = .01$ Min $F'(1,20) = 5.73$, $p < .03$	$F_1(1,21) = .52$, $p = .48$ $F_2(1,8) = 1.12$, p = .32 Min F' is n.s.	$F_1(1,21) = 1.50$, $p = .23$ $F_2(1,8) = 1.45$, $p = .26$ Min F' is n.s.	$F_1(1,21) = 2.94$, $p = .10$ $F_2(1,8) = 9.6$, $p = .02$ Min $F'(1,29) = 2.25$, $p < .15$
Conditions				
Congruent display Near competitor	T (113 ms), C (134 ms) $F_1(1,21) = .59$, $p = .45$ $F_2(1,8) = .11$, p = .75 Min F' is n.s.	T (213 ms), C (230 ms) $F_1(1,21) = .15$, $p = .71$ $F_2(1,8) = .01$, $p = .94$ Min F' is n.s.	T (275 ms) > C (180 ms) $F_1(1,21) = 4.58$, $p = .04$ $F_2(1,8) = 4.23$, $p = .07$ Min $F'(1,22) = 2.20$, $p = .15$	T (357 ms) > C (123 ms) $F_1(1,11) = 21.60$, $p < .001$ $F_2(1,8) = 27.19$, $p < .001$ Min $F'(1,25) = 12.04$, $p < .002$
Congruent display Far competitor	T (166 ms), C (49 ms) $F_1(1,21) = 17.52$, $p < .001$ $F_2(1,8) = 6.42$, $p = .04$ Min $F'(1,14) = 4.70$, $p < .05$	T (231 ms), C (139 ms) $F_1(1,21) = 2.43$, $p = .13$ $F_2(1,8) = 4.78$, $p = .06$ Min F' is n.s.	T (238 ms) > C (194 ms) $F_1(1,21) = .38$, $p = .55$ $F_2(1,8) = .83$, $p = .29$ Min F' is n.s.	T (276 ms) > C (165 ms) $F_1(1,21) = 4.53$, $p < .045$ $F_2(1,8) = 1.75$, $p = .22$ Min F' is n.s.
Incongruent display Near competitor	T (98 ms), C (96 ms) $F_1(1,21) = .01$, $p = .91$ $F_2(1,8) = .00$, $p = .97$ Min F' is n.s.	T (188 ms), C (199 ms) $F_1(1,21) = .47$, $p = .50$ $F_2(1,8) = .03$, $p = .86$ Min F' is n.s.	T (232 ms) > C (208 ms) $F_1(1,21) = .01$, $p < .91$ $F_2(1,8) = .91$, $p = .37$ Min F' is n.s.	T (260 ms) > C (192 ms) $F_1(1,21) = 1.47$, $p = .24$ $F_2(1,8) = 1.42$, $p = .27$ Min F' is n.s.
Incongruent display Far competitor	T (69 ms), C (128 ms) $F_1(1,21) = 2.4$, $p = .14$ $F_2(1,8) = .1.62$, $p = .24$ Min F' is n.s.	T (212 ms), C (187 ms) $F_1(1,21) = .09$, $p = .76$ $F_2(1,8) = .28$, $p = .61$ Min F' is n.s.	T (263 ms) > C (182 ms) $F_1(1,21) = 3.52$, $p < .08$ $F_2(1,8) = 2.81$, $p = .13$ Min F' is n.s.	T (303 ms) > C (166 ms) $F_1(1,21) = 8.93$, $p = .007$ $F_2(1,8) = 3.31$, $p = .11$ Min $F'(1,14) = 2.41$, $p < .15$

ANOVAs of Object × Display × Near/Far competitor for each of four 500-ms windows, with means for gaze at T and C, and T versus C contrasts in each condition of interest; significant or marginal effects highlighted.

of the matchers', but rotated to face the matcher; what was to the director's right on her display was to the matcher's right on his.

We also addressed two other concerns with the design of Experiment 2. First, even though the Congruent display/Far competitor condition showed an immediate advantage for the target shape versus the competitor in the first 500 ms after the color word, this advantage then disappeared; a target versus competitor advantage showed up again 1000 ms later in the Near competitor condition. It is possible that we recorded two different kinds of eye gaze behavior on the part of the matcher: one that reflects an immediate and automatic use of the speaker's eye gaze cue that can then aid in resolution, and another that reflects that ambiguity has been truly resolved. We were successful in eliciting long gazes at the target from the director with our materials, since they needed to count the number of dots on the shapes in preparation for speaking. However, counting was perceptually difficult (the dots were arranged randomly) and tended to take place in the middle of the referring expression, after or as the color and shape were produced. Therefore, it often caused a pause or slowdown in the instruction. Sometimes matchers, especially when the competitor was far away, looked back and forth between the target and competitor while waiting for disambiguating information during this pause. Therefore, in order to make shapes easier to distinguish, both for the directors during production planning and for matchers during reference resolution, we removed the dots and replaced them with letter pairs. Instead of asking the matcher to *move the blue circle with five dots*, directors asked the matcher to move *the blue circle with a "bf" on it*.

Second, it is interesting that the spatially Non-congruent displays did not disrupt or delay reference resolution as much as we had predicted with Far competitors. There are several possible reasons for this. First, matchers were shown the director's side of the board before each block of displays; it is possible that some of them noticed the details of how the circular arrangement related to their own horizontal arrangement and were able to re-map the director's eye gaze onto the corresponding locations in their own display. This is possible but not likely, however, given that the display was shown briefly, and the matcher would have had to distinguish looks that were near each other and oriented vertically and on the diagonal rather than horizontally. What is more likely is that when the target was in one of the central two locations in the matcher's display and the competitor was in one of the non-central locations, as they were on a third of the trials, the director's eye gaze to any of the locations in their centrally located display (coincidentally) facilitated reference resolution. That is, while the circular displays were spatially Non-congruent, they did not cause directors' eye gaze to be unavailable or wholly uninformative. Therefore, in order to remove the possibility that centrally oriented eye gaze was clearly disambiguating for spatially Non-congruent displays, we relocated the critical targets and competitors to non-central positions in the display.

