ɔter has emphasized, both person perception and impression man-
ediated by language use. The information we receive about others
from our interactions with them, just as the stage for our attempts
management are those same interactions. And of course, these
e largely verbal interactions. So, how we talk, what we say, and
are an important source of information for the person-perception
; it is an important resource used for managing the impressions oth-
;. Clearly, the study of language contributes substantially to our
of these phenomena.

89-120

4

Conversational Structure

It is a rare conversation that consists of a single utterance. Of course, a single utterance is not a conversation at all; a minimum requirement for a conversation is that it is comprised of at least two turns of talk. Consequently, our utterances always occur, not in isolation, but in the context of the utterances of other people. Because of this, one's talk is not a result of a single individual—one can't just say whatever one wants whenever one wants to say it—but rather the result of two or more people jointly engaging in talk. What we do as we converse is constrained by what is done by our interlocutors. And of course their contributions are constrained by us.

An amazing and yet obvious feature of conversations is how highly structured they are. People take turns speaking, their utterances generally address the same topic, misunderstandings are often handled quickly and easily, and so on. But equally interesting is the fact that prior to its occurrence, the course of a conversation is completely unpredictable; it is difficult to specify in advance how a conversation will turn out. Structure and order are apparent when viewing a completed conversation, but not predictable at the beginning.

What, then, is the nature of conversational orderliness and structure? How is this accomplished? What are the mechanisms that allow it to be accomplished? Speech act theory and politeness theory really do not help much in this regard.

Now, it is possible to extend these approaches in various ways to account for the sequential properties of talk. Some attempts have been made in this regard, and they will be briefly discussed. But there are other approaches—perspectives that eschew concepts of intention and politeness—that provide some insight into conversational structure. Foremost here is a subdiscipline termed conversation analysis, and most of this chapter will be devoted to research within that tradition.

Conversational structure has important implications for many of the phenomena discussed in previous chapters. Speech acts, for example, are highly contextualized; illocutionary force does not exist independently of the conversational context within which the utterance occurs. Similarly, politeness is often manifested over a series of moves, and face-management processes might be missed if the focus is on single utterances. Of course, the interpersonal implications of language use—its role in person perception and especially impression management—are based on the sequential properties of talk, or how people accommodate to their conversational partner, or negotiate topic changes in order to disclose something positive or avoid conveying face-threatening information. The importance of conversational structure was implicit in the previous chapters; its importance is explicitly recognized in this chapter.

In the first part of this chapter, I provide a brief overview of the genesis and methodological perspective of conversation analysis. This is followed by an extended discussion of the role played by adjacency pairs (e.g., questions and answers) in conversations. Adjacency pairs have a number of interesting properties, and they are capable of accounting for much conversational structure. Then, the systems used for managing turn-taking and repair are considered. In the final section, I discuss the manner in which conversational coherence is achieved, and this will include a brief consideration of some alternatives to conversation analysis.

CONVERSATION ANALYSIS

Conversation analysis is a unique approach to the study of language use. It was developed in the 1960s by the sociologist Harvey Sacks and his colleagues Emanuel Schegloff and Gail Jefferson. Many of Sacks's original ideas were articulated in his lectures and circulated only in mimeo form for a number of years. It is only relatively recently that his lectures have been fully transcribed and published (Sacks, 1992).

Conversation analysis represents a sociological (rather than linguistic) orientation to language; it seeks conversational regularity in terms of the social order rather than in linguistic acts. But as a sociological approach it was quite unlike—and in fact was directly opposed to—most other sociological traditions. Conversation analysis has its roots in ethnomethodology, an offbeat branch of sociology developed by Harold Garfinkel (1967). Ethnomethodology was largely a reaction

(both substantive and methodological) to mainstream functionalism. In a functionalist view (e.g., Parsons, 1937), society is an entity that is external to the individual; individuals become socialized and become part of the social order, via the internalization of societal norms. In contrast, Garfinkel adopted a primarily phenomonological perspective (derived from Schutz, 1970) and argued that the social order, rather than being independent of and external to individuals, is continuously created by people as they use language and make sense of their behavior and that of others. Garfinkel thus advocated a shift in emphasis from macrosociological forces to microsociological phenomena, specifically, the commonsense knowledge and reasoning people use as a means of creating a social order (cf. Goffman, 1967).

Ethnomethodology involves a unique methodological stance. A functionalist methodology is primarily quantitative; categories and coding schemes are developed by researchers and then applied to the social phenomenon of interest. For Garfinkel, such coding schemes are not to be trusted because they are based on the interpretive (sense-making) strategies of the researchers who develop them. For Garfinkel, an objective social science was not possible. What is possible, however, is the exploration of the means by which people make sense of their world. It is for that purpose that Garfinkel developed his famous breaching experiments. In these experiments, student experimenters were instructed to enter into interactions with others and violate basic, taken-for-granted routines (e.g., greeting exchanges) and observe the resulting breakdown, which inevitably occurred. These experiments were presumed to illuminate the taken-for-granted knowledge that allows people to engage in social interaction.

Conversation analysis, like ethnomethodology, focuses on the common, everyday competencies that make social interaction possible. Like ethnomethodology, there is a putative rejection of a priori categories and theorizing; the emphasis is on *participants'* (not researchers') sense making. Unlike Garfinkel, however, conversation analysts assume that it is possible to develop a science of social life. And in fact, conversation analysis represents a rigorously empirical approach to social interaction, albeit one that is essentially inductive. Rather than examining rule violations as a means of documenting commonsense understandings (Garfinkel's breaching experiments), conversation analysis seeks to uncover these understandings as they are *revealed* in normal everyday conversational exchanges. A person's understanding of another's conversation turn is assumed to be revealed in the subsequent turn that she produces (see also Clark, 1996a). In this way, an analyst works only from what is available to the interactants themselves—their talk.

The general strategy in conversation analysis is to examine in great detail actual verbal interactions. And they must be actual interactions; contrived examples or speech generated in the laboratory are not acceptable. Actual interactions are audiotaped, with much of the early research relying on recorded telephone conversations. These recordings are then transcribed, keeping as much detail (e.g.,

overlap talking, pauses, breathiness, word and syllable stress) in the transcription as possible. A fairly standardized transcript system, developed primarily by Gail Jefferson, exists for this purpose. Transcripts are, of course, not entirely neutral, and in the case of conversation analysis they reflect the central concern these researchers have with turn taking and speech delivery (particularly insofar as it is relevant for turn taking).

Working from both the transcript and the actual recording, the analyst then proceeds to examine the interaction for regularities, in particular, patterns in the sequential organization of talk. Again, conversation analysts claim that no a priori concepts or theories are applied to the data at the outset. Instead, the approach is to articulate a category only if the interactants themselves in some way display an orientation to that category. It is thus an ethnomethodological stance; the focus is on how interactants themselves make sense of their activity. If the participants do not display an orientation to a particular conversational occurrence, then the analyst is presumably in no position to claim any particular status for that occurrence. After analysts have identified an interesting conversational phenomenon, they will then collect a number of instances of that phenomenon and ascertain whether these other instances can be described with their account. This, then, is a purely inductive strategy with the aim of providing a formal description of a large set of data.

Despite the emphasis on the accomplishment of interactional work through talk, it is important to note that absolutely no reference is ever made to the internal states (e.g., goals, expectancies, motives) of the interactants. It is an approach that takes talk, and only talk, as the object of study. Thus, many of the internal concepts that lie at the heart of speech act theory (e.g., illocutionary force, intention) and politeness theory (e.g., positive face, negative face) are not relevant, even though, as we will see, many of the phenomena uncovered in conversation analysis can be interpreted within these frameworks. Nor do traditional sociolinguistic variables such as power relations, gender, and formality have an a priori relevance; such variables are relevant only if the interactants themselves display some orientation to them.

The primary aim of conversation analysis is to bring to light the structural properties of talk. Hence, there is an emphasis on conversational context and sequence. The organizational features uncovered with this approach are viewed as resources that interactants use to manage talk, and these resources are available to anyone. In this way, conversational interaction is viewed as a system—a system existing apart from any specific individual. The structural regularities that exist are a means by which the conversational system is maintained. Hence, the focus is on the "system" properties of talk and not on any of its interpersonal or ritual (e.g., Goffman, 1976) properties. In conversation analysis, there really is no meaningful distinction between the conversational system and conversational ritual; the system *is* the ritual.

Research in conversation analysis has, over the past 30 years, documented the existence of a number of sequential properties of talk. This is why conversation

analysis is central to a consideration of the structure of conversations; conversation analysis has a near monopoly on investigating this feature of talk. Some of the major findings are described next, beginning with the smallest unit—adjacency pairs—and working outward to the overall structure of conversations.

ADJACENCY PAIRS

People can't say just anything during a conversational exchange; their utterances are constrained in various ways by the context, most notably the utterances of other interactants. If Bob makes a request of Andy, then Andy's subsequent turn at talk is constrained. He cannot, for example, perform the act of greeting Bob. Of course, he might perform a greeting, but if he does so he displays, through his greeting, his orientation to the preceding turn; indicating a lack of understanding, a refusal to comply with the request, a joke, and so on. In other words, the significance of Andy's action, as part of a jointly produced interaction, is constrained by Bob's prior action.

