Research Article

The Art of Conversation Is Coordination

Common Ground and the Coupling of Eye Movements During Dialogue

Daniel C. Richardson,[1] Rick Dale,[2] and Natasha Z. Kirkham[3]

[1]University of California, Santa Cruz; [2]Universityof Memphis; and [3]Stanford University

ABSTRACT—*When two people discuss something they can see in front of them, what is the relationship between their eye movements? We recorded the gaze of pairs of subjects engaged in live, spontaneous dialogue. Cross-recurrence analysis revealed a coupling between the eye movements of the speaker and listener. In the first study, we found their eye movements were coupled across several seconds. In the second, we found that this coupling increased if they both heard the same background information prior to their conversation. These results provide a direct quantification of joint attention during unscripted conversation and show that it is influenced by knowledge in the common ground.*

When people talk, they coordinate whose turn it is to speak (Sacks, Schegloff, & Jefferson, 1974). They also implicitly agree upon names for novel objects (Brennan & Clark, 1996; Clark & Brennan, 1991), align their spatial reference frames (Schober, 1993), and use each other's syntactic structures (Branigan, Pickering, & Cleland, 2000). Their accents become more similar (Giles, Coupland, & Coupland, 1992), they sway their bodies in synchrony (Condon & Ogston, 1971; Shockley, Santana, & Fowler, 2003), and they even scratch their noses together (Chartrand & Bargh, 1999). These acts of coordination serve many purposes, such as making sure that conversation flows easily and intelligibly (Garrod & Pickering, 2004) and that conversants are well disposed toward each other (Dijksterhuis & Bargh, 2001). In the present work, we studied another conversational art that is a form of "joint activity" (Clark, 1996): the coordination of visual attention.

Address correspondence to Daniel C. Richardson, 273 Social Sciences 2, University of California, Santa Cruz, Santa Cruz, CA 95064, e-mail: dcr@ucsc.edu.

During a conversation, a lecture, or an argument in a restaurant, people might use many tools to direct each other's attention—an index finger, waving hands, a laser pointer, slide animations, even a wielded steak knife. The most basic and universal of these attentional implements is a pointing finger (Kita, 2003). In Bangerter's (2004) study, subjects described pictures that were arrayed on an easel. Their pointing increased as the distance between them and the easel decreased, with pointing behavior replacing verbal descriptions of target locations. Bangerter concluded that pointing was used fluidly and opportunistically as a composite part of speech (Bavelas & Chovil, 2000).

It has been argued that coordinating joint attention is essential for successful communication (Brown-Schmidt, Campana, & Tanenhaus, 2004; Clark, 1996; Clark & Brennan, 1991; Schober, 1993), and may even be the basis for prelinguistic learning between infants and caregivers (Baldwin, 1995). In previous studies, researchers have focused on the actions that serve to coordinate attention, such as gestures and pointing (Bangerter, 2004; Clark & Krych, 2004). The studies we report here are distinct in two ways. First, we measured the coordination of attention itself, by tracking the eye movements of conversants discussing a shared visual display and by quantifying the temporal coupling. Second, the conversants could not see each other, and hence could not directly control each other's attention by pointing or gesturing. The hypothesis tested in our first study was that visual attention would be coordinated on the basis of verbal contact alone.

Our second hypothesis concerned the possible basis for attentional coupling. According to Clark (1996) and other investigators (for a recent review, see Lee, 2001), conversation is understandable only against a background of *common ground*. Common ground has two components. The first is the knowledge, beliefs, and assumptions shared by conversants. The second is

UNCORRECTED PROOF

 Copyright © 2007 Association for Psychological Science

PSCI 1914 | PSCI Journal Name | 1914 Manuscript No. | B | Dispatch: 12.4.07 Journal: PSCI CE: Blackwell | Author Received: No. of pages: 7 PE: TMS

US PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)

their mutual experience of the interaction. For example, a museum guide who is describing a battlefield painting to you might rely on some skeletal knowledge of the English civil war, as well as on the fact that you are looking at a particular central figure as she refers to Cromwell as "he." In our second study, we examined the effects of the first component of common ground on the second. Prior to discussing a painting in front of them, the subjects in each pair heard either different pieces of factual information (about the artist or painting) or the same information. We measured whether a greater shared understanding would result in closer visual coordination within the shared visual scene.

