

This article was downloaded by: [Janet Bavelas]

On: 21 September 2011, At: 06:49

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



International Journal of Listening

Publication details, including instructions for
authors and subscription information:

<http://www.tandfonline.com/loi/hijl20>

The Listener as Addressee in Face-to-Face Dialogue

Janet Beavin Bavelas^a & Jennifer Gerwing^a

^a Department of Psychology, University of Victoria

Available online: 20 Sep 2011

To cite this article: Janet Beavin Bavelas & Jennifer Gerwing (2011): The Listener as Addressee in Face-to-Face Dialogue, *International Journal of Listening*, 25:3, 178-198

To link to this article: <http://dx.doi.org/10.1080/10904018.2010.508675>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The Listener as Addressee in Face-to-Face Dialogue



Janet Beavin Bavelas



Jennifer Gerwing

Janet Beavin Bavelas and Jennifer Gerwing
Department of Psychology
University of Victoria

Face-to-face conversation is a unique listening setting, with a particular kind of listener; the person the speaker is directly addressing is the *addressee*. Our research program has included several experiments involving detailed, reliable examinations of the subtle yet crucial behaviors that addressees use to collaborate with the speaker in face-to-face dialogue. We have found that addressees respond to speakers using either *generic* back channels (e.g., “m-hm” or nodding) or responses that specify what the addressee has understood (e.g., opening eyes wide to show surprise). Addressees timed these *specific* responses to precise moments in the

speaker's narrative, and they tailored their responses to that moment (e.g., wincing when the speaker described something painful). Distracting addressees with a task that prevented them from following the speakers' narratives made these addressees unable to contribute specific responses, which, in turn, had a deleterious effect on the speakers' storytelling. Further research showed that addressees who were not distracted used a wide variety of behaviors to contribute to dialogue without interrupting the speaker, such as brief vocalizations, facial displays, and even gestures. Speakers and addressees regulated the timing of addressee responses using an interactive pattern of gaze. Addressees also indicated understanding by their *formulations*, which summarized or paraphrased what the speaker had said. However, our analysis showed that these formulations were not neutral. The analysis of addressees in face-to-face dialogue generates a deeper understanding of the listening process and has implications for listening in applied settings, such as psychotherapy or health care interactions.

THE LISTENER AS ADDRESSEE IN FACE-TO-FACE DIALOGUE¹

As those who study the phenomenon of *listening* know well, listening research encompasses a diversity of interests, theories, and methods. Simply in terms of who is listening, one can focus on everything from a crowded auditorium to a single listener in a dialogue. Questions as to why they listen, how they do, and what outcomes are relevant and measureable all lend themselves to investigations from different disciplines, many of which might be otherwise disconnected from each other. Therefore, listening research can be an unwieldy body of research to organize within one framework. However, with a framework in place, an opportunity arises for areas of literature to emerge as being relevant and interesting.

In their review of listening research, Bodie, Worthington, Imhof, and Cooper (2008, e.g., p. 112, figure 1) provided an organizing framework with three main components: *listening presage* (person factors and listening context), *listening process* (mental processes and overt behaviors), and *listening product* (knowledge, relationship, affect). Our research program focuses on interactive behaviors listeners use in face-to-face dialogue, so it fits within the listening-process component of Bodie et al.'s (2008) framework. It will soon be clear that this focus demands theoretical and methodological choices that contrast with listening research from other disciplines. However, these choices may be interesting or even appealing precisely because of these differences.

In the past two decades, research in experimental psychology and psycholinguistics has identified a unique and highly influential listening role, that of the

¹This article extends a chapter written for M. Mizzau and L. Lugli (Eds.). (2010) *L'ascolto* [Listening]. Bologna, Italy: Il Mulino (in Italian). The program of research reported here is supported by grants from the Social Science and Humanities Research Council of Canada.

addressee in a face-to-face dialogue. Schober and Clark (1989) distinguished between two kinds of listeners: an *addressee* is the person the speaker is addressing directly and who can respond to and interact with the speaker in a dialogue. There are listeners in many other settings (e.g., television, radio, lecture halls, courtrooms, parliaments), but they cannot interact with the speaker. They are *overhearers*; only in dialogue is there an addressee. Unlike an overhearer, the addressee is an immediate and essential part of any dialogue because the speaker is talking specifically to and with that particular addressee. Moreover, in Clark's (1996) *collaborative model* of language use, the addressee is a full partner in creating the dialogue.

In a face-to-face dialogue, the addressee and speaker have visible as well as audible means of communication available at all times. (There are speakers and addressees in other settings, such as telephone dialogues, but they cannot see each other.) When listening face-to-face, addressees use facial displays, gaze, head and hand gestures, and brief vocalizations to respond constantly and specifically to the speaker—yet without interrupting or taking up the speaking turn (Bavelas, Coates, & Johnson, 2000, 2002). In this article, we will review our ongoing program of research focused on discovering the details of what addressees do that makes them essential in face-to-face dialogue. First, however, we will provide an example of a speaker and an addressee in a face-to-face dialogue that illustrates the research findings covered in this chapter.

THE "SLEEPER": EXAMPLE OF A SPEAKER AND ADDRESSEE DURING A NARRATIVE²

The participants were two female university students who had volunteered to be in one of our studies, where they met each other for the first time. By random assignment, one participant was the speaker and the other was the addressee. The speaker's task was to tell a personal close call story, that is, an experience that could have been dangerous but that turned out all right in the end. The gist of this speaker's story was that she had fallen asleep and left her bedside lamp on, and although the heat from the lamp began to burn her pillow while she continued to sleep, she awakened before getting hurt. The addressee to whom this speaker was telling her story had no instructions other than simply to listen to the story. When reading the transcript, it is important to know that the speaker used more than her words to tell this story. She was a lively storyteller, constantly using her gestures,

²It would be best to see a video of this dialogue, but only a transcript is possible in this article. However, the video is available on a CD that was included as part of a special issue of the *Journal of Communication* (2002, Vol. 53, No. 2).

face, and vocal variation to illustrate and animate her words. Also, like most of the other speakers in this study, she emphasized the humor as well as the drama of her story, depicting the possible danger but also laughing at it because she was not harmed.

