Chapter 23

Referring as a Collaborative Process

Herbert H. Clark and Deanna Wilkes-Gibbs

Conversation is the fundamental site of language use. For many people, even for whole societies, it is the only site, and it is the primary one for children acquiring language. From this perspective other arenas of language use—novels, newspapers, lectures, street signs, rituals—are derivative or secondary. How, then, do speaking and understanding work in conversation? For psychologists this ought to be a central question, but surprisingly, it has not been. The main attempts to answer it have come instead from philosophy and sociology.

Among philosophers the study of conversation grew out of an analysis of what speakers mean and what listeners understand them to mean. The idea was that, when speakers utter sentences, they do so with certain intentions toward their addressees. They assert, request, promise, and perform other illocutionary acts, and their interlocutors are expected to recognize these intentions (Austin 1962; Grice 1957, 1968; Schiffer 1972; Searle 1969). In 1967 Grice argued that, for this scheme to work, people in conversation must be cooperative. Speakers must try to "make their contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which [they] are engaged" (Grice 1975, 45). Only then can their partners go beyond what is "said" to infer what is conversationally "implicated" (Grice 1975, 1978).

Among sociologists the issue has been how people direct the course of conversation and repair its inherent troubles. As this work has shown, people in conversation manage who is to talk at which times through an intricate system of turn taking (Sacks, Schegloff, and Jefferson 1974). Further, when one person speaks, the others not only listen but let the speaker know they are understanding—with head nods, *yes*'s, *uh huh*'s, and other so-called back-channel responses (Duncan 1973; Goodwin 1981; Schegloff

1981; Yngve 1970). When listeners don't understand, or when other troubles arise, they can interrupt for correction or clarification (Schegloff, Jefferson, and Sacks 1977). The participants also have techniques for initiating, guiding, and terminating conversations and the topics within them (Schegloff 1968; Schegloff and Sacks 1973).

In both traditions a central issue is coordination: How do the participants in a conversation coordinate on the content and timing of what is meant and understood? The issue, however, cannot be resolved within either tradition alone. In the first tradition conversation is idealized as a succession of illocutionary acts—assertions, questions, promises—each uttered and understood clearly and completely (Gazdar 1979; Kamp 1981; Stalnaker 1978). Yet from the second tradition we know that many utterances remain incomplete and only partly understood until corrected or amplified in further exchanges. How are these two views to be reconciled?

In this paper we propose a resolution for an essential use of language: how people in conversation coordinate in the making of a definite reference. Our concern is not with semantic reference, but with speaker's reference—not, for example, with what the phrase *the clown with the red nose* means, but with what the speaker does in referring, say, to a clown as part of an assertion that the clown is funny (Donnellan 1978; Kripke 1977; Searle 1969). Our premise is that making such a reference is a collaborative process requiring actions by both speakers and interlocutors. To some it may appear self-evident that the process is collaborative, but it is one thing to assume it is and quite another to understand why it is and how it works. The goal here is important, since, if conversation is fundamental, its processes are likely to underlie or shape those in other uses of language as well.

In the first section of this paper, then, we offer evidence for the premise itself and outline what we will call a collaborative model for the process of reference. In the second and third sections we describe an experiment on referring and use it to corroborate and fill in details of the model. In the final section we return to the general issue of coordination and note problems still to be resolved.

1 Referring in Conversation

Traditionally, philosophers, linguists, and psychologists have presupposed what might be called a *literary model* of definite reference. Speakers refer as if they were writing to distant readers. When Elizabeth selects the noun phrase *the clown with a red nose* in talking to Sam, the assumption is that she intends it to enable him to identify the clown uniquely. She satisfies her intentions by issuing the noun phrase. Her act of referring is cotemporal with that noun phrase, beginning with *the* and ending with *nose*. Further,

she retains complete responsibility and control over the course of this process. Sam hears the definite description as if he were reading it and, if successful, infers the identity of the referent. But his actions have no bearing on hers in this reference.

The literary model makes these tacit idealizations. (1) The reference is expressed linguistically with one of three standard types of noun phrase—a proper noun (for instance, *Napoleon, King George*), a definite description (*this year, the man with the moustache*), or a pronoun (*he, this, they*). (2) The speaker uses the noun phrase intending the addressee to be able to identify the referent uniquely against their common ground. (3) The speaker satisfies her intention simply by the issuing of that noun phrase. And (4) the course of the process is controlled by the speaker alone.

A conversational model of the process, however, ought to look quite different for three reasons. First, in conversation unlike writing, speakers have limited time for planning and revision. They need to overcome this limitation, and in doing so they may exploit techniques possible only in conversational settings. Second, speech is evanescent. The listener has to attend to, hear, and try to understand an utterance at virtually the same time it is being issued. That requires a type of process synchronization not found in reading. And third, listeners in conversations aren't mute or invisible during an utterance. Speakers may alter what they say midcourse based on what addressees say and do.

Indeed, once we look at actual conversations, we find that the four idealizations of the literary model are very wide of the mark. To see this, let us turn to eight types of examples that fail on one or more of these assumptions.

1.1 Eight Problems

1. *Self-corrected noun phrases.* Consider this attested utterance: *She was giving me all the people that were gone this year I mean this quarter y'know* (from Schegloff, Jefferson, and Sacks 1977, 364, in simplified notation). The speaker began the referential process by uttering *all the people that were gone this year*, but corrected the last two words to *this quarter* in what Schegloff, Jefferson, and Sacks (1977) called a *self-initiated repair.* The referential process, clearly, isn't cotemporal with one particular noun phrase, since two noun phrases were uttered in succession. It is more naturally described as a process in which the speaker decided midcourse to repair the initial noun phrase, indicated here change with *I mean,* and then uttered *this quarter* (see Levelt 1983). The process began with *all the people* and was completed with *y'know.*

2. *Expanded noun phrases.* Although the first noun phrase a speaker utters may be technically correct, the speaker may still judge it insufficient and change course, as here (from Cohen 1985):

> S: Take the spout—the little one that looks like the end of an oil can—
>
> J: Okay.
>
> S: —and put that on the opening in the other large tube. With the round top.

S began with *the spout*. But when he saw that it was insufficient for J to pick out the referent, he expanded on it with the parenthetical noun phrase. Ordinarily, parenthetical phrases are nonrestrictive—not needed for identifying the referent. Here, the parenthetical phrase *was* deemed necessary, and S changed course midutterance to add it.

3. *Episodic noun phrases.* For similar reasons, once S completed *the other large tube*, he judged that to be insufficient as well and added the restrictive phrase *with the round top* under a separate intonation contour, as part of a new tone group. He produced a single noun phrase, but intonationally, he divided in into two information units. We will call this an *episodic noun phrase*, and it is another nonstandard type.

4. *Other-corrected noun phrases.* The process becomes more complicated when the addressee makes the repair, as with A's reference to Monday in this example (from Schegloff, Jefferson, and Sacks 1977, 369):

> B: How long y'gonna be here?
>
> A: Uh- not too long. Uh just til uh Monday.
>
> B: Til- oh yih mean like a week f'm *tomorrow*.
>
> A: Yah.
>
> B: [Continues]

A initiated the referential process by uttering *Monday*. Uncertain of the intended referent, B offered a correction, which A accepted, all before B proceeded. The process took place over several turns and was participated in by both A and B.

In the four cases so far, then, the speakers changed the course of their reference after uttering an initial noun phrase. They did so in reaction to both their own and their addressee's judgments of inadequacy or error. But speakers are not merely reactive. At other times they bring addressees into the referential process by the very design of their utterance. Consider the next four classes of examples.