EYE MOVEMENTS AND MONOLOGUES

What is the support for our prediction that eye movements will be coupled during a conversation on the basis of verbal contact alone? Research has revealed the time course of language processing in the eye movements of listeners (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Cooper, 1974; Hanna, Tanenhaus, & Trueswell, 2003; Henderson & Ferreira, 2004; Kamide, Altmann, & Haywood, 2003; Richardson & Matlock, in press; Tanenhaus, Spivey Knowlton, Eberhard, & Sedivy, 1995) and the time course of language production in the eye movements of speakers (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998). In these studies, individual subjects heard or spoke a short sentence while looking at a visual scene. Eye movements were closely linked to the timing of the words. If these results are general, when a listener listens to a speaker, there should be a systematic link between their eye movements.

In a test of this hypothesis (Richardson & Dale, 2005), conversational partners looked at a scene that was the topic of the discussion. The situation was analogous to two people discussing a diagram on a chalkboard, figuring out a route on a map, or talking during a movie. In the first study, the speech and eye movements of one set of subjects were recorded as they looked at pictures of six cast members of a TV sitcom (either "Friends" or "The Simpsons"). They spoke spontaneously about their favorite episode and characters. One-minute segments of their monologues were then played back unedited to a separate set of subjects. The listeners looked at the same visual display of the cast members, and their eye movements were recorded as they listened to the segments of speech. They then answered a series of comprehension questions.

Listeners' and speakers' eye movements were coded as to which of the six cast members was fixated every 33 ms. Pairs of eye movement sequences were analyzed using recurrence analysis (Zbilut, Giuliani, & Webber, 1998; see "Design and Analysis" in the Method section for a brief explanation). *Recurrence* is expressed as a percentage and conveys the degree to which speakers' and listeners' eye positions overlap. Recurrence was calculated with varying time lags between the speaker's and the listener's eye movement sequences, and averaged across all speaker-listener pairs. This speaker-by-listener distribution was compared with a speaker-by-randomized-listener distribution, produced by first shuffling the temporal order of each listener's eye movement sequence and then calculating the average recurrence across these speaker/randomized-listener pairs. This served as a baseline of looking "at chance" at any given point in time, but had the same overall distribution of looks to each picture as in the real listeners' data.

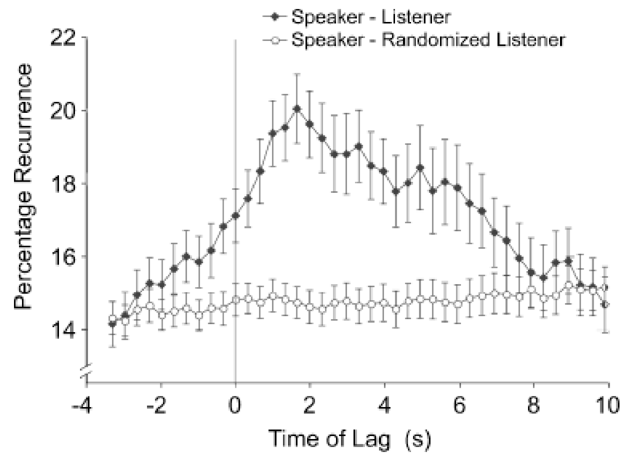

Fig. 1. Results from Richardson and Dale's (2005) study of the overlap between speakers' and listeners' eye movements. Speaker-listener recurrence at different time lags is compared with a randomized baseline.

As Figure 1 shows, from the moment a speaker looked at a picture, and for the following 6 s, a listener was more likely than chance to be looking at that same picture. The length of this time frame suggests that speakers and listeners may keep track of a subset of objects in the visual scene that are relevant moment by moment. Such behavior was suggested by the findings of Beun and Cremers (1998) and Brown-Schmidt et al. (2004). In these studies, speakers referred to "the red block" even when there were several red blocks in sight. Brown-Schmidt et al. found that listeners were rapidly able to fixate the correct block, because what speakers had said previously had implicitly identified a set of objects that included only one of the red blocks. In this way, linguistic context can circumscribe the referential domain.