We expected to replicate many of the results of Experiment 1. We predicted that reference resolution would be speeded by eye gaze when the displays were spatially congruent in a mirror image, especially when the competitor was far away from the target. We also predicted that eye gaze can serve as an automatic orienting cue, such that the very first looks that matchers made to their display would be to the same side that the director was fixating. Finally, if eye gaze serves as a cue only in this automatic manner, we expected that reference resolution would be delayed past the linguistic point of disambiguation when the directors' and matchers' displays were reversed. However, if eye gaze can also be used more flexibly, matchers should be able to take that into account and still show early reference resolution with reversed displays.

Experiment 2

Method and design

The task and physical setup were the same as in Experiment 1. Directors' and matchers' displays held identical copies of the same objects, located either in a mirrored, spatially congruent arrangement as before (Mirror displays) or else in a reversed, spatially Non-congruent arrangement, so that this time what was to the director's right was to the matcher's right (Reverse displays). Display types were blocked; however, because our results did not change dramatically with the different display orders in the previous experiment, all pairs experienced Mirror displays first and then Reverse displays.

Materials and procedure

To reduce looking back and forth between target and competitor shapes (to make them easier to distinguish), we modified the objects by printing two letters on them (e.g., a blue circle with the letters "*bx*"). Therefore, on critical trials, in addition to a target shape (e.g., blue circle with *bx*), there was a same-color competitor shape (e.g., yellow circle with *fj*) either next to it (Near competitor) or at least three spaces away (Far competitor), or the other shapes were a different color (No competitor). Target and competitor locations were again counterbalanced, but were restricted to the outer two shelf positions on the left and the right to prevent the matcher from using the director's gaze to distinguish targets and near competitors in the central locations. Target locations in the filler trials all had central shelf positions in order to counterbalance the experimental trials. The number of Tangram filler trials was reduced to eight in order to shorten the experiment.

Participants

Twenty-four Stony Brook University undergraduates (13 men, 11 women) volunteered for research credit or were paid for their participation. Participants were paired on the basis of availability; three of the pairs knew each other. All were native speakers of English and were naive to the purpose of the experiment.

None of the participants became aware of the experimental manipulation during the experiment, although upon debriefing, 21 of them (12 matchers, 9 directors) reported noticing that where the other person was looking could help them figure out which shape was being referred to with the mirror image display, and that the reverse display was harder.

Coding

In Experiment 2, we started recording each trial when the experimenter placed the schematic instruction card in front of the director; this allowed us to capture the director's early visual orientation, once she had the goal of the trial in mind but before she started speaking. We coded directors' and matchers' looks to the display objects and partner as in Experiment 1, including their first looks so these could be analyzed separately. We calculated several key points in time to serve as reference points for our analysis of object fixations: (1) the visual point of disambiguation, or VPOD, which was the director's first look toward the target object once the trial began, as long as she did not look away before beginning to speak; (2) the onset of the color word within the referring expression; and (3) the linguistic point of disambiguation, or LPOD, which was the point at which the referring expression was linguistically unambiguous. In trials without a competitor, the LPOD occurred *at* the onset of the color word. In trials with competitors, the onset of the color word was the point at which the director's referring expression was ambiguous only between the target and competitor, as these two objects were the same color; this ambiguity was not resolved by spoken information until, on average, 1191 ($SD = 426$) ms later, at the LPOD. Note that this LPOD was earlier than the LPOD in Experiment 1, since competitors and targets were easier for directors to distinguish in Experiment 2 (they distinguished same-color objects by reading letters rather than by counting dots). We discarded the 13 cases in which directors produced the wrong information or repaired utterances in such a way that the time course of interpretation was not comparable or could not be analyzed. Two additional trials were lost due to a bad track and missing sound, adding up to a loss of less than 1% of the data.

Data analysis

Analysis was similar to Experiment 1. To consider the early impact of speakers' gaze upon addressees' visual orienting well before the onset of the color word, we computed matchers' looks to the side of the display board containing the target object vs. the other side of the display, over three 500-ms windows starting at 500 ms after the visual point of disambiguation (VPOD). Fig. 3 shows proportions of matchers' looks to the target and other side of the display zeroed and aligned at the VPOD for all conditions. For these side looks, we conducted an omnibus ANOVA (Competitor × Display × Window) and Competitor × Display ANOVAs for the three individual windows (501–1000, 1001–1500, and 1501–2000 ms).

To test differences among gazing patterns, we computed the time matchers spent looking at the various objects over three windows, starting at the onset of the color word: 0–500, 501–1000, and 1001–1500 ms. The third window overlapped with the mean linguistic point of disambiguation for the competitor conditions; as before, any looks in this interval still reflect early processing, as it takes about 200 ms for a saccade to be launched after being programmed.[2] Fig. 4 shows proportions of matchers' looks zeroed and aligned at the onset of the color word for all conditions. We computed a Competitor × Display × Window ANOVA of matcher looks to the target, similar to Experiment 1; the relevant findings from these results are reported in the text. To examine competition between same-color objects, we did Object × Competitor × Display ANOVAS for each of the three 500-ms windows, as well as contrasts for each (Display/Competitor) condition (also similar to Experiment 1). Significance tests from these competitor analyses are reported in Table 3.