5. *Trial noun phrases.* Some noun phrases are uttered with a rising intonation, or *try marker* (Sacks and Schegloff 1979), imposed on them, as in this example (from Cohen 1985):

> S: Okay now, the small blue cap we talked about before?
>
> J: Yeah.

S: Put that over the hole on the side of that tube—

J: Yeah.

S: —that is nearest to the top, or nearest to the red handle.

With *the small blue cap we talked about before*? S asks J to say whether or not he has understood S's reference. The process begins when S utters this phrase and ends only with J's *Yeah*. If J hadn't understood, the process would have continued as here (from Sacks and Schegloff 1979):

A: ... well I was the only one other than than the uhm tch *Fords*?,
Uh Mrs, Holmes Ford? You know uh =
⌈ = the the cellist?

B: ⌊ Oh yes. She's she's the cellist.

A: Yes. Well she and her husband were there.

When A received no reply to *Fords*?, she offered the expanded noun phrase *Mrs. Holmes Ford*? and then went to *the cellist*? before B implicated that she had identified the referent. The referential process was continued until A said *Yes*, confirming that B's display of understanding was correct.

6. *Installment noun phrases.* Speakers can also utter noun phrases in *installments*, as we will call them, and invite addresses to affirm their understanding of each installment. In the earlier example, S began *the hole on the side of that tube*, paused for confirmation from J, and then completed the noun phrase with *that is nearest to the top, or nearest to the red handle.* As with his trial noun phrase, S made the course of his reference contingent on the addressee's midcourse response.

7. *Dummy noun phrases.* Speakers sometimes initiate the referential process with terms like *what's-his-name, whatchamacallit, whatzit,* or *thingamabob,* which we will call *dummy* nouns or noun phrases. Consider: *If he puts it into the diplomatic bag, as um—what's-his-name, Micky Cohn, did, then it's not so bad* (from Svartvik and Quirk 1980, 35). The speaker recognized from the start that *what's-his-name* was inadequate as a definite description. Yet, pressed for time, he used it to initiate the referential process until he could replace it with an adequate noun phrase, *Micky Cohn.* Dummy noun phrases are *not* standard, and when speakers use them, they do *not* intend them to enable their addresses to identify the referent uniquely. Dummy noun phrases are uttered only as part of a more extended process.

8. *Proxy noun phrases.* In some circumstances, the speaker makes it clear that a noun phrase is to come next, but the addressee actually utters it. Here is one many spontaneous examples recorded by Wilkes-Gibbs (unpublished):

A: That tree has, uh, uh ...

B: Tentworms.

A: Yeah.

B: Yeah.

A initiated the referential process by halting at a place where he needed a noun phrase and uttering two *uh*'s. B helped out by offering a proxy, or stand-in, noun phrase she thought appropriate. A confirmed the proxy with *Yeah*, and then B responded to A's full assertion. B took part in the process from the very beginning.

As all eight examples make plain, a conversational model of the referential process must be quite different from the literary model. First, many noun phrases are distinctly nonliterary in form or nonstandard in intonation. These include trial, episodic, installment, dummy, and proxy noun phrases. Second, the process takes a very different course in conversation than in literature. In all eight examples, speakers went beyond the issuing of standard noun phrases; in three examples they deliberately drew the addressees into the process; and in three they began by knowingly issuing a questionable or inadequate noun phrase. What characterizes these examples is that the speaker and addressee put in extra effort, generally together, to make sure the reference has been understood. To understand the process of referring, we need to know how this works.

1.2 Establishing Understanding

Suppose A, a man, is speaking to B, a woman, and refers to a dog. In making the reference, according to most theories, A intends the identity of the dog to become part of A's and B's mutual knowledge or beliefs (see Clark and Marshall 1981). Establishing such mutual knowledge or belief is a stringent requirement. To meet it, A must convince himself that the identity of the dog is truly going to become part of their common ground. If at any moment in making the reference he thinks it won't, he should change or expand on what he has done so far. The same requirement applies to B, since she is trying to understand A's reference. To meet it, she should find ways of letting A know, as she listens, whether or not she is understanding him. Indeed, A should suppose that she is cooperating in precisely this way.

For each reference, then, A and B should have procedures for establishing the mutual belief, at some level of confidence, that B has identified A's reference. We have already seen evidence in our examples that they do. These procedures, we will argue, are inherently collaborative.

The evidence is clearest when B believes she may *not* have identified A's referent and attempts to repair the problem, as in our earlier example:

B: Til- oh yih mean like a week f'm *tomorrow*.

A: Yah.

Copyrighted Material

These turns constitute a *side sequence*, a block of exchanges embedded within or between anticipated contributions to the conversation (Jefferson 1972; Schegloff 1972). So although the side sequence was initiated by B, it was completed by A before the conversation was allowed to proceed. That was needed for them to *mutually* believe that B has now understood A's reference correctly.

More often, A and B have to establish that B *has* understood the reference, and for this, B can use a simple expedient: allowing the next contribution to continue. Suppose the conversation had continued this way:

> B: How long y'gonna be here?
> A: Uh- not too long. Uh just til uh Monday.
> B: Oh that's too bad.

By asserting *Oh that's too bad*, B would be passing up the opportunity to correct a possible misunderstanding and would thereby be implicating that she understood A's reference. "Regularly, then," as Sacks, Schegloff, and Jefferson (1974, 728) put it, "a turn's talk will display its speaker's understanding of a prior turn's talk, and whatever other talk it marks itself as directed to" (see also Goffman 1976). Note that going on wouldn't necessarily mean B had truly understood. She might not recognize her misunderstanding, or she might want to claim she had understood when she hadn't. But in either case, going on is a signal that B believes she has understood. In the first case she is making a mistake; in the second she is using the signal to deceive.

The same mutual belief can be established more directly by what Schegloff (1981) has called *continuers*, as in his example from a radio call-in show (p. 80):

> A: Now, I wanna ask you something, I wrote a letter. (pause)
> B: Mh hm,
> A: T'the governor
> B: Mh hm::,
> A: -telling 'im what I thought about i(hh)m!
> B: (Sh:::!)
> A: Will I get an answer d'you think.
> B: Yes.

By inserting the continuers *mh hm* and *sh:::* while A's turn was still underway, according to Schegloff, B was showing, first, that she was paying attention and realized that A was in the middle of an extended unit of talk. At the same time, she was explicitly signaling that she was passing up the

Copyrighted Material

opportunity to initiate a repair on the turn so far and, by implication, that she understood the turn so far. With the second *Mh hm::*, for example, she was claiming to understand the phrase *t'the governor* and, therefore, the definite reference it contained. The same holds for the other definite references.

B may even be intended to interrupt A as soon as she believes she has identified the referent, as in this example (from Sacks, quoted in Jefferson 1973, 59):

> A: I heard you were at the beach yesterday. What's her name, oh you know, the tall redhead that lives across the street from Larry? The one who drove him to work the day his car // was-
>
> B: Oh *Gina!*
>
> A: Yeah Gina. She said she saw you at the beach yesterday.

A indicated he would go on until B identified the referent. Indeed, he stopped at B's interruption and completed the process by confirming B's identification with *Yeah Gina.*

Taken together, this evidence suggests that A and B accept mutual responsibility for each definite reference. Roughly speaking, they try to establish the mutual belief that B has understood A's reference before they go on. So far we have only informal examples of how they do this. The challenge is to characterize the system and the logic behind it.

