COGNITIVE SCIENCE **13**, 259–294 (1989)

Contributing to Discourse

HERBERT H. CLARK
EDWARD F. SCHAEFER
Stanford University

For people to contribute to discourse, they must do more than utter the right sentence at the right time. The basic requirement is that they add to their common ground in an orderly way. To do this, we argue, they try to establish for each utterance the mutual belief that the addressees have understood what the speaker meant well enough for current purposes. This is accomplished by the collective actions of the current contributor and his or her partners, and these result in units of conversation called contributions. We present a model of contributions and show how it accounts for a variety of features of everyday conversations.

People take part in conversation in order to plan, debate, discuss, gossip, and carry out other social processes. When they do take part, they could be said to *contribute* to the discourse. But how do they contribute? At first the answer seems obvious. A discourse is a sequence of utterances produced as the participants proceed turn by turn. All that participants have to do to contribute is utter the right sentence at the right time. They may make errors, but once they have corrected them, they are done. The other participants have merely to listen and understand. This is the view subscribed to in most discourse theories in psychology, linguistics, philosophy, and artificial intelligence.

A closer look at actual conversations, however, suggests that they are much more than sequences of utterances produced turn by turn. They are highly coordinated activities in which the current speaker tries to make sure he or she is being attended to, heard, and understood by the other participants, and they in turn try to let the speaker know when he or she has succeeded. Contributing to a discourse, then, appears to require more than just uttering the right words at the right time. It seems to consist of collective acts performed by the participants working together.

In this paper we describe a model of contributions as parts of collective acts. We first describe the need for such a model, next present the model

We thank many colleagues for discussions on the issues presented here and Eve V. Clark, Florence H. Edwards, and several reviewers for suggestions on the manuscript. The research was supported in part by Grant BNS 83-20284 from the National Science Foundation.

Correspondence and requests for reprints should be sent to Herbert H. Clark, Department of Psychology, Jordan Hall, Stanford University, Stanford CA 94305.

itself, and then show how it accounts for the commonest devices people use in contributing to conversations. As evidence for the model, we appeal to a large corpus of everyday conversations called the London-Lund corpus (Svartvik & Quirk, 1980). The empirical claim is that the model accounts for the bulk of the successful talk in these conversations.

THE COURSE OF DISCOURSE

Models of discourse differ greatly depending on whether they originate in philosophy (e.g., Kamp, 1981; Lewis, 1979; Stalnaker, 1978), linguistics (Heim, 1983), artificial intelligence (Grosz & Sidner, 1986; Reichman, 1978; Polanyi and Scha, 1985), or psychology (Clark & Haviland, 1977; Johnson-Laird, 1983; van Dijk and Kintsch, 1983). Still, in one way or another, most of them make three assumptions. (1) *Common ground:* The participants in a discourse presuppose a certain common ground. (2) *Accumulation:* In the course of a discourse, the participants try to add to their common ground. (3) *Unilateral action:* The principal means by which the participants add to their common ground is by the speaker uttering the right sentence at the right time. To take Kamp's proposal as an example, the content of a discourse is accumulated in a Discourse Representation Model, or DRM, which is tacitly assumed to be common ground for the participants. With each new utterance, new structures get added to the DRM. These structures are simply assumed to be what the speaker intended; there are no special provisions for making certain they are. The first two assumptions, we will argue with certain qualifications, are necessary for any model of discourse. The third, however, is insufficient to handle a broad class of discourse phenomena that have been systematically excluded from consideration.

When people take part in a conversation, they bring with them a certain amount of baggage—prior beliefs, assumptions, and other information. Part of that baggage is their *common ground,* which Stalnaker (1978) described this way: "Roughly speaking, the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of the conversation. . . Presuppositions are what is taken by the speaker to be the *common ground* of the participants in the conversation, what is treated as their *common knowledge* or *mutual knowledge*" (p. 320, Stalnaker's emphases).[1] Each participant in a conversation, of course, makes his or her own presuppositions. But as Stalnaker noted, "it is part of the concept of presupposition that the speaker assumes that the members of his audience presuppose everything that he presupposes" (p. 321). In actual conversations, the presuppositions vary from one participant to the next, though usually not too drastically.

[1] Here Stalnaker refers explicitly to the technical notions of common and mutual knowledge proposed by Lewis (1969) and Schiffer (1972); for more on the need for this criterion, see Clark & Marshall (1981) and Cohen (1978).

The common ground of the participants in a conversation changes as the conversation proceeds. As Lewis put it, "Presuppositions can be created or destroyed in the course of a conversation. This change is rule-governed, at least up to a point. The presuppositions at time t' depend, in a way about which at least some general principles can be laid down, on the presuppositions at an earlier time t and on the course of the conversation (and nearby events) between t and t'" (1979, p. 339). But even when presuppositions are destroyed, the participants know they have been destroyed, and that knowledge itself becomes part of their common ground. So we can say that the common ground of the participants *accumulates* in the course of a conversation.

Assertions offer an example of how common ground accumulates. Suppose Ann tells Bob she is leaving. As Stalnaker argued, "the essential effect of an assertion is to change the presuppositions of the participants in the conversation by adding the content of what is asserted to what is presupposed. This effect is avoided only if the assertion is rejected" (p. 323). Initially, Ann takes it *not* to be common ground that she is leaving. So as she makes her assertion, Ann and Bob accumulate one more piece of common ground.

Other speech acts add to common ground in other ways. When Ann asks Bob what he is doing, the effect is to add the proposition that she wants him to tell her what he is doing. When Ann promises Bob to hire him, the effect is to add her commitment to hire him. Many things Ann tells Bob require presuppositions for them to be acceptable. In saying, "Even Connie has read *Ulysses,*" Ann takes it for granted that people other than Connie have read Joyce, and that Connie wasn't expected to have. What if those presuppositions are lacking? As Lewis noted, "Say something that requires a missing presupposition, and straightway that presupposition springs into existence, making what you said acceptable after all" (p. 339). So Ann adds to their common ground not only that Connie has read the novel, but also that other people have and that Connie wasn't expected to have. Adding these extra presuppositions has been called *bridging* (Clark & Haviland, 1974, 1977) and *accommodation* (Lewis, 1979). This process is ubiquitous.

But the models of discourse mentioned so far all skirt an essential requirement for the accumulation of common ground—namely, that the participants establish that each utterance has been understood as intended. Suppose that Ann utters "She's leaving" in trying to assert that Connie is leaving her job. That act doesn't automatically add the content of what is asserted to what is presupposed. What if Bob is distracted and doesn't hear Ann? What if he thinks she has uttered "She's sleeping"? What if he thinks she is referring to Diane and not Connie? What if he thinks Connie's leaving her husband and not her job? In these and other cases, Ann's beliefs about their common ground will change in one direction, and Bob's in another. Instead of accumulating, their beliefs will diverge, setting the stage for fur-

ther divergences. Ann and Bob must take positive steps to see that the content of her assertion is added to common ground.

Most models of discourse deal with this issue via the following tacit idealization: Each participant assumes that the content of each utterance is automatically added to common ground. Once Ann has uttered "She's leaving" in the right context, she and Bob are done. He will have understood it as intended, and the two of them can mutually believe this. A few models (e.g., Grosz & Sidner, 1986; Litman & Allen, 1987; Stalnaker, 1978) tacitly make a weaker, conditional idealization: Each participant assumes that the content of an utterance is added to common ground *unless there is evidence to the contrary.* Ann and Bob repair any troubles they encounter with Ann's utterance, and once they have done that, they add its contents to common ground. They don't require or offer *positive* evidence of understanding.

These two idealizations, of course, are just that—idealizations—and in detail incorrect. They work only for conversations with certain features removed.[2] But, as we will argue, many of these features are essential to the process by which common ground actually accumulates. Excluding them misrepresents not only the process of accumulation, but also, ultimately, such phenomena as illocutionary acts, definite reference, repairs, and certain processes of producing and understanding utterances.

What Ann and Bob need are systematic procedures for establishing the mutual belief that Bob has understood what Ann meant. Our proposal is that they do this via a collective process we call *contributing* to discourse (Clark & Schaefer, 1987; Clark & Wilkes-Gibbs, 1986; Isaacs & Clark, 1987; Wilkes-Gibbs, 1986).

CONTRIBUTIONS TO DISCOURSE

Suppose that one person, a *contributor,* wants to contribute something to a conversation with other participants, his or her *partners.* By our proposal, making such a contribution requires two things. One is that the contributor try to specify the content of his or her contribution, and the partners try to register that content. Ann tries to tell Bob that Connie is leaving, and he tries to register that information. This process we will call *content specification.* The second requirement is that the contributor and partners together try to reach the following criterion:

> *Grounding criterion:* The contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for current purposes.

---

[2] Many models of discourse are designed around entirely artificial examples, and the rest, around natural examples sanitized in various ways. The question is *how* sanitized. The London-Lund transcripts too fail to represent certain features, but they include more than most.

Ann and Bob try to establish the mutual belief that he has understood that she believes Connie is leaving. This process we will call content grounding or, simply, *grounding*. It is this process, we propose, that enables common ground to accumulate in an orderly way. It satisfies the common ground and accumulation assumptions and replaces the tacit idealization that is otherwise needed. Together these two processes create a unit of conversation we will call a *contribution*.