[2] Once again, there were no reliable differences in mean LPODs across the competitor conditions in Experiment 2 (ANOVAS by-subjects and by-items for Near vs. Far competitor × Display showed no differences). As in Experiment 1, we computed correlations of LPODs with target looks for each of the four competitor conditions, in each of the three windows (the 12 contrasts on the lower half of Table 3). The only reliable contrast in Table 3 for which LPOD was correlated with target gaze was for the Mirror display/Near competitor condition in Window 3 ($r = -.351$, $p = .045$, $N = 33$).

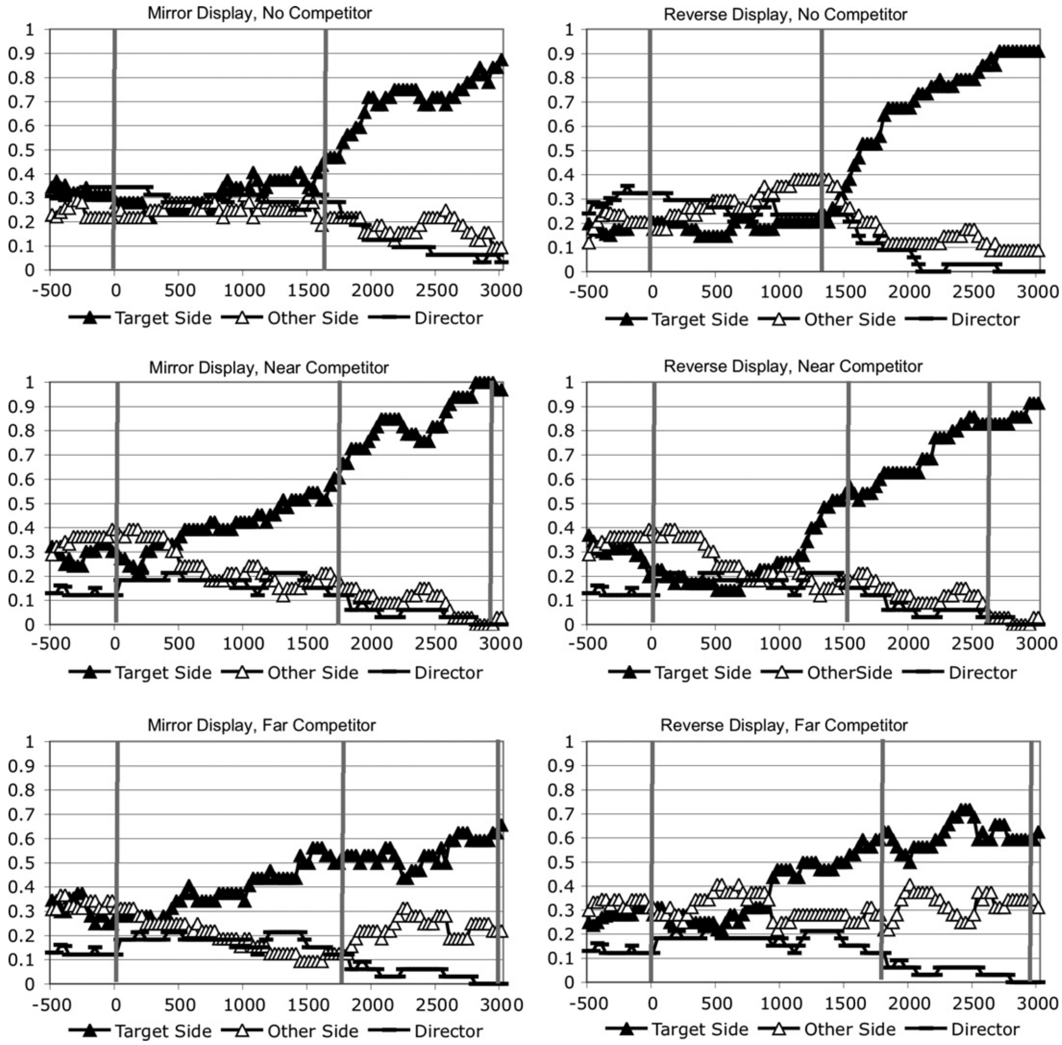

Fig. 3. Proportion of matchers' looks to the sides of the display (the target side and the other side) and to the director over time in Experiment 2 for the No competitor (top panels), Near competitor (middle panels), and Far competitor (bottom panels) conditions, when the displays were Mirror (left panels) and Reverse (right panels). The zero point is the visual point of disambiguation and is marked with a vertical line. The mean onset of the color word is marked with the second vertical line; this is also the linguistic point of disambiguation (LPOD) for the No competitor conditions. The mean LPOD is marked with the third vertical line in the Near and Far competitor conditions.

Results

The information available in a speaker's eye gaze

First, we established the extent to which directors' looks were potentially informative to matchers. After seeing the schematic card, directors' first looks (VPODs) were toward the target side of the display 97% of the time, and they gazed at the target object as they began to say the noun phrase 83.4% of the time, consistent with controlled studies of object naming in which speakers typically fixate an object while accessing and articulating a label for it (e.g., Griffin, 2001; Griffin & Bock, 2000; Meyer et al., 1998; van der Meulen et al., 2001). This underscores that this cue is potentially informative to matchers in disambiguating ambiguous referents. Directors' looks were uninformative 14.1% of the time (when they gazed down at their instructions while initiating the referring expression) and misleading only 2.5% of the time (when they gazed at an object on the other side of the display while initiating the referring expression). Directors were marginally more likely to be looking at the target object while naming it when there was a competitor present in the display than when there was No competitor, $F_1(1,11) = 4.29$, $p = .06$; $F_2(1,8) = 12.93$, $p = .007$. Min $F'(1,17) = 3.22$, $p < .10$.