1.3 Mutual Acceptance

The idea behind the view of reference we are taking is this: A and B must *mutually accept* that B has understood A's references before they let the conversation go on. Conversations proceed in an orderly way only if the common ground of the participants accumulates in an orderly way (see Clark 1985; Clark and Carlson 1981; Gazdar 1979; Stalnaker 1978). A and B must therefore establish the mutual belief that B has understood, or appears to have understood, A's current utterance before they go on to the next contribution to the conversation. They establish that belief, we argue, through an acceptance process.

The two basic elements in this process are (1) a presentation and (2) an acceptance. Suppose A wants to refer to a mutually identifiable dog. To do so, he *presents*, as we will put it, the standard noun phrase *the dog that just barked*. With this presentation A presupposes a number of things. First, he believes B is now paying attention, is able to hear and identify the words, and understands English. Second, he believes B can view the referent as fitting the description "dog that just barked." That is, he believes that referent r can be viewed under description d. And third, he believes B will be able to pick out r uniquely with this description d along with the rest of their common ground.

Once A has made this presentation, B must accept it, and A and B must mutually recognize that acceptance. We propose that B has two main methods of accepting it. First, she can *presuppose acceptance*, as illustrated earlier, by continuing on to the next contribution or by allowing A to continue. Letting the next contribution begin is mutually recognized as an acceptance of the last presentation. Second, she can *assert acceptance*, as with continuers, *yes, right, I see*, and head nods. These, too, are mutually recognized as acceptances of the last contribution.

But B may have reasons for *not* accepting A's presentation. She may not have heard it fully; if so, she might respond, *What?* or *The dog that just what?* She may not accept *d* as a description of *r*; then she might respond, *That's a toy not a dog*. Or she may not accept that *d* is sufficient with their common ground to pick out *r* uniquely; then she might respond, *Which one?* When B doesn't accept the presentation, A must deal with B's implicit or explicit questions until B does accept it. That may take several exchanges.

As our examples show, however, A's presentation can take more complicated forms. It can be a trial or installment noun phrase, which B can accept only by assertion. It can be a dummy noun phrase, which B isn't intended to accept until amended. It can be a self-corrected or expanded noun phrase, which B is to accept only as amended. It can even be a proxy noun phrase made by B, which A is then intended to accept.

These informal examples, though suggestive, still do not specify precisely how the acceptance pocess works. For that we need more systematic evidence.

2 References in an Experimental Task

In search of such evidence we turned to a communication task originally devised by Krauss and Glucksberg (Krauss and Weinheimer 1964, 1966, 1967; Krauss and Glucksberg 1969, 1977; Glucksberg, Krauss, and Higgins 1975; see also Asher 1979). In our version two students were seated at tables separated by an opaque screen. In front of each student were 12 cards, each showing one of the so-called Tangram figures in figure 23.1. For the person we will call the *director*, the cards were already arranged in a target sequence of two rows of six, and for the person we will call the *matcher*, the same figures lay in an identical matrix but in a random sequence. (For ease of exposition, we will talk as if the director were male and the matcher female, even though both sexes took both roles in our task.) The director's job was to get the matcher quickly and accurately to rearrange her figures to match the target ordering. They could talk back and forth as much as they needed, but the director was to go through the positions in the array sequentially (numbered 1 to 6 on the top row and 7 to 12 on the bottom). After they had matched their arrangements, the
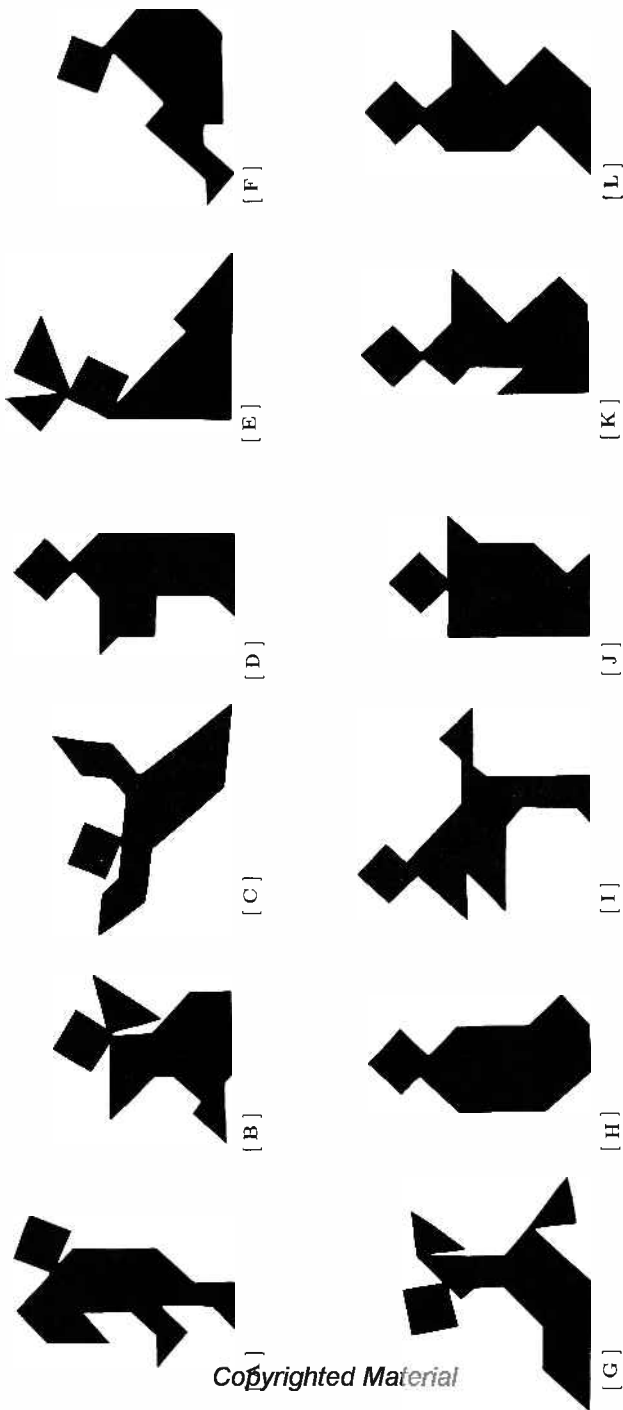
Figure 23.1
The 12 Tangram figures arranged by directors and matchers.

director's and matcher's figures were placed in two new random orders, the director's new sequence became the new target, and the procedure was repeated. They carried out the task six times, for six trials.

The collaborative view of reference makes several global predictions about this task. First, it should take the two partners many words to reach acceptance the first time they encounter a figure since they will often need nonstandard techniques such as episodic, installment, or expanded noun phrases. Later references to the same figure should be shorter since they can appeal to prior acceptance of a related description and succeed more often with standard noun phrases, which are typically shorter. This reasoning would account in part for Krauss and Weinheimer's (1964) original finding that, as people referred repeatedly to the same figure, they tended to shorten their noun phrases, although only if their listeners could speak in return. The collaborative view also predicts that, since the later references are more likely to be standard noun phrases, they should require fewer turns. For this prediction there is no evidence. We will defer more detailed predictions about the acceptance process itself.

2.1 Method

Eight pairs of partners each arranged 12 figures on each of six trials. The 12 figures, each formed from different arrangements of seven elementary shapes, were selected from a book with 4000 such figures collected by Elffers (1976) from the ancient Chinese game of Tangram. These 12 were chosen because their varying abstraction and similarity seemed to provide a good range of difficulty. Two copies of each figure were cut out of black construction paper and pasted individually on white 15 cm by 20 cm cards. The identifying letters in figure 23.1 did not appear on the stimuli.