The overlap between speakers' and listeners' eye movements in our previous study (Richardson & Dale, 2005) peaked at about 2,000 ms. In other words, the time when the listener was most likely to be looking at the same cast member as the speaker was 2 s after the speaker looked at that cast member. The timing of this peak corresponds to the sum of fixation latencies by speakers and listeners. Speakers fixate objects 800 to 1,000 ms (Griffin & Bock, 2000) before naming them, and listeners typically take 500 to 1,000 ms to fixate an object after the word onset (Allopenna et al., 1998). The coupling between speakers' and listeners' eye movements held throughout the speech stream, not just when speakers used the names of the cast members. Planning diverse types of speech influenced the speakers' eye movements, and a few seconds later, hearing this speech

UNCORRECTED PROOF

PSCI 1914
US PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)

influenced the listeners' eye movements. It also indicates that listeners might have been adopting the "co-presence heuristic" (Clark & Marshall, 1981), assuming they shared a visual context with the speaker.

The coupling of eye movements between speaker and listener was not merely an epiphenomenon of language processing. It played a causal role in comprehension. The degree of recurrence between speakers' and listeners' eye movement patterns reliably predicted how many of the comprehension questions the listener answered correctly. This correlation was supported by a follow-up study that experimentally manipulated the relation between speakers' and listeners' eye movements. In visual perception and problem solving (Grant & Spivey, 2003; Pomplun, Ritter, & Velichkovsky, 1996), a low-level perceptual cue can cause one person's eye movements to correspond more closely to another's and, as a consequence, affect the cognitive state of the person whose eye movements are manipulated. Similarly, we found that flashing the pictures in time with the speakers' fixations caused the listeners' eye movements to follow the speakers' more closely (Richardson & Dale, 2005). Such listeners answered comprehension questions more quickly than those who had seen a randomized sequence of flashed pictures.

DIALOGUES

Our previous work demonstrated that the close coordination of visual attention between speakers and listeners plays a causal role in comprehension of spontaneous monologues (Richardson & Dale, 2005). These findings are limited, however, because the studies missed an essential element of conversation: real-time interaction. Does the opportunity to interrupt and query a speaker when misunderstandings arise mean that the listener no longer has a need to ground the speaker's words in the visual display? Also, in a dialogue, a listener can plan and produce his or her own utterances. Perhaps the eye movement patterns during a frequent alternation of speaker and listener roles differ from the eye movement couplings of a mute, obedient listener following the words (and therefore the gaze) of a speaker.

In contrast to this line of reasoning, we hypothesized that during an interactive dialogue, eye movements continue to be coupled. If conversation is a joint activity with many levels of coordination (Clark, 1996; Garrod & Pickering, 2004), participants in a dialogue should act to coordinate their attention around the visual common ground. Moreover, we predicted that their ability to do so depends in part on the knowledge that they share. To test these hypotheses, we carried out two studies during a single session with the same pairs of subjects.

GENERAL METHOD

Subjects and Apparatus

Forty Stanford undergraduates participated in exchange for course credit. Subjects were randomly assigned to pairs. Four pairs were discarded from analyses because of problems calibrating the eyetracker to one of the subjects. In Study 2, two additional pairs were excluded because of equipment malfunction and experimenter error.

Members of a pair arrived for their session at the same time, reporting to two eye-tracking labs on different floors of the same building. In Lab 1, an ASL 504 remote eye-tracking camera (Applied Science Laboratories, Bedford, MA) was positioned at the base of a 17-in. LCD display. Subjects sat unrestrained approximately 30 in. from the screen. The display subtended a visual angle of approximately $26^{\circ} \times 19^{\circ}$. The camera detected pupil and corneal reflection position from the right eye, and the eye-tracking computer calculated point of gaze in terms of coordinates on the stimulus display. This information was passed every 33 ms to a PowerMac G4 that controlled stimulus presentation and collected looking-time data. Lab 2 used an identical setup, except that the display was a 48-in. × 36-in. back-projected screen, and subjects sat 80 in. away (this lab was designed for infants under a year old). This display subtended a slightly larger visual angle of approximately $33^{\circ} \times 25^{\circ}$.

Experimenters in the two labs operated the computers and communicated with each other using an instant-messaging application. Subjects communicated with each other through hands-free headsets that were connected via the intercom feature on a set of 2.4-GHz wireless phones. The speech of the subjects was recorded by microphones at the base of the displays.

Design and Analysis

Before the experimental session proper, the experimenters ran a 9-point calibration routine that typically took 1 or 2 min. Once the gazes of both subjects were successfully tracked, the experimenters agreed upon a precise time at which to start each study. Each computer was synchronized with an external time server, ensuring that the trials and data streams began simultaneously. In each study, the 2 subjects were presented with exactly the same visual display with the same predefined regions of interest (ROIs).