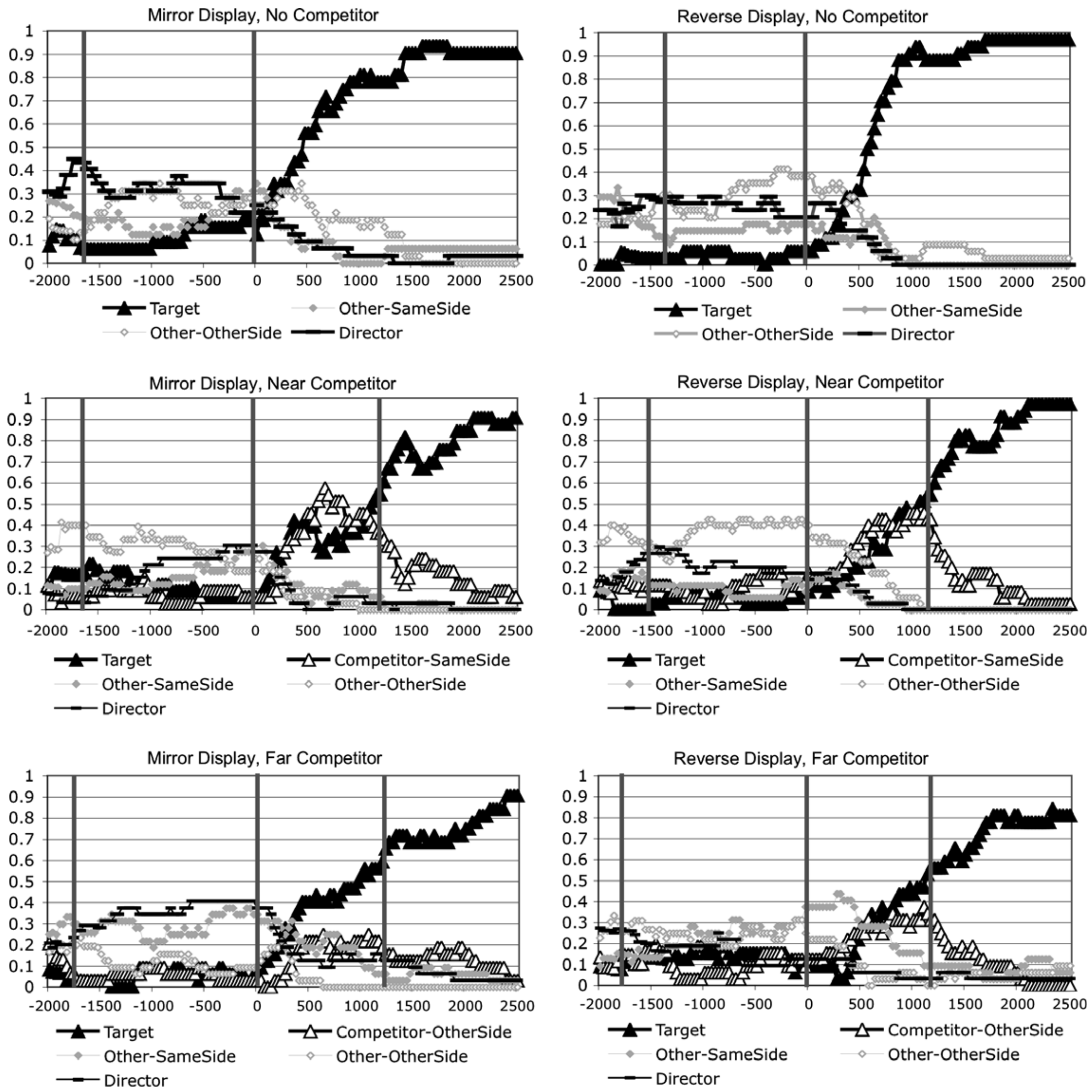

Fig. 4. Proportion of matchers' looks to the display objects and to the director over time in Experiment 2 for the No competitor (top panels), Near competitor (middle panels), and Far competitor (bottom panels) conditions, when the displays were Mirror (left panels) and Reverse (right panels). Objects other than the target are labeled as being on the same side of the display as the target, or the other side. The zero point is the onset of the color word and is marked with a vertical line; this is also the linguistic point of disambiguation (LPOD) for the No competitor conditions. The mean visual point of disambiguation is marked with the first vertical line; the mean LPOD is marked with the third vertical line in the Near and Far competitor conditions.

The use of a speaker's eye gaze as an early orienting cue

Next we considered matchers' use of this cue. With spatially coincident (Mirror) displays, the location of the target from the matcher's perspective was on the same side as where the director was looking; for Reverse displays, it was on the opposite side. With Mirror displays, matchers looked first at the target side rather than the other side 62.9% of the time, more often than chance, $t_1(11) = 3.01$, $p = .012$; $t_2(8) = 2.10$, $p = .069$. (Matchers' first looks were after the VPOD and were often before the directors began speaking). This suggests that early on in the trial, matchers and directors achieved a shared visual focus on the target side of the display when they had congruent visual perspectives. With Reverse displays, matchers looked first at the target side 53.9% of the time, no different from chance, $t_1(11) = .76$, n.s.; $t_2(8) = .84$, n.s.