What type of unit is a contribution? Traditionally, intentional acts have come in at least two types—*individual* acts and *collective* acts (Clark & Carlson, 1982; Grosz & Sidner, 1989; Searle, 1989). When Ann shakes a stick, plays the piano, or paddles a kayak, she is performing an individual act—an act performed by an individual. When Ann and Bob shake hands, play a duet, or paddle a canoe together, the pair of them are performing a collective act—an act done by a collective, two or more people acting in ensemble. Yet in the hand shake, we can identify three distinct acts:

(a) the collective act of Ann and Bob shaking hands;
(b) Ann's individual act as part of (a);
(c) Bob's individual act as part of (a).

The individual acts in (b) and (c), however, are of a special type. When Ann shakes a stick, her act is *autonomous,* something she could do without coordinating with anyone else. But when she shakes Bob's hand *as part of* (a), her act is something she can achieve only as part of the collective act of shaking hands. What she does just isn't the same as pumping Bob's hand when he is unconscious or not cooperating. The first requires Bob to do his part, and the second doesn't. We can therefore distinguish two types of individual acts: *participatory acts* are those that an agent performs as parts of collective acts; and *autonomous acts* are those that an agent performs on his or her own. Every collective act is performed by means of participatory acts.

The proposal here is that contributions are participatory acts. When Ann tells Bob that she is busy, she is performing an individual act: She is contributing an assertion to the discourse. But this is something she can only do as part of a collective act in which Bob also does his part. Again we can identify three acts:

(a) the collective act of Ann and Bob adding what Ann meant to their common ground;
(b) Ann's individual act of contributing to the discourse as part of (a);
(c) Bob's individual act of registering Ann's contribution as part of (a).

Many units of conversation—words, phrases, clauses, sentences, tone units, and the like—are created by a speaker acting autonomously. But contributions are created by the participants acting collectively. They emerge only as the contributor and partners coordinate actions in just the right way. We

will use the term contribution to refer both to Ann's participatory act and to the collective unit of conversation formed by it.

An Example

Consider this attempt by one man, A, to ask another, B, how much time Norman gets off (from the London-Lund corpus):[3]

A. is it . how much does Norman get off --
B. pardon
A. how much does Norman get off
B. oh, only Friday and Monday
A. m
B. [continues]

Traditionally, one would describe A as having asked B his question by uttering the sentence *How much does Norman get off.* But that description won't do. Although A utters this sentence, he clearly recognizes that he hasn't *succeeded* in asking B his question. Our interest is in how A succeeds.

The process is initiated when A issues the utterance "how much does Norman get off." The trouble is, B apparently doesn't hear it, so he says "pardon" to get A to re-present it, and this A does. This time B responds "oh." In doing so, he asserts his new awareness of what A is doing. He then proceeds to answer A's question by uttering "only Friday and Monday." Note that this answer gives further evidence of B's understanding. If he had replied "about six hundred pounds a month," he would have displayed a misunderstanding, and that would have taken him and A several turns to clear up before going on. As it is, A accepts B's reply with "m," what Schegloff (1982) has called a *continuer,* and that conversation goes on.

What A tries to contribute is a question. For it to be a genuine contribution, A must do more than *try* to ask it: He must believe that he has *succeeded* in asking it. That requires A and B to mutually believe that B has understood it. More precisely, A must come to believe he and B have satisfied the grounding criterion, and so must B. At what point does this occur? Clearly not after A's first "is it . how much does Norman get off," for B has to ask A for a repeat. Nor is it right after A's second "how much does Norman get off." Apparently, B thinks he has understood by that point, but feels the need to let A know by saying "oh." In making this claim public, B establishes the mutual belief that *he* thinks he has understood. But that isn't enough to satisfy the grounding criterion, for A might not accept B's claim. Here, however, A does accept it, for he lets B give his answer. He could still

[3] In our examples we will retain the following symbols from the London-Lund notation: "." for a brief pause (of one light syllable); "-" for a unit pause (of one stress unit or foot); "," for the end of tone unit, which we mark only if it comes mid-turn; "(laughs)" or single parentheses for contextual comments; "((words))" or double parentheses for incomprehensible words; and "*yes*" or asterisks for paired instances of simultaneous talk.

have rejected B's claim of understanding by saying that B's answer was inappropriate, but he doesn't. So A and B reach the mutual belief in B's understanding only with the completion of B's answer, "only Friday and Monday." A's contribution takes A and B four turns and five moves to complete.

This process makes essential use of all the mechanisms available for repair in conversation. As Schegloff, Jefferson, and Sacks (1977) have argued, repairs are organized according to the participants' opportunities for making them. In our example, the first opportunity comes within A's first turn, where the repair would ordinarily be initiated and made by A. As it happens, A makes such a repair. He begins with "is it," interrupts himself, and starts anew with, "how much does Norman get off." The second opportunity comes in the space immediately after A's turn, where the repair would ordinarily be initiated by B and made by A. In our example there is a repair here too. It is initiated by "pardon" and ends with "oh." The third opportunity comes after B's response to A's question—after "only Friday and Monday" —where A could initiate and make the repair, but A doesn't do this.

The model of contributions we are proposing cannot be reduced simply to a system of repairs. The main reason is that the model relies not only on evidence of troubles that need repairing, but on positive evidence of understanding. One of the participants' goals is to reach the grounding criterion, and to do that, they must not only repair any troubles they encounter, but take positive steps to establish understanding and avoid trouble in the first place. Repairs are a necessary ingredient in the model, but they are not sufficient. We will here take for granted what is known about repairs.

Presentation and Acceptance

How do A and B achieve A's contribution to discourse? That is dictated in part by their joint goal—reaching the mutual belief that B has understood A well enough for current purposes. Most contributions begin with an action by A, the contributor. After all, he is the one who will be held responsible for asking the question, so it is usually up to him to initiate it. What happens next depends on both A and B, because they have to arrive at a mutually satisfactory interpretation of that action. The process of contributing divides conceptually into two phases:

> *Presentation Phase:* A presents utterance u for B to consider. He does so on the assumption that, if B gives evidence e or stronger, he can believe that B understands what A means by u.
>
> *Acceptance Phase:* B accepts utterance u by giving evidence e' that he believes he understands what A means by u. He does so on the assumption that, once A registers evidence e', he will also believe that B understands.

We will speak of A *presenting* an action for B to consider, and of B *accepting* that action as having been understood. If these two steps are done right, A

and B will each believe they have arrived at the mutual belief that B understands what A meant by his action. That, in turn, is what it takes for A to contribute to the discourse.

Ordinarily, the presentation and acceptance phases are identified with particular moves by A and B in the conversation. In our example, they are as follows:

Presentation Phase:
A. is it . how much does Norman get off --

Acceptance Phase:
B. pardon
A. how much does Norman get off
B. oh

The presentation phase is completed with A uttering "how much does Norman get off," and the acceptance phase with B's "oh." The way we will view it, A's contribution ends with the initiation of B's answer, "only Friday and Monday." This, however, is something A and B can determine only retrospectively—after B has given his answer without A's objection. Only then can A and B think back and say, "Ah, that was the point at which we completed A's question and B initiated his answer."

In the simplest cases, the presentation phase consists of A uttering a full or elliptical sentence (like "how much does Norman get off") or just a word or phrase (like "pardon" or "only Friday and Monday"). In more complex cases, as we will see, it may consist of two or more contributions, which creates a hierarchy of contributions. In our scheme, every signal that one person directs toward another, whether verbal or nonverbal, is presented for the other person to consider.[4] This way, every utterance and every nonverbal signal belongs to the presentation phase of some contribution.

A should try in this phase to present an utterance that B can understand, and that isn't easy. The main problem is that A has to do this in real time, which often leads to mischosen words, premature commitments, and other errors. To be comprehensible, A has to detect and repair these errors as soon as possible (Schegloff et al., 1977). The result is often a convoluted presentation, such as A's here:

A. I I'm . we're not prepared, to go on being part, I'm not prepared to go on being part of Yiddish literature
B. yeah
A. we must ha- we're . big enough to stand on our own feet now
B. yes

[4] By signal, we mean any act by which the speaker means something in Grice's (1957) sense of nonnatural meaning; that is, it must involve what Grice called m-intentions.

A's self-repairs must themselves make clear what is being revised and how, and that also takes care. As Levelt (1983) has shown, speakers make self-repairs as soon as they detect a problem, and they almost always succeed in making them structurally unambiguous. The point is, a presentation is more than the uttering of a sentence. It is the creation in real time of a spoken structure from which the partner can identify the words, phrases, and sentences that the contributor intended as final.

Still, the acceptance phase is where most complications arise. It is generally initiated by the partner B indicating his state of understanding at that moment. There are two main cases to consider—when B indicates understanding, and when he indicates trouble understanding.

Evidence of Understanding

The acceptance phase is usually initiated by B giving A evidence that he believes he understands what A meant by *u*. B's evidence can be of several types. He can say that he understands, as with "I see" or "uh huh." Or he can *demonstrate* that he understands. One way is by showing what it is he understands, as with a paraphrase, or what it is he heard, as with a verbatim repetition. Another is by showing his willingness to go on. The least obvious way is by showing continued attention. When B reveals no change in his attentive demeanor or eye contact, he implies that he hasn't detected any problems—that he believes he is understanding well enough for current purposes. The five main types of evidence, then, are these;

1. *Continued attention.* B shows he is continuing to attend and therefore remains satisfied with A's presentation.
2. *Initiation of the relevant next contribution.* B starts in on the next contribution that would be relevant at a level as high as the current one.
3. *Acknowledgement.* B nods or says "uh huh," "yeah," or the like.
4. *Demonstration.* B demonstrates all or part of what he has understood A to mean.
5. *Display.* B displays verbatim all or part of A's presentation.

