Journal of Communication, 52, 566-580

Listener Responses as a Collaborative Process: The Role of Gaze

By Janet Beavin Bavelas, Linda Coates, and Trudy Johnson

The authors examined precisely when and how listeners insert their responses into a speaker's narrative. A collaborative theory would predict a relationship between the speaker's acts and the listener's responses, and the authors proposed that speaker gaze coordinated this collaboration. The listener typically looks more at the speaker than the reverse, but at key points while speaking the speaker seeks a response by looking at the listener, creating a brief period of mutual gaze called here a gaze window. *The listener was very likely to respond with "mhm," a nod, or other reaction during this period, after which the speaker quickly looked away and continued speaking. This model was tested with 9 dyads in which 1 person was telling a close-call story to the other. The results confirmed the model for each dyad, demonstrating both collaboration in dialogue at the microlevel and a high degree of integration and coordination of audible and visible acts, in this case, speech and gaze.*

While listening to someone tell a story, listeners regularly make brief appropriate responses (sometimes called "back channels," after Yngve, 1970). The goal of the research reported here was to understand the timing of these responses, that is, why listeners respond when they do. Is this a haphazard event, is it determined by characteristics of the listener, or is it related to actions of the speaker (and if so, how)? We have pursued these questions first through an inductive qualitative analysis, then by statistical tests of the hypotheses we had developed. The data were videotapes of unacquainted dyads in which one person told a close-call story to the other.

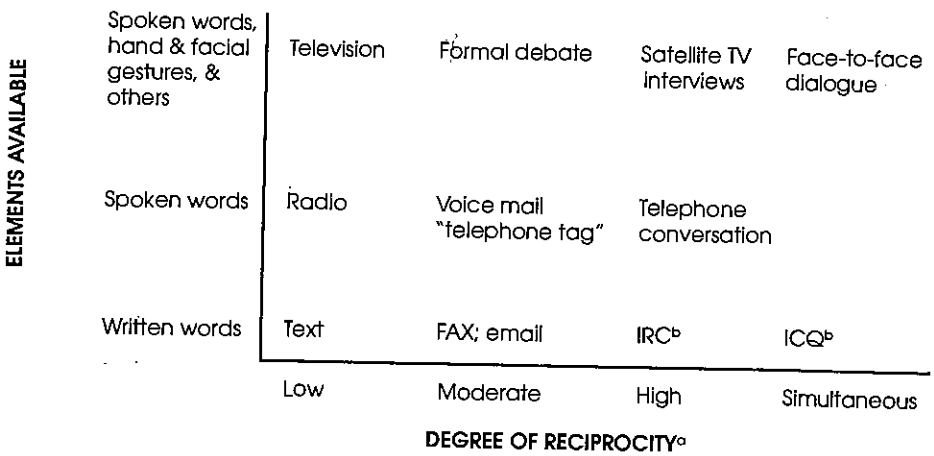
Janet Beavin Bavelas (PhD, Stanford) is a professor in the Department of Psychology, University of Victoria, BC, Canada. Linda Coates (PhD, University of Victoria) is an assistant professor in the Department of Criminology, St. Thomas University, Fredericton, New Brunswick. Trudy Johnson (MA, MPA, University of Victoria) is a senior research analyst for R. A. Malatest & Associates, Victoria, BC. These studies were supported by research grants from the Social Sciences and Humanities Research Council of Canada. The title is a tribute to Herb Clark and Deanna Wilkes-Gibbs's (1986) first collaborative study. The authors would also like to thank Sarah Hutchinson, Daniel Noel, and Christine Kenwood, their gaze analysts, and Danielle Prevost, who helped them locate the past literature on gaze. Correspondence regarding this article should be addressed to Janet B. Bavelas, Department of Psychology, University of Victoria, P. O. Box 3050, Victoria, BC, V8W 3P5, Canada; email: jbb@uvic.ca.

The assumptions guiding this research are different in many respects from familiar approaches to interpersonal communication, so they will be outlined briefly here. Our focus was on face-to-face dialogue, with an emphasis on two important features that combine to make it a unique form of communication (see Figure 1). As we will see, both of these features are intimately involved in the timing of listener responses.