Because the speaker and the addressee were strangers, the addressee would never have heard the story of the bedside lamp that burned the pillow. There would seem to be nothing that the addressee could contribute to the story, but this turned out to be far from true. Table 1 is a transcript of the story with the addressee's responses inserted in square brackets immediately below the underlined words that the speaker was saying at that moment. Table 1 shows that the addressee was far from passive, responding with hand and facial gestures (*in italics*) as well as words (in quotation marks). For example, precisely while the speaker was saying "left it on" (foreshadowing what the close call would be), the addressee began to look concerned: She stopped smiling, raised her eyebrows, and rounded her mouth. Indeed, most of the addressee's responses were simultaneous with the speaker's words. Only when the addressee spoke up at the end of the story did their interaction resemble a turn-taking dialogue. The addressee's actions in Table 1 illustrate many of the addressee behaviors that we have studied in our research on addressees in dialogue. Therefore, we will return to examples from Table 1 frequently in this chapter, using excerpts to enliven the findings of our experiments. These findings are best introduced with some information about their theoretical and methodological context.

FOCUSING ON FACE-TO-FACE DIALOGUE

Our research group focuses on the unique features of face-to-face dialogue. In the view of many scholars, face-to-face dialogue is the basic form of language use (e.g., Bavelas & Chovil, 2000, 2006; Bavelas, Hutchinson, Kenwood, & Matheson, 1997; Chafe, 1994; Clark, 1996; Fillmore, 1981; Garrod & Pickering, 2004; Goodwin, 1981; Levinson, 1983; Linell, 2005). For example, Chafe (1994) assumed "that there is one particular use of language—ordinary conversation—whose special status justifies treating it as a baseline from which all other uses are deviations" (p. 41). In addition to being the setting of ordinary conversations, we would point out that face-to-face dialogue is a child's first language arena developmentally, and it must have been the first site of language use for early humans as they evolved. Yet in many ways, the setting of face-to-face dialogue is not a primary focus of communication research. Instead, speakers and listeners often inhabit different studies, especially in quantitative experiments.

Our theoretical commitment to the unique status of face-to-face dialogue leads to two methodological imperatives: Our unit of study is a dyad consisting of two real participants (i.e., neither is a confederate or the experimenter), and our data

TABLE 1
The "Sleeper" Story

1	S: "Uh, I have a single bed with a headboard <u>on the back of it.</u> [<i>"mm"; "hm"; nodding; looking attentive</i>]
2	And I got a light <u>for Christmas,</u> [<i>slight nod</i>]
3	a lamp that you clamp on <u>to the headboard.</u> [<i>slight nod</i>]
4	And it's got like a, um, you know, a reading lamp <u>or whatever.</u> [<i>"m-hm" + slight nod</i>]
5	And, for at night when <u>you're in bed and you don't feel like getting out of bed.</u> [<i>continuous small nods</i>]
6	<u>It's just attached right to your headboard.</u> [<i>continuous slight nodding</i>]
7	And I guess I <u>left it on.</u> [<i>stops smiling, raises her eyebrows; looks concerned</i>]
8	And it's got a really, <u>really strong,</u> [<i>bites lip</i>]
9	<u>hot light, like the, the light</u> [<i>smiles, ducking her head puts her fingers on her chin; fades to closed-mouth smile</i>]
10	really heats up. You, you can put your hand really close <u>to it</u> [<i>"m-hm"</i>]
11	and feel heat coming off the <u>light.</u> [<i>"m-hm"</i>]
12	I guess it was on for, I don't know how <u>long it was on for.</u> [<i>raises her eyebrows, widens her eyes</i>]
13	<u>But it was facing down towards my pillow.</u> [<i>maintains previous expression</i>]
14	It started <u>burning a hole in my pillow.</u> [<i>straightens her head; "Oh my goodness!"</i>]
15	<u>And I--my head is, like, THIS far away from the light, right?</u> [<i>maintains position, fingers still on chin</i>]
16	Burning a hole, <u>burning a h-hole.</u> [<i>eyes are very wide; ducks her head slightly</i>]
17	Then it starts to catch <u>on fire.</u> [<i>looks more alarmed: moves hand down to chest and "freezes" there with open mouth and wide eyes</i>]

(Continued)

TABLE 1
(Continued)

18	And <u>I still don't</u> , [slight smile]
19	I--like <u>I'm still sleeping</u> . [resettles her hand on her chest; smiling]
20	And it's <u>like, you know, right like</u> . [starts to twist her collar]
21	you know, <u>my head's on the pillow like this</u> [hand back up to her mouth; gasps, which becomes a laugh]
22	and it's just right <u>there</u> . ["m-hm"]
23	And my mom smelled the smoke <u>in her room</u> . [tilts her head, looks puzzled]
24	And she came in and she's going—"
25	A: [interrupting and spreading her hands out to depict the whole room] "So it--the room was starting to get full of smoke?"
26	S: "Yeah. [laughs] [looks away, then back; rolls her eyes; "Oh my goodness"]
27	And I didn't even wake up. <u>And like it could have gotten on fire, might—</u> " [stares with mouth open, head forward]
28	A: [interrupting; raises her hand to side of her face; puzzled expression] "You couldn't feel the heat or anything?"
29	S: "No. I just sleep <u>really, soundly—</u> " [shaking with laughter; raises her hand to mouth].