These types are graded roughly from weakest to strongest.

But what type of evidence *should* B present? Most presentations carry some indication of the strength of evidence A expects in order to convince him that B has understood. The presentation of a telephone number may project a verbatim display of that number (Clark & Schaefer, 1987). Other presentations may project an acknowledgement like "uh huh." Still others may project merely continued attention. Generally, the more complicated A's presentation, or the more demanding the current purpose, the more evidence should be needed to convince A that B has understood. What evidence is needed for which presentations is an empirical issue that we will examine.

Note that the acceptance process is recursive. B's evidence in response to A's presentation is itself a presentation that needs to be accepted. But where does the recursion stop? Suppose A presents "I'm leaving tomorrow." Why isn't it possible for B to accept that by presenting "m," which A accepts by presenting "m," which B accepts by presenting "m," and so on ad infinitum? What keeps the process from spinning out indefinitely? The answer, we propose, is this:

> *Strength of evidence principle:* The participants expect that, if evidence e_0 is needed for accepting presentation u_0, and e_1 for accepting the presentation of e_0, then e_1 will be weaker than e_0.

B may accept A's presentation by uttering "m," but they expect something weaker to be able to accept that "m." The upshot is that every acceptance phase should end in continued attention or initiation of the next turn, the weakest evidence available.

If this rule is correct, recursion should rarely go beyond two or three cycles, and it rarely does. Here is an acceptance phase with three cycles (again from the London-Lund corpus):

A. F . six two
B. F six two
A. yes
B. thanks very much

A presents a book identification number "F six two." B accepts the number by displaying it verbatim. A in turn accepts the display by the weaker evidence of "yes." Finally, B accepts the "yes" by proceeding to the next contribution (see Clark & Schaefer, 1987). The final step completes the acceptance phase of A's original contribution.

Evidence of Trouble in Understanding

There will be times, of course, when B doesn't hear or understand A's presentation entirely, as in the Norman example, and then B should initiate the acceptance phase by giving evidence of that trouble. Now, for any u' that is part of A's presentation, B could believe he is in any one of four successively stronger states of understanding (Clark & Schaefer, 1987):

State 0: B didn't notice that A uttered any u'.
State 1. B noticed that A uttered some u' (but wasn't in state 2).
State 2. B correctly heard u' (but wasn't in state 3).
State 3. B understood what A meant by u'.

Ordinarily, state 3 presupposes state 2, which presupposes state 1, though B may sometimes believe he understands what A meant without knowing precisely what A presented. And, of course, B may be in different states for

different parts of A's presentation. A and B's goal is to reach the mutual belief that B is in state 3 for the entire presentation.

The route by which A and B reach that goal depends on the partner's initial assessment of his understanding. In the Norman example, B indicates (with "pardon") that he is in state 1 for A's entire initial presentation, and that leads A to repeat it. After the repeat, B indicates (with "oh") that he is in state 3, and that allows him to go on to his answer. If B had initially indicated something else—for example, "Norman who?"—the acceptance phase would have taken a different course (see Clark & Schaefer, 1987). As with positive evidence, each part of the acceptance phase is itself a contribution. B's "pardon" is the presentation phase of a request; A's repeat of "how much does Norman get off" is the presentation phase of a response to it; and B's "oh" is the presentation phase of a type of assertion. Each of these utterances initiates a contribution that is hierarchically subordinate to A's original contribution—his question.

One assumption of the model is the principle of least collaborative effort (Clark & Wilkes-Gibbs, 1986). The idea is that the participants in a contribution try to minimize the total effort spent on that contribution—in both the presentation and the acceptance phases. Each time A initiates a contribution, he tries to anticipate how much effort it will take him and B, and he designs his presentation to minimize it. That means, for example, that A should repair his own errors as he goes along rather than leave them for B to deal with. Self-repairs by A usually take less total effort than repairs initiated by B. Empirically, indeed, self-repairs are preferred to repairs initiated by others (Schegloff et al., 1977). Generally, the more effort spent on designing the right presentation, the less effort is needed for acceptance. The problem is how to distribute the effort in order to minimize it, and that depends on both systematic and accidental features of the situation.

Patterns of Contributing

The model of contributions proposed here is mainly a logic for the process of adding to a discourse. How the process actually gets played out should depend on several factors. One is the evidential devices available. Face to face, people can nod, smile, and display mutual gaze; on the telephone, they cannot. Face to face and on the telephone, people can exploit the precise timing of their utterances, as with brief interruptions (Jefferson, 1973); on computer terminal hookups, they cannot (Cohen, 1984). The course that contributions take should depend on the availability, effectiveness, and efficiency of devices such as these. It should also depend on the nature of the discourse at hand. Task-oriented dialogues, for example, may require stronger evidence of understanding on average than casual discussions (Clark & Wilkes-Gibbs, 1986; Cohen, 1984). In telephone calls to directory enquiries, for example, the caller and operator set a high criterion to establish precise

names, addresses, and telephone numbers (Clark & Schaefer, 1987). Other factors should be important too.

Still, the model of contributions leads to three general predictions. First, if contributions are necessary for successful conversations, their presentation and acceptance phases should be identifiable in such conversations. Second, the forms the two phases take should depend on the evidential devices available and the requirements placed on that evidence. And third, presentation and acceptance phases should emerge as hierarchical structures reflecting the recursive process by which they are created. For evidence, we will look to the London-Lund corpus, including previous studies of the corpus by Oreström (1983), Stenström (1984), and Thavenius (1983). This corpus is a vast collection of casual British conversations surreptitiously tape recorded in and around university settings. The transcripts are marked for pauses, overlapping speech, intonation, and loudness, but not for gestures or eye contact. Since we cannot test for the use of continued attention or eye gaze (see Goodwin, 1981), our analyses must remain incomplete in this respect. All of the examples cited in this paper are taken from this corpus.

Two patterns of contributions dominate this corpus. One occurs every time there is a relevant, orderly change in turns, and the other, every time a partner adds a "yes" or "uh huh" or "m" in the background. We will examine the logic behind these two patterns first. Yet the less common patterns of contributing are important in their own right, so we will examine some of those as well. We will argue that these patterns, taken together, account for most though not all patterns that occur in everyday conversation.

CONTRIBUTIONS BY TURNS

The commonest form of contribution coincides with the turn. Almost every time a speaker starts a new turn, he or she either (a) accepts what the last speaker has just said or (b) initiates a repair of the problem they ran into in accepting it. In this way, a new contribution is initiated with each relevant change in turn. The question is, how is this done?

Many turns in conversation are organized in adjacency pairs (Schegloff & Sacks, 1973). The prototype is the question-answer pair, as in the first two turns of this example:

A. how far is it from Huddersfield to Coventry .
B. um . about um a hundred miles -
A. so, in fact, if you were . living in London during that period, . you would be closer - .

Adjacency pairs consist of two ordered utterances, the first and second pair parts, produced by two different speakers. The two parts (here, the question and answer) come in types that specify which is to come first and which second; the form and content of the second part (here, the answer) depends on

TABLE 1
Types of Adjacency Pairs

Type of Adjacency Pair		Example	
First Part	Second Part	A's Utterance	B's Response
Question	Answer	Where is Connie?	At the store.
Request	Compliance/Refusal	Please pass the horseradish.	[B passes horseradish.]
Request	Acceptance/Rejection	Please pass the horseradish.	Okay.
Proposal	Acceptance/Rejection	Here is your change.	[B takes it.]
Offer	Acceptance/Rejection	Would you like some coffee?	Yes, thanks.
Invitation	Acceptance/Rejection	Come to dinner Sunday	Okay.
Apology	Acceptance/Rejection	Sorry.	Oh, that's all right.
Thanks	Acceptance/Rejection	Thank you.	You're welcome.
Assessment	Agreement/Disagreement	That film was terrible	Yes, it was.
Compliment	Agreement/Disagreement	Your new coat is beautiful	Yes, it's nice.
Summons	Answer	Hey, Ben	Yes?
Greetings	Greetings	Hi, Ben.	Hi, Ann.
Farewell	Farewell	Goodbye.	Goodbye.

the type of the first part (the question). One crucial property is conditional relevance. Given a first pair part, a second pair part is conditionally relevant, that is, relevant and expectable, as the next utterance. Once A has asked the question, it is relevant and expectable for B to answer it in the next turn. Other types of adjacency pairs are illustrated in Table 1.

Conditional Relevance

These features of adjacency pairs are systematically exploited by people contributing to discourse. Take A's question, which he initiates by presenting, "How far is it from Huddersfield to Coventry?" How should B initiate the acceptance phase? If she thinks she understands A's presentation, she can reach her goal most efficiently by initiating her answer immediately. In doing that, she gives three types of evidence of understanding at once: (1) By passing up the chance to ask A for a repair, she indicates that she believes she has understood A's contribution. (2) By initiating an answer, a second pair part, she shows that she recognizes that A has asked a question, a first pair part. She does this by exploiting the conditional relevance of a second pair part given the first. (3) By formulating the answer she does, she displays part of her understanding of what particular question was asked. Her "Um about um a hundred miles" is consistent with A having asked a WH-ques-

tion about distance. So when A accepts B's answer as an appropriate one, he also accepts 1, 2, and 3, B's evidence that he has understood what A meant. The crucial point is this: Giving an answer can be used to accept the presentation of a question by virtue of its conditional relevance. That holds for the second part of any adjacency pair.