The eye-tracking computers recorded separate streams of data specifying which (if any) ROI each subject was fixating every 33 ms. Our analyses concerned the degree to which the 2 subjects were looking at the same thing at the same time. We quantified the answer to this question by generating categorical cross-recurrence plots between the speaker's and listener's time series of fixations (for a detailed explanation, see Dale & Spivey, 2006; Richardson & Dale, 2005). Such plots permit visualization and quantification of recurrent patterns of states between two time series (Shockley et al., 2003; Zbilut et al., 1998).

Points of recurrence are simply the times at which the two data streams have the same value; in our case, this meant that the 2 subjects were fixating the same ROI. For a pair of time series, we added up all the points of recurrence and divided by the total

UNCORRECTED PROOF

US PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)
PSCI 1914

number of possible points to get a percentage. One of the data streams was then lagged, so that 0 ms on one data stream was aligned with 33 ms on the other. Again, all the points of recurrence were calculated to determine the degree to which 1 subject was looking at the same thing as the other subject 33 ms later. The full recurrence analysis consisted of calculating the recurrence for all possible alignments, or lag times, of the two data series. In our previous study (Richardson & Dale, 2005), we used this technique on monologue data to find exactly what temporal lag between the listener and the speaker would produce the greatest degree of recurrence, or overlap, between the eye movement patterns. Figure 1 shows the average recurrence for 49 dyads at different lag times. This plot reveals that speakers' and listeners' eye movements were coupled at above-chance levels from when there was no lag between them up to when the listeners' data streams were 6,000 ms behind the speakers'.

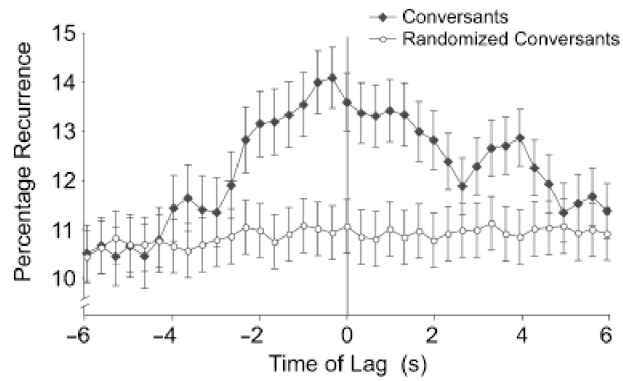

Fig. 2. Results from Study 1: average percentage of eye movement recurrence at different time lags during a dialogue. Conversants' recurrence is compared with a randomized baseline with the same distribution of eye movements.

STUDY 1

In Study 1, we explored whether eye movement couplings of conversants in an interactive dialogue might differ from the couplings we had observed for one-way monologues. The task and stimuli were identical to those of Study 1 in our previous investigation (Richardson & Dale, 2005), except that subjects took part in a live, two-way conversation. Subjects saw pictures of six cast members from "Friends" or "The Simpsons." Using the same prompts that elicited the monologues, we asked them to discuss their favorite characters or episodes from the show. An ROI was defined around each cast member's picture. The subjects were allowed to say as much as they liked, but typically, conversations lasted for 1 to 5 min.

In the original monologue study, there was a peak of recurrence when the listeners' eye movements followed the speakers' at a lag of roughly 2,000 ms. We hypothesized that in this dialogue study, there would be a similar peak, reflecting a similar process of grounding language in the visual context. We predicted that this peak would be centered around 0 ms on average, because the subjects took turns in speaking and, consequently, in leading the eye movement coordination.

Figure 2 shows the average recurrence between subjects' eye movements at different time lags, averaged over 16 dyads. As in Figure 1, the randomized baseline provides a comparison of looks with the same distribution as the subjects' eye movements, but with the temporal structure removed. Also as in Figure 1, there is a window of roughly 6 s in which subjects' eye movements are clearly coupled at above-chance levels. Unlike the monologue data, though, the dialogue data show a peak in recurrence at around 0 ms.

