

HOW DO WE ACCOUNT FOR THE MINDFULNESS OF FACE-TO-FACE DIALOGUE?

JANET BEAVIN BAVELAS AND LINDA COATES

Langer, Blank, and Chanowitz (1978) distinguished between mindfulness, where "people attend to their world and derive behavioral strategies based on *current* incoming information" (p. 635) and mindlessness, where "new information actually is not being processed. Instead, prior scripts, written when similar information was once new, are stereotypically reenacted" (p. 636). Simple motor acts are often so overlearned that performance is automatic and mindless. The provocative proposal is that apparently complex social interaction may also be more often mindless than mindful.

We will not join sides in this debate. Rather, we will identify the puzzle that natural dialogue presents for both sides, that is, for all who accept a distinction between mindful and mindless behavior and, indeed, for proponents of any cognitive model of language. In our view, conversation occurs so quickly and yet so skillfully that it cannot be treated as either mindful or mindless as these terms are currently used.

Conversation is the most common medium of social life. Even in technological societies—in spite of letters, telephones, and E-mail—most of our communication is dialogue in face-to-face interaction. Conversation is arguably "the fundamental site of language use" (Clark & Wilkes-Gibbs, 1986, p. 1). There is every reason, therefore, to expect conversation to be so highly practiced as to become mindless, that is, overlearned, stereotyped, and automatic, with minimal processing of new information. Certainly, everyone can easily recall conversations in which he or she apparently responded automatically, while thinking of something else. Still, every day brings dozens of new conversations. Our thesis is that, while many of these conversations are trivial in content at a macro level, they are conducted with great precision at the micro level.

EXAMPLES

Dialogue is not like written language. There are two key, additional features: (a) the necessity to respond to and coordinate with another person, on-line in real time and (b) the use and integration of precise nonverbal as well as verbal elements (cf. Bavelas, 1990; Bavelas & Chovil, 1992). Let us enter the micro world of conversation with some typical examples captured on videotape, in

Janet Beavin Bavelas is professor of psychology at the University of Victoria, where her research group conducts experiments on all aspects of dialogue. This research overlaps communication, social psychology, and discourse analysis, and she publishes in all three disciplines. She is the author of three books, including Pragmatics of Human Communication (with Watzlawick and Jackson, 1967) and Equivocal Communication (with Black, Chovil, and Mullett, 1990) and over 30 articles or chapters. She has received about a dozen research grants from the Social Sciences and Humanities Research Council of Canada. Linda Coates is a doctoral student in psychology at the University of Victoria. Her master's thesis on irony in dialogue was the top student paper in Language and Social Interaction at the International Communication Association 1992 conference. She collaborates on the Victoria Group's gesture and dialogue studies and is co-authoring a manuscript on that work, "Gestures specialized for dialogue." In 1991, the Social Sciences and Humanities Research Council awarded her a fellowship for her doctoral studies.

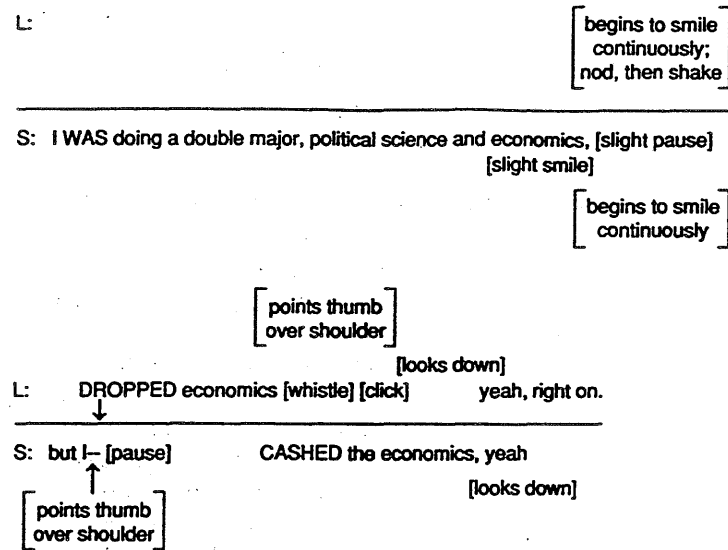


FIGURE 1.

A 6-SECOND EXCERPT OF CONVERSATION

order to understand how different spontaneous dialogue is from a Shakespear-ean playscript.