Matchers did not need to look directly at directors in order to detect which side they were gazing at, and in fact, they did so rarely (directors often turned their heads in the direction they were looking, which could be seen peripherally). When matchers looked at

Table 3
Matcher gaze at target (T) versus same-color competitor (C) after onset of color term in Experiment 2

	Window 1: 0–500 (ms)	Window 2: 501–1000 (ms)	Window 3: 1001–1500 (ms)
Factors			
Object gaze (Target vs. Competitor)	$F_1(1,11) = .10$, $p = .76$ $F_2(1,8) = .19$, $p = .67$. Min F' is n.s.	$F_1(1,11) = .03$, $p = .86$ $F_2(1,8) = .09$, $p = .77$ Min F' is n.s.	$F_1(1,11) = 24.1$, $p < .001$ $F_2(1,8) = 37.47$, $p < .001$ Min $F'(1,19)$=14.67, p= .001
Object × Display	$F_1(1,11) = 6.31$, $p = .03$ $F_2(1,8) = 3.54$, $p = .10$ Min $F'(1,16) = 2.27$, $p = .15$	$F_1(1,11) = .39$, $p = .55$ $F_2(1,8) = .09$, $p = .77$ Min F' is n.s.	$F_1(1,11) = .89$, $p = .37$ $F_2(1,8) = .43$, $p = .53$ Min F' is n.s.
Object × (Near/Far) competitor	$F_1(1,11) = 27$, $p = .61$ $F_2(1,8) = .00$, $p = .96$ Min F' is n.s.	$F_1(1,11) = 2.60$, $p = .14$ $F_2(1,8) = 4.66$, $p = .06$ Min $F'(1,19) = 1.67$	$F_1(1,11) = .05$, $p = .82$ $F_2(1,8) = .19$, $p = .66$ Min F' is n.s.
Interaction: Object × Display × (Near/Far) Competitor	$F_1(1,11) = 1.00$, $p = .34$ $F_2(1,8) = .49$, $p = .51$ Min F' is n.s.	$F_1(1,11) = .92$, $p = .36$ $F_2(1,8) = 2.39$, p = .16 Min F' is n.s.	$F_1(1,11) = 3.38$, $p = .09$ $F_2(1,8) = .74$, $p = .41$ Min F' is n.s.
Conditions			
Mirror display Near competitor	T (123 ms), C (99 ms) $F_1(1,11) = .20$, $p = .66$. $F_2(1,8) = .55$, p = .48 Min F' is n.s.	C (240 ms), T (172 ms) $F_1(1,11) = 1.11$, $p = .32$ $F_2(1,8) = 1.34$, $p = .28$ Min F' is n.s.	T (304 ms) > C (158 ms) $F_1(1,11) = 5.54$, $p = .04$ $F_2(1,8) = 9.66$, $p = .01$ Min $F'(1,19) = 3.52$, $p < .08$
Mirror display Far competitor	T (117 ms), C (39 ms) $F_1(1,11) = 3.35$, $p < .10$ $F_2(1,8) = 2.26$, $p = .17$ Min F' is n.s.	T (219 ms), C (102 ms) $F_1(1,11) = 4.58$, $p = .056$ $F_2(1,8) = 4.23$, $p = .07$ Min $F'(1,18)$=2.20, $p < .16$	T (317 ms) > C (87 ms) $F_1(1,11) = 25.27$, $p < .001$ $F_2(1,8) = 9.86$, $p = .01$ Min $F'(1,14) = 7.09$, $p < .02$
Reverse display Near competitor	C (97 ms), T (71 ms) $F_1(1,11) = 1.15$, $p = .31$ $F_2(1,8) = .38$, $p = .56$ Min F' is n.s.	C (200 ms), (174 ms) $F_1(1,11) = .36$, $p = .56$ $F_2(1,8) = .11$, $p = .75$ Min F' is n.s.	T (318 ms) > C (154 ms) $F_1(1,11) = 9.05$, $p = .01$ $F_2(1,8) = 6.92$, $p = .03$ Min $F'(1,17) = 3.92$, $p = .06$
Reverse display Far competitor	C (84 ms), T (47 ms) $F_1(1,11) = .71$, $p = .42$ $F_2(1,8) = 1.21$, $p = .30$ Min F' is n.s.	T (182 ms), C (146 ms) $F_1(1,11) = .01$, $p = .95$ $F_2(1,8) = .26$, $p = .62$ Min F' is n.s.	T (274 ms) > C (129 ms) $F_1(1,11) = 1.91$, $p = .19$ $F_2(1,8) = 6.51$, $p = .03$ Min F' is n.s.

ANOVAs of Object × Display × Near/Far competitor for each of three 500-ms windows, with means for gaze at T and C, and T versus C contrasts in each condition of interest; significant or marginal effects are highlighted.

directors' faces, they did so early in the trial rather than later (linear trend), $F_1(1,11) = 13.76$, $p = .003$; $F_2(1,8) = 62.62$, $p < .001$, Min $F'(1,15) = 11.28$; $p < .01$. Although matchers knew when their displays were reversed (this was blocked), there was no evidence from their fixations that they *avoided* looking at directors when their displays were reversed; $F_1(1,11) = 1.27$, n.s.; $F_2(1,8) = .98$, n.s.; Min F' is n.s.

Fig. 3 shows matchers' general orienting to target vs. non-target sides of the display, zeroed at the point where directors began looking at the target object before beginning to speak (VPOD). With Mirror displays, speakers' direction of gaze was a compelling attractor for matchers' visual attention, especially when there was a same-color competitor: matchers' looks to the target side began rising in the second 500-ms window after the VPOD (2 lower left panels). This rise began about 500 ms later with Reverse displays (2 lower right panels), quantified by a significant Display × Window interaction, $F_1(1,11) = 10.25$, $p < .01$; $F_2(1,8) = 4.18$, $p = .075$, Min $F'(1,14) = 2.97$, $p < .11$. Note that the rise in looks to the target side for both displays occurred well before the onset of the color word. Mirror displays led to more looks to the target side than did Reverse displays in the window from 501 to 1000 ms after the VPOD, $F_1(1,11) = 10.76$, $p < .01$; $F_2(1,8) = 7.37$, $p < .03$, Min $F'(1,17) = 4.37$, $p = .05$; the congruence of the display did not matter in matchers' steadily rising looks to the target side in the two windows after that. We interpret this as evidence that, on average, matchers began to orient in the same direction as directors 501–1000 ms after the VPOD, and then, from 1001 to 1500 ms, began to adjust to their knowledge that displays were reversed.