We examined the differences between the dialogue and monologue data by analyzing the two experiments together. A 2 (experiment: monologue vs. dialogue) × 41 (lag times) mixed-effects analysis of variance (lag as a repeated measures factor) revealed a significant main effect of experiment, $F(1, 79) = 15.6$, $p < .001$, $p_{rep} = .997$, $\eta_p^2 = .16$, and a main effect of lag, $F(40, 3160) = 8.3$, $p < .0001$, $\eta_p^2 = .09$. Most important, there was a significant interaction between the factors, $F(40, 3160) = 3.1$, $p < .0001$, $\eta_p^2 = .04$, showing that the two-way communication in the dialogue experiment changed the temporal structure of the eye movement coupling.

We ruled out the possibility that such differences were produced by changes in overall gaze activity. It was conceivable that the dialogue situation was more engaging to the subjects, so that they simply looked around the screen more in this study than in our previous study. Further investigation showed that the rate at which subjects switched their gaze from one picture to another did not differ between the monologue and dialogue conditions, $t(80) < 1$.

The main effect of experiment, and a glance at Figures 1 and 2, shows that there was a difference between the overall levels of recurrence in the two experiments. Recurrence peaks at 20% in Figure 1 and 14% in Figure 2. This difference has two possible explanations. There might be less coordination of attention during a dialogue than during a monologue. Or it could be that tracking eye movements of two people simultaneously was considerably more challenging than tracking eye movements of one person at a time, so that there was more noise and lost data in the dialogue experiment. Lost data would produce lower overall recurrence between the eye tracks at all points in time. There is good reason to think that the second explanation accounts for the differences in recurrence values between the experiments, because the conversational results and randomized baseline were reduced by comparable amounts.

The results from this first study support our hypothesis that the eye movement coupling found in monologue communication extends to dialogues. Even though subjects were able to interact with each other verbally, and make use of all the common verbal back channels in communication that signal assent, understanding, or a need for more information (Clark, 1996), subjects

UNCORRECTED PROOF

US PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)

PSCI 1914

still visually coordinated their attention as they conversed. This coupling peaked at 0 ms, meaning that subjects were most likely to be looking at the same thing as each other at the same point in time.

STUDY 2

The term common ground refers to much more than the visual context shared by conversants. It also refers to the many beliefs, opinions, and facts that they share (Clark, 1996; Lee, 2001). In the second study, we tested the hypothesis that manipulating the amount of knowledge subjects shared would affect their ability to coordinate their attention in the visual common ground.

Subjects were asked to talk about a painting by Salvador Dali, *Nature Morte Vivante* ("Still life, fast moving"). Before their conversation, they were told that they would each hear a short discussion of Dali's art and that they would hear either the same or different information. The subjects then listened to 90-s passages that discussed either the history, content, and meaning of the specific painting (e.g., "The still life objects in the original canvas have separated from the table and float in the air, and even the particles of paint have broken loose from the canvas.") or Dali's personality and theory (e.g., "The paranoiac critical method entailed the creation of a visionary reality from elements of dreams, memories, and psychological or pathological distortions. At times, Dali would stand on his head to induce hallucinations."). The subjects were allowed to talk for as long as they wanted while their gaze was recorded. The defined ROIs on Dali's painting corresponded to six of the main objects or elements. Our prediction was that pairs of subjects who had heard the same information about Dali would show greater coordination of their eye movements than those who heard different passages.

For each dyad, we quantified the amount of recurrence within a window of ±3,000 ms (see Fig. 3 for examples of recurrence plots from two pairs of subjects). In other words, we looked at the overlap between subjects' eye movements when each lagged the other by up to 3,000 ms. This window was chosen because in Studies 1 and 2 of our previous investigation (Richardson & Dale, 2005), subjects' eye movements were coupled at above-chance levels in a roughly 6-s window. Using this window, we focused on times when the eye movements were indeed coupled and looked specifically at the effects of common ground.

A one-way analysis of variance on the average recurrence for each dyad within this window yielded a significant effect of common-ground condition, $F(1, 12) = 4.9$, $p < .05$, $p_{rep} = .881$, $d = 1.15$; dyads who heard the same information had recurrence levels more than a third higher than those who heard different information (Fig. 4). We conclude that a simple manipulation changing the information subjects shared about a painting directly affected the coordination of their visual attention.