Example 1

In Figure 1, we have transcribed a very brief excerpt from a "getting acquainted" conversation between strangers, filmed in our video lab (Lawrie, 1988). The two men had been talking just over a minute when one asked the other about his program of studies. The speaker (S) replied that he had been doing a double major, in political science and economics. In the key segment, he said "but I . . ." and completed his sentence with a gesture in which he (metaphorically; cf. McNeill & Levy, 1982) tossed away his economics courses. Simultaneously with this gesture, the listener (L) completed the speaker's sentence with a verbal predicate, "dropped economics." Then they rapidly switched media: L re-enacted S's tossing gesture (accompanied by appropriate sound effects), while S confirmed L's statement by completing his original sentence verbally ("CASHED the economics, yeah"). His partner then acknowledged not only their mutual following but their implicit agreement about dropping economics ("yeah, right on").

The first question this excerpt raises is, How was L able to complete S's sentence so quickly and accurately? (Indeed, L could have let S continue without participating.) In the preceding minute, S had mentioned, along with a great deal of other information, that he was majoring in political science. So, combined with the stress on "WAS," L could know that S had chosen political science over economics. However, at the pause after "and economics," any number of continuations were still possible; for example,

but I CHOSE political science because . . .

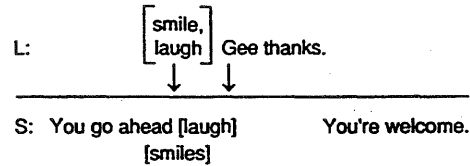


FIGURE 2.

A 3-SECOND EXCERPT OF CONVERSATION

However, the smile that accompanied “and economics” hinted that S’s emphasis was on the negative aspects of economics. Their immediate mutual smiling suggested that they agreed about this, if only because there is nothing else to smile at in S’s statement. S then continued his sentence with “but I” and a gestural predicate. Within .17 second of “but I,” L provided the verbal predicate.

The next unanswered question is how they were then able to switch verbal and nonverbal roles so quickly, with L using S’s gesture (with an added whistle and click) while S switched to verbal expression, along with an acknowledgment that his point had already been made by L (“yeah”). L immediately conveyed his agreement, and they both wound down the exchange by looking away from each other. The entire episode was 5.88 seconds long.

Example 2

In Figure 2, two women, also strangers, had been given instructions for their discussion task, which required each of them to give an opinion (Coates, 1991). As the door closed behind the experimenter, the first speaker (S) offered L the turn with a laugh. We understand this offer to be sarcastic or ironic—an inversion in which S offers what she is sure L does not want, namely, the “privilege” of speaking first. At that point, L had only a peripheral view of S’s face because both were looking down at the instruction cards, yet she smiled and laughed at S’s offer before S herself had laughed. Moreover, within .30 second, she initiated an appropriate (ironic) continuation by thanking S for the offer, to which S immediately responded in kind.

Thus, within 3.5 seconds, two complete strangers enacted a standard offer/thanks/welcome sequence *with completely inverted meaning*: They understood that neither would want to go first; S offered a humorous inversion based on this understanding; L responded both with appreciation of the joke and a continuation of its ironic form; and S completed the sequence with a formulaic but ironic “You’re welcome.” As Coates (1991) has noted, inverting meaning is risky for strangers; a literal interpretation could be seen as stupid or hostile. Both participants must signal their initial understanding (e.g., no one wants to go first), and the inverted play on this understanding must be registered by both participants. Coates’ data showed that strangers can manage to signal and enact inversion very quickly; moreover, they also mutually signal the return to normal meaning, closing the inversion frame.

CHARACTERISTICS OF DIALOGUE

The precision and rapidity of responses in the above examples is commonplace. In our experience, the most inconsequential conversations contain numer-

ous responses with three characteristics that are simultaneously impressive and theoretically troublesome.

First, these responses are *improvised*, because the participants cannot know exactly what the other person will present them with. At the precise level of meaning and syntactic structure that the participants are generating and responding to, no formula exists. Each person must provide a spontaneous, creative response. Second, the participants produce *precise*, perfectly fitted responses, often so exactly tuned to both the immediate and larger context that we as outsiders require microanalysis and repeated viewing to appreciate the subtle rightness of a particular reply. Third, these responses are *immediate*; all of this is happening now, in real time, often below simple reaction time. There is no time to stop and think in conversation. To be socially skillful, it is both essential and usual to respond appropriately in micro time.