First, the participants can use visible as well as audible acts[1] to convey their meaning. They can combine gestures, facial displays, and other symbolic depictions with their words. Bavelas and Chovil (2000; see also Bavelas, 1994; Bavelas & Chovil, 1997; Bavelas, Hutchinson, Kenwood, & Matheson, 1997) have outlined an "integrated message model" in which certain specific acts work closely with words to convey meaning. To distinguish this subset from the much larger domain of all nonverbal behaviors, we have called them *visible acts of meaning*. They are limited to symbolic acts that are closely integrated with the simultaneous words in both timing and meaning, that occur in the presence of a receiver, and that are understood by the receiver. They consist primarily of conversational hand gestures and facial displays but also include nods, shrugs, and other meaningful acts that meet these criteria. We propose that this subset of nonverbal acts does not serve as a separate channel conveying different information from words (e.g., emotional or relationship information). Rather, the audible and visible acts of meaning together form an integrated message, which has also been called "mixed syntax" (Slama-Cazacu, 1976) or a "composite signal" (Clark, 1996, p. 156; Engle & Clark, 1995). The speaker integrates his or her words with these visible acts, whose meaning may be either redundant or nonredundant with the words but is always complementary and coordinated. In the background, the listener also uses visible and audible acts, together or interchangeably, for example, nodding and saying "yeah."

The other distinguishing feature of face-to-face dialogue is its high degree of *reciprocity*, which we define as the probability and latency of response from the other person. Compared to any other form of communication, the probability of a response is very high, and its latency is extremely short, often simultaneous, because listeners can nod, smile, look surprised, or say "mhm" without taking up the speaking turn. Goodwin (1986) showed that some listener responses occur in the brief pauses between the speaker's phrases, whereas others systematically occur within phrases, overlapping the speaker's words. Far from being disruptive, these insertions and overlapping contributions fit Clark's (1996) collaborative model, which describes language and communication as "joint action," something that two people do together rather than separately—a duet rather than alternating solos. A keystone of collaboration is grounding, in which "speaker and listener go beyond [individual] actions and collaborate with each other moment by moment

[1] In discussing face-to-face dialogue, we consider the terms *audible* and *visible* to be more precise and accurate than *verbal* ("of or pertaining to words") and *nonverbal*. Spoken words are not only verbal; they are oral or, more accurately, audible, because they are inextricably wrapped in paralinguistic features such as stress, pitch, speed, and voice texture. Moreover, to describe other acts as nonverbal only says what they are not. It is more precise to say that they are visible rather than audible, that is, seen rather than heard.



[a] In this metaphorical scale, the distance between "high" and "simultaneous" is very small but, in our opinion, significant. The main difference is whether turns are imposed or whether overlapping and simultaneous acts regularly occur, with the microcollaboration they make possible.

[b] These are computer-mediated communication (CMC) formats: In IRC (Internet Relay Chat; www.mirc.com), each person takes turns composing and then sending, so the receiver cannot see the message as it is being generated. In ICQ (I-Seek-You; www.icq.com), both participants can write and see each other's writing at the same time. See Phillips and Bavelas (2000).

Figure 1. Face-to-face dialogue compared to other forms of communication on two dimensions.

to try to ensure that what is said is also understood" (Schober & Clark, 1989, p. 211). Thus, the listener's moment-by-moment responses to the speaker arguably play an important role in producing the dialogue.

Speakers and Listeners

Bavelas, Coates, and Johnson (2000) have applied these two principles (integrated messages and collaboration) to highly asymmetrical dialogues, in which one person did virtually all of the speaking and the other was ostensibly only a listener: One stranger told a personal close-call story to another. Our interest was in the listener, who could not have known anything about the incident being described and therefore would seem to have virtually nothing to contribute. However, microanalysis revealed that these listeners were very active. In normal conditions, they responded (either visibly, audibly, or both) to the narrator's story an average of every 3.5 seconds. The majority of their responses were what we called *generic* listener responses, such as nods, "yeah," and "mhm," all of which conveyed attentiveness and understanding but were not specifically designed for what the listener was saying at the moment. In addition, there were many *specific* listener responses, such as winces, frowns, or supplying words, all of which were tightly connected to that precise moment in the speaker's narrative. We proposed that these specific responses went beyond indications of understanding and contributed to the narrative; the listener briefly but frequently became a co-narrator.

The broader theory outlined above guided our experiments, one of which provided the data for the present analysis; that is, we looked at both speaker and listener, and we treated audible and visible acts of meaning as integrated wholes. A generic response was often a package of "mhm" plus a nod (as well as each of those separately), just as a specific response could combine a horrified facial display with "oh no!" We found that this integrated approach was both highly reliable (i.e., high interanalyst agreement) and necessary to adequately represent the listener's meaning.

In two experiments, we created conditions that distracted the listener from the narrative, with two parallel effects: The rate of listener responses (especially specific responses) dropped significantly, and the quality of the speaker's storytelling did as well. That is, compared with the normal control conditions, the speakers with distracted and unresponsive listeners could not seem to finish their stories effectively, measurably faltering at what should have been the dramatic ending. We interpreted this as evidence of collaboration: Speakers need their listener's feedback to be able to tell their stories well.