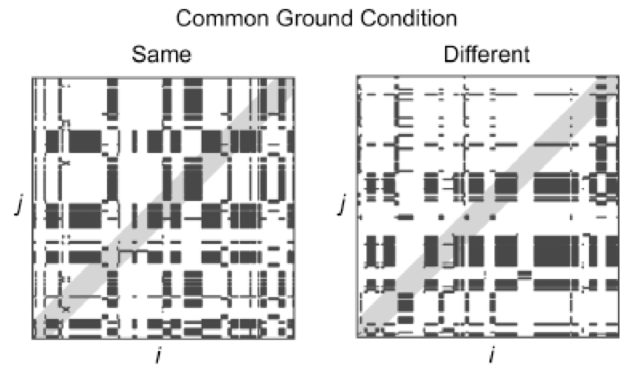

Fig. 3. Example recurrence plots from two pairs of subjects in Study 2. Eye movement recurrence at 0 ms of lag is plotted along the line $i = j$, where each axis represents the time course for 1 conversant. Parallel diagonals represent the recurrence at different degrees of lag (for more information on this application of recurrence analysis, see Richardson & Dale, 2005). The gray area represents the band of ±3,000 ms of lag that was used to calculate average recurrence between pairs in each condition (see Fig. 4).

CONCLUSION

The present studies provide the first quantification of coordinated joint attention between two individuals engaged in an unscripted conversation. In a spontaneous, interactive dialogue relating to a common visual scene, conversants' eye movements were tightly coupled. The recurrence was greatest at 0 ms, but was at above-chance levels for lags of up to ±3,000 ms. Picking any two points in the subjects' eye movements that were within 3,000 ms, one would find they were more likely than chance to be looking at the same thing. The likelihood would be greatest for eye movements at the same point in time.

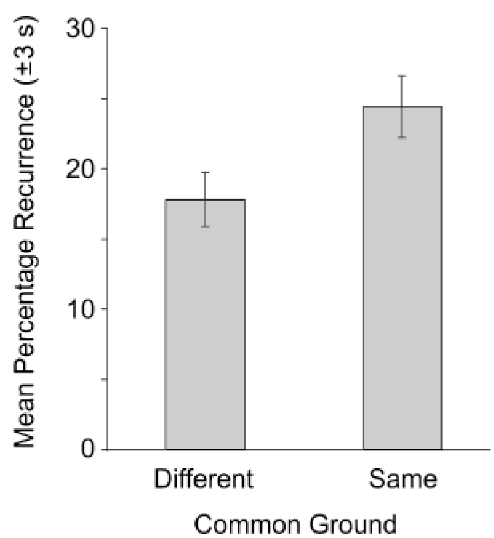

Fig. 4. Results from Study 2: average percentage of eye movement recurrence within a window of ±3,000 ms as a function of whether conversants had heard the same or different information prior to discussing the painting).

UNCORRECTED PROOF

AUS PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)

PSCI 1914

We found that this eye movement coupling is sensitive to the knowledge that conversants bring to their conversation. In our study, conversants' eye movements had a significantly tighter coupling with each other if the conversants heard the same background passage, rather than two different ones. Changes in one type of common ground, shared knowledge, can have a direct effect on subjects' ability to exploit another type of common ground, the shared visual context. This result raises several interesting issues for further research. Did conversants use the shared information as a way to refer to elements of the surrealist painting, in the same way that people can come to "conceptual pacts" when naming novel objects (Brennan & Clark, 1996; Clark & Brennan, 1991)? Is the advantage of sharing information due solely to the fact that conversants actually know the same information, or is it also important that they know that they each know the same information?

Sharing information does not, by itself, improve the coordination of attention; presumably, the information must be related to the visual display.[1] But how close does this relationship have to be? We imagine that if paired subjects in Study 2 had both heard about the 2006 World Cup, this shared knowledge would not have affected how they coordinated attention to Dali's painting. But what if they had heard a passage explaining how wavelengths of light combine to produce color, or a passage about the literature that was written during Dali's time? If this type of information were in the common ground, it could conceivably shape subjects' discussion and the coordination of their attention toward Dali's painting, even though it does not relate directly to the painting. Although these are speculative thoughts, we note that there was no significant difference in the increase in recurrence brought about by sharing the passage that was specifically about the painting and sharing the passage that was more broadly about the artist's life and methods ($F < 1$). This raises the question: What sort of raw material can subjects use to construct a common reference frame for their visual attention?