The time course of using a speaker's eye gaze in disambiguation

Experiment 2 provided even stronger evidence than Experiment 1 that matchers could use directors' eye gaze to disambiguate referring expressions. Fig. 4 shows the proportions of time matchers spent gazing to the target, competitor, and director over four and a half seconds, zeroed and aligned at the onset of the color word for all conditions. The rise in looks to the target began at the onset of the color word in all conditions. The top two panels show that with No competitor, mapping the referring expression to the target happened very quickly. For Mirror displays with No competitor (top left panel), there was an immediate rise in looks to the target beginning at the LPOD, which corresponded in this condition to the onset of the color word. For Reverse displays with No competitor (top right panel), the graph shows a slight delay in this rise; the slope of this rise differs across the three time windows (linear trend of the interaction of Display × Window), $F_1(1,11) = 8.11$, $p < .02$; $F_2(1,8) = 9.77$, $p = .01$; Min $F'(1,19) = 4.43$, $p < .05$.

In the competitor conditions, looks to the target did not reliably exceed looks to the competitor until the third window (individual contrasts for each of the three windows, Table 3). The impact of the director's gaze appeared to be strongest in Mirror displays when competitors were Far (on different sides) from targets (bottom left panel of Fig. 4): matchers looked marginally more at targets (219 ms) than at competitors (102 ms) during the second 500-ms window, and reliably more at targets than at competitors (317 vs. 87 ms) during the third 500-ms window. This was not the case when competitors and targets were adjacent; in fact, as Fig. 4 illustrates for Mirror displays with Near competitors (middle left panel), competitors attracted on average 68 ms *more* gaze than did targets during the second 500-ms window, 240–172 ms. In a contrast of looks to competitors during this second window, there were more looks to Near than Far competitors, $F_1(1,11) = 13.36$, $p = .004$; $F_2(1,8) = 4.97$, $p = .056$; Min $F'(1,14) = 3.62$, $p < .10$.

Taken together, these results suggest that directors' gaze was a disambiguating cue for matchers, especially when it was highly diagnostic, that is, when displays were spatially correspondent and when competitors were far away from targets.

The flexible use of speakers' eye gaze

Finally, matchers were able to use directors' eye gaze to find the target object before the linguistic point of disambiguation, *even when this required re-mapping gaze cues to a reversed display*. With no competitor (two top panels in Fig. 4) there was a slightly late but steep rise in looks to targets with Reverse displays (right panel) compared to Mirror displays (left panel). With same-color competitors (bottom four panels of Fig. 4), there was substantial competition during the first 1000 ms for all conditions except Far competitors in Mirror displays (bottom left panel); a marginal difference in target versus competitor looks emerged for this condition by Window 2 (Table 3). In all four bottom panels in Fig. 4, looks to same-color competitors dropped sharply in Window 3, which was before disambiguating linguistic information could have begun driving the target advantage, since eye movements generated by the LPOD would begin to show up 200 ms later, towards the end of this window.

We considered matchers' looks to target objects in the No competitor conditions as well as the Near and Far competitor conditions for Mirror and Reverse displays by computing a (3) Competitor × (2) Display ANOVA for each 500 ms window. The congruence of displays mattered only during the first 500 ms. Targets attracted more looks in Mirror displays than in Reverse displays during the first window, $F_1(1,11) = 8.76$, $p = .01$; $F_2(1,8) = 5.70$, $p = .04$, Min $F'(1,17) = 3.45$, $p < .10$, but not in the second, $F_1(1,11) = .41$, n.s.; $F_2(1,8) = .42$, n.s., or the third, $F_1(1,11) = .02$, n.s.; $F_2(1,8) = .03$, n.s. This suggests that matchers began to flexibly re-map directors' direction of gaze onto reversed displays by the second 500 ms window after the color word. Recall Table 3, in which there was little competition from the same-color competitor by the third window, regardless of whether the display was Mirror or Reverse. We consider this to be a *flexible* use of gaze cues, since some studies have claimed that gaze—even a disembodied pair of cartoon eyes—serves as an automatic orienting cue, even when it is uninformative or misleading (Friesen & Kingstone, 1998). However, this re-mapping of eye gaze was not immediate (and it did not appear to work like a toggle, even though display mappings were blocked); the three Reverse displays (right panels of Fig. 4) consistently show a 150–250 ms delay in the rise of looks to the target compared to the three Mirror displays (left panels of Fig. 4).

Finally, similar to our results in Experiment 1, directors' gaze seems to have provided useful cues even when competitors were not very spatially distinct from targets. Although we had intended for the Near (adjacent) competitor conditions to be entirely uninformative, matchers appeared to be better at using directors' gaze cues than we expected, as shown with the target versus competitor advantage by the third window for both displays (Table 3). It is possible that matchers picked up on relative differences in directors' alternating looks to two same-color shapes in the non-central regions of the display (e.g., noticing a look to the outermost position versus one that was less extreme).