However, these findings were based on separate analyses of the speakers and listeners and did not tell us how their acts were connected in real time. In the present study, therefore, we investigated precisely when and how each of the listener's responses occurred in relation to the speaker's actions. We began this exploration inductively and qualitatively, by examining a pilot sample of videotapes intensively for anything that had happened just before a listener response. The several possibilities included the speaker's words ("y'know?"); paralinguistic features (pauses, changes in intonation); conversational gestures (especially interactive gestures; Bavelas, Chovil, Coates, & Roe, 1995); conversational facial displays (Bavelas & Chovil, 1997); and gaze. Although all had some influence (and we would dismiss none of them as entirely irrelevant), one candidate stood out: Speaker gaze seemed to have the strongest and most consistent relationship to a listener response.

The Gaze Window

The typical gaze pattern in face-to-face dialogue is asymmetrical: The listener looks for fairly long intervals at the speaker, while the speaker looks at the listener for frequent but much shorter periods. Because of this asymmetry, the speaker's glances at the listener tend to determine whether mutual gaze occurs. When the speaker and listener change turns, there is often a brief period of mutual gaze and a change of gaze roles, so the former speaker now looks at the partner more rather than less (Argyle & Cook, 1976; Duncan & Fiske, 1977; Kendon, 1967).

In the present data, there was only one person in the speaking role, so mutual gaze should not lead to a turn exchange. Instead, when the speaker looked at the listener, this started a brief period of mutual gaze, a *gaze window,* in which we noticed that a listener response was very likely to occur. Moreover, this was not simply a stimulus-response system in which the speaker evoked a response from the listener. Rather, it was the listener's response that seemed to terminate the speaker's gaze and therefore end the gaze window. The speaker and listener then resumed their previous gaze pattern. It is this termination without an exchange of roles that distinguishes the gaze window from the pattern associated with a turn exchange.

We proposed the following microcollaboration between speaker and listener: The speaker typically seeks a response from the listener by looking at him or her, which begins a brief period of mutual gaze. When the listener responds within this gaze window, the speaker quickly looks away, terminating the window and continuing to hold the turn. Figure 2 shows an example from one of our dyads. The speaker was describing a close call that happened while traveling in Europe with her sister. Our transcription is threefold, linking the speaker's words with the gaze windows and with the listener's responses. In this excerpt, three of the four listener responses occurred within a gaze window that ended shortly after the listener responded.

(We also noticed some context-specific variations on this pattern: On some occasions, the speaker might continue to gaze after the first listener response while also continuing to explain his or her point, and the listener would, for example, continue to nod until the speaker looked away. On other occasions, when the listener did not respond or responded tentatively, the speaker might repeat or elaborate the information until the listener indicated understanding.)

Previous Research on Gaze in Dialogue

In his 1977 review, Cook optimistically pointed out that "from the mid-sixties the volume of research on gaze expanded rapidly, and in the decade 1965 to 1975 a great deal was discovered on the topic" (1977, p. 328). Unfortunately, Cook's summary hardly needs to be updated a quarter-century later. The expansion of research on gaze seems to have been virtually limited to the period Cook described (e.g., Argyle & Cook, 1976; Davey & Taylor, 1968; Duncan & Fiske, 1977; Ellsworth & Carlsmith, 1968; Ellsworth & Ludwig, 1972; Kendon, 1967; Kendon & Cook, 1969; Kleck & Nuessle, 1967; Knapp, Hart, Friedrich, & Shulman, 1973; Mobbs, 1968; Scherwitz & Helmriech, 1973) with the last major review in 1986 (Kleinke, 1986). By then, a bifurcation common to research on nonverbal communication was already obvious, dividing the research into two distinct approaches: (a) A minority of studies focused on the function of gaze in the dialogue itself, using microanalysis to examine the relationship of gaze to the immediately surrounding audible and visible behaviors (e.g., Goodwin, 1981); and (b) the majority of studies examined gaze in relation to other variables external to dialogue, for example, correlating the amount of gaze to interpersonal attitudes, emotions, or personality differences (e.g., Argyle & Cook, 1976) or studying the effects of staring at strangers in public (Ellsworth, Carlsmith, & Henson, 1972). Our interest is in the former approach, the functional microanalysis of gaze in dialogue.

Kendon (1967) opened up the field with a classic, detailed exploration of the

First of all I have to tell you that my sister calls me Surefoot Charlotte cause I always trip.
(asterisks over "sister calls me Surefoot Charlotte cause I always trip.")