The smallest structural unit exhibiting this quality is termed an adjacency pair [Goffman (1976) used the term *couplet* for a similar unit]. Adjacency pairs figured prominently in early conversation analytic writings (Schegloff & Sacks, 1973; Schegloff, 1968) and continue to receive attention due to their fundamental importance in structuring talk. According to Schegloff and Sacks (1973), adjacency pairs have the following general features.

1. They consist of two utterances, a first-pair part and a second-pair part.
2. The two utterances are spoken by different speakers.
3. The utterances are paired so that a first-pair part must precede a second-pair part.
4. The first-pair part constrains what can occur as a second-pair part.
5. Given the first-pair part, the second-pair part becomes conditionally relevant.

Examples of adjacency pairs include question-answer, apology-minimization, offer-acceptance, greeting-greeting, summons-answer, request-promise, and so on. Adjacency pairs reflect obvious regularities in peoples' talk; people return greetings, answer questions, accept offers, and so on. But the concept, and what it reveals about talk, extends far beyond this simple superficial description.

First, the emphasis is on how people display their understanding of one another's talk and how they ground their conversational contributions (Clark & Schaefer, 1989; see chap. 5 for more detail on Herb Clark's approach to conversational structure). This understanding is observed not through an assessment of the interactants' internal states, interpretations, emotional reactions, and so on. Instead, understanding is to be seen in the talk itself, and this view becomes quite clear with adjacency pairs. The occurrence of an answer in response to a question displays the second interactant's understanding of what was accomplished with the

first speaker's utterance, that the first speaker has asked a question. If an answer is not forthcoming, the second speaker, through a failure to answer, displays a lack of understanding of the first speaker's utterance (although this view has problems, as discussed later). It is this emphasis on the structural and sequential regularities as a reflection of understanding that is the hallmark of this approach.

Second, the second-pair part of an adjacency pair need not immediately follow the first-pair part. This is because embedded sequences of talk, termed insertion sequences (Schegloff, 1972) or side sequences (Jefferson, 1972), can occur between the first- and second-pair part. These sequences involve issues raised by the first-pair part and can themselves be ordered as adjacency pairs. Consider the following example from Merritt (1976):

(1) A: Can I have a bottle of Mich?
B: Are you twenty-one?
A: No
B: No

Rather than providing the second-pair part, B at Turn 2 instead begins an insertion sequence by performing the first-pair part of a second adjacency pair. In Turn 3, A provides an answer to B's question, thereby ending the insertion sequence and denying a condition that would allow B to comply with the request. The second-pair part of the original adjacency pair is then provided by B in Turn 4.

Insertion sequences can become quite lengthy, with adjacency pairs becoming embedded within other adjacency pairs. As a result, the first- and second-pair part of the original adjacency pair might be rather far apart. Consider the following service encounter (simplified from Levinson, 1983, p. 305):

(2) 1 A: How many tubes would you like sir?
2 B: U:hm (.) what's the price now eh with VAT do you know eh
3 .....
4 A: Three pounds nineteen a tube sir
5 B: Three nineteen is it=
6 A: Yeah
7 B: E::h (1.0) yes u:hm (dental click) (in parenthetical tone) e:hjus-justa think, that's what three nineteen That's for the large tube isn't it
8 A: well yeah it's the thirty seven c.c.s.
9 B: Er, hh I'll tell you what I'll just eheh ring you back I have to work out how many I'll need. Sorry I did-wan't sure of the price you see.

The first-pair part (Turn 1) of the initial adjacency pair is not answered until Turn 9 (and then it is deferred, with an account and apology). Between the first- and

second-pair parts of the original adjacency pair are three additional adjacency pairs (Turns 2 and 4, 5 and 6, and 7 and 8) each a question-answer pair dealing with issues raised by the first-pair part of the original adjacency pair. The relevance of these side issues for the original question helps maintain the expectation that the second-pair part will be forthcoming. Clearly, then, adjacency pairs can structure relatively long sequences of talk.

Third, adjacency pairs represent a template for the production of talk. Given the occurrence of a first-pair part, the second-pair part becomes *conditionally relevant* (Schegloff, 1968); it is expected, and if not forthcoming, its absence is noticeable. Thus, greetings usually follow greetings, acceptances follow offers, and so on. But not always. As we've seen, the second-pair part may be delayed with an insertion sequence. And, in fact, the second-pair part may never occur. Adjacency pairs are not to be viewed as statistical regularities; for example, that answers will follow summons 92% of the time (Heritage, 1984). Rather, adjacency pairs are templates for both the production and the interpretation of talk; the conditional relevance of the second-pair part serves as a template for interpreting what follows a first-pair part. The absence of a second-pair part is noticeable and grounds for making inferences about the speaker, for initiating a repair sequence, and so on. For example, failing to return a greeting can result in additional (and louder) greeting attempts and/or inferences that the person is rude or boorish.

There is an obvious similarity here to Grice's (1975) theory of conversational implicature. In Grice's view, flouting a conversational maxim instigates an inference-making process; the recipient will reason about what the speaker really means with the utterance (see chap. 1). Similarly, failing to provide the expected second pair part of an adjacency pair will result in the making of inferences about the person who has failed to perform the expected action. So, not answering a question is both a flouting of the relation maxim and a failure to provide an expected second-pair part, and both serve as grounds for making inferences about the speaker and the meaning of the utterance.

But there are some important differences here. First, the adjacency pair concept has an advantage over Grice's maxims in that it represents a more precise specification of when conversational inferencing will occur (as noted in chap. 1, Grice's maxims are somewhat ambiguous). Second, in conversation analysis the examination of inference making is limited to those inferences that are revealed in a subsequent turn at talk. And quite often inferences are revealed in this way, as when one person comments on the boorishness of another who has failed to return a greeting. But inference making beyond that revealed in talk generally is not considered in conversation analysis. In contrast, the Gricean approach, at least as used by psycholinguists, allows for and in fact focuses on internal inference making.

Adjacency pairs structure action; the first-pair part of an adjacency pair constrains what can follow. But some adjacency pairs are most constraining than others (Levinson, 1983). For example, the greeting-greeting pair is tightly constrained and ritualistic; people generally return a greeting with a greeting, there

are few options here. But what about question-answer pairs? There are numerous acceptable responses to a question, acceptable in the sense that the first speaker would display acceptance of the move in subsequent turns. For example, a person can say that she does not possess the requested information, that another person should be consulted, and so on. Although the second-pair part is not provided, the speaker accounts for this, and in this way displays an orientation to the adjacency pair action template. But what about indirect replies? Virtually *any* subsequent utterance (even silence) can serve as an answer to a question in the sense that the recipient will interpret it within an adjacency pair template (see chaps. 1 and 2). The first speaker's subsequent turn at talk may reveal that it has been interpreted in just this way. But are one's interpretations *always* revealed in subsequent turns at talk? Might not some interpretations be made and yet not revealed? If this is so and if any subsequent utterance can serve as an answer to a question, then the explanatory value of the question-answer adjacency pair is decreased.

Preference Agreement

What counts as a second-pair part, then, is something of an issue. The range of utterances that can serve as a second-pair part is quite large and in some instances unlimited. But there is an interesting feature of adjacency pairs that goes some way toward resolving this issue. The feature is this. Not all second-pair parts are of equal status; some second-pair parts are preferred over other second-pair parts. For example, agreements are preferred over disagreements, acceptances over refusals, answers over nonanswers, and so on. The former are referred to as preferred turns, or seconds, and the latter as dispreferred turns, or seconds.

The preferred-dispreferred construct is *not* defined in terms of peoples' attitudes towards others' utterances. Instead, dispreferred turns are defined as turns that are marked (in the linguistic sense) in some way; preferred turns are unmarked turns. The validity of this concept stems in part from the fact that the manner in which dispreferred seconds are marked is remarkably constant over all adjacency pairs (as well as other phenomena such as conversational repair; see as follows). These markings (some of which are described) have been documented for disagreements (Pomerantz, 1984; Holtgraves, 1997b), blamings (Atkinson & Drew, 1979), rejections (Davidson, 1984), and many other actions as well. The preferred-dispreferred construct appears to be a very general one. Table 4.1 displays preferred and dispreferred turns for several types of adjacency pairs.

Dispreferred seconds are marked in a variety of ways. Quite often they are delayed in some way; their preferred counterparts are not. To illustrate, compare the responses to the following two invitation sequences (both from Atkinson & Drew (1979, p. 58):

TABLE 4.1
Preferred and Dispreferred Adjacency Pair Seconds

First-Pair Part	*Second-Pair Part*	
	Preferred	*Dispreferred*
Assessment	Agreement*	Disagreement
Request	Acceptance	Refusal
Offer	Acceptance	Refusal
Blame	Denial	Admission
Question	Answer	Nonanswer

*Unless the assessment is a self-deprecation.