In many everyday interactions, conversants make use of pointing, placing, and gesturing to take charge of each other's visual attention. What was surprising in our studies was that there was a coupling of attention even though there were no such means for subjects to interact visually. Within each conversant, there is an intimate connection between discourse processes and visual processes (Richardson & Matlock, in press; Spivey & Richardson, in press). Between conversants, behavior is coordinated on many levels, and communication is grounded in the context they share. Our results fill in the picture of language as a joint activity (Clark, 1996; Garrod & Pickering, 2004) and provide another case of cognitive activity situated in the external world (Barsalou, 2003; Hutchins, 1995; Robbins & Aydede, in press; Suchman, 1987).

[1]We thank Robert Krauss for raising this point.

Acknowledgments—The authors are indebted to Herbert H. Clark and Michael Spivey for insightful comments on this manuscript, to Michael Ramscar and Teenie Matlock for many inspiring discussions, and to the members of the Kirkham Learning Lab who donated their time and talents to the eye tracking for this study: Lisa Smythe, Carl Dambkowski, Sasha Filippova, Debbie Kim, Lauren Rimoin, and Rosemary Reidy. Rick Dale was supported by a Paller-Dallenbach Fellowship from Cornell University.

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.

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* (pp. 131–158). Hillsdale, NJ: Erlbaum.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, *15*, 415–419.

Barsalou, L.W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18*, 513–562.

Bavelas, J.B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163–194.

Beun, R.J., & Cremers, A.H.M. (1998). Object reference in a shared domain of conversation. *Pragmatics & Cognition*, *6*, 121–152.

Branigan, H.P., Pickering, M.J., & Cleland, A.A. (2000). Syntactic coordination in dialogue. *Cognition*, *75*, B13–B25.

Brennan, S.E., & Clark, H.H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482–1493.

Brown-Schmidt, S., Campana, E., & Tanenhaus, M.K. (2004). Real-time reference resolution by naïve subjects during a task-based unscripted conversation. In J.C. Trueswell & M.K. Tanenhaus (Eds.), *Approaches to studying world-situated language processing: Bridging the language-as-product and language-as-action traditions* (pp. 153–172). Cambridge, MA: MIT Press.

Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Clark, H.H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

Clark, H.H., & Brennan, S.E. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.

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

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, England: Cambridge University Press.

Condon, W., & Ogston, W. (1971). Speech and body motion synchrony of the speaker–hearer. In D. Horton & J.J. Jenkins (Eds.), *The*

UNCORRECTED PROOF

PSCI 1914

US PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)

perception of language (pp. 150–184). Columbus, OH: Charles E. Merrill.

Cooper, R.M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107.

Dale, R., & Spivey, M.J. (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, *56*, 391–430.

Dijksterhuis, A., & Bargh, J.A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 1–40). San Diego, CA: Academic Press.

Garrod, S., & Pickering, M.J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8–11.

Giles, H., Coupland, N., & Coupland, J. (1992). Accommodation theory: Communication, context and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation* (pp. 1–68). Cambridge, England: Cambridge University Press.

Grant, E.R., & Spivey, M.J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*, 462–466.

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

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 integration of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Kamide, Y., Altmann, G.T.M., & Haywood, S.L. (2003). The timecourse of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133–156.

Kita, S. (Ed.). (2003). *Pointing: Where language, culture, and cognition meet*. Mahwah, NJ: Erlbaum.

Lee, B.H.P. (2001). Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics*, *33*, 21–44.

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

Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception*, *25*, 931–948.

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.

Richardson, D.C., & Matlock, T. (in press). The integration of figurative language and static depictions: An eye movement study of fictive motion. Cognition.

Robbins, P., & Aydede, M. (Eds.). (in press). *The Cambridge handbook of situated cognition*. Cambridge, England: Cambridge University Press.

Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, *50*, 696–735.

Schober, M.F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*, 1–24.

Shockley, K., Santana, M.V., & Fowler, C.A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 326–332.

Spivey, M., & Richardson, D. (in press). Language embedded in the environment. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition*. Cambridge, England: Cambridge University Press.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. New York: Cambridge University Press.

Tanenhaus, M.K., Spivey Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995 June). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Zbilut, J.P., Giuliani, A., & Webber, C.L. Jr. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters*, *246*, 122–128.

(Received 5/30/06; Revision accepted 8/28/06; Final materials received 8/31/06)

UNCORRECTED PROOF

US PSCI 1914.PDF 12-Apr-07 20:51 159867 Bytes 7 PAGES n operator=ananthabs)