I'm not, I'm not clumsy but I just, if there's a place to trip I'll find it. So we're exploring in the
Mm + nod + smile (under "I'm not,") ... *nod + smile* (under "I'll find")

castle and there's this (slight pause) **tall ladder going up to a window, and I was going to**
nod (under "and")

climb up to it.
nod

The speaker's words are in boldface. The asterisks above indicate periods of mutual gaze (gaze windows). The listener's responses appear in italics directly below the words they accompanied.

Figure 2. Excerpt illustrating the relation of gaze windows to listener responses.

functions of gaze in face-to-face dialogues (which he ingeniously filmed in split-screen with one camera by using a mirror), and many of his observations have subsequently been replicated by others. Kendon was the first to document the patterns summarized above, namely, that a person "tends to look at [the other] more while he is listening than while he is speaking"; that the speaker's glances at the other person "tend to be shorter than those observed during listening" (p. 37); and that mutual gaze played a role in smooth turn exchanges.

In addition and of greatest relevance to the present research, Kendon (1967) also observed mutual gaze *within* long turns and speculated about a "regulatory" role of such gaze in eliciting listener responses, which he called "accompaniment signals, . . . the short utterances that the listener produces as an accompaniment to a speaker, when the speaker is speaking at length" (p. 43). These short utterances are not turns. Rather, Kendon argues:

> During the course of a long utterance, [the speaker's] glances at [the listener] come at the points at which he receives an accompaniment signal from him, and so may function not only as checks on [the listener's] behavior, but as signals to [the listener] that [the speaker] wants confirmation that what he is saying is getting across. (p. 56)

In an earlier part of his article, Kendon (1967, pp. 31–32) illustrated his general method with a diagram and description of speaker and listener actions in one conversational excerpt. Although he did not enter the accompaniment signal on the diagram, the text makes it clear that there was an instance of what we would call a gaze window with a listener response at the end, just as our model predicts. We also found diagrams containing a gaze window terminated by a listener response in reports by Cook (1977, p. 329) and Duncan and Fiske (1977, p. 213).

In summary, we propose that the timing of the listener response is a collaborative process, accomplished by joint action: Speaker gaze creates the opportunity for a listener response, and the response then terminates that gaze. Neither of the individuals alone controls when and where a listener response occurs. In the research reported here, we tested this hypothesis by analyzing a new set of narrations, identifying gaze windows and listener responses and then evaluating whether they co-occurred by chance or fit our model.

Method

Participants

The 24 participants were first-year psychology students who signed up for extra course credit; they formed 12 dyads who were strangers to each other. They were all in one condition of a larger study (Bavelas et al., 2000, Experiment 2). The gender mix was counterbalanced in both conditions, creating equal numbers of male-male, female-female, and male-female dyads. All participants had consented to being videotaped in the Psychology Department's Human Interaction Lab. They also viewed their tape after the experiment and gave permission for its subsequent analysis.

For the present study, we selected the 12 dyads in the control condition because the listener was listening normally. (As described earlier, in the experimental condition the listener was deliberately distracted from the speaker's story, significantly reducing the rate of listener responses and producing a highly unusual interaction; see Bavelas et al., 2000, pp. 947–949, for details.) After assessing gaze, we excluded three dyads from analysis because of atypical speaker gaze: One speaker never looked directly at the listener but instead maintained a pseudo-gaze slightly down and off to the side; another speaker only looked at the listener 5% of the time; a third looked 83% of the time. All three were extreme outliers below or above the other nine, who had a mean total gaze of 31% (range = 15 to 62%; *SD* = 14%). These parameters are consistent with those given by Kendon (1967), Duncan and Fiske (1977), and Argyle (1967). It is worth noting that the first two pairs, whose gaze was unusually low, were male-male dyads, whereas the third, whose gaze was unusually high, was female-female. In summary, our sample consisted of the 18 participants in 9 dyads[2] whose amount of gaze fell in the statistically normal range.

Equipment

Our Human Interaction Lab has four remotely controlled Panasonic WD-D5000 color cameras and two special effects generators (a Panasonic WJ-5500B overlaid

[2] We were not concerned about this relatively small *N* for several reasons. First, each dyad provided an average of 17 listener responses, for a total of 154. Second, it is more difficult to reach statistical significance with a small sample, so the small sample does not favor our hypotheses. Third, we also tested our model on each dyad separately, a harder test that ensures that the overall effect was not due to just a few dyads.

on a customized Panasonic four-camera system). We used two cameras to videotape both the narrator and the listener in a VHS split-screen layout that recorded a face-on view of each of the participants and a time signal on the tape in minutes, seconds, and hundredths of seconds. For analysis, we used either a Sharp 2500S or a JVC BR-S605-UB VCR and a 19-inch color Sony or Electrohome monitor.