(3) A: Why don't you come up and see me some // times?
B: I would like to

(4) A: Uh if you'd care to come over and visit a little while this morning then I'll give you a cup of coffee.
B: hehh. Well that's awfully sweet of you, I don't think I can make it this morning .hh huhm I'm running an ad in the paper and—and uh I have to stay near the phone.

The acceptance (a preferred turn) in (3) is immediate, it actually overlaps with the invitation. The refusal (a dispreferred turn) in (4) is delayed (hehh). In addition, (4) illustrates the use of prefaces in dispreferred turns, in this case the inclusion of a marker (Well) and appreciation (that's awfully sweet of you). The following (from Pomerantz, 1984) also illustrates the delay and "well" marker of a dispreferred turn, in this case the nonanswer to a question.

(5) A: An how's the dresses coming along. How d'they look.
B: *Well uh* I haven't been uh by there.

Other frequently used prefaces include token agreement (e.g., "Yeah, but . . .") when disagreeing and apologies when refusing requests. The former is illustrated here (from Holtgraves, 1997b).

(6) A: And it's, I think, it's just like murder.
J: *Yeah, but*, what about, I mean it's an unliked child, you know, because how many kids are beaten, because, you know, they . . .
A: *Yeah, but* you know how many parents are are with well couples without children are, dying to have babies.

Frequently, the speaker of a dispreferred turn will provide an account so as to explain why it is a dispreferred rather than preferred action that is being

performed, as can be seen in (4) and in (7). Dispreferred turns often will be indirect as in (7); preferred turns, in contrast, tend to be direct. Finally, dispreferred turns often will be marked in multiple ways as in (4) and (7). The following (from Wooten, 1981) illustrates the inclusion of a preface, account, appreciation, and indirectness:

(7) B: She says you might want that dress I bought, I don't know whether you do.
A; Oh thanks (well), let me see I really have a lot of dresses.

The systematic marking of dispreferred turns is a very basic feature of conversational interaction and provides independent criteria for categorizing different second-pair parts. But these markings have another very interesting feature; they illustrate the essentially collaborative nature of conversational interaction. Specifically, the delay or preface or account of a dispreferred turn indicates to the speaker of the first-pair part that the turn in progress is dispreferred. This delay, preface, or account thus provides an opportunity for the producer of the first-pair part, on recognizing that a dispreferred turn is underway, to revise or reformulate the first-pair part before the dispreferred second is completed.

In the following example (from Davidson, 1984), the first speaker, P, reformulates the offer immediately on hearing A's "well" preface.

(8) P: Oh I mean uh: you wanna go to the store or anything over at the Market [Basket or anything?]
A: hhhhhhhhhhhhhhhhh =
A: =Well ho [ney I-]
P: Or Richard's?

A speaker of the first-pair part can facilitate this process by building into his turn postpositioned turn components, spaces that anticipate and orient to the possible rejection by the other. In the next example (Heritage, 1984, p. 277), there are two such places where the second speaker could provide a preferred turn (acceptance), once after "anything we can do," and again in his second turn after "shopping for her." In both instances, the acceptance is not forthcoming and H immediately expands or clarifies the offer in lines 3 and 6 in order to make it more concrete.

(9) 1 H: And we were wondering if there's anything
2 we can do
3 to help.
4 S: [Well 'at's]
5 H: I mean can we do any shopping for her
6 or something like that?

In the following example (from Levinson, 1983; p. 320), C first provides a brief pause and then a two-second pause, both indicating an orientation to a possible dispreferred second. Consequently, C reformulates following the brief pause ("by any chance") and following the two-second pause ("probably not").

(10) C: So I was wondering would you be in your office on Monday (.) by any chance?
(2.0)
C: Probably not
R: Hmm yes =

People are clearly attuned to possible dispreferred turns, and they structure their utterances to avoid them if possible. In this way, one's turn at talk is impacted by the other as the turn unfolds. If there are indications that the next turn is dispreferred, then the current speaker can quickly change an utterance before it is completed. The speaker of the first-pair part can facilitate this process—increase the likelihood of receiving information regarding the possibility of a dispreferred turn—by building in postpositioned turn components and attending closely to the other's possible reactions. The orientation of people to dispreferred turns and the manner in which this orientation impacts language use illustrates clearly the coordination involved in conversation. More important, it illustrates how conversations are truly collaborative. Turns at talk are not isolated entities; their course can be influenced by the other person as the turns unfold. Conversations are joint products constructed on a moment-to-moment basis.

In conversation analysis, the nonequivalence of preferred and dispreferred turns is assumed to be independent of interactant's internal states; even if one very much wants to disagree with another person, the disagreement will still be often marked in some way. It is as if preference agreement has a life of its own. It is a general feature of the conversation system, a resource used by interactants as they converse. Because of this, no reference is made to the psychological states of the interactants.

But why do conversations have the preference structure that they seem to have? It seems reasonable that preference organization must be motivated in large part by concerns with face, as Brown and Levinson (1987), Heritage (1984), Lerner (1996), and others have argued. Note in this regard that the marking of dispreferred seconds bears a strong resemblance to many of Brown and Levinson's (1987) positive politeness strategies. For example, the expression of token agreement ("Yes, but . . ."), the use of "well" prefaces, mitigators, and indirectness in general are all instance of positive (and sometimes negative) politeness strategies. Although preference organization resides in the talk itself and not the people who produce it, it seems likely that the origin of this preference is the rudimentary respect for face that is required if an exchange is to take place at all (Goffman, 1967). All preferred turns are affiliative; they address positive face wants. If this were not the case, then one would expect an agreement to be the preferred

response to all assessments. But this is not the case. If the first turn is a self-deprecation, then the preferred response is a disagreement, and an agreement is the dispreferred response (Pomerantz, 1984). To agree with a speaker's self-deprecation is nonaffiliative and threatens positive face. This preferred/dispreferred reversal illustrates how face concerns may motivate the form that preference agreement takes. Although preference agreement may be viewed as reflecting general properties of the conversational system, these regularities might also reflect interpersonal concerns.

The linking of actions with adjacency pairs provides a means of structuring talk throughout a conversation. Adjacency pairs are a conversational resource, and they are a resource that can be used by people for various conversational purposes. For example, as shown in the next section, adjacency pairs play an important role in the initiation and termination of conversations.

Opening and Closing Conversations

How do people begin a conversation? It seems relatively easy—at least with people we know. But it is not. Consider a few of the issues that must be dealt with in some manner. First, potential interactants must indicate to one another their availability and willingness to enter into a conversation. Second, there must be some recognition of with whom one will be conversing; the parties must recognize each other. Third, a topic must be brought up and sustained. And so on.

These problems are handled neatly and economically with a particular type of adjacency pair termed a summons-answer pair (Schegloff, 1968). A summons—the first-pair part—obliges an answer from the person to whom it is directed. A person who answers a summons has indicated availability for talk; refusing to respond indicates a lack of availability (and may be grounds for making inferences about the person who has failed to respond). If there is some doubt about availability, the summons may be repeated after a short delay. Note that a telephone ring is a summons, and its ring-pause-ring sequence simulates the verbal summons-pause-summons sequence.

Summons-answer sequences are different from other adjacency pairs in that something always follows them; they have implications beyond the two turns of which they are comprised. This is because the person issuing the summons is obligated to respond to the answerer. Thus, the summons-answer sequence, at a minimum, has a three-turn format. A common variation here is for the second-pair part—the answer—to be phrased as a question (e.g., "What is it?") in which case it simultaneously serves as the first-pair part of another adjacency pair, thereby obligating a subsequent move from the other person. In telephone conversations, the answer to a summons (the telephone ring) is most often the first-pair part of a greeting exchange, thereby making a return greeting conditionally relevant.

Telephone conversations, because of the lack of physical copresence, require some sort of identification procedure so that interactants know with whom they are talking (Schegloff, 1979). Mutual recognition is required; both the caller and answerer must identify each other. Consider how this is accomplished. Many times the caller's first turn speaking will be a return greeting, and often this greeting serves to indicate to the other party that she has been identified, as in (11). Also, it is in this slot that a caller may identify herself, if she chooses to do so. But callers frequently do not identify themselves, with the failure to identify serving as a claim that the other should be able to recognize the speaker from the voice sample alone. Examples such as the following (from Schegloff, 1979) are not uncommon:

(11) (ring)
A: Hello*::*
C: *Hi:::.*
A: oh: *hi*:: 'ow are you Agne::s,

Of course if these claims are problematic (i.e., the answerer does not recognize who is calling), then this will usually be indicated quickly with a pause, request for clarification (" 'Who is this?' "), and/or failure to return the greeting, as in the following (from Schegloff, 1979):

(12) (ring)
A: Hello?
C: Hello
(1.5)
A: Who's this?

If the caller is unsure of the answer's identity, then the caller's first turn will often take the form of a try-marked address term (Schegloff, 1979), as in (13).