Answers, of course, are also contributions, so they too should have presentation and acceptance phases. The commonest way to accept an utterance as an answer is to exploit a slightly different form of conditional relevance. Take B's answer that it is about a hundred miles from Huddersfield to Coventry. This is not the first pair part of an adjacency pair. And yet, once it is on record, it is relevant and expectable that A will proceed to the use he wants to make of that information. That is, after the second part of an adjacency pair, it is conditionally relevant immediately to initiate the next contribution at the same level as those two parts.

We can see how this works in our example. Once B has uttered, "Um about um a hundred miles," it is conditionally relevant for A to use this information, and he does. He initiates his acceptance by proceeding to the next contribution at the same level as B's answer, "So in fact if you were living in London during that period you would be closer." In this way, A gives the same three types of evidence that B did with her answer: (1) he passes up the opportunity to ask B for a repair on her utterance; (2) he shows his recognition that B has answered his question; and (3) he displays part of his understanding of B's answer (by drawing a reasonable conclusion from it). As Sacks et al. (1974, p. 728) noted, "Regularly, then, a turn's talk will display its speaker's understanding of a prior turn's talk, or whatever other talk it marks itself as directed to" (see also Goffman, 1976). So A accepts the presentation of B's answer via much the same rationale as B uses in accepting the presentation of A's question.

Contributions like A's question, B's answer, and A's next question can be represented in what we will call *contribution trees*. The tree for this example, shown in Figure 1, illustrates several features of contributions. First, every contribution (*C*) has a presentation phrase (*Pr*) and an acceptance phase (*Ac*). Second, every utterance belongs to the presentation phrase of some contribution. Third, as a result, most contributions are ultimately completed by the partner initiating some next contribution. (The rest are completed by evidence of continued attention.) We have denoted this by drawing a slanting line from *Ac* of the first contribution to the presentation of the next. And fourth, a contribution C_2 belongs to the acceptance phase of a previous contribution C_1 only if it directly addresses the hearing or understanding of *Pr* of C_1. Although Figure 1 has no embedded contributions, we will present other trees that do.

A word of caution. Contribution trees are not fixed beforehand. They emerge piece by piece as the participants construct them in collaboration.

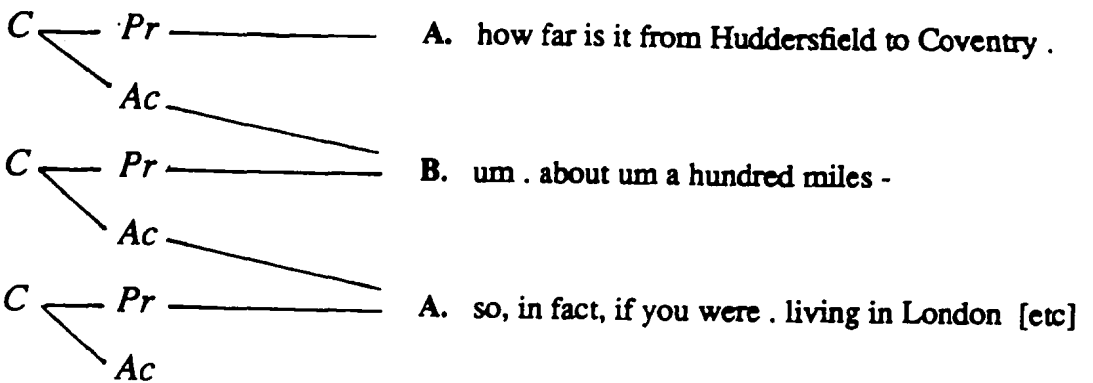

C
Pr
Ac
C
Pr
Ac
C
Pr
Ac
A. how far is it from Huddersfield to Coventry .
B. um . about um a hundred miles -
A. so, in fact, if you were . living in London [etc]

Figure 1.

They are often revised en route, and the function of certain utterances is determined only retrospectively. Revisions and retrospective identifications are impossible to capture in static trees. Yet we have tried to represent some of the emergent properties by placing on each line the contribution, presentation, or acceptance that the current speaker believes he or she is working on at the moment. We cannot always be sure even of these beliefs, and in some trees they have plausible alternatives.

Another way to initiate the acceptance phase of an answer, or of the second part of most adjacency pairs, is by explicit assertions of understanding. Take this example:

A. are you going to America
B. yes
A. m
A. Iz . tried to go to America earlier this year [continues]

A initiates his acceptance of B's answer by uttering "m," explicitly asserting that he believes he has understood it, and then immediately goes on. Stenström (1984) has called these moves *follow-ups.* In her analysis of questions and answers in the London-Lund corpus, answers had follow-ups 41% of the time—36% of the time in face-to-face conversations and 50% of the time in telephone conversations—so they are common. The contribution tree for this example is shown in Figure 2.

On occasion, however, the questioner will understand an answer but reject it as inappropriate because it reveals a misunderstanding of the question, as in this example:

B. k who evaluates the property ---
A. uh whoever you ask((ed)), . the surveyor for the building society
B. no, I meant who decides what price it'll go on the market -
A. (- snorts) . whatever people will pay --
B. but why was Chetwynd Road so cheap ---

A's answer in line 2, "uh whoever you asked -- the surveyor for the building society," shows B that she has misinterpreted the word *evaluates.* So B rejects her acceptance with "no" and rephrases what he meant, "who decides what price it'll go on the market." This time A presents as her answer "whatever people will pay," which B accepts by going on. B's "no, I meant..." is often called a third turn repair.

Consider the acceptance phase of B's question. Ordinarily, two partners treat the initiation of an appropriate answer as completing the acceptance of the question. But, as we said, they can only do that retrospectively, since they have to wait on the questioner accepting the answer as evidence of correct understanding. In this example, B rejects A's first answer as inappropriate. So A's answer, instead of initiating the next contribution, is now

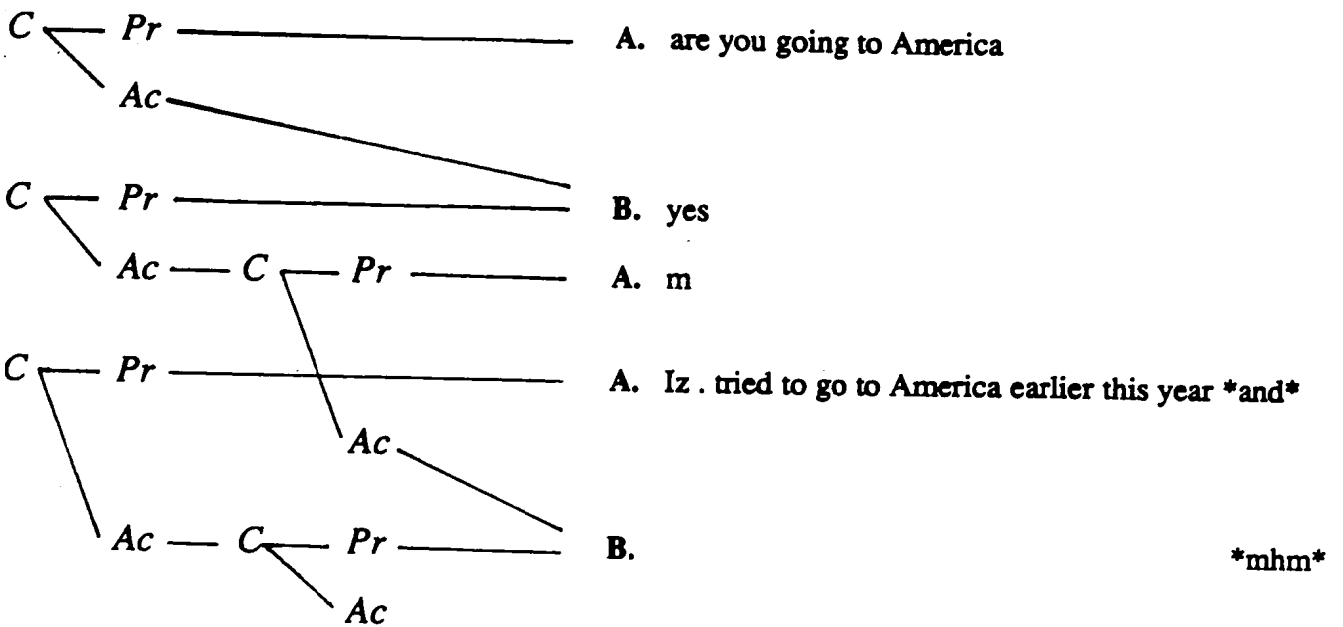

C
Pr
Ac
A. are you going to America
C
Pr
B. yes
Ac
C
Pr
A. m
C
Pr
A. Iz . tried to go to America earlier this year *and*
Ac
Ac
C
Pr
B. *mhm*
Ac

Figure 2.

treated as the first move in the acceptance phase of B's question. The acceptance phase gets completed only with the initiation of A's second answer, "whatever people will pay," which B does accept. The contribution tree that results is shown in Figure 3. It shows how A first accepts B's question by trying to answer it, but when B rejects A's interpretation of the question, the two of them leave A's answer high and dry, abandoning it altogether.

On still other occasions, a participant's attempt to contribute will fail when his or her presentation is ignored altogether. Consider this example:

C. well . I've got um . a boy ex Gordonstoun - .
B. I say
C. who sticks out like a *sore thumb*
B. *what* what's his name then Charlie --
C. and I've got . several flower people
B. ooh uh tha- that's nice.
C. oh it isn't actually, cos I've been giving them dictation - in English, . because their spelling's hopeless, their punctuation's worse you know
B. yeah

When B says "What's his name then Charlie?" he appears to initiate a contribution asking C for the name of the ex-Gordonstoun boy. As it happens, C ignores B's presentation, and B lets it drop. So although B tried to ask a question, there is no evidence that he succeeded. He simply failed to contribute to the discourse. The conclusion is important: Acceptance requires more than just going on to a next contribution. It must be a *relevant* next contribution—such as an answer.