Procedure

After introductions, the experimenter randomly chose one member of the dyad to tell a close-call story to the other person:

> I'd like you to tell something about a close call or near-miss incident. A close call is something that happened where someone was almost hurt, or something bad almost happened, but in the end everything turned out okay. Make sure that you tell something you're comfortable telling. And if you can't think of something that happened to you, then you can tell about a close call that happened to a friend. Just to give you some ideas, other people have told stories about skiing accidents, horseback-riding accidents, and nearly losing a term paper in the computer. I would like you to tell your story in as much detail as you can. So don't just describe it in a couple of sentences.

After these instructions, the narrator left the room to think about his or her story, and the experimenter gave the listener instructions that created the experimental or control condition. In the present data (the control condition), the instructions were simply to "listen to the story so that, if you had to, you could summarize the gist or main point of it to someone else." We intended these instructions to approximate the way in which listeners would ordinarily listen to a stranger's story, that is, with polite attention.

The average story length for the nine dyads analyzed was 2 min, 44 s (range = 1 min 5 s to 6 min 42 s). For the present study, we analyzed the first and last minute of each story. The first minute began with the start of the story itself, excluding irrelevant preliminary comments (e.g., "OK, I've thought of one"). The actions of either participant could define the ending point. Usually the speaker announced the ending (e.g., "That's it"; "That's my story") or marked the ending with a discourse shift marker (e.g., "Anyway"). Sometimes the listener began to speak (e.g., commenting on the story), so he or she was no longer in the listening role. Two expert analysts identified these points by consensus. For four dyads, the time between beginning and ending was less than 2 minutes, which meant we analyzed the entire story. For the others, we used the video time signal to locate 60 seconds forward from the beginning and backward from the ending.

Measures

Listener Responses

First we identified each listener response by the time of onset and transcribed both the response and the speaker's words that surrounded it. We defined listener responses as follows:

> actions [both audible and visible] that indicate the person is attending, following, appreciating, or reacting to the story. They include (but are not limited to) nodding, "mhm," "yeah," smiling, laughing, motor mimicry,[3] gesturing the content of the story, supplying words or phrases, dramatic intakes of breath, and displays of excitement, fear, or alarm.

This definition excluded noncommunicative listener behaviors such as adaptors and any behaviors that were too ambiguous to analyze. Among the latter were smiles that occurred on their own, with no other communicative acts.

Consistent with our integrated message model, what we called a listener response was not necessarily just one isolated physical act. In many cases, the listener combined two or more behaviors together in a "package" that conveyed his or her meaning, for example, a nod simultaneous with "yeah." We treated simultaneous or very closely sequential behaviors as a single listener response as long as their meaning was the same or complementary. However, if different behaviors had different meanings, these were treated as separate responses. For example, in tracking the story closely, the listener might nod and then immediately look alarmed; these would be two different listener responses. The interanalyst reliability for two experts who packaged or separated the listener responses was above 90%. Then two new analysts, who were unaware of our hypotheses and of the experimental conditions, distinguished between generic and specific listener responses; their agreement was over 95%.

The time of onset of the listener response, which is particularly important for this research, was rechecked and refined by an expert analyst, using the jog/shuttle function on the VCR and the time signal on the videotape.

Gaze

We timed the onsets and offsets of mutual gaze that created the gaze windows. Although we were working with a split-screen image in which both participants appeared to be looking at us rather than each other, it was possible to identify mutual gaze reliably, using several guidelines. First, we made a similar videotape of two of us in exactly the same configuration, systematically varying our gaze while announcing whether we were in mutual gaze (e.g., "Now I'm looking at your eyes, your ears, etc."). This tape provided a standard for inferring mutual gaze from individual gaze directions in this setting. Second, there are subtle but noticeable changes associated with mutual gaze. For example, the speaker's head often moves as well as his or her eyes, for example. Eye movement typically stops momentarily during mutual gaze; Argyle (1967) noted that, although ordinary visual scanning typically involves fixations of .25 to .35 seconds, gaze duration is typically much longer, in the range of 1 to 7 seconds. Finally, the participants' facial expressions may display heightened attentiveness to the other during mutual gaze.

[3] We had defined motor mimicry in our earlier research as actions by an observer that are appropriate to the situation of the other person but not to the observer's own situation (e.g., wincing at the other's injury; Bavelas, Black, Lemery, & Mullett, 1986).