(13) (ring)
A: Hello?
B: Connie?
A: Yeah Joanie

In this example, the try-marked term is verified by the answerer. Simultaneously, in that turn the answerer has indicated identification of the caller, thereby preempting the need for the caller to self-identify.

In general, there appears to be a preference for the identification procedure to function without interactants paying explicit attention to the process; other-recognition is preferred over self-identification (Schegloff, 1979). Why? Well, there is the matter of economy—conversational resources (which are scarce) will not

need to be used for recognition. This process is possible because simple one-word turns can perform multiple acts. Consider again a prototypical (made up) telephone opening:

(14) 1. Phone rings
2. Hello?
3. Hi.
4. Hi.

This sequence has brought two people in contact, each displaying their availability to talk and indicating their recognition of one another. These minimal units accomplish all of this through their sequential patterning and the fact that multiple adjacency pairs are overlaid on one another. Note also that the interactional work that must be performed at the beginning of a conversation makes multiple acts more likely here. So, Turn 2 is an answer to a summons (adjacency pair 1), a greeting (adjacency pair 2), and a display for recognition. Turn 3 is a return greeting (adjacency pair 2), claim that the other has been recognized, and a display for recognition by the other.

The use of adjacency pairs provides an economical means for beginning a conversation. But it may not be exclusively a matter of economics. There are also strong interpersonal reasons for this system, just as there are for the structure of preference agreement. Mutual, implicit recognition is obviously more economical than explicit identification, but it is also a marker of relationship closeness (positive politeness). To begin a conversation with each person recognizing the other without any explicit identification process implicates a relatively high degree of familiarity between the parties. Again, conversational structure reflects both system and interpersonal processes.

Closing a conversation presents obstacles similar to those involved in opening a conversation, and these also must be handled with care. The final portion of a conversation is usually a terminal exchange (e.g., bye-bye). Again, an adjacency pair is being used, and again the second-pair part serves to indicate the speaker's awareness of what is being accomplished by the first speaker. But one can't just end a conversation. This can happen, of course, but when it does it is often marked with an account of some sort (e.g., "I hear the baby; I've got to run."). The general tendency is to gradually close down the conversation. And, as with openings, there are both economical and interpersonal reasons for this.

First, people may still have topics that they want to talk about. To simply end a conversation with a terminal exchange would prevent participants from having the opportunity to mention some topic that had not yet been talked about. A device that provides an opportunity for this possibility is a preclosing move, or more accurately, a possible preclosing (Levinson, 1983). A preclosing (e.g., "well", "So . . . ," "OK") occupies an entire turn and serves as a "pass," indicating that the speaker has nothing new (at this point) to add. Sometimes the preclosing

will be initiated by mentioning the initial topic or reason for the call ("Well, I just wanted to see how you were doing"), a move that serves to frame the encounter as a separate unit (Sacks & Schegloff, 1979). A preclosing breaks with a prior topic, thereby suspending any rules regarding topical coherence and allowing the next speaker to bring up an entirely new topic. And she may do just that, with the conversation then proceeding for a number of turns. Or, she may acknowledge the preclosing with an acknowledgment of her own (e.g., "Yeah," "OK, so . . ."). This serves to indicate an understanding of the first speaker's move and acceptance of it, and hence permission to initiate the terminal exchange. Thus, the prototypical (made up) closing looks something like (15):

(15) A: OK
B: OK
A: Bye
B: Bye

A second motivation for the negotiated manner in which conversations are brought to a close is interpersonal. Simply ending a conversation can be offensive to the other person. It represents a threat to the positive face of the other person, because it is a dramatic severing of the momentary bond constituted by the current exchange. Preclosings and adjacency pair terminations allow this predicament to be handled mutually, so that neither interactant feels particularly threatened by the closing of the talk. Things can go wrong, of course. One may preemptively end an exchange. But to do so is impolite and grounds for inferences of brusqueness on the part of the person who does so. Note also that the interpersonal variables discussed in chapter 2 may play a role here. There has been little research in this vein because conversation analytic researchers purposefully exclude consideration of such variables. But one might reasonably expect, for example, a higher status participant to be more likely to unilaterally end a conversation, and to do so with greater impunity, than a participant who is relatively lower in status.

Presequences

Adjacency pairs represent a fundamental, if not *the* fundamental unit of conversations. They capture, in a very basic way, the interlocking nature of a stretch of talk. In its simplest and most basic form an adjacency pair represents two contiguous turns at talk. But they can be expanded, as we've seen, with the inclusion of an insertion sequence. Adjacency pairs also underlie longer sequences of talk, sequences that display many of the same properties of adjacency pairs. Levinson (1983) refers to these as presequences, and they have a number of interesting properties.

We have already seen how adjacency pairs are used in preclosings to negotiate the termination of a conversation. Another common type of presequence is a

preannouncement, a tightly structured set of adjacency pairs that prefigure an extended turn (for the telling of a story, news, gossip, etc.) by one of the interactants. Preannouncements can be viewed as two superimposed adjacency pairs, with the second part of the first pair simultaneously functioning as the first-pair part of the second pair. Consider the following exchange (reported in Levinson, 1983, p. 349):

(16) D: Heh, you'll never guess what your dad is lookih-is lookin' at.
R: What are you looking at?
D: A radar range.

Turn 1 is a check on the newsworthiness of the potential announcement and is the first part of the first adjacency pair. Turn 2 both validates the newsworthiness (the second part of the first adjacency pair) and requests that the information be told (the first part of the second pair). Turn 3 provides the announcement and is the second part of the second adjacency pair.

Note the interlocking nature of the sequence, with Turn 2 orienting both backwards to the prior turn and forward to the subsequent turn. The sequence structures a stretch of talk, and not just the presequence but the talk that will follow. Preannouncements are a means of securing an extended turn at talk (see turn taking following). They are also motivated by a general preference for not communicating information that is already known by one's interlocutors; preannouncements thus serve as a check on the Gricean (1975) maxim of quantity (do not be overinformative).

Other presequences include prerequests and preinvitations, both of which display the same structure. Position 1 is a check on the precondition for the act to be performed, and Position 2 provides an answer regarding that precondition. Position 3 is the prefigured act (i.e., request or invitation), contingent on the preconditions holding in Position 2. Finally, Position 4 is the acceptance/compliance. Examples of a preinvitation and prerequest, respectively, follow.

(17) (From Atkinson & Drew, 1979, p. 253)
A: Whatch doin'?
B: Nothin'
A: Wanna drink?

(18) (From Merritt, 1976, p. 357)
C: Do you have hot chocolate?
S: mmhmm
C: Can I have hot chocolate with whipped cream?
S: Sure (leaves to get)

Prerequest sequences provide an alternative approach to the issue of how best to characterize indirect requests (as discussed in chap. 1). Following Levinson

utilized to check on the potential obstacles that would prevent the recipient from complying with the request (see also Gibbs, 1986; Francik & Clark, 1985); their use is motivated by a desire to avoid an action that would be followed by a dispreferred second. Thus, the first position of the sequence is usually an attempt to assess the likelihood of the request succeeding, as in (18). In this way, the first speaker can avoid making a request that would be likely to result in refusal (a dispreferred second). Alternatively, Position 1 may elicit an offer from the other interactant, a preferable situation insofar as it eliminates the need to even make the request. In this view, the issue of whether or not indirect requests have literal and indirect forces does not arise. Rather, the turns in a prerequest sequence simply have literal meanings that prefigure a set of alternative responses. But note that face management is the likely motivation for a prerequest sequence, just as it is for indirect requests.

Adjacency Pairs and Speech Act Theory

The conversation analytic view of adjacency pairs is a view of language use as action, and more importantly, as joint action. As an action-based view of language, it both shares certain features with speech act theory, and extends it in important ways. In speech act theory the emphasis is on illocutionary force, or the action that the speaker performs with an utterance. In conversation analysis the utterances making up adjacency pairs are also viewed as actions. But with adjacency pairs, unlike speech acts, the emphasis is on *joint* (rather than individual) action. A single individual cannot perform an adjacency pair, it must be performed by two different people who are orienting to each other's actions. In contrast, the emphasis in speech act theory has been on the action performed with a single utterance by a single individual (though see Searle, 1990). But does not the successful performance of a speech act require some ratification by the recipient that the intended act has been performed? If I say "I bet $5 the Yankees win the series" and you ignore me, have I successfully placed a bet (even if all the preconditions for betting are met)? One of the most important features of adjacency pairs is the emphasis on the second-pair part as an indicator of the recipient's understanding (or lack thereof) of the act performed with the first-pair part.

But note that the speech act and adjacency pair concepts are not incompatible; for some adjacency pairs, the felicity conditions underlying the performance of the first-pair part link up quite nicely with the felicity conditions underlying the second-pair part (Clark, 1985). Consider, for example, the request-promise adjacency pair depicted below.

First-pair part (Request)	*Second-pair part (Promise)*
Bob: Can you type my term paper?	Tom: I'll do it tonight.
Preparatory: Bob believes Tom is able to type the term paper.	Tom is able to type the term paper.