Note. The addressee's responses are in brackets and italics below the underlined words they occurred with.

are the video, not the transcript. Face-to-face dialogue consists of words, prosody, hand gestures, facial displays, and gaze—all tightly synchronized in both timing and meaning (Bavelas & Chovil, 2000, 2006). For example, when the speaker in the Sleeper story said "my head is, like, THIS far away from the light" (line 15), she demonstrated the distance by holding her hands about 15 cm apart precisely as she was saying "THIS far away from the light." The timing and meaning of her gesture were completely synchronized with her words. The addressee gave evidence later that she had understood their integrated meaning: When the addressee later asked, "You couldn't feel the heat or anything?" (line 28), she simultaneously held her hand about the same distance from the side of her own face.

Our essential tool for studying face-to-face dialogue is *microanalysis*, which is the detailed and reliable examination of observable communication sequences as they proceed moment by moment in the dialogue (Bavelas, McGee, Phillips, & Routledge, 2000). The time scale of microanalysis is small, focusing on events that often occur within a one- or two-second interval. This time scale is necessary because of the speed and precision of the participants' actions. For example, the entire Sleeper story (Table 1) was 77 seconds long, and most of the addressee's responses were around one second or less, yet each was precisely timed to appropriate points in the speaker's narrative. At this level of observation, an inductive approach (e.g. Bavelas, 1987) leads to many new insights.

Many of the interesting phenomena in face-to-face dialogue are intrinsically qualitative in nature. In the Sleeper story, the addressee made a wide variety of responses (e.g., "m-hm," smiling, nodding, looking concerned, or gesturing), none of which are easily suited to a parametric scale. However, quantification is useful in two ways. First, it provides a means to test inter-analyst reliability. We require high agreement between independent analysts, preferably around 90%. Second, observations at the nominal level of measurement easily convert into frequencies or rates per minute for statistical analysis. Such methods can straddle the quantitative/qualitative divide (cf. Bavelas, 1995, 2005). The rest of this article describes what can be learned with these methods.

HOW IS AN ADDRESSEE DIFFERENT FROM AN OVERHEARER?

Clark's (1996) *collaborative model of dialogue* has become an important framework for our program of research. Recall that Schober and Clark (1989) introduced the term *addressee* in order to distinguish between someone with whom the speaker is interacting (an addressee) and someone who can simply hear the speaker but who cannot interact with the speaker (an overhearer). They proposed that the interaction between a speaker and addressee is a special kind, because the two of them have not just the opportunity but the obligation to collaborate in order to ensure mutual understanding. Speaker and addressee participate together in a moment-by-moment process that Clark and Schaefer (1987) called *grounding*. With each of their utterances, they collaborate to ensure that the addressee has understood the speaker sufficiently for their current purposes (see also Clark, 1996, Ch. 8). Grounding is a more or less sequential process that is occurring constantly in all dialogues. A typical grounding sequence is the following:

1. the speaker presents information;
2. the addressee indicates that he or she has understood (or not); and
3. the speaker acknowledges (explicitly or implicitly) that the addressee has understood correctly (or not).

For example, in line 1 of the Sleeper story, the speaker presented information about her bed and headboard. The addressee conveyed that she understood well enough for the speaker to continue (she said, “mm,” “hm,” and she nodded). In line 2, the speaker implicitly accepted the addressee’s signal of understanding by saying new information that presupposed understanding of the bed and headboard (“and I got a light for Christmas”). This new information thus had two functions: It ended one grounding sequence while starting a new one. The addressee indicated her understanding of this new information with a slight nod, and again the speaker built on the grounded information (in line 3), and so on throughout the story. At any point, if either of them detected a misunderstanding, they could pause to repair it.

Thus, an important distinction between addressees and overhearers is that only an addressee can participate in grounding; an overhearer cannot. If grounding ensures mutual understanding, then addressees should have an advantage compared to overhearers. Specifically, an addressee should understand more of what a speaker says than an overhearer would. The next section describes Schober and Clark’s (1989) experimental test of whether the ability to participate in grounding really makes that difference.

Addressees’ Versus Overhearers’ Performance with the Same Information

To test their theory that addressees were a special kind of listener, Schober and Clark (1989) conducted a pair of experiments that demonstrated the difference between addressees and overhearers. In these experiments, the participants were separated by a partition so they could hear but not see each other. The speakers’ role was to give verbal directions about how to do a task correctly. Each speaker had two different listeners who were doing the same task using the speaker’s directions; there was an *addressee* (who could interact with the speaker) and an *overhearer* (who could not). The independent variable was whether the person receiving the speaker’s directions was able to interact with the speaker or not. The dependent variable was how accurately the addressees and overhearers were able to complete the task. If all that mattered were the quality of the speaker’s directions, then there would be, on average, no difference between how well the addressees versus the overhearers did the task because they had heard the same information. Alternatively, if the ability to collaborate through grounding were important, then the addressees should do better than the overhearers.

Both experiments showed that the addressees did significantly better on their task than the overhearers did. The addressees’ advantage could not have been due to the quality of their speaker’s directions or to simply listening carefully to their speaker, because these factors were the same for the overhearers as they had been for the addressees. The same directions that led the addressees to achieve

near-perfect scores did not work as well for the overhearers. Note that each over-hearer not only heard the same directions as their matched addressee, but they even heard the grounding sequences between the speaker and the addressee. However, they could not participate in their own grounding sequences with the speaker. Without the ability to ground, the overhearers had several specific disadvantages: They could not indicate to the speaker what they did or did not understand. They could neither identify nor clarify misunderstandings. They could not indicate to the speaker what information would be most helpful to them. They could not tell the speaker how they were acting on the directions they heard, so the speaker could not correct them if they had misunderstood. The addressees had all of these advantages.