Still, a fine judgment was required, so interanalyst reliability was essential. In the first minute of dialogue for two dyads, two independent analysts identified mutual gaze with 100% agreement. They identified the same onset and offset of speaker gaze within .1 second of each other with 83.9% agreement. For the second minute of another dyad, two different analysts also agreed 100% on mutual gaze, and they had 86.75% agreement on onset and offset within .2 second. Subsequently, one of the second pair of analysts located speaker gaze in the remainder of the data and also verified that the listener had been looking at the speaker at the same time, that is, that mutual gaze (a gaze window) had occurred.

We then added .5 second to each gaze window because our own reenactments made it clear that the speaker could still see the listener's response for a brief period after direct mutual gaze was over. That is, by sitting face-to-face, we determined (as can the reader) that the speaker is able to see listener responses, even tiny nods, as he or she starts to look away from the listener. So the *effective gaze window* is slightly longer than the length as measured by direct eye contact. It is important to point out that, although this extension allowed additional listener responses to be included in gaze windows, our statistical model prevented this increase from favoring our hypothesis. As will be seen below, the increased gaze window made it harder to confirm our hypothesis.

Results

Confirming our model required demonstrating that the placement of listener responses within gaze windows was not simply a coincidence. That is, if both listener responses and mutual gaze were very frequent events, then they would often occur together by chance. So it was necessary to demonstrate that the occurrence of listener responses within a gaze window exceeded the level that could be expected to occur randomly. The familiar parametric statistics will not evaluate this possibility, so it was necessary to set out our logic and find a suitable method.

We chose a one-tailed ("upper tailed") binomial test (Conover, 1999, pp. 126–127). As implied above, if listener responses occurred randomly, the number observed to fall within gaze windows (O) would be a function of the proportion of time that gaze windows were available (p) and the number of listener responses (n). On the other hand, if our hypothesis is correct, then more listener responses should fall in gaze windows than can be accounted for by a random process.

There are two forms of the binomial test, depending on the number of events (listener responses) that occur (Conover, 1999, pp. 124–125). When the number of responses is less than 20, a table of the binomial distribution gives exact significance levels for all combinations of n, p, and O (e.g., Conover, 1999, Table A3). When the total number of responses is 20 or more, the binomial distribution approximates the normal distribution, especially with a correction for continuity, so we can use the following formula:

$$z = \frac{O - E - .5}{\sqrt{npq}}$$

where

n = total number of listener responses[4]
p = proportion of time in mutual gaze (+ .5 s)
$q = 1 - p$
O = the number of listener responses that occurred within in a gaze window
$E = pn$, the number expected to occur in a gaze window by chance
- .5 = correction for continuity

Table 1 presents the results of these tests. An omnibus test across all nine dyads showed that the probability of these listener responses occurring during mutual gaze by chance alone was almost incalculably small. Next, we applied the same binomial test to each of the nine dyads. This is a much more stringent test, because it both reduces the n and, unlike most statistics, requires that the pattern hold for each case, not just on average. For each of the nine dyads, listener responses were significantly more likely to occur within gaze windows than chance could account for.

These dyads were about equally divided between stories shorter and longer than 2 minutes (i.e., in which the first and last minute did and did not overlap), so the pattern held throughout the narrative and regardless of its length. Because we had shown (Bavelas et al., 2000) that specific responses occur significantly later, on average, than do generic responses, this finding suggests that the two kinds of responses do not differ in their relation to mutual gaze. A more direct test of whether generic versus specific responses fell inside or outside gaze windows was also nonsignificant; $\chi^2(4, N = 173) = .30, p > .05$).

Finally, we tested our corollary hypothesis that a listener response tended to terminate the gaze window, that is, once the listener had responded, the speaker would look away, indicating that this was not a turn exchange. If we were correct, listener responses should occur in the latter part of the gaze window, whereas by chance they would occur, on average, in the middle. To be conservative, we analyzed the direct gaze window (and not the slightly longer effective gaze window used above). That is, we excluded the listener responses that occurred in that extra .5 sec interval, because all of these responses would automatically corroborate our hypothesis. On average, listener responses occurred .69 through the direct gaze window. A t-test comparing this average placement with .5, the value expected by chance, supported our prediction; $t(58) = 6.16, p < .001$.

Discussion

These results strongly supported our gaze-window hypothesis: The listener tended to respond when the speaker looked at him or her, and the speaker tended to

[4] Some readers may be concerned that our data violate the assumption of independent events because each dyad contributes many more than one listener response. Recall, however, that the classic examples of binomial probability use n tosses of the *same coin*. Statistical independence means that the probability of A is identical to the conditional probability of A given that B occurred (e.g., Conover, 1999, p. 18).