So, accepting the definitions set out at the beginning, one would have to conclude that conversation is mindful: The interlocutors respond precisely to current incoming information and not to stereotypic scripts. There is no previous similar information, and their responses are not stereotypic. Full attention is demanded.

COGNITIVE MODELS

The problem with treating communicative behavior in conversation as mindful is that existing models of mindful behavior do not deal well with such precisely and rapidly improvised behavior. Here, we will briefly examine two standard models.

Bower and Cirilo (1985) put together a cognitive model for *language comprehension* from existing literature. (While these and other authors focused mainly on written text, they treated spoken language as equivalent.) The model traces the hypothesized step-by-step comprehension of a sentence: The receiver comprehends a word by first detecting the word as a stimulus event and then laying down the perceived pattern as an icon in a sensory buffer. Then a set of feature detectors begin "extracting significant features from the stimulus" (p. 74). These features (i.e., lines or sound) form an initial description of the stimulus word, which is classified according to the best-fitting prototype from long-term memory. The accuracy of this match will depend on the quality of sensory input, number of possible alternative prototypes, and the context in which the word occurred (i.e., a word will be identified more easily if it is expected or probable within the context).

The cognitive processes involved in comprehending a sentence are similar. After a sentence has been detected, short-term memory organizes the sentence into its surface constituents (e.g., word, noun phrase, verb phrase), and the functions of these constituents are identified. Then the underlying propositions of a sentence are extracted and searched. The "parser" steadily builds a representation of propositions embedded in the sentence that is in short-term memory, resulting in semantic interpretation of the sentence, which is then deposited into working memory.

So far the listener/receiver has only comprehended a sentence. He or she has yet to respond in dialogue. There are few models of *language production*, but we can draw on Clark and Clark's (1977) synthesis. They described how speech

production would be planned and executed, in steps, through discourse plans (identifying the global context of the discourse); sentence plans (selection of speech acts, such as literal or ironic); constituent plans (choosing words and idioms); an articulatory program (which holds the planned constituents in buffer); and finally, articulation (execution of the plan), which includes the physical movement and timing of the speech apparatus. The schematic description of these five stages takes over 30 pages for planning and another 30 pages for execution (Clark & Clark, 1977, pp. 224–257, 260–291).

It is obvious that the above models, involving many steps for comprehension and more for production, cannot account for the rapidity of response in the examples given. Thus, *neither* mindful nor mindless models can account for dialogue. The problem is that dialogue, unlike reading and writing, does not permit a leisurely pace, much less re-reading and editing. It is driven by the real-time pace of interaction.

Clearly, reductionist models with their stages and substeps take too much time; a radical break from traditional models is required. It must be that speech acts are comprehended globally and immediately, with incomplete information—just as an Impressionist painting is comprehended before its fragmentary details have been assessed. Similarly, the response flows quickly and globally—just as a skillful squash player does not sequentially analyze the positions of self, ball, and opponent and then decide how to strike. Our own research interest is in documenting the intricacies of conversation and theorizing at the systemic level. Here, we have paused in that endeavour to point out the puzzle for colleagues who seek its cognitive mechanisms and who, it seems to us, have not fully appreciated the problem. The answer to the question in our title is: At present, we cannot.

REFERENCES

- Bavelas, J.B. (1990). Nonverbal and social aspects of discourse in face-to-face interaction. *Text*, 10, 5–8.
- Bavelas, J.B., & Chovil, N. (1992). *An integrated message model of language in face-to-face dialogue*. Manuscript in preparation.
- Bower, G.H., & Cirilo, R.K. (1985). Cognitive psychology and text processing. In T.A. Van Dijk (Ed.), *Handbook of discourse analysis* (Vol. 1, pp. 71–105). Orlando: Academic Press.
- Clark, H.H., & Clark, E.V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt, Brace, Jovanovich.
- Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Coates, L.J. (1991). *A collaborative theory of inversion: Irony in dialogue*. Unpublished master's thesis, University of Victoria, Victoria, B.C.
- Langer, E., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action. *Journal of Personality and Social Psychology*, 36, 635–642.
- Lawrie, D.A. (1988). *Discourse-oriented hand gestures with a proposal for a new category: The semantic emphasize*. Unpublished honour's thesis, Department of Psychology, University of Victoria, Victoria, B.C.
- McNeill, D., & Levy, E. (1982). Conceptual representations in language activity and gesture. In R.J. Jarvella & W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics* (pp. 271–295). Chichester: Wiley.