The two students in each session drew lots for director and matcher roles. They were told they had identical figures and would play the game six times while timed and tape-recorded. A timer was started on each trial when both students were ready, and stopped when they were satisfied they had finished. After each trial the two orderings were checked and the students were told of the positions of any mismatches. The error rate was only 2%. The six trials took about 25 minutes. The students, seven men and nine women, were Stanford University undergraduates fulfilling a course requirement.

One of us transcribed the conversations, including changes of speaker, back-channel responses, parenthetical remarks, interruptions, hesitations, false starts, and basic intonational features; the other checked the transcripts, especially for intonation. The transcripts contained 9792 words, reflecting the positioning of 576 figures (12 figures on six trials by eight pairs of students).

2.2 General Patterns

For a broad picture of what occurred, consider this very simple series of utterances by one director for figure I on trials 1 through 6:

1. All right, the next one looks like a person who's ice skating, except they're stricking two arms out in front.
2. Um, the next one's the person ice skating that has two arms?
3. The fourth one is the person ice skating, with two arms.
4. The next one's the ice skater.
5. The fourth one's the ice skater.
6. The ice skater.

As this series illustrates, directors generally referred to the location (e.g., *the fourth one*) and then asserted something about the Tangram figure to be placed in that location. On trial 1 directors always *described* the figure, generally with an indefinite reference (e.g., *a person who's* ...). On trials 2 through 6, in contrast, they *referred* to the figure with a definite description (e.g., *the ice skater*). Directors tended to use nonstandard noun phrases in the early trials (e.g., this director's trial and episodic noun phrases in trials 2 and 3) and standard noun phrases later (e.g., *the ice skater*).

Partly because of these features, this director took many more words to secure acceptance of his presentation on trial 1 than on trial 6. As predicted, this pattern held in general. Figure 23.2 shows that directors used an average of 41 words per figure in trial 1 but only 8 words per figure in trial
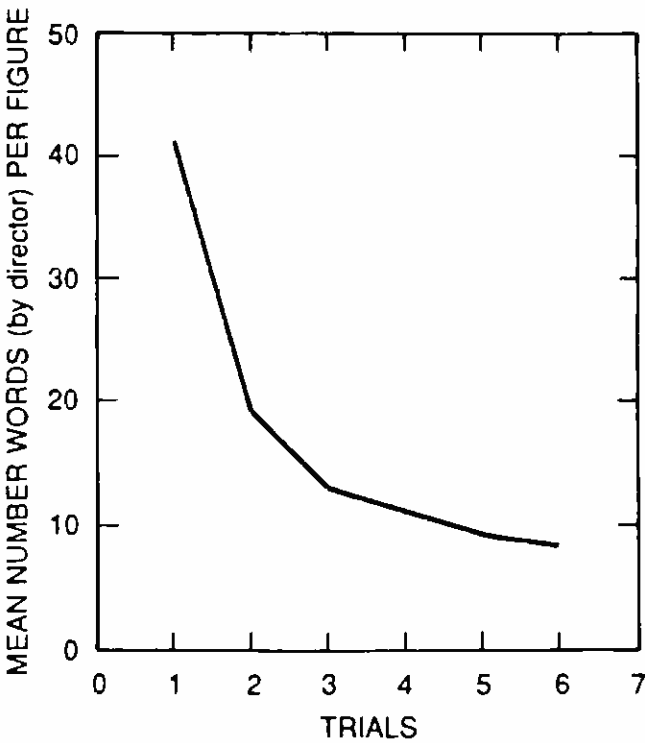
Figure 23.2
Average number of speaking turns per figure taken by directors on each trial.

6. This decline is highly significant, $F(1,35) = 44.31$, $p < .001$. The decline was steepest from trial 1 to trial 2 and had almost disappeared by trial 6.

The example we have cited, however, is atypical in that the director took only one turn on each trial for this figure; it is also incomplete in that we have omitted the matcher's single turns. More often, the two partners took many turns for a single placement, and as predicted, the number of turns they needed declined from trial 1 to trial 6. Figure 23.3 shows that the director averaged 3.7 turns per figure on trial 1 but only about one per figure by trial 6. This trend was also highly significant, $F(1,35) = 79.59$, $p < .001$. So figure 23.2 includes the director's words not just from his first turn on each figure but from *all* of his turns on that figure.

The director and matcher became more efficient not only from one trial to the next but also from the beginning to the end of each trial. Figure 23.4 plots the number of words per figure over the 12 spatial positions in the arrangements for trials 1, 2, and 6. Since the figures were randomly assigned to the positions on each trial, there is some confounding of figures with positions, but the pattern is still clear. On trial 1, there was a steep decline in word count (4.6 words per position) as the two partners worked from position 1 to position 12 ($F(1,77) = 40.01$, $p < .001$). On trials 2 and 6, there were successively smaller declines (1.0 and .4 words per position), both also significant ($F(1,77) = 5.83, 7.16$, $p < .05$). Number of turns per figure shows a similar pattern.
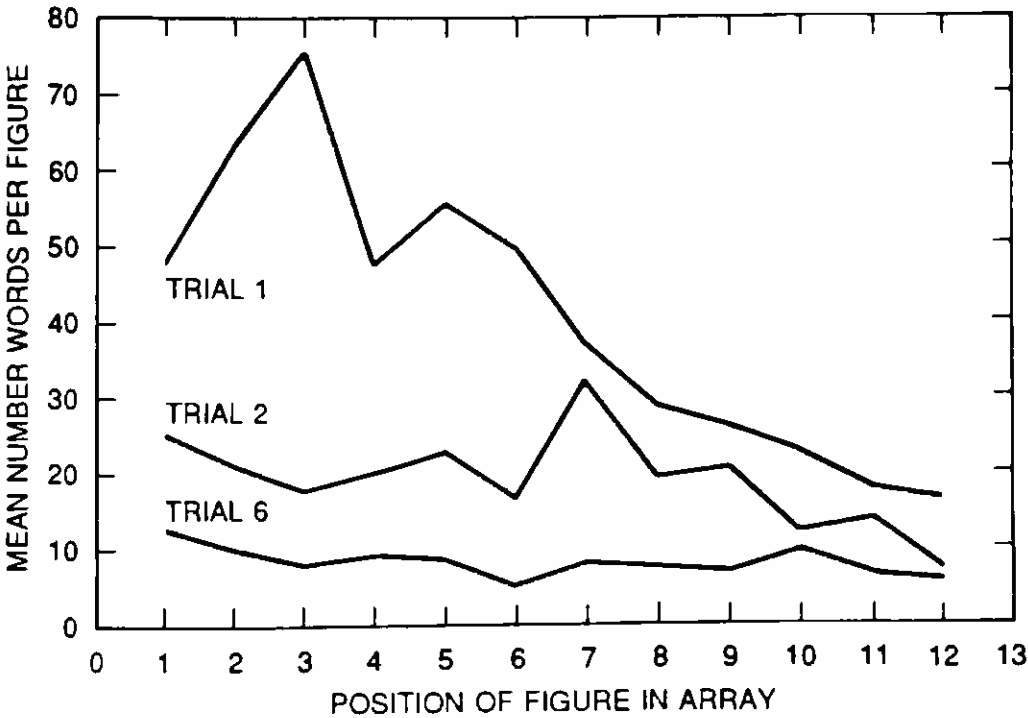
The general decline in number of words used from position 1 to position 12 is predicted by the collaborative view but also by others. By any
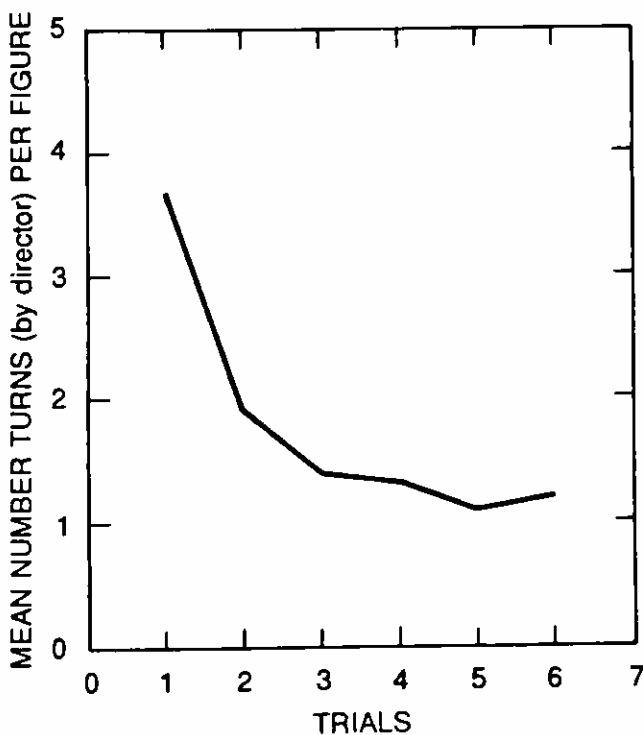
Figure 23.3
Average number of speaking turns per figure taken by directors on each trial.