Schober and Clark's (1989) findings were seminal, but several features of their research design could limit the ability to generalize their findings to other kinds of dialogues. First, the task they used was inherently collaborative: The speaker and addressee were working together toward a shared goal, which was to get the addressee to do the task correctly. Second, the addressee had information that the speaker needed but did not have, namely, what the addressee was doing in response to the speaker's directions. In contrast, the speaker and overhearer shared neither this goal nor this information. The speaker was not tailoring his or her directions for the overhearer's success, and the overhearer could not provide the speaker with any information about what he or she was doing with those directions. One could argue that the difference that Schober and Clark found between addressees and overhearers would only apply to collaborative tasks in which each person had some of the necessary information to complete that task. We doubted this limitation and tested the importance of grounding in a pair of experiments with an addressee who was in a much more passive role.

"Passive" Addressees

In the experiments by Bavelas, Coates, and Johnson (2000), the speakers told an addressee whom they had just met about a personal close call they had experienced in the past (e.g., the example of the burning pillow). This task differed from Schober and Clark's task in the two ways described above. First, there was no shared goal; the speaker was to tell a story, and the addressee was simply to listen. Second, because the two participants did not know each other, the addressee could not contribute any valid information to the story. Although these addressees were free to respond to the speaker, and nothing prevented them from participating in grounding, these addressees were as passive as we could make them.

We randomly assigned other addressees to an experimental condition that distracted them from the story. We gave these addressees a cognitive task that precluded them from grounding with the speaker. For example, in one experiment,

addressees in the experimental condition had to count the number of words the speaker said that began with the letter *t*. These distracted addressees were listening intently, but not to the meaning of the narrative. They had a different task, which was listening for and counting particular words. Distracted addressees were essentially overhearers.

The distracted addressees had two measurable effects on their speakers. First, the speakers with distracted addressees did not tell their stories well. Independent raters who were not aware of the experimental condition judged these stories to be significantly more poorly told than the stories told to addressees who were not distracted. Second, a subsequent microanalysis of the story endings revealed that, at what should have been the climax of their close call stories the speakers with distracted addressees faltered. These speakers were significantly more likely to become disfluent and to justify or overexplain their stories than were the speakers with addressees who were listening without being distracted (see example in Bavelas et al., 2000, p. 949).

These results extended Schober and Clark's (1989) findings beyond overtly collaborative tasks. Even though a narrative to a stranger might at first seem to be a monologue, it still requires that the stranger is an addressee who is paying attention to the meaning of the story. We propose that because the distracted addressees were disengaged from the narrative and were not contributing to it, they were reduced to being overhearers. Taken together, the two replicated experiments by Schober and Clark and the two by Bavelas et al. (2000) demonstrated the difference between addressees and overhearers. The rest of this article describes what we found the addressees were doing that had such a unique influence on the speaker.

WHAT ADDRESSEES DO

Generic and Specific Responses

Before conducting the two experiments just described (Bavelas et al., 2000), we watched many videos of storytelling similar to the Sleeper story. We sought to understand, through inductive observation, what the addressees were doing. We analyzed virtually every response the addressees made, although we did exclude two kinds of responses from the analysis. First, for obvious reasons, we excluded instances when the addressee became the speaker (e.g., lines 25 and 28 at the end of the Sleeper story). Second, we did not include smiles when they occurred alone, that is, when a smile was not combined with other actions. There are so many different kinds of smiles that distinguishing among them would be a project in itself (e.g., Ekman, 1985). Furthermore, the sheer number of smiles could overwhelm all other addressee responses.

Still, there was a wide array of other addressee responses to examine, and we soon noticed an interesting difference. Some responses were familiar and ubiquitous; for example, saying “m-hm,” “yeah,” or nodding are stereotypic addressee responses in English Canada. We called these *generic* responses because they were not uniquely tied to the meaning of any particular narrative or point in a narrative. In the Sleeper example in Table 1, the addressee made generic responses such as nodding and variations on “m-hm” in lines 1–6, 10–11, and 22.

In contrast, some addressee responses were highly specific to particular, precise points in the speaker’s narrative. For example, in the Sleeper story, after the speaker began to hint at what the danger would be (starting at line 7), the addressee’s face and gestures began to depict concern, alarm, or amusement. At the precise moment when the speaker described the light as “really strong” (line 8), the addressee bit her lip. When the speaker went on to say that it was a “hot light,” the addressee smiled and looked alarmed at the same time (line 9), conveying that she had begun to anticipate what the close call might be. We called these *specific* responses; they would definitely not fit just anywhere. None of the addressee responses described above would have made sense if they had occurred earlier or at other points in the story. For example, when the speaker was simply providing background information, such as “I got a light for Christmas” (line 2), imagine how bizarre it would have been for the addressee to bite her lip or look alarmed—there is nothing inherently alarming about getting a light for Christmas! Each specific addressee response appeared to be shaped for the particular point in the speaker’s narrative where it occurred. Other observers had also noticed both kinds of responses, in both experimental data (Krauss, Garlock, Bricker, & McMahon, 1977; Krauss & Weinheimer, 1966) and nonexperimental data (Goodwin, 1986; Yngve, 1970), although they gave them different names. For example, Goodwin (1986) had used the terms *continuers* and *assessments* for what we called generic and specific responses. He also noticed that they tracked the speaker’s overall narrative (Goodwin, 1986, pp. 214–215).

We developed detailed operational definitions for distinguishing between generic and specific responses and applied these to new experimental data. Using these definitions and working independently, two new analysts who were not aware of hypotheses or experimental conditions agreed on 95% of their distinctions between generic and specific addressee responses.

We found that distracted listeners made significantly fewer specific responses than listeners who were not distracted from the speakers’ narratives. Although distracted listeners’ rates of both generic and specific responses were significantly lower than in the control condition, the reduction was much greater for specific responses. Distracted listeners made virtually no specific responses. We proposed this is because specific responses must be tailored and timed to the precise content of the narrative at that moment, and listeners who were distracted from being able to follow the meaning of the narrative could not shape their

responses accordingly. Using data from the control condition in both experiments, we tested Goodwin's observation that generic and specific responses tracked the speaker's narrative. We confirmed that specific responses occurred significantly later in the narrative than generic responses did. The Sleeper example in Table 1 illustrates this pattern. The addressee made exclusively generic responses from the beginning of the story, when the speaker was setting the scene, until the first hint of what the close call would be (at "left it on"; line 7). After that point, virtually all of her responses were specific, providing vivid illustrations of both the danger and the humor in the climax of the story.