In addition, a constraint-based account may explain why there was not as much competition in the Near competitor conditions as expected. Consider the cues available to the matcher with Near competitors: the director's eye gaze provided evidence in favor of objects on one side of the display, and the color word provided evidence in favor of two objects on the same side of the display (whether Mirror or Reverse). Now consider the cues available to the matcher with Far competitors: the director's eye gaze provided evidence in favor of objects on one side of the display, but the color word actually provided evidence in favor of one object on the same side of the display and one object on the other side of the display. In a sense, the two sides of the board competed for matchers' attention more in the Far competitor conditions than in the Near competitor conditions. This might explain the unexpected speed of reference resolution in the Near competitor conditions, especially when displays were reversed and matchers had no shapes that matched the color term on the side of the board the director was looking at. This account is consistent with the target versus competitor advantage in Window 3 for the Reverse/Near competitor condition that is not present for the Reverse/Far competitor condition (Table 3).

General discussion

The results of Experiments 2 replicated our findings from Experiment 1: eye gaze produced by a speaker can be used by an addressee to resolve temporary ambiguity, and it is a cue that can be used rapidly in face-to-face conversation. Directors' looks to the target in spatially Congruent displays had an immediate orienting effect on matchers, whose initial looks were usually to the same side. In addition, when there was a same-color competitor far away from the target, matchers were able to use the directors' visual orientation to distinguish the target from the competitor earlier than when the competitor was nearby, and significantly earlier than the linguistic point of disambiguation.

Both experiments suggest that eye gaze is a conversationally based source of information that can be used to rapidly constrain reference resolution. Moreover, the methods of these experiments depart radically from most of the work investigating the time course of spoken language comprehension, in that they balance naturalness with control. Effects of eye gaze were demonstrated using naive participants with unscripted utterances in order to obtain spontaneous looking behaviors as well as relatively naturalistic task-based conversational interaction.

Of course, eye gaze cues may have both benefits and costs; to the extent that where a speaker is looking constrains the domain of interpretation for an addressee, resolution may be speeded, but to the extent that a speaker looks back and forth between similar objects, or perceptual difficulty or confusion is reflected in a speaker's gaze, these cues may themselves be ambiguous. Our results suggest that the level of ambiguity in the directors' eye gaze was low enough in this task that the benefits dominated, but scaling up our results to more naturalistic situations will mean taking into account these complexities. In addition, our matchers were actually somewhat slower to reach for objects than in other eyetracking experiments, and this might reflect a cost of monitoring a partner's eye gaze, even when the director's gaze was not ambiguous. It is still unclear how detailed the non-verbal cues of a speakers' orientation need to be in order to be useful to addressees, since in our experiments these cues consisted of both eye gaze and head/body orientation. The congruity between these cues is likely to have some effect, given the findings of congruity effects between pointing gestures and head/gaze orientation (e.g., Langton & Bruce, 2000); we do not know the degree to which our effects were driven by eye gaze alone.

The results of both experiments are consistent with the idea that eye gaze has an automatic, reflexive orienting effect on attention (e.g., Friesen & Kingstone, 1998; Friesen et al., 2004; Langton et al., 2000); matchers' attention was immediately oriented to the spatial location in the directors' line of sight. When the matchers' and directors' displays were congruent—as is usually the case in the real world, when people are engaged in face-to-face conversation about visually co-present objects—this orienting effect contributed a constraint on the domain of interpretation that speeded reference resolution. However, while eye gaze may have had its effects in part due to an initially automatic mechanism, it may not be a reflexive or hard-wired one, as we also found evidence that it is a flexible cue that can rapidly be re-mapped. Matchers were able to use their knowledge of the spatial non-congruency of the displays to re-map the eye gaze cues when the displays were reversed; in these conditions they still showed speeded reference resolution, although it was somewhat delayed in comparison to when the displays were spatially congruent. This suggests that eye gaze is a source of information whose use can be modified according to the communicative context.

These results are in line with much other work demonstrating the effects of pragmatic and other conversational contexts on the earliest moments of language processing (e.g., Brown-Schmidt et al., 2005; Chambers, Tanenhaus, & Magnuson, 2004; Hanna & Tanenhaus, 2004; Metzing & Brennan, 2004), and contribute to the increasing evidence for constraint-based models (e.g., MacDonald, 1994; Tanenhaus & Trueswell, 1995). From a constraint-based perspective, information that might reveal the attention and intentions of an interlocutor, such as where she is looking, should be integrated simultaneously and continuously with other lexical, structural, and discourse-based constraints to provide probabilistic evidence for alternative interpretations. The degree to which an addressee will monitor the actions, goals, and likely knowledge of an interlocutor should vary with the salience and accessibility of that information and its importance for the goals of the addressee. In the current experiment, the eye gaze of the director was a simple, perceptually available cue that had a relatively straightforward link to the goals of the task. This kind of non-verbal cue can be an important source of information about interlocutors' perspectives. Our demonstration that this information plays an immediate role in constraining linguistic interpretation argues against models that claim that initial comprehension is egocentric and does not take into account others' perspectives (e.g., Keysar et al., 2000). One crucial prediction that follows from our view of eye gaze as an immediately available, yet flexible constraint, is that with enough practice, participants should be able to learn a re-mapping well enough to no longer show the automatic orienting effect. This would indicate that while the low-level processing of eye gaze is relatively automatic, it is still a constraint whose influence can be weighted differently given the nature of the communicative context, much like other conversationally based sources of information.

In closing, we note again that while eye gaze is often an instrumental behavior on the part of the gazer, it is also a communicative one. It can say a great deal, whether the gazer intends to or not. Consider this famous example, described by Case Western University's News Center, 2004: "During the 1992 debate, President George H.W. Bush was feeling the heat from Arkansas Governor Bill Clinton and Texas businessman Ross Perot. Bush wasn't fond of debating in the first place, and made this clear as he glanced at his watch, painstakingly counting the minutes left in the debate. To the President's dismay, this moment was caught on-camera and reinforced Democratic vice presidential nominee Al Gore's mantra: "It's time for them to go." Voters agreed with Gore, showing that body language sometimes says it all."