Propositional Content: Bob predicates a future act of Tom.	Tom predicates a future act of Tom.
Essential: Utterance counts as an attempt to get Tom to type the paper.	Utterance counts as an obligation to type the paper.

Each of the felicity conditions underlying the request has a complimentary felicity condition underlying the promise.

TURN TAKING AND REPAIR

One of the most amazing features of conversations is how orderly they are. Usually, one speaker speaks at a time, and speaker transitions occur with little or no overlap. Moreover, this orderliness occurs regardless of the number of speakers, the size of the turns, the topic of the conversation, and so on. How is this possible?

In a seminal article, Sacks, Schegloff, and Jefferson (1974) proposed a model to account for turn-taking regularity. In general, their model can be viewed as a system for allocating in the fairest and most economical manner a scarce conversational resource, turns at talk. There are two components in the model. The first is termed the turn-constructional component, a feature that specifies what constitutes a turn. Turns can vary in length, from a single word or partial word to a complete utterance, but they all have the characteristic of projectability; once begun, it generally becomes known what type of turn is underway and what it will take to complete it. In general, projectability is assumed to be determined by the intonational and syntactic properties of the turn. That talk has this feature can be seen by the fact that sequentially appropriate starts by the next speaker occur after single-word turns with no significant gap. Interactants apparently have the ability to identify the end of a turn (termed a transition-relevance place).

The second component of the model is the turn allocation component, or the techniques used for the allocation of turns within a conversation. Turn selection is accomplished by either "self-selection" or "current speaker selects next speaker." This process is governed by three rules along the following lines:

1. If at the first possible transition point, the current speaker uses the "current speaker selects next" technique, he selects the next speaker who is then obligated to take the turn at that point (as with adjacency pairs).
2. If, at the first possible transition point, the current speaker does not use the "current speaker selects next" technique, then the "self-selection" technique becomes operative allowing the first speaker to hold the floor.
3. If, at the first possible transition point, the current speaker does not use the "current speaker selects next" technique, then she may, but need not, continue unless another self selects.

These rules are ordered such that Rule 1 takes precedence over Rule 2, and Rule 2 over Rule 3. They are also cyclical such that if Rule 3 applies and current speaker continues talking, then Rules 1 through 3 reapply at the next transition relevance place.

This model allows for variation in turn size and number of parties. In other words, it is managed both locally (not specified in advance) and interactionally (it arises out of the interaction process itself). These rules certainly describe some of the most basic facts about turn taking. To wit, usually only one speaker speaks at a time (thus, some rule system is being followed), but when overlap does occur, it occurs at places that would be predicted by the rules. For example, overlap most frequently occurs as competing first starts (Rule 2 is operative), and at possible transition relevance places (supporting the idea of a turn constructional unit).

There is some debate about the nature and significance of these turn-taking rules. For example, are people generally aware of them? No. Similar to the rules of grammar, they are, for the most part, outside of conscious awareness. Are they causal, then, in the sense that grammatical rules are causal? Searle (1992) argues that they are not. In fact, he argues that they are not rules at all, having no intentional component. Turn taking is clearly patterned and systematic (it is not random), but not because people are following these particular rules. In other words, these rules might *describe* turn-taking regularity (but so could any number of other rule systems), but they do not *explain* it.

Conversation analytic theorists maintain, however, that these rules are more than merely descriptive. Schegloff (1992a) argues that although the rules may not cause or bring about actions, the rules are displayed or discoverable in the action. That is, as people converse, they display an orientation to these rules, they are sensitive to their violation, and they undertake actions designed to get the turn-taking system back on track.

This, then, is an ethnomethodological conception of rules; rules are discoverable in the actions of people. But because of this, it is very difficult to empirically test this model. For example, according to the model, turn taking should occur at transition relevance places, and transition relevance places are assumed to be identifiable in advance (turns project to their completion points). But people sometimes start talking before another has finished. To account for this, it is actually claimed only that people are orienting to *possible* transition relevance places, rather than actual ones (Schegloff, 1992a). If stated in this way, this aspect of their rule system is not really falsifiable. But, then, producing a falsifiable model was not really the developers' intent.

It should be noted that there are alternative approaches to turn taking (e.g., Argyle, 1973; Duncan & Fiske, 1977; Kendon, 1967). Many of these approaches are based on the idea that turn taking is regulated via a signaling procedure. Speakers are assumed to signal their intent to relinquish a turn; other participants are assumed to signal their intent to occupy the next turn. The signal most frequently invoked here is gaze. For example, a speaker will signal that he intends to relinquish the speaker role by returning his gaze to the next speaker (Argyle, 1972). This

position is undermined somewhat by the fact that turn taking seems to be quite orderly even when signaling via gaze is eliminated, as in telephone conversations. Although it is clear that gaze does play a signaling role in conversational exchanges, signals alone do not seem capable of explaining the orderliness of turn taking.

Repair

One of the ways in which people are said to display their orientation to the turn-taking system is by how they repair the system when it momentarily breaks down. Consider, for example, instances of overlapping talk. This might seem to indicate a failure on the part of the turn-taking rule system. But, as Jefferson (1986) has shown, overlapping speech is itself quite orderly. In the first place, much overlapping talk does not occur at random points in a conversation, but instead at transition relevance places. Moreover, when overlapping talk does occur, the repair system is quickly engaged; either one speaker will rapidly drop out, or if this does not happen, an allocation system will kick in whereby the speaker who upgrades the most (e.g., increased volume) wins the floor. Regardless of how this is settled, the overlapped turn will then be repeated by the remaining speaker.

The notion of repair has received considerable attention from conversation analytic researchers, an attention that is warranted by the fact that the conversation repair system appears to be generalizable over conversational problems (e.g., mishearing, misunderstanding, incorrect word selection, etc.). Conversational repair, like turn taking, preference structure, and adjacency pairs, is part of a conversational system, a resource that is available to conversationalists. In another important paper, Schegloff, Jefferson, and Sacks (1977) outlined a number of important features of this system. First, they made a distinction between the repair itself and the initiation of the repair. The latter, termed a next-turn repair initiator (NTRI) refers to phrases (e.g., "huh?"), questions (e.g., "What do you mean?"), repetitions of the preceding turn, nonverbal gestures, and so on that may serve as prompts for the undertaking of a repair.

Second, repairs and repair initiations can be undertaken either by the speaker or by the other interactant. These two distinctions result in a fourfold scheme of repair types: self-initiated self-repair, self-initiated other repair, other-initiated self-repair, other-initiated other repair. Finally, repairs can be classified in terms of when they occur. Repairs can occur in the turn that is problematic (Position 1 repair), in the transition space or turn following the problematic turn (Position 2 repair), or two turns after the problematic turn (Position 3 repair). Note that turns need not be adjacent here; insertion sequences may separate them. Repairs later than the third position are extremely rare (but see Schegloff, 1992b on fourth-position repairs) and for good reason. The system is structured so that trouble will be repaired as quickly as possible in order to decrease the likelihood of subsequent organizational problems (e.g., the difficulty involved in backtracking in order to deal with an earlier problem).

In addition to early repair, the system clearly favors self-repair over other repair. In other words, conversational repair displays a preference structure, with the least-preferred options being marked in various ways. The most preferred is self-initiated self-repair in first or second position, followed by other-initiated self-repair in Position 2, and finally, other-initiated other repair in Position 2 (Levinson, 1983). Note that the structural preference for early repair increases the likelihood of self-repair (early repair can only be undertaken by the speaker). If self-repair does not occur, the recipient will often provide a brief delay, thereby providing the speaker one last opportunity to repair. Finally, even when other repair is possible, there is a preference for the recipient to use a NTRI and hence only initiate a repair, allowing the speaker to perform the actual repair to correct the problem. In the following example (from Schegloff, Jefferson, & Sacks, 1977, p 370), the recipient B delays her response for one second, thereby providing A an opportunity to self-correct. When this fails to occur, B then initiates repair, which A undertakes.

(19) A: Hey the first time they stopped me from sellin' cigarettes was this morning.
(1.)
B: From selling cigarettes?
A: From buying cigarettes.

Conversational repair, like the other structural regularities discussed in this chapter, is generally viewed by conversation analytic researchers as constituting a conversational system, a system designed to facilitate communication. Again, there are also strong interpersonal reasons for the form that repairs take. Both self-repair and early repair are more efficient than other repair and later repair. But self-repair also minimizes or eliminates both the negative and positive face threat that can be implied when correcting another person. And early repair (Position 1 or 2), because it can only be undertaken by the speaker, can also be interpreted in this way.

Extended Turns at Talk

During the course of a conversation people will sometimes hold the floor for an extended period of time, telling jokes and stories and gossip, and so on. Like openings and closings, taking the floor for an extended period of time presents certain problems for interactants, and again, these problems are both system oriented and interpersonally oriented.