Table 1. Statistical Tests for Occurrence of Listener Responses Within Gaze Windows

Dyad	Responses in gaze (*O*)	Response total (*n*)	Proportion of gaze + .5 s (*p*)[a]	*z*	One-tailed *p* value
		Binomial tests ($n < 20$)[b]			
1	6	7	0.41		<.002
2	7	8	0.54		<.009
3	8	9	0.36		<.0001
4	11	15	0.55		<.05
5	12	15	0.36		<.0001
6	11	16	0.22		<.0000
	Approximation to normal distribution, with correction for continuity ($n > 20$)				
7	19	25	0.40	3.69	<.0003
8	24	27	0.46	4.5	<.00003
9	30	32	0.73	2.63	<.008
Total	128	154	0.45	9.43	$<.01 \times 10^{-10}$

[a] We measured each direct gaze window in tenths of seconds, then added .5 s (see text) to give the duration of each effective gaze window. The sum of these durations for the dyad was then divided by the total number of seconds analyzed, yielding *p*, the proportion of time an effective gaze window was available.

[b] Using Table 3a in Conover (1999).

look away soon after the listener responded. Together, speakers and listeners created and used the gaze window to coordinate their actions. They demonstrated an efficient and precise use of gaze, not only to regulate turn exchanges as already proposed in the literature, but also to seek and provide listener feedback without signaling a turn exchange.

These results extend the collaborative model in several ways. First, it might not be surprising to observe collaborative acts when both participants have equal standing in the dialogue, that is, when both contribute through speaking turns. However, in these data, the listeners remained listeners with no role other than very brief interjections, yet the evidence suggests that they made even these minimal responses in close collaboration with the speaker.

Second, a collaborative model is necessarily a model of microprocesses. Although it can be supported by aggregated data such as outcomes (e.g., Bavelas et al., 2000; Schober & Clark, 1989), it must in addition examine the local processes of collaboration (e.g., Clark & Schaefer, 1987; Roberts & Bavelas, 1996). The present data complement qualitative studies of these processes with a quantitative, statistical analysis applied at the microlevel.

Third, collaboration is not only a verbal process. In the present instance, the listeners' responses were both audible and visible, often both. The speaker's actions also combined visible and audible acts, namely, changes in gaze direction while continuing to speak. The question of why and when listeners respond cannot be answered by looking at only one channel or the other.

There are also implications for our integrated message model. As it was used in this specific context, speaker gaze fits the four criteria required for a visible act of meaning (Bavelas & Chovil; 2000):

1. It is sensitive to the availability of a receiver in that the speaker uses the listener's typical gaze pattern to create the gaze window.

2. It is serving a symbolic function; the speaker is not looking at the listener merely to see him or her but in order to request a response.

3. We have evidence that the listeners understood the meaning of act because they responded immediately and appropriately.

4. The gaze was fully integrated with the words and other visible acts. In this case, the relationship between the speaker's words and gaze was primarily nonredundant and complementary rather than redundant. With regard to other acts, our initial inductive analysis showed that the speaker's gaze was often redundant with his or her concomitant pauses, intonation contours (e.g., rising pitch), interactive gestures, or facial displays (e.g., raised eyebrows).

Finally, we can return to our broader theory of face-to-face dialogue with some thoughts on units of analysis, which the social psychologist Kurt Lewin (1935) proposed may be the most important choice the theorist or researcher makes. The typical scientific approach is reductionist, seeking the smallest possible units. The assumptions underlying this practice are that smaller units are easier to study, that larger units are simply additive assemblages of the basic units, and that the units thus isolated can later be added to or combined with other units, which were also studied in isolation. It is perfectly sensible to act on these assumptions as long as they are valid, but it is equally important to remember that their validity is an empirical matter.

Reductionist assumptions dominate many approaches to interpersonal communication and even to face-to-face dialogue, which separate verbal from nonverbal acts or focus on the individual rather than the dyad (or both). We propose that the smallest possible unit for studying dialogue must necessarily be larger, along two dimensions. The elements available to the participants for conveying their moment-by-moment meaning are not separable into audible and visible but must be taken as a whole. Similarly, the appropriate unit of analysis is not each of individuals involved but rather their dialogue with all of its reciprocity and collaboration. As shown in the present study, even these two dimensions cannot be studied separately. The process of collaboration on listener responses combined integrated audible and visible messages with the principle of reciprocity; had we chosen to study only one or the other, the process would have eluded us. The good news is that a larger unit of analysis may not be too complex after all. We found that, once we abandoned the limits imposed by reductionism, the process revealed itself in a straightforward and simple form.