Acknowledgments

This material is based upon work supported by the National Institutes of Health under National Research Service Award 1F32MH1263901A1, and by the National Science Foundation under Grants No. ISS-0527585, ITR-0082602, and ITR-0325188. We are grateful to Alex Christodoulou, James Sedoruk, and Randy Stein for their assistance with data collection and coding. We also thank Roger van Gompel, Ellen Bard, and an anonymous reviewer for their insightful comments and suggestions. Some of the research in this paper was described at the Fifteenth Annual CUNY Conference on Human Sentence Processing (March, 2002) in Philadelphia, PA, the Seventeenth Annual CUNY Conference on Human Sentence Processing (March, 2004) in College Park, MD, the 43rd Annual Meeting of the Psychonomic Society (November, 2002) in Kansas City, MO, the 44th Annual Meeting of the Psychonomic Society (November, 2003) in Vancouver, Canada, and the 46th Annual Meeting of the Psychonomic Society (November, 2005) in Toronto, Canada.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bloom, P. (2002). Mindreading, communication and the learning of names for things. *Mind and Language, 17*, 37–54.

Boyle, E., Anderson, A., & Newlands, A. (1994). The effects of visibility on dialogue and performance in a cooperative problem solving task. *Language and Speech, 37*, 1–20.

Brennan, S.E. (1990). Speaking and providing evidence for mutual understanding. *Unpublished doctoral dissertation*. Stanford University, Stanford, CA.

Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95–129). Cambridge, MA: The MIT Press.

Brennan, S. E., Chen, X., Dickinson, C., Neider, M., & Zelinsky, G. (in press). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*.

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Developmental Psychology, 38*, 958–966.

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (2005). Real-time reference resolution by naive participants during a task-based unscripted conversation. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use*. Cambridge, MA: The MIT Press.

Brown-Schmidt, S., Gunlogson, C., Watson, D., Faden, C., & Tanenhaus, M.K. (2006, March). Perspective matters during on-line production and interpretation of questions and replies in unscripted conversation. Paper presented at the 19th Annual CUNY Conference on Human Sentence Processing. New York, NY.

Butterworth, G., & Grover, L. (1990). Joint visual attention, manual pointing, and preverbal communication in human infancy. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control* (pp. 605–624). Hillsdale, NJ: Erlbaum.

Case Western University's News Center (September 20, 2004). <http://www.case.edu/news/2004/9-04/tv_debates.htm/>.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 687–696.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*, 62–81.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5*, 490–495.

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 319–329.

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *American Journal of Psychology, 76*, 386–394.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*, B1–B14.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*, 274–279.

Hanna, J.E. (2001). The effects of linguistic form, common ground, and perspective on domains of referential interpretation. Unpublished doctoral dissertation, University of Rochester, Rochester, NY.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: evidence from eye movements. *Cognitive Science, 28*, 105–115.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*, 43–61.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91–117.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic disambiguation. *Cognitive Science, 20*, 137–194.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26*, 11–63.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: the role of mutual knowledge in conversation. *Psychological Science, 11*, 32–37.

Kraljic, T., & Brennan, S. E. (2005). Using prosody and optional words to disambiguate utterances: For the speaker or for the addressee? *Cognitive Psychology, 50*, 194–231.

Langton, S. R. H., & Bruce, V. (2000). You must see the point: automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 747–757.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4*, 50–59.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*, 157–201.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects in the comprehension of referring expressions. *Journal of Memory and Language, 49*, 201–213.

Metzing, C., & Brennan, S. E. (2004). When conceptual pacts are broken: partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language, 49*, 201–213.

Meyer, A., Sleiderink, A., & Levelt, W. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition, 66*, B25–B33.

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

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition, 85*, 53–78.

Pusch, M., & Loomis, J. M. (2001). Judging another person's facing direction using peripheral vision [Abstract]. *Journal of Vision, 1*, 288.

Repacholi, B. M. (1998). Infants' use of attentional cues to identify the referent of another person's emotional expression. *Developmental Psychology, 34*, 1017–1025.

Richardson, D. C., & Dale, R. (2005). Looking to understand: the coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*, 1045–1060.

Soproni, K., Miklósi, Á., Topál, J., & Csányi, V. (2001). Comprehension of human communicative signs in pet dogs (Canis familiaris). *Journal of Comparative Psychology, 115*, 122–126.

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology, 37*, 276–309.

Stein, R., & Brennan, S.E. (2004). Another person's eye gaze as a cue in solving programming problems. *Proceedings, ICMI 2004, Sixth International Conference on Multimodal Interfaces* (pp. 9–15), Penn State University, State College, PA.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 632–634.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Speech, Language, and Communication* (pp. 217–262). San Diego, CA: Academic Press.

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: influences of context-based expectations. *Journal of Memory and Language, 27*, 597–632.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.

Trueswell, J. C., & Tanenhaus, M. K. (Eds.). (2005). *Approaches to studying world-situated language use*. Cambridge, MA: The MIT Press.

Tschudin, A., Call, J., Dunbar, R. I. M., Harris, G., & van der Elst, C. (2001). Comprehension of signs by dolphins (Tursiops truncatus). *Journal of Comparative Psychology, 115*, 100–105.

van der Meulen, F., Meyer, A., & Levelt, W. (2001). Eye movements during the production of nouns and pronouns. *Journal of Memory and Cognition, 29*, 512–521.