First, to hold the floor for an extended period of time requires a temporary suspension of the turn-taking system. Everyone involved must understand that an extended turn is underway, such that bids for the floor at transition relevance places are put on hold until the extended turn is completed. This does not mean that

others cannot speak during the extended turn; they do participate in various ways, particularly in terms of providing feedback and demonstrating appreciation. Then, once an extended turn is complete, the turn-taking system must be reengaged. That this routinely happens suggests that interactants can tell when a story is complete. Stories appear to have the feature of projectability, just like single turns at talk.

Second, conversation turns can be viewed as a scarce resource, and thus an extended turn at talk represents a monopoly of this resource. Consequently, an extended turn at talk will often involve a justification of some sort. Accordingly, speakers often attempt to make a story relevant in some way for the conversation. Of course, this is an issue for any turn at talk (see as follows), but it is particularly important for stories because of their potential extended duration. There appears to be a distinct preference for the introduction of stories at points in the conversation where they would fit (Schegloff & Sacks, 1973). Stories, it seems, are usually "locally occasioned," triggered by something in the conversation or the environment (Jefferson, 1978; Schenkein, 1978).

So, the potential speaker of an extended turn needs to make it clear that a story is underway, and hence that the turn-taking system is temporarily suspended, that the story is in some way connected with their conversation, and so on. Interactants must align themselves into the teller and recipient roles (Sacks, 1992) and mutually agree that a joint project is underway (Clark, 1996a). Often this is accomplished with some version of a preannouncement, as discussed earlier. Preannouncements are a set of adjacency pairs that prefigure an extended turn. Frequently they deal with the issue of whether the recipients have heard the story, joke, or gossip. In general, there is a concern about not telling people things they already know (e.g., Sacks & Schegloff, 1979), a concern that makes sense in terms of attempting to hold the floor for an extended period of time. A crucial point here is that interactants must *mutually* agree to the extended turn. Lack of agreement (e.g., that the recipient has already heard it) will short-circuit the attempt, as in the following example (from Clark, 1996a, p. 347):

(20) C: Did I tell you, when we were in this African village, and *—they were all out in the fields,—the*
I: yes you did, yes—yes*
C: babies left alon,—
I: yes

An extended turn is usually followed by a response sequence, including moves designed to show appreciation of the story (e.g., commenting on its significance), the joke (laughter), the gossip (e.g., indignation), and so on. Structurally, such moves are designed, in part, to move the conversation forward. For example, in some conversations the telling of a story by one person occasions the telling of a story by another person (Jefferson, 1978; Ryave, 1978); the conversation consists of a round of stories. Generally, the stories that people tell will

bear some relationship to each other. Sometimes they will be on the same topic and share a common referent (e.g., stories about bosses). Other times, they may be topically different, but related in terms of the significance or point of the story (Ryave, 1978).

Finally, note how the structural features of extended turns also have implications for the interactants' face. Holding the floor for an extended turn is presumptuous; it presumes the speaker has a right to this scarce resource and thus threatens the face of the other interactants by restricting their opportunity to speak (their freedom is threatened). The alignment procedures used to take an extended turn (e.g., justifying the turn) are probably motivated, in part, by face-management concerns. The response sequence following an extended turn also activates interpersonal issues. Consider, for example, reactions to jokes. Obviously, one is supposed to laugh, but the timing of one's laughter can be problematic (Sacks, 1974). Timing is crucial here, as inferences of stupidity may result if laughter is delayed.

CONVERSATIONAL COHERENCE

The tendency of stories to be related reflects a very general feature of conversations. This is that most conversations appear to cohere on a topic or a set of topics. Interestingly, conversations are generally not planned in advance, and hence, conversational coherence is an emergent property. Obviously, people can bring agendas and topics they want to talk about to a conversation, but the specific manner in which these things will get talked about emerges on a turn-by-turn basis as the conversation unfolds. In other words, one cannot specify in advance what people will say during the course of a conversation. So, what happens? How is it that (reasonably) coherent conversations are usually achieved?

The notion of conversation topic is ambiguous; there is really no agreed-on definition regarding the manner in which utterances can be viewed as being related. There is, of course, a large literature on text grammar dealing with the manner in which sentences within a discourse are related or connected to each other (e.g., Givon, 1989; Grimes, 1975; Halliday & Hassan, 1976). This research has dealt with both referential coherence (the manner in which referents are referred to again in a discourse) and topical coherence (the manner in which a set of sentences in a discourse can be said to be "about" a common topic), as well as the manner in which coherence is linguistically realized. There is also a large and influential literature within cognitive psychology dealing with the cognitive operations involved in the comprehension of narrative or expository text. The emphasis in this research is on the inferential processing required of readers in forming a coherent representation of a text (Gernsbacher, 1990; Graesser et al., 1994; Kintsch & van Dijk, 1978).

Certain aspects of discourse comprehension and text grammar (e.g., the given-new distinction) are relevant for understanding conversational coherence. But in

the end, conversational coherence is a very different animal than cohesiveness in a text. This is because conversational coherence is achieved by two people (rather than a single writer) interacting in real time in a specific context. Coherence is something that is created on the spot and may be visible only to those involved in the conversation (due to the use of elliptical expressions, etc.).

In one sense conversational coherence is not really a problem at all. If conversational relevance is assumed (Grice, 1975), then all conversation utterances will be heard by participants as being somehow relevant to what has gone before (e.g., see Sperber & Wilson, 1986). Thus, regardless of how incoherent and off-the-wall a contribution might seem, the recipient will usually interpret the utterance as if it was relevant.

Still, conversationalists do seem to be oriented to some sort of topic continuation rule. For example, many times people seem to position their topic introductions so that a connection with prior turns can be made by others. This type of topic shading allows a prior topic of talk to serve as a resource for the raising of a new topic, such that the new topic will be heard as relevant to the prior topic (Schegloff & Sacks, 1973). Sometimes shifts of this sort will be marked in various ways, as with embedded repetitions (e.g., "Speaking of the dean, I heard that . . .") and disjunct markers (e.g., "Oh, that reminds me . . .") (Jefferson, 1978). Sometimes topics discussed earlier will be reintroduced, and these reintroductions will often be marked as well (e.g., "Anyway . . . ;" "Like I was saying . . ."; Keenan & Schiefflin, 1976). And when completely new topics are introduced, they too are quite likely to be marked, for example, with disjunct markers (e.g., "One more thing . . . ;" "Incidentally . . .").

Experimental evidence in support of a topic continuation rule also exists. For example, Vuchinich (1977) recorded the conversations of two students in a get-acquainted session. One of the students, however, was actually an experimental confederate who, at various times during the conversation, produced abrupt, incoherent topic shifts that were not marked by the speaker in any way. The other (real) participants displayed their orientation to a topic continuation rule in their responses. The unmarked shifts often resulted in the initiation of a repair sequence: "Huh?" "What?" "Why are you saying that?" Also, such shifts often resulted in a noticeable gap, with a mean latency of 2.8 seconds versus a mean latency of .6 seconds following a cohesive turn.

Participants were probably attempting, with the pause, to give the speaker an opportunity to make the relevance of the contribution clear. Noncohesive turns were also more likely to result in topic termination than were cohesive turns. Finally, in postexperimental interviews all participants indicated they had been aware of the topic shift, even when their conversational responses did not display an orientation to the shift.

Conversation coherence is an extremely elusive construct. It is one of those things that no one can define, yet everyone can recognize. People agree quite

highly, for example, regarding how to segment conversations into topics (Planalp & Tracy, 1980). Yet no one has really been able to define conversational coherence in a formal manner, and it has stymied researchers in conversation analysis, artificial intelligence, linguistics, philosophy, and so on. This is clearly an important and wide-open avenue for future research.

Researchers in disciplines other than conversation analysis have investigated various aspects of conversation structure, including conversation coherence. Most notable in this regard is a set of topics and techniques referred to as discourse analysis. Although there is some overlap between discourse analysis and conversation analysis, methodologically they are quite different (Levinson, 1983). Unlike their conversation analytic counterparts, discourse analysts will undertake the detailed examination of stretches of talk using a variety of sources of information (e.g., the social context, talk occurring later in the conversation, conversationalists' possible motivations) that are not part of the talk itself. Discourse analysis is rather eclectic (unlike conversation analysis) and includes a wide variety of approaches and techniques that have yielded important findings regarding conversation and written discourse (e.g., see van Dijk, 1998).

Much of this research has examined conversational structure within a particular conversational activity. Thus, there are studies of the organization of classroom talk (Sinclair & Coulthard, 1975), therapy talk (Labov & Fanschel, 1977), arguments (Jacobs & Jackson, 1982; Rips, Brem, & Bailenson, 1999), marital communication (Gottman, 1979), and so on. Although some of the research emphasizes how genre determines structure (e.g., the structure of talk in a classroom is heavily influenced by the situational context), other research programs have implications beyond the specific verbal activity examined. Two examples will be considered briefly here: Labov and Fanschel's (1977) work on therapy talk and the artificial intelligence approaches of Schank (1977) and Reichman (1985).