In the next sections, we will look even more closely at specific responses, examining the many resources that addressees have available for their active and essential role in a face-to-face dialogue.

Facial Displays

In the Sleeper story, the addressee's specific responses were often facial actions (lines 7–9, 12–13, 16–19, 23, and 27), such as raising her eyebrows, dropping her mouth open, or widening her eyes. A member of our research group, Nicole Chovil, conducted the first systematic analysis of facial actions in dialogue (Bavelas & Chovil, 1997; Chovil, 1989, 1991–1992). Chovil developed and conducted a reliable microanalysis of the facial actions of both the speaker and the addressee in dyads who were doing a variety of conversational tasks. (For reasons similar to those given in the previous section, she also excluded smiles.) Although the speakers' facial displays were more numerous and varied than those of the addressees, Chovil discovered that the addressees' faces still served several functions in the dialogue. Chovil found three common functions of addressees' facial displays: First, facial displays provided *back channel* information (Yngve, 1970) to the speaker, indicating whether and how well the addressee was understanding the speaker. For example, the addressee could look attentive, interested, puzzled, or "blank," and each of these provided different back-channel information to the speaker. (Brunner [1979] had shown that smiling, which we have not studied, can also be a positive sign of understanding.) Second, the addressee's face could convey the addressee's *personal reaction* to what the speaker was describing, for example, displaying disgust when the speaker described a food that the addressee disliked. Third, the addressee could also portray his or her understanding of the *speaker's reaction* to what the speaker was describing, for example, wincing when the speaker described being injured. The latter response, in which the addressee displayed a response that was appropriate to the speaker's situation rather than his or her own situation, was historically called *motor mimicry* (Bavelas, Black, Lemery, & Mullett, 1986; see also Bavelas, 2007).

The face is an excellent resource for the addressee for several reasons. First, participants in a dialogue apparently do not consider the addressee's facial

actions as interruptions or as taking over the turn. Indeed, the fact that addressees' facial contributions are usually simultaneous with the speaker's speech raises interesting questions about the utility and viability of the concept of "turn taking." For example, the speaker in the Sleeper story appears to have one long turn, but the addressee's facial displays constantly contribute details of her own reaction and provide the speaker with ongoing feedback. These simultaneous contributions do not fit the traditional definition of a turn. Second, this simultaneity with the speaker's speech can be quite precise because of the potential speed of facial actions. The addressee can start and stop a facial action in synchrony with a single phrase or even a single word. For example, when the speaker first described the light in her lamp, she said it was a "really strong, hot light" (lines 8 and 9). The addressee bit her lip in a grimace with "really strong." Then, as the speaker said "hot light" (less than one second later), she began to smile.

Notice that we are describing the information that a facial display can provide to the speaker rather than any emotion that it might express. Both Chovil (1989, 1991–1992) and Ekman (1997) distinguished between facial actions that express an individual's emotional state and those that display information to the other person in a dialogue (see also Bavelas & Chovil, 1997; Bavelas & Gerwing, 2007). To stress this distinction, we have adopted Kraut and Johnston's (1979) terminology: *Facial displays* function to convey information to another person in the dialogue, whereas *facial* (or *emotional*) *expressions* reveal (not necessarily deliberately) what the individual is feeling. One of the main differences between facial displays and facial expressions is how each is timed. As shown in the Sleeper example, a conversational facial display is synchronized with the speaker's speech, changing quickly and precisely in relation to what the speaker is saying at a particular moment. For example, the addressee looked alarmed at line 17 when the speaker said, "Then it starts to catch on fire." Then in lines 18 and 19, when the speaker started to make a joke of sleeping through the fire, the addressee immediately started smiling. An emotional expression would be timed to the individual's emotional state, which could be initiated by the speaker's words but would thereafter be independent of them. It is unlikely that emotional states change as frequently and rapidly as the facial actions of the participants in a dialogue. In this example, if the addressee was actually feeling real fear or alarm when the speaker said "catch on fire," then that emotion would presumably run its course. It would not change to a smile within two seconds, as it did.

Chovil (1989, 1991) also conducted an experiment that demonstrated the social nature of addressees' facial displays in dialogue. She randomly assigned the participants in her experiment to listen to a close-call story in one of four different conditions: (a) face to face, (b) on the telephone, (c) in the same room but separated by a partition, or (d) from a tape recording. Addressees made significantly more facial displays in person (i.e., when the speaker would see their display) than in the other three conditions, when the speaker would not see it.

Chovil also obtained independent rankings of how social each situation was. The face-to-face condition was considered the most social, followed by the telephone, the partition, and finally the tape recording, which was the least social condition. As the conditions became less social, the frequency of facial displays also declined. The frequency of addressees' facial displays was significantly inversely related to how social the situation was.

Research on facial displays in dialogue is still relatively rare (cf. Bavelas & Gerwing, 2007), perhaps submerged by interest in emotional expressions. Yet these displays are an important part of the addressee's repertoire. To ignore them is to ignore how closely and constantly the speaker and addressee are interacting.

Gaze Patterns

Once we started studying what addressees do in dialogue, we became curious about the timing of their responses (Bavelas, Coates, & Johnson, 2002). Why did they respond at one time rather than slightly before or after? An examination of the pattern in Table 1 suggests that the addressee usually responded when the speaker had just presented a new piece of information, that is, as part of a grounding sequence. However, there were exceptions, such as the periods of continuous nodding early in the story (lines 5 and 6). Also, the pieces of information that the speaker was presenting did not correspond precisely with grammatical sentences or clauses (e.g., lines 9 and 10), so syntax was not the main cue. We approached this question inductively, examining the precursors of all addressee responses in several of the close-call stories that formed our pilot study. The answer was surprising: The speaker and addressee seemed to be coordinating the addressee's responses using their gaze.