Copyrighted Material

Figure 23.4
Average number of words per figure used by directors on trials 1, 2, and 6 plotted by position of the figure in the array.

reasonable theory of information or reference (for example Olson 1970), the fewer figures there are in the array, the less information it should take to distinguish the target from the remainder. In the limit, the figure in position 12 needs only a minimal description since it is the only one left—for example, *Number 12 is the last one*. Indeed, sometimes it wasn't even mentioned. The number of turns should decrease by the same reasoning, as it did.

The decline from position 1 to position 12, however, got smaller from trial 1 to trial 6, and that is predicted by the collaborative view but not by general theories of information. By the collaborative view, as we will justify later, the two partners come to rely on descriptions mutually accepted on previous trials, forming shorter noun phrases accepted in fewer turns until they arrive at optimal descriptions. This is nicely illustrated in the example cited. But as the descriptions become optimal, they should be less influenced by the physical context. The decline from position 1 to position 12 should be largest on trial 1, when reaching acceptance takes many words, and smallest on trial 6, by which time the two partners have perferred descriptions. This is precisely what occurred. The difference in slopes between trials 1, 2, and 6 was significant ($F(2,284) = 15.49$, $p <$ .001). By information theory, in contrast, going from one position to the next should reduce the array size as much on trial 6 as on trial 1; hence, the slopes should remain the same. This prediction is disconfirmed.

Finally, the 12 figures also varied in difficulty, $F(11,77) = 5.94$, $p <$ 0.001. Figure B, the most difficult one, averaged 26.5 words per trial, eliciting 39.6 words on trial 1. Figure C, the easiest, averaged only 9.7 words per trial, with 24 words on trial 1.

Having sketched the global performance in this task, we now turn to details of the referential process itself.

3 Collaborating on References

Our primary goal here is a process model of how speakers and addressees collaborate in the making of a definite reference. The collaborative model, as we will call it, must do more than list the devices used—trial noun phrases, interruptions, continuers, and the like. It must spell out how the process of mutual acceptance gets initiated, carried through, and completed. The process usually begins with the speaker issuing a noun phrase. But these noun phrases come in many types, as already noted, and do no more than initiate the process. The model must show how these noun phrases are organized as a system and how they enter in a uniform way into the referential process as a whole. We must resist the temptation, engendered by the literary model of reference, to treat standard noun phrases as genuine and all others as aberrations, for that doesn't explain the role of any of the noun phrases in the process.

Definite references to the Tangram figures, as noted earlier, occurred only on trials 2 through 6. In the simplest pattern, the director would refer to a position (for example, *Number 4*) and assert which figure appeared there (*is the guy leaning against the tree*), and the matcher would signal she had placed it with *okay, all right, got it,* or *right,* as in this exchange:

A: Number 4's the guy leaning against the tree.

B: Okay.

Sometimes the matcher responded with two moves, as in *Okay, I've got it,* or with a brief confirmation of the description plus an okay, as in *Dancer, okay.* The director would then go to the next position. These all constitute what we will call the *basic exchange.*

Our main interest is in the director's use of the noun phrase for the figure as a whole, here the *the guy leaning against the tree.* By the collaborative view, he presents it as a means for the matcher to identify the intended figure, and she is expected to accept it. In the basic exchange, indeed, the matcher uses her *okay* to assert (1) that she believes she has identified the figure correctly, and (2) that she has placed the figure in the right location. In doing so, she presupposes (3) that she accepts the director's presentation, including his perspective on the referent. Sometimes the matcher handled these components separately. One matcher signaled her identification (1)

and acceptance (3) but signaled trouble with (2), finding and placing the figure: *Okay, um. Wait, just a sec, just a sec. I can't find it again. God ... Okay, okay.* So in the basic exchange, the acceptance process is canonical: the director presents a noun phrase, and the matcher presupposes her acceptance.

The basic exchange should only be possible when the matcher can accept the director's initial presentation without refashioning it. If so, basic exchanges should have occurred seldom on early trials, but often on later trials, where they could be based on prior mutually accepted descriptions. The percentages of basic exchanges on trials 1 through 6 were 18, 55, 75, 80, 88, and 84. This trend is highly significant, $F(1,55) = 84.19$, $p < .001$. Since the basic exchange requires fewer words and turns than most other exchanges, this accounts for much of the decrease in word count and turns in figure 23.2 and 23.3.

Within the structure of the basic exchange, we can now examine the three processes by which the two partners reached mutual acceptance of each reference—initiating, refashioning, and evaluating presentations.

3.1 Initiating a Reference

Suppose the director has just uttered *Number 4 is ...*, intending the next noun phrase to pick out a particular figure. It is at this moment that the referential process gets initiated. We will call the first full noun phrase uttered at the point the *initial* presentation. These noun phrases fall into at least six distinct types.

1. *Elementary noun phrase.* The director utters this type of noun phrase in a single tone group, such as *the guy leaning against the tree*. Presumably, he believes the matcher can accept it canonically. This is the type of noun phrase that usually occurred in basic exchanges.

2. *Episodic noun phrase.* The director utters this type of noun phrase in two or more easily distinguished episodes or tone groups, as in *Number 7's the goofy guy that's falling over, with his leg kicked up*. The first episode ends with *over* and is immediately followed with more of the same noun phrase in a second episode.

3. *Installment noun phrase.* The director utters this type of noun phrase in episodes too, but gets explicit acceptance of each installment before going on, as in this exchange:

A: And the next one is the one with the triangle to the right ...

B: Okay.

A: With the square connected to it.

The director doesn't end the first installment with a try marker but does indicate by his intonation that he intends to go on. His pause is effective in getting the matcher to respond.

4. *Provisional noun phrase.* Often, the director presents a noun phrase he comes to realize is inadequate—a provisional noun phrase—and immediately expands on it without prompting, as in *And the next one is also the one that doesn't look like anything. It's kind of like the tree?* Note that the expansion is *not* part of the initial noun phrase, but comes in a new clause.