Labov and Fanschel's Therapeutic Discourse

Labov and Fanschel's detailed analysis of 15 minutes of a psychotherapy session is a prototype of discourse analysis and notable for its attempt to incorporate within speech act theory both interpersonal and sequential considerations. These researchers assume that utterances involve the simultaneous performance of multiple acts, a specific speech act, and depending on the context and type of speech act, an interpersonal act (or acts).

Talk is viewed as being structured both vertically and horizontally. Vertical structure involves the relation between utterances and actions and is described in terms of rules of production and interpretation. There are, for example, rules for interpreting indirect requests, and these are quite similar to the rules described in

chapter 1 for performing indirect requests (e.g., Gordon & Lakoff, 1975). Another rule specifies when and how assertions function as questions. Specifically, an assertion made by the speaker about a topic known only to the hearer will be interpreted by the hearer as a request for confirmation about the topic in question. Some of the proposed rules deal with interpersonal acts. For example, a request that is repeated will be interpreted as a challenge to the recipient's competence. This is obviously similar to a face-management view of requests. But with a difference. The challenge in this case comes not from the request so much (although a single request can also function as a challenge if it involves an act that the recipient should have performed) but from the fact that the request is repeated. Thus, the threat made to the recipient's face emerges over a sequence of moves.

Horizontal structure involves the relation between chains of utterances and the actions they perform, and these are described in terms of sequencing rules. It is the actions (what is done) that are structured via these rules rather than the utterances themselves (what is said). Thus, understanding the structure of conversations involves both rules of interpretation (utterance → action) and sequencing rules (action → action). This is in direct contrast to conversation analysis, where the exclusive focus is on the surface meaning of utterances (or what is said).

Consider, again, a repeated request. An interpretation rule specifies that a repeated request will be heard as a challenge to the recipient's competence. According to sequencing rules, the options open to the recipient are either to acquiesce to this characterization or attempt to defend against it. Or consider an indirect request (the result of a specific interpretation rule). One can comply with the request, of course. But there are other possibilities. Sequencing rules, for example, allow people to put off requests by requesting additional information, relaying the request to someone else, challenging the requester's right to make the request, and so on. Importantly, these means of putting off requests are related to the conditions presupposed in the request itself (e.g., the recipient's ability), as well as to the interpersonal underpinnings of the act (e.g., the speaker's right to make the request). In this way, one gets sequential chains of speech acts based on the meshing of the felicity conditions and the interpersonal implications of the act.

Labov and Fanschel's research demonstrates the role of interpersonal actions in structuring talk and, hence, how those implications are revealed in a stretch of talk. Note that in contrast to politeness theory (chap. 2), the emphasis here is on a sequence of talk rather than a single utterance. On the other hand, the results of this research are limited and involve only a very restricted set of actions, primarily requests and the responses to them. The methodology itself can be criticized for being overinterpretative; the researchers use information in their analysis that was not available to the interactants themselves (e.g., Levinson, 1983). In fact, whether these rules and categories have any psychological reality is an open question. In many ways, this represents an important but unfulfilled approach to conversation structure.

Artificial Intelligence Approaches to Conversational Structure

For a number of years, researchers in artificial intelligence have attempted to simulate human language use. Much of this research has been concerned with modeling sentence comprehension. But some researchers, Roger Schank and Rachel Reichman in particular, have been concerned with modeling larger stretches of talk.

Schank (1977), for instance, has attempted to model conversation coherence with a set of production rules that specify what counts as an appropriate response to a prior utterance. These rules exist for both objects and actions. An example of an object rule is as follows: if an utterance introduces a new object, then one appropriate response is to ask where the object came from. If a person says "I just bought a new car," an appropriate reply would be "Where did you get it?" An example of an event rule is: if a person describes an event, then one appropriate response is to ask what happened next. So, a reasonable reply to "I was arrested last night" would be something like "Did you spend the night in jail?"

Shank's approach is intuitively reasonable, but obviously of limited value. It specifies some properties of concepts that can explain how two utterances about that concept are related. But it does not appear to be an exhaustive set of rules. Also, the approach is completely asocial (the interpersonal implications of utterances are ignored) and deals only with adjacent turns (the relation between nonadjacent turns is ignored).

A broader and somewhat more promising approach was developed by Rachel Reichman (1978, 1985). Although similar to other artificial intelligence approaches to modeling task-related conversations, her model is a general one and presumably relevant for any conversational activity. The most important feature of her model is the treatment of conversations as having a hierarchical structure. This assumption allows for capturing an aspect of conversational structure that is largely missed in the major approaches discussed so far (i.e., speech act theory, politeness theory, conversation analysis).

In Reichman's model, conversation utterances can either embellish and continue what has previously been said, or they can begin a new communicative act. The latter are referred to as conversational moves, and they are often marked with cue words specifying their discourse function such as support (e.g., because . . . ; like when . . .), restatement or conclusion (e.g., so . . .), interruption (e.g., by the way . . .), indirect challenges (e.g., yes, but . . .), direct challenges (e.g., No, but . . .). And so on. Now, conversations consist of a sequence of utterances, but this temporal sequence represents only a surface structure. At a deeper level, conversations consist of a hierarchically structured set of "context spaces," or a group of utterances referring to a single episode or issue. All context spaces contain the utterances (and inferences if necessary) that define the context space, as well as slots specifying context type, context function (the method used to perform the

move and its relation to the context space), context status (whether the context space is active [or current]), and so on.

The underlying structure of a conversation is specified by the set of relations existing between the context spaces. It is assumed that conversational structure is based on a set of discourse rules in much the same way that sentence generation is based on a linguistic rule system. The overarching rule here is Grice's (1975) maxim of relation (be relevant), which underlies the possible relations that can exist between context spaces. Relevance can be achieved in various ways. One can contribute to the same topic, or one can introduce a new concept, idea, and so on as long as it bears some sort of structural relation (e.g., supports, contradicts, illustrates, etc.) to the active context space. It is important to note also that an utterance can be relevant for a context space occurring earlier in a conversation, rather than the currently active space (although a surface marking of this relation would be required).

Reichman's model does reasonably well at modeling certain types of talk; whether it describes how humans actually produce and comprehend conversational utterances is unknown. It does seem that her model applies best to certain types of conversations, in particular arguments in which speakers take extended turns to develop their points; in certain respects her model parallels the work of others on the structure of arguments (Jacobs & Jackson, 1982). Also, like Schank's approach, there is no concern with the interpersonal features of talk and how such features might structure conversations.

CONCLUSION

That conversations are structured is obvious; how this structure comes about is something of a mystery. Capturing the manner in which conversation turns are sequentially related has proved to be a very difficult task (Slugoski & Hilton, in press). Currently there does not exist a grammar of conversation in the sense that there exists a grammar for individual sentences. Clearly, a linear sequence (or finite state grammar) is not appropriate, given the possibility of embedding, side sequences, topic recycling, and so on. Conversations appear to be too contextualized and idiosyncratic to allow for a formal grammar.

But researchers working within the discipline of conversation analysis have made some headway with this issue. Like speech act theory, conversation analysis views language use as action. Unlike speech act theory, the actions performed with talk are not characterizable in terms of interactants' abstract intentions. Intentions are irrelevant. Language use is action (and it is social action) in the sense that its occurrence has implications for what follows—implications for one's own actions and those of one's interlocutors.

This emphasis on sequential implicativeness is one of the most, if not *the* most, important contribution of conversation analysis. And as such it represents an

important correction to the primarily single-utterance-based approaches (speech act theory and politeness theory) covered so far. In converation analysis, an utterance performs an action based not so much on any inherent characteristics of the utterance itself but on its placement in a sequence of utterances. For example, an answer cannot be defined in terms of syntactic, phonological, logical, or semantic features—only by its placement in a sequence of utterances.

The principle of sequential implicativeness has implications for many of the phenomena discussed earlier in this book. The adjacency pair construct, for example, provides an important addition to Grice's (1975) ideas regarding conversational implicature. According to Grice, violation of the cooperative principle results in conversational inferencing. But as we've seen, what constitutes a violation of the cooperative principle is somewhat ambiguous. With adjacency pairs, the impetus for conversational inferencing is clearer—failure to provide an expected second-pair part. More importantly, when this happens, the failure to provide the expected part is usually marked, and these markings might play a role in comprehension, facilitating recognition of a nonliteral meaning. For example, when people provide an indirect reply to a question, the nonliteral meaning of the reply can be recognized more quickly if it is marked with a "well" preface or if it is preceded by a delay (Holtgraves, 2000). In this way, a dispreferred marker provides the recipient with evidence that she should search for an ulterior meaning. And because dispreferred turns are often face threatening, they may guide the recipient to a face-threatening interpretation of the utterance. The structure of adjacency pairs and marking of dispreferred turns specify more clearly how the conversational inferencing process described by Grice (1975) might actually work.