So far, then, we have two powerful methods for accepting presentations. Partners can accept a presentation—almost any presentation—by proceeding to a relevant next contribution at the same level as the current one. Or they can explicitly assert that they understand with a "yes," "right," or "I see."

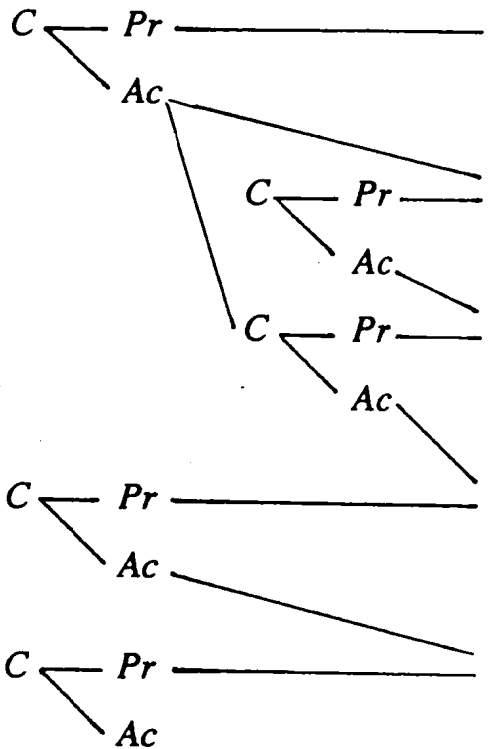
Side Sequences

In adjacency pairs, when one partner doesn't accept the first pair part, he or she will usually initiate a repair sequence, as here:

A. ((where *are* you))
  B. m? .
  A. where *are* you.
B. well I'm still at college.
A. [continues]

B apparently doesn't hear A's question and asks for a repeat. Or consider this example:

A. well wo uh what shall we do about uh *this* boy then
  B. Duveen?



B. k who evaluates the property - - -

A. uh whoever you ask((ed)), . the surveyor for the building society

B. no, I meant who decides what price it'll go on the market -

A. (- snorts) . whatever people will pay - -

B. but why was Chetwynd Road so cheap - - -

Figure 3.

A. m
B. well I propose to *write,* uh saying . I'm very sorry I cannot, uh teach . at the *institute* . [continues]

B seems unclear about A's reference—which boy?—and initiates a repair to clear it up. In both examples, A and B step aside for two turns (the indented ones) to clear up the problem before B initiates his answer.

The device A and B use in the two indented turns are *side sequences* (Jefferson, 1972).[5] Once A has presented "Well wo uh what shall we do about uh this boy then?" B initiates the acceptance phase by opening up a side sequence to clear up A's reference to the boy. He asks a question (with "Duveen?"), which A tries to answer (with "m"). But that doesn't complete the acceptance phase of A's question. B needs to accept both A's original presentation and A's "m." A and B, however, each recognize that, once the side sequence is completed, it is again conditionally relevant for B to answer A's question. So B can complete his acceptance of A's question by initiating that answer, "Well I propose to write uh saying I'm very sorry I cannot teach at the institute." In doing this, he also gives evidence he has understood A's "m," closing the side sequence. Unless A rejects B's answer, A's contribution is complete. The tree that results is shown in Figure 4.

Side sequences can be initiated not just after first pair parts, but after almost any presentation. What makes them so useful is that they allow the partners to focus on precisely those features of a presentation that are troublesome. They can focus on general hearing, as with "What?" or on highly specific information, as with "Duveen?" (see Clark & Schaefer, 1987; Schegloff et al., 1977). Also, side sequences can be extended until the problem is cleared up, as in this five-turn side sequence:

A. what film have you been to see .
B. film .
A. I thought you went . you were going to the National - Theatre - National Film Theatre
B. no no, . um . that was at the weekend, - .
((we were discussing)) the weekend, remember
A. *oh yes* - . yes . yes
B. I'm going to see it's uh -- (- sighs) ((it's called)) il Posto - it's uh - . Olmi ((I think))

Side sequences like these are closed by the participants proceeding on to the next contribution at a level as high as the contribution of which they are a part. They are one of the commonest and most versatile grounding devices available.

[5] Specialized types of side sequences have been studied under the terms insertion sequence by Schegloff (1972), clarification and correction subdialogues by Litman and Allen (1987) and debugging explanations by Grosz and Sidner (1986).

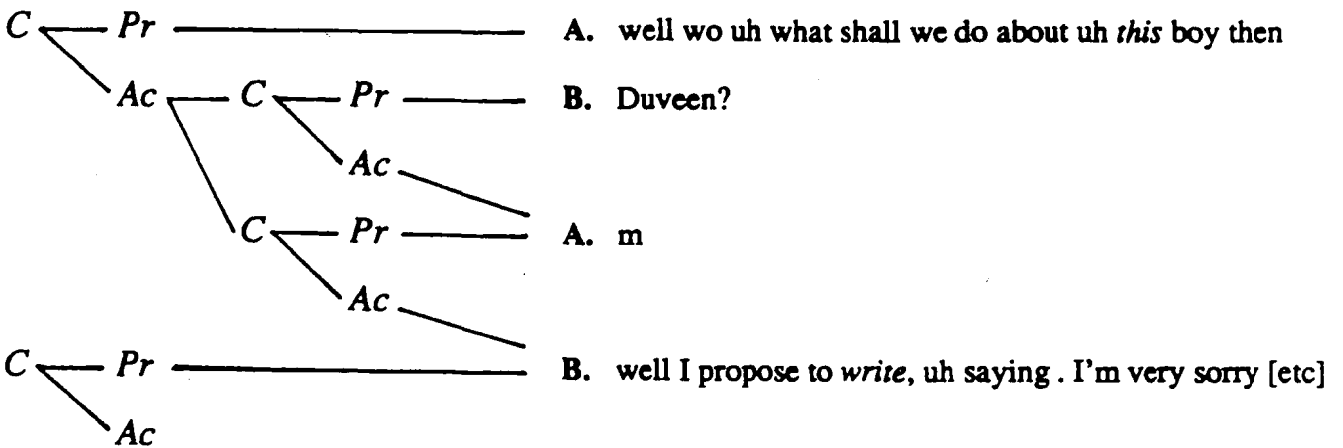

C
Pr
A. well wo uh what shall we do about uh *this* boy then
Ac
C
Pr
B. Duveen?
Ac
C
Pr
A. m
Ac
C
Pr
B. well I propose to *write*, uh saying . I'm very sorry [etc]
Ac

Figure 4.

All this leads to a fundamental generalization about contributions: *A new contribution is initiated with every cooperative change in turns.* Such a contribution closes the current one when it is relevant and at a level as high as the current one—subject to a veto by the current contributor. Or it opens a side sequence that is closed when a relevant new contribution is initiated at a level above the side sequence. Although contributions created this way can be of any length, they tend not to go on too long. In Oreström's (1983) analysis of the London-Lund conversations, there was a new turn after a median interval of 13 words; two-thirds of all turns were less than 20 words long. So contributions formed this way are of a practical length. They are long enough for a contributor to perform complete illocutionary acts like assertions, questions, and offers. They are also short enough to lie within the participants' memory limitations and allow troubles to get repaired before they snowball.

CONTRIBUTIONS WITHIN TURNS

When a speaker takes an extended turn—as in an anecdote, long description, or involved instruction—the partners ordinarily accept separate portions of it by means of what we will call *acknowledgements.* These are expressions such as *mhm, yes,* and *quite* that are spoken in the background, or gestures such as head nods and smiles. Consider this example:

B. but you daren't set synthesis again you see, . you set analysis, and you can put the answers down, and your assistant *examiners will work them,*
A. *yes quite, yes, yes*
B. but if you give them a give n them a free hand on synthesis, and they'd be marking all sorts of stuff, because they can't do the stuff *themselves,*
A. *quite m*
B. I must watch [continues]

As B comes to the end of certain thoughts, A acknowledges them with *yes, quite,* and *m* in various combinations. How do these work?

Acknowledgements fall into two major categories—*continuers* and *assessments.* People use continuers such as *m, yes, quite,* and *I see,* according to Shegloff (1982), to display continued attention and to indicate they recognize that the primary speaker is in the middle of an extended unit of speech. They use them to signal that they are passing up the opportunity to initiate a repair on the turn so far and, by implication, that they believe they have understood it so far. Continuers are signals for the primary speaker to continue. With assessments such as *gosh, really?, oh,* and *good God,* people do all this as well as offer brief assessments of what is said. Acknowledgements generally do *not* constitute turns. When A says, "yes quite, yes, yes," he is speaking without taking the floor (Schegloff, 1982).