The gaze coordination we saw extended what other researchers had previously reported were typical gaze patterns in interaction. Kendon (1967), Cook (1977), Duncan and Fiske (1977), and others found that, at least in North America and Europe, the addressee tends to look fairly constantly at the speaker, but the speaker does not stare back. More often, the speaker looks slightly away from the addressee, with occasional glances at the addressee. Because the addressee is likely to be looking at the speaker, it is the speaker's glances that create mutual gaze (eye contact). In our pilot data, virtually all of the addressees' responses occurred during these brief periods of mutual gaze, which we called *gaze windows*.

Using the data of nine of the undistracted dyads from the *t*-counting experiment described above, we tested this hypothesis formally (Bavelas, Coates, & Johnson, 2002). First, independent analysts located all gaze windows. They then located the precise onset and offset times of each gaze window. Interanalyst agreement on both dependent variables was high: 100% and 84%–87%, respectively. Then we compared the timing of the generic and specific listener responses identified in the original experiment with the timing of these gaze

windows and calculated the probability that these would coincide by chance. A probability test was necessary because if either the gaze windows or the listener responses occurred frequently, then they could occur at the same time simply by chance. In fact, they occurred together far more often than chance would predict. The addressees' responses were significantly more likely to occur within a gaze window than outside it. When the speaker glanced at the addressee, the addressee responded with a specific or generic response. This result was true not only for the whole sample but also for each dyad; it was a strong, reliable finding. We did not indicate the gaze windows in Table 1, but the pattern was true for this story as well; all of the addressee's responses occurred in gaze windows.

The other finding was that, although the speaker initiated the addressee's response by looking at the addressee and opening the gaze window, the addressee had a role in closing it. Statistically, the addressees' responses did not occur just anywhere within the gaze window; they were significantly more likely to occur in the latter half. We interpreted this shift to mean that once the addressee had responded, the speaker glanced away again. Less often, the speaker continued looking at the addressee, and the addressee continued to respond until the speaker looked away. For example, in lines 5 and 6, the speaker continued to look at the addressee as she explained the reason for the lamp's placement, and the addressee nodded continuously.

In summary, our analysis confirmed that the speakers and addressees collaborated very closely, using their eye gaze to coordinate the precise placement of their respective actions. This does not mean the pattern is universal. It is likely that this pattern would change in different contexts, such as when dyads are seated side by side, when they both need to look at something they are talking about, or when their culture regulates gaze differently (e.g., labels it as rude). Any of these contexts would be an interesting new line of research.

Formulations

Recall that our analyses of addressees' responses so far have not included the occasions when the addressee begins to speak up. As Yngve (1970, p. 568) pointed out, distinctions between having the turn and not having the turn (i.e., between being the speaker or being the addressee) are fluid rather than fixed. For example, at the end of the Sleeper story (lines 25 and 28), the addressee changed the pattern of her behavior and interrupted the speaker, apparently to ask two questions. We propose that these are not simply questions but *formulations*, which Garfinkel and Sacks (1970, p. 350) defined as utterances that describe, explain, characterize, explicate, translate, or summarize what the speaker has said.

We propose that formulations are another way for addressees to play a part in the second step of the grounding sequence. At the end of the Sleeper story, the

addressee indicated that she understood what the speaker had said by stating her understanding explicitly. In line 23, the speaker had said, “And my mom smelled the smoke in her room.” Shortly afterward, the addressee interrupted to say, “So it—the room was starting to get full of smoke?” The speaker had not said explicitly that the burning pillow had filled the room with smoke. The addressee’s formulation explicated this dramatic point, verbally stating her understanding of what the speaker had only implied. The speaker completed the grounding sequence by confirming the addressee’s understanding (“Yeah” and laughing, line 26). She then continued with her story, “And I didn’t even wake up. And like it could have gotten on fire.” Again, the addressee formulated the implication of what the speaker was saying by stating it explicitly: “You couldn’t feel the heat or anything?”, and the speaker confirmed this inference. Although formulations often have the intonation pattern of a question, the addressee is not asking for *new* information; instead he or she is asking whether the formulation is correct. In these examples, it is interesting to note also that the addressee’s formulations added dramatic details to the speaker’s story—details that the speaker had not included. The addressee at these points became the conarrator of a story she had never heard.

Addressees formulate in many other situations as well. For example, when giving directions, speakers often pause for the addressee to summarize or rephrase each new piece of information before the speaker continues. When the speaker receives only generic responses, he or she is not receiving the same level of evidence about whether the addressee is following the new information.

One particular context in which formulation is important is in psychotherapy or conflict resolution. For example, Tomori (2004; Tomori & Bavelas, 2007) showed that experts in client-centered therapy used mostly formulations and few questions, while solution-focused experts used both. In these contexts, formulations are often referred to as “reflecting” or “paraphrasing” the client’s words and are therefore seen as a way of responding neutrally. However, Heritage and Watson (1979) pointed out that formulations are not neutral. They inevitably preserve some of the speaker’s information, delete other information, and thereby transform the original statement, none of which are neutral processes. One of our research group, Bruce Phillips (1998, 1999) demonstrated these nonneutral effects of formulations by comparing them in two different approaches to conflict resolution. He found that the formulations of the traditional, “neutral” mediator focused more on problems than on strengths and successes. Also, one client received more *open* formulations (which created an opportunity to confirm or correct the formulation), and the other client received more *closed* formulations (with no opportunity for grounding or correction). The solution-focused counselor’s formulations were the same for both clients. These formulations were almost all open and focused on strengths and successes rather than on problems.