5. *Dummy noun phrase.* A speaker usually utters this type of noun phrase, such as *the whatchamacallit,* as a stand-in until he or his partner can produce a more complete noun phrase. We found no dummy noun phrases in our transcripts, though, as we noted, they are found elsewhere.

6. *Proxy noun phrase.* If the director pauses long enough, and if the matcher has some confidence she knows what he is about to say, she can present all or the final part of a noun phrase by proxy, as here:

A: And number 12 is, uh, . . .
B: Chair.
A: With the chair, right.
B: Got it.

In some cases, speakers actively solicits proxy noun phrases with *what's the word?* or *you know,* or by the way they pause or gesture. We found only five clear initial proxy noun phrases in our transcripts, although elsewhere we have documented their existence in great detail (Wilkes-Gibbs, unpublished).

Any of these six types of noun phrases can end with a try marker, as in *Um, the next one's the person ice skating that has two arms?* With it, one partner asks the other for an explicit verdict on the noun phrase, or installment, before they go on. Note that try markers don't turn assertions into questions; this utterance doesn't mean "Is the next one the person ice skating that has two arms?" The noun phrase is the only element within the scope of the try marker. With it the speaker queries whether the noun phrase is acceptable as it stands.

Try markers should be used for noun phrases the director is less certain will be accepted. In our task, as it happened, it was impossible to distinguish try markers on initial noun phrases, which almost always came at the ends of utterances, from rising intonation for the utterances as wholes. Our directors often used rising intonation to mark utterances as members of a list, with the final member getting a falling intonation. Indeed, as the directors went from trial 2 through trial 6, they used the list intonation to end steadily more of the utterances containing their initial noun phrases, from 41% through 78%.

Each type of noun phrase is generally marked by the speaker for its *status,* which reflects the speaker's confidence in the noun phrase being produced. Episodic, installment, and provisional noun phrases almost always had distinctive intonation patterns in our data; dummy noun phrases

Copyrighted Material

Table 23.1
Rules of projection for next move.

Type of noun phrase	Projected next more	
	Unmarked	With try marker
Elementary	Implicated acceptance	Explicit verdict
Episodic	Implicated acceptance	Explicit verdict
Installment	Explicit acceptance	Explicit verdict
Provisional	Self-expansion	Self-expansion
Dummy	Self-expansion	Proxy
Proxy	Explicit acceptance	Explicit verdict

have distinctive lexical content, as with *what's-her-name;* and proxy noun phrases are identifiable by the change in speakers and often by the first speaker's hesitation as well. Each of these noun phrases can be modified by a try marker, by which the speaker implies there is some possibility of a negative verdict. Truly elementary noun phrases are identifiable by their lack of special features.

These status markings, we propose, are used by speakers to project the next move in the acceptance process. For an analogy, consider questions and answers as a type of adjacency pair (Schegloff and Sacks 1973). When A asks B a question, it is "expectable" that B answer it in the next utterance. B's next utterance, of course, need not be an answer, but it is interpreted by its relation to what is expected. The answer is the preferred response. Likewise, an installment of a noun phrase by A projects an explicit acceptance by B; with an added try marker, it projects an explicit verdict: accept or not accept. These projections, however, are unlike true adjacency pairs, in which the first and second parts are always produced by different speakers. A provisional noun phrase by A projects an expansion by A and *not* by B. The moves that we propose are projected by each noun phrase are shown in table 23.1. They are consistent with our data, though they need more support. If confirmed in further work, they become excellent evidence that the two partners tacitly recognize they are engaged in an acceptance process.

In selecting a noun phrase, the director presumably aims at several ideals. He prefers uttering the initial noun phrase himself. He prefers it to be elementary—not an episodic or installment noun phrase; to be adequate, free of errors, and uttered fluently—not in need of refashioning; and to be no more prolix than necessary (Grice 1975). Elementary noun phrases should therefore be the most preferred, and proxy noun phrases the least. Our data are consistent with this ordering though hardly definitive. Table 23.2 lists the percentages of initial references on trials 2 through 6 that belonged to each category; the descriptions listed are those utterances in which a figure was described rather than identified, for example, *Okay, number 7 is*

Table 23.2
Percentages of six types of initial noun phrases for trials 2 through 6.

Type of noun phrase	Trial				
	2	3	4	5	6
Elementary	52	68	69	80	72
Episodic	11	10	8	6	5
Installment	0	0	0	0	1
Provisional	17	14	8	2	6
Dummy	0	0	0	0	0
Proxy	0	1	2	1	1
Description	17	7	12	9	14
Unclassified	3	0	1	2	1

$N = 96$ per column.

like she's dancing. The head is tilted. As the table shows, there were too few installment, dummy, and proxy noun phrases to test. But, as predicted, episodic and provisional noun phrases, which were used only when necessary, declined and by trial 6 had mostly disappeared (linear trend, $F(1,28) = 9.02, p < .01$) What remained were the preferred elementary noun phrases, which increased significantly over trials, $F(1,28) = 17.02, p < .01$.

3.2 Refashioning a Noun Phrase

An initial noun phrase that isn't acceptable must be refashioned. This is accomplished in three main ways.

1. *Repair.* In planning and uttering each noun phrase, speakers monitor what they are doing and, on detecting a problem, set about repairing it (Laver 1973; Levelt 1983; Schegloff, Jefferson, and Sacks 1977). These self-repairs were legion in our transcripts, as in *Um, next one is the guy, the person with his head to the right but his legs are, his one leg is kicked up to the left.* There were also many of what Levelt (1983) has called *covert* repairs, as in *Okay, number, uh, 4 is the, is the kind of fat one with the legs to the left—er, I mean, to the right.* In repeating *is the,* the director might well have been repairing something he was about to say even if we have no way of determining what. The numbers of self-repairs on trials 1 through 6 were 85, 30, 20, 8, 7, and 6; the instances of repeated words were 47, 14, 10, 4, 7, and 1. These declines contribute to the decrease in word count in figure 23.2.

Repairs could also be initiated by the addressee, but all of these in our data could be classified in one of the next two categories.

2. *Expansion.* Once the director has completed a noun phrase, he or the matcher may judge it to be inadequate for the purposes at hand and in need of a phrase, clause, or sentence of expansion. If the initial noun phrase is provisional, the director will expand on it without prompting, as in these two examples:

> Okay, number 1 is the just kind of block-like figure with the jagged right-hand side. *The left side looks like a square.*

> Okay, number 6 is the guy, uh, sitting down with his legs to the left, *and he's kind of leaning his head over.*

Note that the clauses in italics are *not* part of the initial noun phrases, but expansions added to improve on them. If we call the description in the initial noun phrase x and that in its expansion y, then what the director and matcher end up mutually accepting is the compound description $x + y$.

Self-expansions like these should be needed less often the more clearly the director can formulate his initial noun phrases, and they were. The percentages of figure placements with self-expansions, under a strict criterion, were 25, 17, 11, 6, and 10 on trials 2 through 6. This decline also helps account for the decrease in word count in figure 23.2.

When the matcher didn't find the director's initial noun phrase x clear enough, she could signal the need for an expansion y, as in this example:

A: Okay, the next one is the rabbit.

B: Uh—

A: That's asleep, you know, it looks like it's got ears and a head pointing down?

B: Okay.