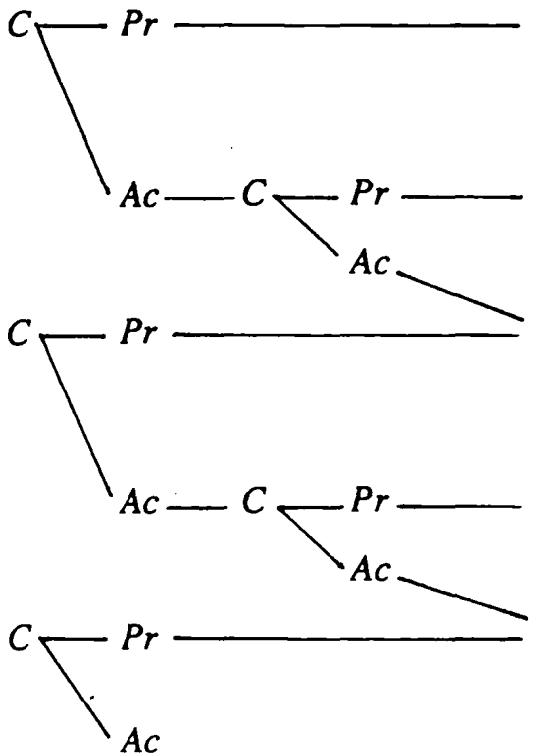
Acknowledgements are generally placed at or near the ends of major grammatical constituents. In the London-Lund corpus, continuers occur right at grammatical boundaries 77% of the time and near such boundaries most of the rest of the time (Oreström, 1983). And they tend to overlap with the primary speaker's talk. In the London-Lund corpus, they do so 45% of the time overall—and even 33% of the time when they occur right at grammatical boundaries (Oreström, 1983). Assessments, which are much rarer, also tend to occur at or near grammatical boundaries, though they are generally engineered to occur without overlap (Goodwin, 1986).

Acknowledgements are placed where they are in order to mark the scope of their acceptance. In our example, A places "yes quite, yes, yes" over the last phrase of B's utterance to indicate acceptance of B's presentation up through that phrase. He places "quite m" over the last phrase of B's next utterance to indicate acceptance of everything from the last acknowledgement through this phrase. Each acknowledgement marks the final boundary of his scope, though not always with the precision shown in our example.

But an acknowledgement must itself be understood and accepted as evidence of understanding of the material in its scope. B can mishear or misunderstand it—"Pardon me?" or "What?" Even if he understands it, he may reject the claim A is making with it—that he has understood B's presentation. B might respond, "Are you really listening?" or "Do you understand?" (Think of half-listening spouses or children.) A has to understand B's acknowledgement and accept it as adequate evidence.

This may be a problem in principle, but it is rarely a problem in practice. Partners generally use acknowledgements only when they are quite confident that they understand and that the contributor isn't expecting strong evidence. That helps explain why acknowledgements are so brief and are reduced in loudness and lower in pitch (Oreström, 1983): The partners don't expect them to need or be given much consideration. The contributors, in turn, should find them easy to understand and accept. All they need to do to accept them is show continued attention and proceed without a break to the next contribution at the same level as their last one. The contribution tree for our example is shown in Figure 5.

Acknowledgements such as *yes, quite,* or *m,* therefore, divide extended turns into units that are practical for establishing understanding and correcting misunderstandings. The participants complete a contribution at the end of every major grammatical unit in which the partner offers an acknowledgement. In Oreström's (1983) analysis of turns 30 words or longer in the London-Lund conversations, there was an acknowledgement after a median interval of only nine words; 80% of the time there was at least one acknowledgement every 15 words. And since these conversations were face-to-face, there were probably also head nods, smiles, and other non-verbal acknowledgements. The contributions created this way are about the same size as



B. but you daren't set synthesis again you see, .
you set analysis, and you can put the answers down,
and your assistant *examiners will work them,*

A. *yes quite, yes, yes *

B. but if you give them a give n them a free hand
on synthesis, and they'd be marking all sorts of stuff,
because they can't do the stuff *themselves, *

A. *quite m *

B. I must watch [etc]

Figure 5.

those created by turns. Acknowledgements, in brief, are backgrounded attempts by partners to create contributions from extended turns, and they almost always succeed.

CONTRIBUTIONS VIA SENTENCE PARTS

Most of the contributions discussed so far have been associated with sentences. The contributor presents a full or elliptical sentence, like "how much does Norman get off" or "who evaluates the property" or "oh," and it is accepted as a whole. These contributions are used for asking questions, making assertions, making requests—that is, for performing illocutionary acts à la Austin (1962). When we think of contributions, this is what we normally think of—people contributing to conversation *by means of* questions, assertions, requests, and other illocutionary acts.

Many contributions, however, are associated with only parts of sentences —usually single words or phrases. With these, contributors perform only parts of illocutionary acts—such propositional acts as referring, naming, denoting, and predicating (Searle, 1969). But why contribute anything so small as that? There is usually a special reason. The contributor is uncertain about some piece of information, or needs help on some item, or wants to present an utterance too complex to be understood in one piece. In this way, contributions via sentence parts appear to be less preferred than contributions one or more sentences in size. Still, they are common enough. We will examine only three types of such contributions—installment contributions, trial constituents, and collaborative completions.

Installment Contributions

When speakers have complicated information to present, they often take more explicit steps to make sure they are being understood. One way is by presenting the information in *installments.* In this example, A has just asked B on the telephone, "Could you possibly tell me what Sir Humphrey Davy's address is—Professor Worth thought you might know," and B is answering:

B. Banque Nationale de Liban ---
A. yes
B. nine to thirteen.
A. sorry
B. nine . to . thirteen
A. yeah .
B. King Edward Street --
A. yeah -
B. London .
A. yes
B. NE two P -
A. yes -

B. four AF -
A. F -
B. yes
A. thanks very much.

What precisely is going on here?

B's answer is accomplished in six *installments,* each taking the form of a contribution, with a presentation and an acceptance phase. B divides her presentation into six parts, first "Bank Nationale de Liban," then "nine to thirteen," and on through "four AF." Furthermore, she pronounces these as items in a list, placing a rising or fall-rise intonation on the first five installments and a falling intonation on the last. So when B pronounces "Liban" with a rising intonation and then pauses, she indicates that this is only the first segment of her presentation and invites A to indicate his understanding of it. A accepts that presentation with "yes." As for the second segment, A doesn't understand it the first time around, so B re-presents it, this time with pauses between the words. The contribution looks like this:

Presentation Phase:
B. nine to thirteen .

Acceptance Phase:
A. sorry
B. nine . to . thirteen
A. yeah .

The remaining four installments are accepted separately as well.

These six contributions are themselves parts of a more inclusive contribution, which has its own presentation and acceptance phases. Its presentation consists of the six installment contributions, and its acceptance is achieved by A initiating the next contribution at the same level, "thank you very much." The contribution tree for B's entire answer is shown is Figure 6 (see also Clark & Schaefer, 1987). So dividing a presentation into installments creates hierarchical structure in the presentation phase of a contribution.

Why use installment presentations? One reason is to enable the partners to register the presented information verbatim, perhaps to write it down, as with addresses, telephone numbers, and recipes (see Clark & Schaefer, 1987; Goldberg, 1975). Another reason is to make sure, in instructing addressees, that they have understood each step before going on (Cohen, 1984; Clark & Wilkes-Gibbs, 1986). But installment presentations also crop up in quite ordinary descriptions, as here:

B. how how was the wedding -
A. oh it was it was really good, it was uh it was a lovely day
B. yes
A. and . it was a super place, . to have it . of course
B. yes -

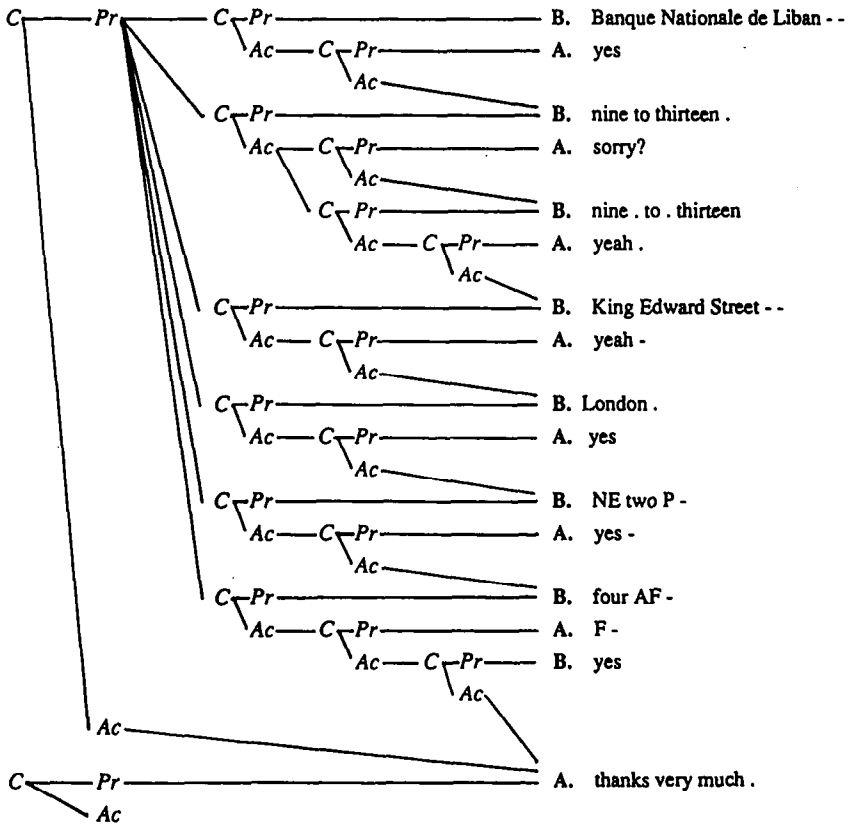


Figure 6.