References

Argyle, M. (1967). *The psychology of interpersonal behaviour.* Harmondsworth, UK: Penguin.

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

Bavelas, J. B. (1994). Gestures as part of speech: Methological implications. *Research on Language and Social Interaction, 27,* 201–221.

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I *show* how you feel": Motor mimicry as a communicative act. *Journal of Personality and Social Psychology, 50,* 322–329.

Bavelas, J. B., & Chovil, N. (1997). Faces in dialogue. In J. Russell & J.-M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 334–346). Cambridge, UK: Cambridge University Press.

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.

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin, 21,* 394–405.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology, 79,* 941–952.

Bavelas, J. B., Hutchinson, S., Kenwood, C., & Matheson, D. H. (1997). Using face-to-face communication as a standard for other communication systems. *Canadian Journal of Communication, 22,* 5–24.

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

Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes, 2,* 19–41.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22,* 1–39.

Conover, W. J. (1999). *Practical nonparametric statistics.* New York: Wiley.

Cook, M. (1977). Gaze and mutual gaze in social encounters. *American Scientist, 65,* 328–333.

Davey, A. B., & Taylor, L. J. (1968). The role of eye-contact in inducing conformity. *British Journal of Social and Clinical Psychology, 7,* 307–308.

Duncan, S., Jr., & Fiske, D. W. (1977). *Face-to-face interaction.* Hillsdale, NJ: Erlbaum.

Ellsworth, P. C., & Carlsmith, J. M. (1968). Effects of eye contact and verbal content on affective response to dyadic interaction. *Journal of Personality and Social Psychology, 10,* 15–20.

Ellsworth, P. C., Carlsmith, J. M., & Henson, A. (1972). The stare as a stimulus to flight in human subjects: A series of field experiments. *Journal of Personality and Social Psychology, 21,* 302–311.

Ellsworth, P. C., & Ludwig, M. (1972). Visual behavior in social interaction. *Journal of Communication, 22*(4), 375–403.

Engle, R., & Clark, H. H. (1995, March). *Using composites of speech, gestures, diagrams and demonstrations in explanations of mechanical devices.* Paper presented at the American Association for Applied Linguistics conference, Long Beach, CA.

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

Goodwin, C. (1986). Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies, 9,* 205–217.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26,* 22–63.

Kendon, A., & Cook, M. (1969). The consistency of gaze patterns in social interaction. *British Journal of Psychology, 60*, 481–494.

Kleck, R. E., & Nuessle, W. (1967). Congruence between the indicative and communicative functions of eye contact in interpersonal relations. *British Journal of Social and Clinical Psychology, 7*, 241–246.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100*, 78–100.

Knapp, M. L., Hart, R. P., Friedrich, G. W., & Shulman, G. M. (1973). The rhetoric of goodbye: Verbal and nonverbal correlates of human leave-taking. *Speech Monographs, 40*, 182–198.

Lewin, K. (1935). *A dynamic theory of personality: Selected papers of Kurt Lewin* (D. K. Adams & K. E. Zener, Trans.). New York: McGraw-Hill.

Mobbs, N. A. (1968). Eye-contact in relation to social introversion/extraversion. *British Journal of Social and Clinical Psychology, 7*, 305–306.

Phillips, B., & Bavelas, J. B. (2000, June). *Comparing collaborative versus autonomous models of conversation in computer-mediated communication.* Paper presented to the annual conference of the International Communication Association, Language and Social Interaction Division, Acapulco, Mexico.

Roberts, G. L., & Bavelas, J. B. (1996). The communicative dictionary: A collaborative theory of meaning. In J. Stewart (Ed.), *Beyond the symbol model* (pp. 135–160). Albany: State University of New York Press.

Scherwitz, L., & Helmreich, R. (1973). Interactive effects of eye contact and verbal content on interpersonal attraction in dyads. *Journal of Personality and Social Psychology, 25*, 6–14.

Schober, M., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211–232.

Slama-Cazacu, T. (1976). Nonverbal components in message sequence: "Mixed syntax." In W. C. McCormack & S. A. Wurm (Eds.), *Language and man: Anthropological issues.* The Hague, Netherlands: Mouton.

Yngve, V. H. (1970). On getting a word in edgewise. *Papers from the sixth regional meeting of the Chicago Linguistics Society* (pp. 567–578). Chicago: Chicago Linguistic Society.