

## Conversation and Dialogue

by Susan E. Brennan

### Definition and Overview

Conversation is a joint activity in which two or more participants use linguistic forms and nonverbal signals to communicate interactively. Dialogues are conversations between two participants (although the terms *dialogue* and *conversation* are often used interchangeably). Face-to-face conversation is universal—engaged in by all human cultures, and providing an interactive context in which children learn their native languages. Conversation may also be mediated, such as when electronic technology is used for speech or text. This entry takes an interdisciplinary approach to defining conversation and its key characteristics.

A conversation is *not* simply a sequence of messages expressed as speaking turns, produced by speakers, and received and decoded by addressees. Conversations are structured into *adjacency pairs*, with first and second parts produced by different speakers as in this example:

Juliet: Art thou not Romeo, and a Montague?  
Romeo: Neither, fair maid, if either thee dislike.  
(*Shakespeare's Romeo and Juliet, Act 2, Scene 2*)

By itself, Juliet's utterance does not yet count as a question; she cannot be sure that man lurking beneath her balcony has heard and understood her until she has the evidence from his response. And Romeo's answer ends up transforming what might have been left to stand as a yes/no question to something highly relevant to their situation, implicating both its interpersonal and familial risks. Utterances in conversation (whether spoken, typed, or produced manually using sign language) are contingent upon one another, such that interpretation depends very much on context.

Real conversation is spontaneous rather than scripted in advance; it is shaped by the coordinated behavior of speakers and addressees. For these reasons, it differs considerably from edited texts. Utterances unfold over time; they are both planned and interpreted incrementally. Consider this excerpt from a telephone conversation between two British acquaintances (*adapted from Svartvik and Quirk's London-Lund corpus*):

Brad: thanks for ringing  
Amanda: right - bye  
Brad: bye bye <pause>  
see you next week  
Amanda: see you  
Brad: see you soon  
Amanda: m  
(both): <laugh>  
Amanda: you're hopeless  
Brad: sorry <pause>  
Amanda: you're hopeless -  
Brad: well <pause> no more than you  
Amanda: <laughs>  
no more than usual either  
Brad: no <pause> more <pause> no more  
than you I said  
not usual  
Amanda: oh I know <pause>  
I said no more than usual  
Brad: ah <pause> I'm sorry  
Amanda: have you got a new job yet  
<etc.>

Compared to Juliet and Romeo's dialogue, this excerpt seems rather disfluent; but it is actually the artifact of an orderly coordination process. Conversations do not begin and end abruptly, but with opening and closing routines with which participants establish that they are willing to begin

interacting, or ready to say goodbye. At first it seems as if Brad is winding down the conversation by initiating a pre-closing routine with ‘thanks for ringing’. This is followed by Amanda’s too-abrupt ‘bye’. Apparently Brad recognizes that Amanda is not serious about hanging up, and so he stays on the line even after responding with ‘bye bye’. They proceed to draw things out over the next few turns, culminating in joint laughter that displays mutual awareness of this joint pretense. Amanda’s attempt to chide Brad (‘you’re hopeless’) fails, apparently because Brad doesn’t hear her (or perhaps can’t believe what he is hearing). This leads to a repair sequence, during which Brad requests repetition, and Amanda complies. Next, Brad (who by now may have figured out that he is being teased) chides her back with ‘no more than *you*’, and Amanda attempts further wordplay (‘no more than *usual...*’). But Brad interprets this as her mishearing him. He attempts a hesitant (and unnecessary) repair, to which Amanda disclaims ‘oh I know’ and then recaps her failed pun. Brad apologizes awkwardly (after an ‘ah’ displaying his belated recognition of her little joke). Then the intrepid Amanda starts up the conversation again. Although only the participants know what they’re thinking moment by moment, even an overhearer can recognize from this transcript that flirting is going on.

## **Empirical Studies of Conversation**

### **Conversation Analysis**

This step-by-step description of how conversation unfolds is inspired by the sociolinguistic approach known as conversation analysis. This approach has been popularized by analysts such as Emmanuel Schegloff, Gail Jefferson, Harvey Sacks, and Charles Goodwin. Conversation analysts provide detailed accounts of the highly coordinated activities that ensue during naturally occurring, everyday conversations (the interpretation of Brad and Amanda’s call

provided here does not begin to do justice to the richness of such accounts). Conversation analysis resists quantitative analysis (sometimes militantly so) but has provided many valuable qualitative insights about the structure and coordination of conversation.

### **The Experimental Tradition**

A contrasting but complementary empirical approach to conversation is the experimental tradition pioneered by Herbert H. Clark and colleagues. Typically, language use by pairs of naive subjects is recorded in a laboratory as they interact in a task assigned by the experimenter. Tasks often involve *referential communication* in which, in order to do the task, two people must come to believe that they are talking about the same thing; the experimenter observes their task-related actions as evidence of what is being referred to. Consider these excerpts from a referential communication experiment in which participants A and B can hear but not see each other. The task requires them to match duplicates of 12 abstract geometric objects; they do this for the first time in Trial 1, and then match the same objects again in Trials 2 and 3 (with many other objects discussed between these excerpted trials).

*Trial 1:*

A: ah boy this one ah boy alright it looks kinda like,  
on the right top there's a square that looks diagonal

B: uh huh

A: and you have sort of another like rectangle shape,  
the- like a triangle, angled, and on the bottom it's  
ah I don't know what that is, glass shaped

B: alright I think I got it

A: it's almost like a person kind of in a weird way

B: yeah like like a monk praying or something

A: right yeah good great

B: alright I got it

*Trial 2:*

B: 9 is that monk praying

A: yup

*Trial 3:*

A: number 4 is the monk

B: ok

The *common ground* that accumulates as a conversation unfolds (whether in a laboratory experiment or an everyday conversation) enables referring to become more efficient over time.