A. and we went and sat on sat in an orchard, at Grantchester, and had a huge tea *afterwards (laughs -)*
B. *(laughs --)*.
A. **uh**
B. **it does** sound, very nice indeed

By presenting her description in installments, A gets B to help her complete her extended answer without interruption (Schegloff, 1982).

Trial Constituents

Another way of grounding mid-presentation is with *trial constituents.* Sometimes speakers find themselves about to present a name or description that they aren't sure is factually correct or entirely comprehensible. They can present that constituent—usually a name or description—with what Sacks and Schegloff (1979) have called a *try marker,* a rising intonation followed

by a slight pause, and get their partners to confirm or correct it before completing the presentation. Consider this example:

A. so I wrote off to . Bill, . uh who ((had)) presumably disappeared by this time, certainly, a man called Annegra? -
B. yeah, Allegra
A. Allegra, uh replied, . uh and I . put . two other people, who'd been in for . the BBST job . with me [continues]

Apparently, A is trying to assert, "A man called Annegra replied, and I . . . ," but he becomes uncertain about the name Annegra. He therefore presents "Annegra" as a trial constituent with rising intonation and a slight pause. B responds "yeah" to confirm that she knows who he is trying to refer to, and then corrects the name to "Allegra." A then accepts the correction by re-presenting "Allegra" and continuing on. The entire correction is made swiftly and efficiently. What we have is a local contribution, complete with its own presentation and acceptance phases, as follows:

Presentation Phase:
A. Annegra? -

Acceptance Phase:
B. yeah, Allegra

This is embedded within A's larger presentation of "a man called Allegra replied," as shown in the contribution tree in Figure 7.

Completions

In initiating a contribution, speakers usually present an utterance for their partners' consideration. Sometimes, however, that utterance is completed by the partners, as here:

A. um the problem is a that you(('ve)) got to get planning consent -
B. before you start -
A. before you start on that part, yes
A. you can do anything internally, you wish
B. but the big stuff is, the external stuff [continues]

A begins, "um the problem is that you've got to get planning consent," and pauses, perhaps searching for a way to express what she wants to say next. This leads B to offer a plausible completion, "before you start." Is the completion appropriate? That is up to A, and indeed she accepts it by repeating "before you start," amending it with "on that part," and asserting acceptance with "yes." Once that is completed A continues with her turn. As Wilkes-Gibbs (1986; see also Lerner, 1987) has argued from an extensive corpus of completions, they tend to follow these steps:

1. A presents a sentence fragment.
2. A may indicate she is having trouble completing it.

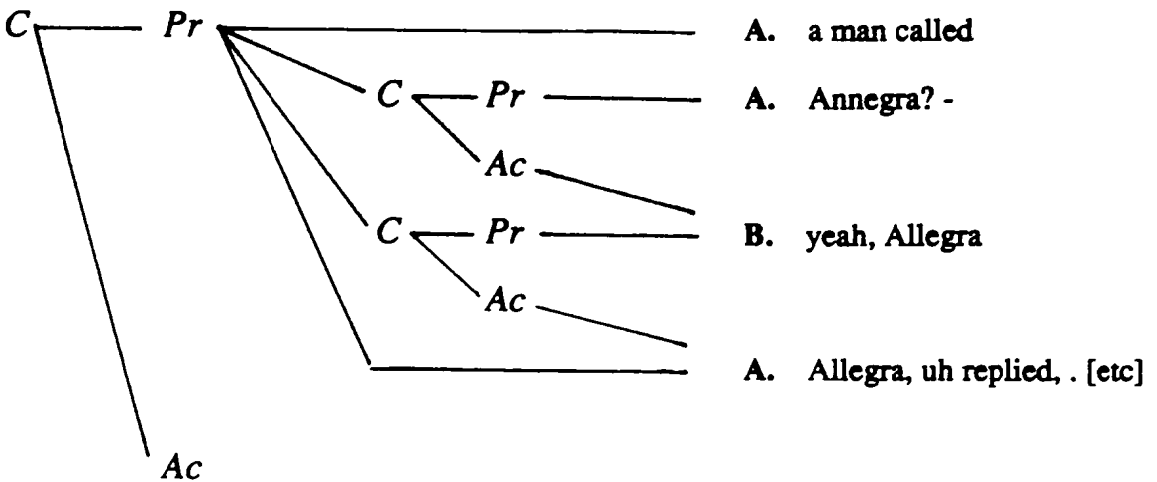

C
Pr
Ac
C
Pr
Ac
C
Pr
Ac
A. a man called
A. Annegra? -
B. yeah, Allegra
A. Allegra, uh replied, . [etc]

Figure 7.

3. B offers a completion, often with a questioning intonation.
4a. A may explicitly accept B's completion ("yes") or reject it ("no").
4b. B may repeat A's completion verbatim, or B re-presents it correctly.
5. The conversation continues.

Explicit rejections at step 4a tend to precede the correction at step 4b, but explicit acceptances tend to follow the repetition, yielding 4b then 4a. And when these two steps don't occur, the completion gets accepted in some other way. In our example, we find at least five of these six steps. Collaborative completions are surprisingly common in everyday conversation.

If B has completed A's sentence, then whose assertion, whose contribution, is it? Actually, there are two contributions. One is A's when she asserts, "the problem is that you've got to get planning consent before you start on that part." Even if she hadn't repeated "before you start," she would have been held accountable for having made this claim; she would have produced the first fragment and accepted B's version of the rest of it. But B has also made a contribution with his completion. In offering it, he accepts her first fragment and helps complete it. B's contribution takes this form:

Presentation Phase:
B. before you start -

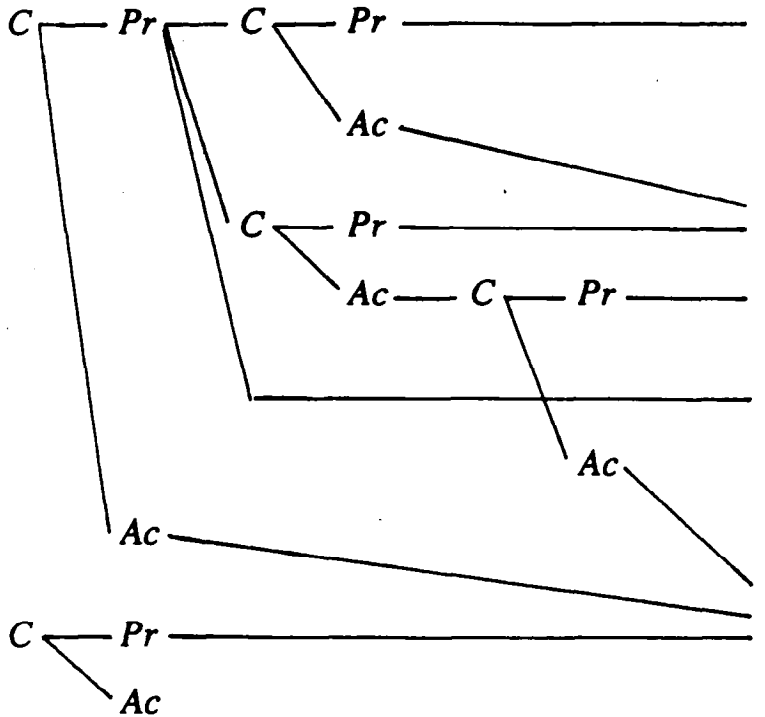
Acceptance Phase:
A. before you start on that part, yes

But this is part of B's whole contribution, as shown in Figure 8.

Completions may become much more extended as the primary speaker and partner search explicitly for a name, as here:

C. and we went to Bridport, and we went to Weymouth one day, - um we went to um - . what was it called .
A. you didn't go inland -
C. um - not very much, -- . oh what's that -
A. *((3 sylls))*
C. *really* lovely place, along the coast, where th- where the swannery is
A. oh um -- Abbot something
C. yes -
A. Abbot Newton -*- Abbotsbury*
C. *no, Abbotsbury,* that's right -
C. it was lovely

C requests completions with the pleas "what was it called," "oh what's that," "where th- where the swannery is," and A does his best to help out. They finally make it when C says, "Abbotsbury that's right," and goes on. So although the acceptance phase runs quite a complicated course, it has essentially the same structure as the simpler example.



A: um the problem is a that you(('ve)) got to get planning consent -

B: before you start -

A: before you start on that part, yes

A: you can do anything internally, you wish

B: but the big stuff is, the external stuff [etc]

Figure 8.

CONCLUSIONS

When people participate in a discourse, they generally try to make a success of it. According to most theories, all they have to do to achieve success is utter the right sentence at the right time. But this leaves too much to chance. Was the utterance heard correctly? Was it interpreted correctly? Do all the participants believe it was interpreted correctly? In actual conversations, we have argued, people hold out for a higher criterion. They try to ground what is said—to reach the mutual belief that what the speaker meant has been understood by everyone well enough for current purposes. In doing this, they create units of discourse called contributions.

Shapes of Contributions

Contributions take the shape they do, we have argued, because they are constructed in two phases. In the presentation phase, the contributor, say Ann, typically presents an utterance *u* for her partner, say Bob, to consider. In the acceptance phase, Bob then provides her with evidence *e* that he believes he understands what Ann means by *u,* evidence she then accepts. But these two phases admit so many options and complications that they take a variety of shapes. Here are some of them.

Ann's presentation can take any of the forms in the left hand column of Table 2 and many other forms as well. It may also contain subordinate contributions. A difficult instruction, for example, can be divided up and presented in installments, each with its own presentation and acceptance phases, and yet the entire instruction is treated as a presentation that requires an acceptance phase. This is one source of embedding in contributions.