In the side sequence here, the matcher used *Uh*—with an extended, level intonation to signal that she needed more description, and the director complied. Requests for expansion like this took many forms, often occurring more than once on a single figure. Many times the matcher signaled uncertainty with a tentatively voiced *um, uh huh,* or *yes?* as if saying, "I'm still uncertain, so please expand on your noun phrase." Other times she displayed silence where a reply could have been expected—such as at a pause after a completed utterance. Still other times she repeated the main part of the director's description with a rising intonation, as in this example:

A: Uh, person putting a shoe on.

B: Putting a shoe on?

A: Uh huh. Facing left. Looks like he's sitting down.

B: Okay.

Prompts of this latter type occurred on 15, 3, 3, 2, 1, and 1% of the figure placements on trials 1 through 6.

Overall, matchers should have had less need to request expansions if they had previously found a mutually acceptable description for a figure. On trial 1, 36% of the figure placements included at least one request for expansion (counting prompts as a subtype); on trials 2 through 6, the

percentages decreased to 12, 8, 3, 1, and 3. So requests for expansions also contribute to the decrease in word count and turns in figures 23.2 and 23.3.

The matcher herself often expanded on the director's noun phrase, almost always in the form of a request for confirmation, as in this example:

> A: Um, third one is the guy reading with, holding his book to the left.
> B: Okay, kind of standing up?
> A: Yeah.
> B: Okay.

The matcher initiated a side sequence by accepting what the director had said so far (x) with *Okay*—a *postponement*, as we will call it—but by asking him to confirm her expansion y. Once he accepted it, the side sequence was complete, and with her next *Okay*, the matcher presupposed acceptance of the amended noun phrase $x + y$. Requests for confirmation, like the other forms of expansion, also declined over trials, occurring in 37, 12, 8, 6, 1, and 2% of the figure placements on trials through 6.

Logically, at least some episodic noun phrases might be considered initiating noun phrases plus self-expansions. In this view the director presents an elementary noun phrase *the goofy guy that's falling over*, immediately judges it inadequate, and then adds the restrictive phrase *with his leg kicked up* in a new tone group, all before allowing the matcher to respond. So he adds an expansion just as he does to a provisional noun phrase, but here the expansion is still part of the initial noun phrase and not a new clause. Consistent with this view, the number of episodic noun phrases declined over trials—11, 10, 8, 6, and 5 in trials 2 through 6 (table 23.2)—just as other forms of expansion did.

3. Replacement. Once the director had finished his noun phrase, the matcher could reject it and present a noun phrase of her own, which we will call a *replacement*. The following is one example:

> A: Okay, and the next one is the person that looks like they're carrying something and it's sticking out to the left. It looks like a hat that's upside down.
> B: The guy that's pointing to the left again?
> A: Yeah, pointing to the left, that's it! (laughs)
> B: Okay.

Since the director's noun phrase x was still unacceptable, the matcher presented a description z from an alternative perspective, which the director then accepted. Indeed, the director took up her replacement on the next trial when he said, *And the next one's the guy pointing to the left*. Replacements are different from expansions. In presenting z, the matcher was rejecting x

and replacing it with z, expressing a different description and not merely an additional one. What the two of them accepted in the end wasn't $x + z$, but simply z.

Most replacements in our transcripts included try markers, as in this example. With the demands of the task, it was rare for a matcher to have a strong enough hypothesis to make a replacement. Also, then, replacements by the matcher shouldn't be that prevalent. They occurred on only, 10, 5, 0, 2, 2, and 0% of the figure placements in trials 1 through 6.

3.3 Passing Judgment on Presentations

A presentation, expansion, or replacement that is put forward needs to be judged acceptable or unacceptable. That can be accomplished by three methods.

1. *Acceptance.* Once one person has presented a noun phrase, the partner can *presuppose acceptance* by continuing on to the next contribution, as when the matcher completes the basic exchange with *Okay.* Or the partner can *assert acceptance*, as in the last example, when the director replied *Yeah* to the matcher's trial replacement. Both types occurred in our transcripts.

2. *Rejection.* A partner can reject a noun phrase either directly or by implication. The clearest rejections are asserted, as in *Oh, the ice skater?* followed by *Y—er, no.* Implicated rejections can also be clear, as when the matcher made the replacement *The guy that's pointing to the left again?* Face to face, a partner can offer other signals, like quizzical looks, which should also be effective.

3. *Postponement.* The partner can also signal that she accepts the presentation so far but is postponing final judgment until it is expanded, as with a tentatively voiced *Okay.*

The matcher can also render verdicts by interrupting the director, but then she is generally signaling more than mere acceptance or rejection. Here is one example:

> A: Okay, our kneeling person with the hook on the—
> B: ⎡Okay.
> A: ⎣—left side.

Although the matcher may simply have suffered from mistiming, she was more likely signaling that she didn't need such an extensive description or any further qualifications (see Jefferson 1973, 59).

3.4 The Acceptance Process

As these results demonstrate, the acceptance process is played out in conversation, as in other human affairs, as a series of steps. It takes at least two such steps—a presentation and its acceptance—but it may take more.

Table 23.3
Mutual acceptance as a recursive process.

Initiating a reference	
To initiate a reference,	present x_1 or
	invite x_1
If an x_1 is invited,	present x_1
Refashioning a noun phrase	
If x_i is inadequate,	present revision x_i' or
	expansion y_i or
	replacement z_i or
	request x_i', y_i, or z_i
If an x_i', y_i, or z_i is requested,	present x_i', y_i, or z_i
If x_i', y_i, or z_i is presented,	let $x_{i+1} = x_i'$, $x_i + y_i$, or z_i
Concluding a reference	
If x_i is adequate,	accept x_i
If x_i is adequate and accepted,	conclude mutual acceptance

With the devices summarized in table 23.3, the possibilities are, indeed, unlimited.

The basic process, which might be called the *acceptance cycle*, consists of a presentation plus its verdict. Let x, y, and z stand for noun phrases or their emendations. A present x and then B evaluates it. If the verdict is not positive, then A or B must refashion that presentation. That person can offer a repair x', an expansion y, or a replacement z. The refashioned presentation, whether x', $x + y$, or z, is evaluated, and so on. Acceptance cycles apply iteratively, with one repair, expansion, or replacement after another, until a noun phrase is mutually accepted. With that, A and B take the process to be complete.

A positive verdict from B alone, however, may not bring the process immediately to completion, since A may not be satisfied that B has understood A's reference. This leads to what we will call *follow-ups*, turns initiated immediately after one partner has accepted the noun phrase, as here:

A: The first one's the one I said looked like a rabbit last time.
B: Okay.
A: You've got that one, right?
B: Yeah.

Sometimes follows-ups seemed to have been initiated because the director couldn't tell whether the matcher's *Okay* meant "I understand you so far" or "I have identified the figure and have placed it in my array." Other times they came on the heels of an error or confusion in the previous trial; the director had good reason for seeking reassurance. Still other times they were initiated because the director didn't seem satisfied with his description,

even though the matcher had accepted it, as here:

A: Okay, the next one looks, is the one with the person standing on one leg with the tail.

B: Okay.

A: Looks like an ice skater.

B: Yeah, okay.

On all later trials, this director referred to the figure as an ice skater.

Follow-up sequences may be a good indicator of the director's confidence in the accuracy of their mutual beliefs about the referent. As this would suggest, the number of follow-ups decreased with successive acceptances for each figure. The percentages of figure placements with follow-ups on trials 1 through 6 were 35, 12, 6, 6, 1, and 5.

A mutual acceptance, once reached, can also later be reconsidered. Recall that the goal of the acceptance process is to establish the mutual belief that the listener has understood what the speaker meant. Once a mutual acceptance has been arrived at, many things can shake those beliefs. The mutual acceptance might have been premature or mistaken, and all it takes to revoke it is some reason for thinking it was in error. Mutual acceptances were reconsidered in several cases in our task.