Our recent project (e.g., Korman, Bavelas, & De Jong, 2011) compared the opening moments of demonstration sessions by experts in three different approaches to psychotherapy. The goal was to examine the extent to which therapists' formulations consisted of material that was preserved, altered, or added to what the client had said. They found that the formulations of two solution-focused therapists preserved significantly more of what the client had actually said than did the formulations of the cognitive-behavioral or motivational-interviewing experts. The latter group added significantly more of their own material, such as their interpretations of what the client had said.

SUMMARY

The studies described here support our proposal that the addressees' moment-by-moment contribution to dialogue, particularly their role in grounding cycles, makes them quite different from other kinds of listeners. The details of these differences become apparent through a combination of experimental design and microanalysis of dialogue. Starting with Schober and Clark's (1989) demonstration that addressees had an advantage over overhearers because of their ability to participate in grounding, our own program of research has gone on to reveal more details about their unique contribution to dialogue: When addressees were distracted from listening to the speaker's close-call story and therefore could not participate in grounding, they became mere overhearers. This change from addressee to overhearer not only influenced their own behavior, but it also impaired the speaker's ability to tell what should have been an exciting story. In contrast, addressees who were not distracted provided their speakers with both generic and specific responses. They made generic responses (e.g., "Mhm" or nodding) at the beginning of the story as the speaker was giving background information. Then addressees used specific responses at the more dramatic points of the story. These specific responses included vocalizations, facial displays, and even gestures, all of which conveyed the addressees' understanding of the meaning and importance of the dangerous and often simultaneously humorous specific moments in the story.

Other studies revealed more about the visible and audible resources available to addressees. Conversational facial displays are an often-overlooked way in which addressees contribute to the dialogue while remaining in a background role. These facial actions are a rapid and versatile resource for providing moment-by-moment feedback to the speaker without interrupting. It is evidence of the communicative function of these displays that they were much more likely to occur when the speaker would see them. Another resource for collaboration is eye gaze. In our data, when the speaker looked toward the addressee, it created a brief period of mutual gaze in which the addressee responded and then the

speaker looked away. The speaker and addressee used these gaze windows to coordinate the timing of the addressee's responses. Finally, addressees sometimes play their role in grounding by taking up the turn to formulate verbally what the speaker has been saying. Because formulations inevitably reshape the speaker's utterance, they can have a major influence on the conversation.

Our current or future interests are in three areas. First, the generic/specific dichotomy may not reflect the nuances of these responses as they actually occur in dialogue, where they may instead form a continuum. For example, prosody can make an ordinary "mhm" closer to a specific response, depending on whether addressee says it enthusiastically, doubtfully, or flatly. A related goal is to assess the quality of feedback that the addressee's responses provide the speaker. What specific factors contribute to the level of certainty that the speaker can have about what the addressee has understood?

A second area of interest is more theoretical. We have hinted above that many subtle addressee behaviors present a problem for any models that propose *turns* as the organizing structure of dialogue. The turn model privileges speakers and treats dialogue as alternating monologues. However, close observation of dialogue makes it clear that addressee responses are an inevitable and essential part of dialogues—but they are not turns. Yngve (1970) raised this issue when he first identified "back channels":

The distinction between having the turn or not is not the same as the traditional distinction between speaker and listener, for it is possible to speak out of turn, and it is even reasonably frequent that a conversationalist speaks out of turn. In fact, both the person who has the turn and [his or her] partner are simultaneously engaged in both speaking and listening. This is because of what I call the back channel, over which the person who has the turn receives short messages such as "yes" and "uh-huh" without relinquishing the turn. (p. 568)

Now that we know more about the role these responses play in grounding, we can conceive of an alternative to the turn model, which is to make *grounding sequences* the organizing structure of dialogue. This unit of analysis would include the collaborative actions of speaker and addressee together. Indeed, we propose that grounding sequences more accurately capture the orderly coproduction of information and development of mutual understanding in dialogue.

Third, any applied setting that includes face-to-face dialogue affords an opportunity to investigate the role of addressees. Psychotherapeutic, counseling, and mediation interviews are such settings, in which the therapist is typically the listener, but—in our view—never neutrally so. We have described work in our group on therapists' formulations; other possible directions could include how clients respond to various therapist formulations. Do they incorporate these transformations of what they have said? Training and professional development for

therapists benefits from a deeper understanding of the role of listening behaviors in dialogue.

Another setting for fruitful investigations of addressee behaviors is in health care consultations between patients and physicians or nurses. For example, because the addressee's responses indicate whether and how much he or she has understood, then these could be important indicators when obtaining informed consent. A physician should be concerned when a patient responds to important information with only generic responses or with only verbal claims of understanding (e.g., "Yes, I understand"); both fail to provide clear evidence of understanding. Future investigations with a truly dialogic focus might then go on to focus on identifying the physician behaviors that seem to elicit generic versus specific responses from patients.

A focus on addressee responses has been fundamental to the development of a quality assurance and training program for Norwegian medical emergency operators, aimed at enhancing their skills when receiving emergency calls from immigrants. In this project, we have learned that the operators' listening responses are vital for conveying unambiguous information to callers about what the operator has or has not understood so far. Similarly, operators who are alert to callers' listening responses can have a clearer picture of what the caller has or has not understood (e.g., when receiving instructions for CPR). In all of these situations, grounding is crucial, and the addressee's responses are as important as the speaker's contribution of information.

Taken together, these studies demonstrate that shifting one's focus to an addressee provides a new understanding of the listening process. Microanalysis of the actions of speaker and addressee in their moment-by-moment relation to each other reveals what is not visible by studying each participant individually, in isolation from the other. The collaboration between speakers and addressees is fascinating in its precision, and there is still a great deal to be discovered about the details.