In the acceptance phase, the evidence that Bob offers depends in part on the type of presentation it is in response to. The right hand column in Table 2 lists the typical form of evidence elicited in our corpus for each type of presentation. Since we had no access to eye contact, head nods, or other gestures, we can only assume that acknowledgments were accepted with continued attention. The other presentation types were accepted with linguistic

TABLE 2
Typical Forms of Evidence Elicited by Presentations

Presentation of Utterance *u*	Acceptance of Evidence *e*
Acknowledgement	Continued attention
Completed turn	Initiation of next relevant turn
Portion of continuing turn	Acknowledgement, often overlapping
Installment of extended turn	Acknowledgement during pause
Installment for rote memory	Verbatim display
Trial constituent	Explicit answer
Incomplete utterance	Completion
Completion	Repeat plus assent

acts. These acts, of course, are themselves contributions, with presentation and acceptance phases, and that makes the acceptance phase inherently recursive. This recursion stops, however, once Ann and Bob reach the weakest type of positive evidence—namely, continued attention—and that generally occurs in one or two cycles. So this is a second source of embedding in contributions.

The acceptance phase takes a different course whenever Bob has had trouble understanding. In that case he ordinarily initiates the acceptance phase by indicating what the trouble is—"Duveen?" or "pardon" or "m?" —and the rest of the acceptance phase is spent clearing it up. Different devices are available for indicating different types of trouble (Clark & Schaefer, 1987; Schegloff et al., 1977), and we have illustrated only a few. This is a third source of embedding in contributions. The important point is that these devices help define what constitutes positive evidence of understanding. Offering a continuer like "uh huh" and initiating the next relevant turn are effective as positive evidence in part because they show that the partner is choosing not to initiate a repair.

Size of Contributions

Contributions come in many sizes. Some are initiated by single words or phrases, and others by clauses, full sentences, or whole turns. What determines the size? If the participants stopped to ground every word, it would take too long to say anything, and yet if they didn't stop often enough, misunderstandings could snowball before they could be repaired. The participants should generally settle for something in between. Indeed, the two commonest devices for completing contributions—new turns and acknowledgements—resulted in contributions with median lengths of 9 to 13 words.

But participants systematically vary the size and make-up of their contributions to suit current purposes. According to the contribution model, Bob is to understand Ann to a criterion sufficient for current purposes. If either of them anticipates that Bob will find some word, phrase, or clause especially difficult or want to register some information verbatim, they can divide the discourse into contributions of the corresponding size. We saw in Figure 6 how a complicated address to be registered verbatim was divided into installments (see also Clark & Schaefer, 1987). Likewise, if it is anticipated that Bob will understand everything easily, they can make their contributions longer. Figure 5 shows how an extended description that was easy to understand was divided into presentations several clauses long. Generally, the more difficult it is anticipated a unit will be to understand well enough for current purposes, the more contributions it will be divided into.

The preferred contribution, nevertheless, appears to be the length of a simple or complex sentence. Most contributions are initiated by uttering a full sentence (e.g., "how far is it from Huddersfield to Coventry"), an elliptical sentence (e.g., "about a hundred miles"), a phrasal sentence (e.g.,

"sorry"), or an atomic sentence (e.g., "yes" or "oh" or "m"). Many of these are accomplished as complete turns and therefore count as complete contributions. And when a turn consists of more than one sentence, it is often broken up by acknowledgements into contributions the size of single sentences.

Our conjecture is that the preferred contribution is one or more illocutionary acts à la Austin (1962). When Ann initiates a contribution by uttering a full sentence, she is trying to make an assertion, ask a question, or make an apology, or do more than one of these at a time. Not that she cannot contribute a single propositional act instead, but she will do that only for special reasons. In the first installment in Figure 6, B contributed only a reference ("Banque Nationale de Liban") because it was crucial for A to get the reference verbatim. Even then, the reference was part of a larger contribution in which B was making an assertion (see Cohen, 1984). The same is true of trial constituents and completions. And these are not the only forms contributions can take.

The process of contributing cannot be fixed beforehand because it is subject to accidental features of the situation. Suppose Ann wants to contribute an assertion to the on-going social process. Although she may begin expecting to do this in one unbroken action, she may discover along the way that she can't. It may happen that Bob gets distracted and mishears her, or she cannot retrieve a name quickly enough, or she misjudges how much he knows about a referent. Any accident like this can force Ann and Bob into a complex acceptance process that may bring in any of the devices we have mentioned. In conversations there are no crystal balls. Unforeseen circumstances can take contributions off in entirely unexpected directions.

Contributions, therefore, are different from most standard linguistic units. They are not formulated autonomously by the speaker according to some prior plan, but emerge as the contributor and partner act collectively. Success depends on the coordinated actions by the two of them. We have tried to show what these actions look like and why.

■ Original Submission Date: April 8, 1988.

REFERENCES

Austin, J.L. (1962). *How to do things with words.* Oxford: Oxford University Press.

Clark, H.H., & Carlson, T.B. (1982). Hearers and speech acts. *Language, 58,* 332–373.

Clark, H.H., & Haviland, S.E. (1974). Psychological processes in linguistic explanation. In D. Cohen (Ed.), *Explaining linguistic phenomena.* Washington: Hemisphere Publication Corporation.

Clark, H.H., & Haviland, S.E. (1977). Comprehension and the given-new contract. In R.O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Hillsdale, NJ: Erlbaum.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual nowledge. In A.K. Joshi, B. Webber, & I.A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.

Clark, H.H., & Schaefer, E.F. (1987). Collaborating on contributions to conversation. *Language and Cognitive Processes, 2,* 1–23.

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

Cohen, P.R. (1978). *On knowing what to say: Planning speech acts.* Unpublished doctoral dissertation, University of Toronto, Toronto.

Cohen, P.R. (1984). The pragmatics of referring, and the modality of communication. *Computational Linguistics, 10,* 97–146.

Goffman, E. (1976). Replies and responses. *Language in Society, 5,* 257–313.

Goldberg, J. (1975). A system for the transfer of instructions in natural settings. *Semiotica, 14,* 269–296.

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

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

Grice, H.P. (1957). Meaning. *Philosophical Review, 66,* 377–388.

Grosz, B. J., & Sidner, C.L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics, 12,* 175–204.

Grosz, B.J., & Sidner, C.L. (1989). Plans for discourse. In P.R. Cohen, J. Morgan, & M.E. Pollack (Eds.), *Intentions in communication.* Cambridge, MA: MIT Press.

Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In R. Bäuerle, C. Scharze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language* (pp. 164–189). Berlin: W. de Gruyter.

Isaacs, E.A., & Clark, H.H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General, 116,* 26–37.

Jefferson, G. (1972). Side sequences. In D. Sudnow (Ed.), *Studies in social interaction* (pp. 294–338). New York: Free Press.

Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica, 9,* 47–96.

Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kamp, J.A.W. (1981). A theory of truth and semantic representation. In J. Groenerdijk, T. Janssen, & M. Stockhof (Eds.), *Formal methods in the study of language, Part 1* (pp. 177–321). Amsterdam: Mathematical Centre Tracts.

Lerner, G. H. (1987). *Collaborative turn sequences: Sentence construction and social action.* Unpublished Ph.D. dissertation, University of California, Irvine, CA.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition, 14,* 41–104.

Lewis, D.K. (1969). *Convention: A philosophical study.* Cambridge, MA: Harvard University Press.

Lewis, D.K. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic, 8,* 339–359.

Litman, D.J., & Allen, J.F. (1987). A plan recognition model for subdialogues in conversation. *Cognitive Science, 11,* 163–200.

Oreström, B. (1983). *Turn-taking in English conversation.* Lund, Sweden: Gleerup.

Polanyi, L., & Scha, R. (1985). A discourse model for natural language. In L. Polanyi (Ed.), *The structure of discourse.* Norwood, NJ: Ablex.

Reichman, R. (1978). Conversational coherency. *Cognitive Science, 2,* 283–327.

Sacks, H., & Schegloff, E. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 15–21). New York: Irvington.

Sacks, H., Schegloff, E.A., & Jefferson, G.A. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50,* 696–735.

Schegloff, E.A. (1972). Notes on a conversational practice: Formulating place. In D. Sudnow (Ed.), *Studies in social interaction.* New York: Free Press.

Schegloff, E.A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk. 32nd Georgetown University Roundtable on Languages and Linguistics 1981* (pp. 71–93). Washington, DC: Georgetown University Press.

Schegloff, E.A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53,* 361–382.

Schegloff, E.A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8,* 289–327.

Schiffer, S. (1972). *Meaning.* Oxford: Clarendon Press.

Searle, J.R. (1969). *Speech acts.* Cambridge: Cambridge University Press.

Searle, J.R. (1989). Collective intentions and actions. In P.R. Cohen, J. Morgan, & M.E. Pollack (Eds.), *Intentions in communication.* Cambridge, MA: MIT Press.

Stalnaker, R.C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics: Pragmatics 9* (pp. 315–332). New York: Academic.

Stenström, A.-B. (1984). *Questions and responses in English conversation.* Malmo: Gleerup.

Svartvik, J., & Quirk, R. (1980). *A corpus of English conversation.* Lund, Sweden: Gleerup.

Thavenius, C. (1983). *Referential pronouns in English conversation.* Lund, Sweden: Gleerup.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic.

Wilkes-Gibbs, D. (1986). *Collaborative processes of language use in conversation.* Unpublished doctoral dissertation, Stanford University, CA.