Preference organization also reveals the essentially collaborative nature of talk, a point that will be discussed in more detail in the next chapter. People orient to possible dispreferred seconds, and any signal that one is forthcoming can prompt the speaker to alter the act being performed in midturn. Speakers, in formulating their utterances, build in opportunities (e.g., brief pauses) for others to provide such signals. This responsiveness to feedback from others suggests the speech act view of an intention giving rise to the performance of an act may be too simple. The act that is ultimately performed may, on occasion, be a function of the interactants' joint communicative intentions.

Although not considered by conversation analysts, certain features of conversational structure have clear implications for the interpersonal causes and consequences of language use. Presequences in general, and prerequest in particular, demonstrate how politeness can be conveyed over a series of moves rather than within a single turn. Moreover, many of the interpersonal variables discussed in chapters 1–3 are probably related to various features of conversational structure. Consider, for example, the relationship between status and topic changes. Changing a topic can be tricky business and is often undertaken with a series of bilateral moves that gently transition the conversation from one topic to another. But if the conversationalists differ in status (think boss-employee), then the higher-

status person is probably more likely to undertake unilateral topic changes (Weiner & Goodenough, 1977). And if status is roughly equal, undertaking a unilateral topic change might be a means of making a claim for a higher status. Similarly, dominance can be claimed or established by interupting others (Ng, Bell, & Brooke, 1993). These effects can be extended to intergroup behavior; topic changes and interruptions become more frequent when interacting with outgroup members (Ohlschleget & Piontkowski, 1997). In this case, attempts at conversational control provide interactants with a means of achieving social distinctiveness. On the other hand, continuing with a topic raised by one's partner, being responsive to her talk, can signal liking and affiliation (Davis, 1982).

Options in structuring talk are a resource that people can use to achieve various social goals. In this way, many social psychological phenomena—intergroup behavior, self-presentation, relationship maintenance, and so on—are played out in our conversations. It is not just cognitive processes that undergird our conversations (e.g., keeping track of what is being talked about), but also our interpersonal goals and concerns.

Despite its contributions to our understanding of talk, there are several ways in which the conversation analytic approach is limited. First, in conversation analysis, much of what occurs in a conversation, at least at the surface level, is explained with a limited set of concepts such as adjacency pairs, preference organization, a turn-taking system, and conversational repair. These concepts are part of the conversational system, a system existing independently of the people who use it. It is thus a system readily adaptable for use by anyone, in any context, with any number of participants.

But why does this conversational system exist? The underlying motivation is generally assumed to be economic; the system allows for efficient communication. And no doubt the turn-taking and repairs systems are models of communicative efficiency. But efficiency alone cannot explain the form that conversations take. As we've seen, virtually all of the structural regularities uncovered by conversation analytic researchers are amenable to a face-management interpretation. Talk would look very different if its structure evolved independently of face concerns. It would be fast and explicit and more efficient than it already is. Because acts could not be face threatening, indirectness would probably not exist, moves would not be necessary to open and close conversations, a clear preference structure would not exist, and so on. Unfortunately, the interpersonal underpinnings of conversational structure have not been investigated heavily. But it does seem likely that conversational structure is, in part, interpersonally motivated, a reflection of conversationalists' attempts to manage face, negotiate power, establish relationships (or dissolve them), present a specific identity, and so on.

Second, one of the hallmarks of conversation analysis is the exclusive focus on talk and only talk. No internal concepts—no thoughts, motivations, intentions—are brought to bear on the analysis. In so doing, the researchers are using only that which interactants have available to them as they engage in a conversation. But are our utterances the only thing we respond to when interacting with others? Do

we produce and interpret utterances based only on linguistic input? Clearly not. Language use is nothing if not contextualized; we produce and interpret utterances based on various features of the context including our shared knowledge (see chap. 5). Conversation analytic researchers would reply that context is obviously important in their analyses, and that when it is important interactants will display that importance in their talk. No doubt this is true. But contextual effects can be quite subtle and impact comprehension in ways that may not be apparent in the interaction itself (Holtgraves, 1994).

Is it reasonable to assume, then, that our utterances display our understanding (or lack thereof) of our interlocutors' remarks? Clearly, utterances display an orientation to other utterances, both those occurring previously and those yet to come. Yet not *everything* is revealed in talk. We don't always communicate everything we've inferred, noticed, or thought about regarding another's utterances. For example, in Vuchinich's (1977) experiment in which one conversationalist produced an abrupt, unmarked topic change, all participants reported noticing the shift, but for a subset of them this awareness was not noticeable in their conversational behavior. Thus, our interpretations of others' talk, our sense of what they mean or intend, may remain covert. We sometimes misunderstand one another, and these misunderstandings might never be revealed (e.g., Weizman & Blum-Kulka, 1992). So, as an approach to talk and only talk, conversational analysis is a reasonable approach. But it is obviously not a psychological model of how people produce and comprehend utterances.

Third, conversation analysis takes a unique methodological stance toward the study of conversation. In addition to eschewing all internal concepts, there is the putative rejection of all a priori concepts and theorizing—a presumably pure inductive strategy. But of course the regularities that emerge are not completely free of the analysts' a priori categories. Some conceptual stance must be adopted if any sense is to be made of conversational interaction. For example, the terms that conversation analysts use—rejection, acceptance, and so on—reflect their conceptions of what constitutes those terms.

Conversation analytic researchers have also demonstrated little concern with the representativeness of the conversations on which their analyses are based. Talk is analyzed if it is available; there is no attempt to randomly sample instances of conversational interaction. Therefore, the generalizability of their findings is not known. And it seems likely that there will be cultural variability in many of the regularities that they have uncovered. Take the turn-taking requirement that only one speaker speaks at time (Sacks et al., 1974). Is this universal? Probably not. Reisman (1974) notes, for example, that in Antigua this requirement is relaxed considerably. The initiation of a turn by a speaker is not regarded as a signal for the current speaker to relinquish a turn. According to Reisman, conversations in Antigua have an almost anarchic quality.

Despite these limitations, conversation analysis has brought to light many of the subtle patterns underlying the organization of talk. And it should be noted that the conversation analytic approach is not necessarily incompatible with

traditional, experimental methodologies. Some scholars have called for the merging of these different research traditions (e.g., Robinson, 1998). And fairly impressive attempts have been made in this regard (e.g., Clark, 1996a; Roger & Bull, 1989). Clark (1996a), in particular, has energetically pursued the examination of the psychological properties of conversational structure informed by conversation analytic research. His research is discussed in more detail in the next chapter.

5

Conversational Perspective Taking

One of the most social aspects of language use is the fact that it involves perspective taking. Take the use of a simple directive such as "Close the door." Successful use of this expression will require the recipient to take the speaker's perspective in order to recognize the intention (what speech act is being performed), what door is being referred to, whether the door should be shut right now or later, and so on. And the speaker must have some awareness of the perspective the recipient is likely to take if the utterance is to be understood as intended. In this regard language use is clearly a collective endeavor; one's actions, and the meaning of those actions, are closely intertwined with the actions and interpretations of others.

Language use and perspective taking are linked in a reciprocal manner. Successful language use affords people the chance to take another person's perspective; it allows individuals existing in private worlds to achieve a measure of intersubjectivity, to create a temporarily shared social world with others (Rommetveit, 1974). In other words, perspective taking is an affordance of language use. But at the same time, language use *requires* perspective taking. To understand and be understood requires some attempt to take the perspective of one's interlocutor(s). In order to construct an utterance that will be understood by a recipient, the speaker must try to adopt the hearer's perspective, to see the world (roughly) the way the hearer sees it, and to formulate the remark with that perspective in mind.

LANGUAGE AS SOCIAL ACTION: SOCIAL PSYCHOLOGY AND LANGUAGE USE

Thomas Holtgraves
Ball State University

LEA
2002
Lawrence Erlbaum Associates, Publishers
Mahwah, New Jersey London

President/CEO:	Lawrence Erlbaum
Executive Vice-President, Marketing:	Joseph Petrowski
Senior Vice-President, Book Production:	Art Lizza
Director, Editorial:	Lane Akers
Director, Sales and Marketing:	Robert Sidor
Director, Customer Relations:	Nancy Seitz
Editor:	Bill Webber
Textbook Marketing Manager:	Marisol Kozlovski
Editorial Assistant:	Erica Kica
Cover Design:	Kathryn Houghtaling Lacey
Textbook Production Manager:	Paul Smolenski
Full-Service & Composition:	Black Dot Group/ An AGT Company
Text and Cover Printer:	Sheridan Books, Inc.

This book was typeset in 10/12 pt. Times Roman, Bold, and Italic. The heads were typeset in Americana, Americana Bold, and Americana Bold Italic.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430

Library of Congress Cataloging-in-Publication Data

Holtgraves, Thomas.
Language as social action : social psychology and language use / Thomas Holtgraves.
p. cm.
Includes bibliographical references and index.
ISBN 0-8058-3140-1 (alk. paper)
1. Sociolinguistics. 2. Speech acts (Linguistics) 3. Interpersonal communication.
4. Conversation analysis. I. Title.
P40 .H663 2001
306.44—dc21 2001033703

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

For Marnell, Noah, and Micah