### *Entrainment in Conversation*

In the previous example, convergence on ‘the monk’ marks that these partners believe they are talking about the same thing and are taking a similar perspective on it. When partners reuse the same forms, this is known as *entrainment*; another pair discussing the same object in a different conversation is likely to come up with quite a different perspective. For example, 13 pairs in one experiment entrained on 13 distinct expressions for the following object:

"a bat"

"the candle"

"the anchor"

"the rocket ship"

"the Olympic torch"

"the Canada symbol"

"the symmetrical one"

"shapes on top of shapes"

"the one with all the shapes"

"the bird diving straight down"

"the airplane flying straight down"

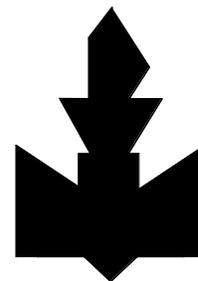
"the angel upside down with sleeves"

"the man jumping in the air with bell bottoms on"

[*Conversation and Dialogue* Figure 1 about here]

Figure 1. Perspectives Vary Across Conversations

Source: Unpublished data, Stellmann & Brennan, 1993



By systematically eliciting evidence such as this, the experimental tradition has established that there is less variability within a conversation than between conversations.

### *Grounding in Conversation*

The previous example illustrates another key process underlying conversation: *grounding*. Conversation involves more than the exchange of messages in sequences of turns; even utterances about everyday objects can be misunderstood. Because individuals cannot read one another's minds, they seek and provide evidence from their partners in order to conclude that they understand one another. A single utterance (be it Juliet's question, Amanda's attempt at a pun, or A's initial description in Trial 1) does not by itself constitute a contribution to a conversation; it must be grounded with evidence of how the addressee has understood it or taken it up before it contributes to common ground. Such evidence can be verbal and linguistic, in the form of a relevant reply, or nonverbal or nonlinguistic, as in a nod, a puzzled look, or an 'uh-huh'. During face-to-face conversation, visual and nonlinguistic signals known as *backchannels* are often produced and processed simultaneously with verbal turns. This means that conversing is a massively parallel activity: Speakers and addressees simultaneously plan their own utterances, process what their partner has said, and monitor one another's reactions. Partners share the responsibility for ensuring that they understand each other well enough for current purposes. They flexibly shift the initiative in the grounding process, providing each other with evidence about their own understanding, or seeking such evidence from their partner (i.e., with a clarification question). In Clark's terminology, linguistically encoded "Track 1" messages pertain to the overt purposes of the conversation, while meta-linguistic "Track 2" signals (e.g., backchannels) provide evidence for grounding.

## *Conversations With and Through Computers*

Human-computer interaction has sometimes been considered to be a kind of conversation, albeit a constrained and task-oriented one; *spoken dialogue systems* are computational agents that can interact with people using speech in limited domains such as information retrieval and travel planning. In order to minimize reasoning or speech recognition errors, such systems often seize most or all of the initiative for directing the dialogue.

When people communicate electronically, be it by telephone, email, or texting, communication media enable different Track 2 signals for achieving and coordinating shared meanings. How partners coordinate the grounding process differs depending on whether they are co-present in time and space, as well as on other characteristics of the medium. For example, speaking is usually less effortful than typing, so spoken conversations tend to be wordier than typed conversations. Because typing to a small screen using the thumbs is so cumbersome, texted messages are usually shorter than email messages. People on the phone cannot see one another, so they produce more verbal backchannels than when speaking face-to-face. Speech is ephemeral, whereas text need not be; email utterances can be easily reviewed, edited, and quoted. Email does not require senders and recipients to be co-present in time and space, so in that medium conversation unfolds over an extended timescale. Email often lands in a mailbox interleaved with other conversations or “threads”, and so email programs facilitate keeping track of conversational context by quoting material from previous messages. Through their distinctive affordances for grounding and the variable costs of using them, communication media play a substantial role in shaping dialogue.

## **Interdisciplinary Approaches**

Task-oriented experiments do not address all of the important joint actions that people do in conversation (such as flirting), but experimental findings have the virtue of being summarizable, replicable, and generalizable. Some researchers have sought to uncover the principles underlying conversation by combining the strengths of both descriptive and hypothesis-testing empirical traditions, via experiments that test predictions developed from sociolinguistic insights. Converging results come from linguistic and computational linguistic analyses of conversational corpora. Finally, innovative headmounted eyetracking techniques (e.g., pioneered by Michael Tanenhaus, Daniel Richardson, and colleagues) can now be used to inobtrusively measure speakers' and addressees' eyegaze as they spontaneously plan, articulate, interpret, and coordinate their utterances moment-by-moment, *on-line*.

- Susan E. Brennan, Stony Brook University

(See also: Production of Language; Gesture and Language Processing; Multimodal Conversation Systems; Joint or Collective Intention; Discourse Processing, Models of)

### **Further Readings:**

Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J.

Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use:*

*Bridging the language-as-product and language-action traditions* (pp. 95-129).

Cambridge, MA: MIT Press.

- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 6, 1482-1493.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: APA.
- Clark, H. H. (1994). Managing problems in speaking. *Speech Communication*, 15, 243-250.
- Clark, H. H. & Schaefer, E. R. (1989). Contributing to discourse. *Cognitive Science*, 13, 259-294.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.
- Goodwin C. (1979). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Richardson, D. C., Dale, R., & Kirkham, N. (2007). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychological Science*, 18, 407-413.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, 696-735.
- Svartvik, J. & Quirk, R. (1980). *A corpus of English conversation*. Lund, Sweden: Gleerup.
- Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.