AUTHOR BIOS

Janet Beavin Bavelas (PhD, Stanford, 1970) is currently an active Emeritus Professor of Psychology at the University of Victoria. Her long-term program of experimental research focuses on the unique features of face-to-face dialogue, with applications in psychotherapy and medical communication. She is coauthor of *Pragmatics of Human Communication* and *Equivocal Communication*, as well as over 70 journal articles or book chapters.

Jennifer Gerwing has a PhD in Psychology from University of Victoria, Canada (2008). Her research has focused primarily on the role of gestures in face-to-face

dialogue. She is currently working on a project at Oslo University Hospital, Ullevål to develop a training and education program to remediate communication difficulties in cross-cultural emergency medical telephone calls between Norwegian operators and foreign language callers.

REFERENCES

- Bavelas, J. B. (1987). Permitting creativity in science. In D. N. Jackson & P. Rushton (Eds.), *Scientific excellence: Origins and assessment* (pp. 307–327). Beverly Hills, CA: Sage.
- Bavelas, J. B. (1995). Quantitative vs. qualitative? In W. Leeds-Hurwitz (Ed.), *Social approaches to communication* (pp. 49–62). New York, NY: Guilford Press.
- Bavelas, J. B. (2005). The two solitudes: Reconciling social psychology and language and social interaction. In K. Fitch & R. Sanders (Eds.), *Handbook of language and social interaction* (pp. 179–200). Mahwah, NJ: Erlbaum.
- Bavelas, J. B. (2007). Face-to-face dialogue as a micro-social context. The example of motor mimicry. In S. D. Duncan, J. Cassell, & E. T. Levy (Eds.), *Gesture and the dynamic dimension of language. Essays in honor of David McNeill* (pp. 127–146). Amsterdam, the Netherlands: John Benjamins.
- 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. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 334–346). Cambridge, England: 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. (2006). Hand gestures and facial displays as part of language use in face-to-face dialogue. In V. Manusov & M. L. Patterson (Eds.), *Handbook of nonverbal communication* (pp. 97–115). Thousand Oaks, CA: Sage.
- Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79, 941–952.
- Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, 52, 566–580.
- Bavelas, J. B., & Gerwing, J. (2007). Conversational hand gestures and facial displays in face-to-face dialogue. In K. Fiedler (Ed.), *Social communication* (pp. 283–308). New York, NY: Psychology Press.
- Bavelas, J. B., Hutchinson, S., Kenwood, C., & Matheson, D. H. (1997). Using face-to-face dialogue as a standard for other communication systems. *Canadian Journal of Communication*, 22, 5–24.
- Bavelas, J. B., McGee, D., Phillips, B., & Routledge, R. (2000). Microanalysis of communication in psychotherapy. *Human Systems*, 11, 47–66.
- Bodie, G. D., Worthington, D., Imhof, M., & Cooper, L. O. (2008). What would a unified field of listening look like? A proposal linking past perspectives and future endeavors. *International Journal of Listening*, 22, 103–122.
- Brunner, L. J. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology*, 37, 728–734.
- Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chovil, N. (1989). *Communicative functions of facial displays in conversation* (Unpublished doctoral dissertation). University of Victoria, British Columbia, Canada.

- Chovil, N. (1991). Social determinants of facial displays. *Journal of Nonverbal Behavior*, *15*, 141–153.
- Chovil, N. (1991–1992). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, *25*, 163–194.
- Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.
- Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, *2*, 19–41.
- Cook, M. (1977). Gaze and mutual gaze in social encounters. *American Scientist*, *65*, 328–333.
- Duncan, S. J., & Fiske, D. W. (1977). *Face-to-face interaction*. Hillsdale, NJ: Erlbaum.
- Ekman, P. (1985). *Telling lies*. New York, NY: Berkeley Books.
- Ekman, P. (1997). Should we call it expression or communication? *European Journal of Social Sciences*, *10*, 333–359.
- Fillmore, C. (1981). Pragmatics and the description of discourse. In P. Cole (Ed.), *Radical pragmatics* (pp. 143–166). New York, NY: Academic Press.
- Garfinkel, H., & Sacks, H. (1970). On formal structure of practical actions. In J. C. McKinney & E. A. Tiryakian (Eds.), *Theoretical sociology* (pp. 337–366). New York, NY: Appleton-Century-Crofts.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Science*, *8*, 8–11.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York, NY: Academic Press.
- Goodwin, C. (1986). Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies*, *9*, 205–217.
- Heritage, J., & Watson, R. (1979). Formulations as conversational objects. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 123–162). New York, NY: Irvington.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63.
- Korman, H., Bavelas, J. B., & De Jong, P. (2011). *Microanalysis of formulations. Part II, comparing SFBT, CBT, and MI*. Manuscript submitted for publication.
- Krauss, R. M., Garlock, C. M., Bricker, P. D., & McMahon, L. E. (1977). The role of audible and visible back-channel responses in interpersonal communication. *Journal of Personality and Social Psychology*, *35*, 523–529.
- Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*, 343–346.
- Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, *37*, 1539–1553.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge, England: Cambridge University Press.
- Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London, England: Routledge.
- Phillips, B. (1998). *Formulation and reformulation in mediation and therapy* (Unpublished master's thesis). University of Victoria, British Columbia, Canada.
- Phillips, B. (1999). Reformulating dispute narratives through active listening. *Mediation Quarterly*, *17*, 161–180.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*, 211–232.
- Tomori, C. (2004). *A microanalysis of communication in psychotherapy: Lexical choice and therapy direction* (Unpublished honors thesis). University of Victoria, British Columbia, Canada.
- Tomori, C., & Bavelas, J. B. (2007). Using microanalysis of communication to compare solution-focused and client-centered therapies. *Journal of Family Psychotherapy*, *18*, 25–43.
- Yngve, V. H. (1970). On getting a word in edgewise. *Papers from the sixth regional meeting of the Chicago Linguistic Society* (pp. 567–578). Chicago, IL: Chicago Linguistic Society.