3.5 Minimizing Collaborative Effort

In classical theories of least effort (for instance, Brown 1958; Brown and Lenneberg 1954; Krauss and Glucksberg 1977; Olson 1970; Zipf 1935), speakers try to utter the shortest noun phrases that will enable their addressees to pick out the referent in context. These theories tacitly assume that speakers work alone, again a literary model of reference with all its problems for conversation. Still there seems to be minimization of effort in conversation. Our proposal is that speakers and addressees try to minimize *collaborative effort*, the work both speakers and addressees do from the initiation of the referential process to its completion. The *principle of least collaborative effort*, as we will call it, is needed to account for many features of the acceptance process.

In the collaborative model there is a trade-off in effort between initiating the noun phrase and refashioning it. The more effort a speaker puts into the initial noun phrase, in general, the less refashioning it is likely to need. Why don't speakers always put in enough effort to avoid refashioning? There are three main reasons.

1. *Time pressure.* Speakers may realize they cannot design the ideal noun phrase in the time allowed. So (1) they may be forced to invite or accept a proxy noun phrase rather than have addressees wait for them to plan their own. Or (2) they may have to use a dummy or provisional noun phrase to

give themselves time to plan a better description, which they offer in an immediate expansion. Or (3) they may utter a noun phrase and, finding it inadequate so far, amend it in a second episode.

2. *Complexity*. Speakers may realize that the noun phrase they are designing is too complex to be easily understood, so they present it in installments.

3. *Ignorance*, Speakers may realize that they don't know enough to decide what addressees would accept anyway, so they are forced into trial and error. They try out a description and leave it to the addressees to refashion if it isn't acceptable. This is one origin of try markers.

The six types of initial noun phrases, each modifiable by a try marker, are therefore devices that enable speakers to deal with these three constraints and yet minimize collaborative effort.

The devices used in refashioning are also designed to minimize collaborative effort. Take repair. As Schegloff, Jefferson, and Sacks (1977) noted, repairs are subject to two strong preferences: speakers prefer to repair their own utterances rather than let interlocutors do it; and speakers prefer to initiate their own repairs rather than let interlocutors prompt them to do it. These preferences have several consequences. One is that speakers repair their own utterances as soon as they detect problems (Levelt 1983). This way they minimize the time a potential misunderstanding is on the floor. Speakers also avert potential exchanges as the interlocutor tries to correct the misunderstanding. That minimizes the number of exchanges needed before mutual acceptance. Together, the two preferences help minimize collaborative effort.

Or take expansion and replacement. As with repairs, speakers prefer to make their own expansions unprompted, as in provisional noun phrases and continuations of episodic noun phrases. As for the addressees, they could in principle respond to every noun phrase they didn't understand with *What?*, but that wouldn't be very informative. For collaborative efficiency they try to pinpoint their problem. When possible, they prompt specific expansions (for instance, *Putting a shoe on?*), offer their own expansions (*Kind of standing up?*), or offer replacements (*The guy that's pointing to the left again?*). They also answer speakers' queries. So addressees minimize collaborative effort by indicating quickly and informatively what is needed for mutual acceptance.

The canonical reference is also predicted by least collaborative effort. In it speakers present an elementary noun phrase and addressees presuppose their acceptance. That is, it consists of a minimal noun phrase (not complex enough to warrant installments) and no extra exchanges. So the canonical reference is preferred because it minimizes effort by both parties.

The principles of least effort and of least collaborative effort, therefore, make very different predictions. Least effort predicts that every reference is

made with (1) a standard (literary) noun phrase that (2) is as short as possible and yet (3) specifies the referent uniquely in that context. Least collaborative effort predicts that references can be made with (1) nonstandard, nonliterary noun phrases, (2) with ones the speaker believes are *not* adequate in context, and (3) with devices that draw addressees into the process. In particular, it predicts trade-offs between effort in initial noun phrases and effort in refashioning. It predicts preferences for self-repair and self-initiated repair. It predicts expansions and replacements, and informative requests for expansion. And it predicts a preference for canonical references. On all these counts the evidence favors least collaborative effort.

4 Speaking Generally

Participants in conversation, we have demonstrated, work together even in such a basic process as the making of a definite reference. Our proposal, more generally, is that they take for granted this principle:

> *Principle of mutual responsibility*
> The participants in a conversation try to establish, roughly by the initiation of each new contribution, the mutual belief that the listeners have understood what the speaker meant in the last utterance to a criterion sufficient for current purposes.

With definite reference their attempts take the form of an acceptance process. The speaker initiates the process by presenting one of at least five types of noun phrases or by inviting a sixth. Both speaker and addressees may repair, expand on, or replace this noun phrase in iterative fashion until they arrive at a version they mutually accept. In this process they try to minimize collaborative effort, presenting and refashioning these noun phrases as efficiently as possible. One result is that the preferred form of reference is the one in which the speaker presents an elementary noun phrase and the addresses presuppose their acceptance of it without taking an extra turn.

The principle of mutual responsibility, however, places two important caveats on this process. The mutual belief is to be established "roughly by the initiation of each new contribution" and "to a criterion sufficient for current purposes." Although our findings don't bear directly on these caveats, we think they are crucial.

4.1 The Criterion Problem
In our proposal the participants aren't trying to ensure perfect understanding of each utterance but only understanding "to a criterion sufficient for current purposes." What are these purposes, and how much is sufficient?

Copyrighted Material

Some conversational purposes are broad and dictate a generally high or low criterion for understanding. Suppose A is telling B where he lives. If B's purpose is to be able to get to his house, she will set her criterion high. If it is merely to break the ice at a party, she will set it low. We have all endured, at a low criterion, people who have talked about each of six children and their families, none of whom we care a whit about. The speaker and addressee may even set discrepant criteria, as when a parent talks to a child. In our task the two partners presumably both set their criteria high, since they were trying to get each figure placed without error before going on. That is one reason they were so diligent in reaching acceptances, often explicit ones.

Even in situations of low or discrepant criteria, however, the ground rules of mutual responsibility are still in force. The participants mutually accept each contribution, at least tacitly, before going on to the next. Granted, they may often be play-acting their parts. Yet even in these conversations we should find coordinating signals such as back-channel responses and try markers (though perhaps distributed more unevenly), however insincere they may be. And speakers should feel they are being understood well enough even when they are not.

Many purposes in conversation, however, change moment by moment as the two people tolerate more or less uncertainty about the listener's understanding of the speaker's references. The heavier burden usually falls on the listener, since she is in the best position to assess her own comprehension. When the speaker utters *I just found the keys*, marking the noun phrase as an elementary (rather than a provisional or trial) presentation, the listener is under strong pressure to accept it. After all, the speaker marked it as elementary, so he must believe it to be adequate for current purposes. If she rejects it, she risks offending him by indicating that it wasn't adequate. She also risks revealing her own incompetence if indeed it should have been adequate. Finally, like the speaker, the listener wants to minimize collaborative effort—to avoid extra steps in the acceptance process—and that, too, puts pressure on her to accept. All this encourages her to tolerate a certain lack of understanding, even to feign understanding when it is not justified. She may do this trusting that the holes will be filled in later, or that they won't have serious consequences.

The listener must tolerate uncertainty anyway. Although the two parties might like to mutually accept each element second by second as they proceed. this ideal is impractical. Certain definite references, for example, cannot be understood until the speaker has completed his utterance. In *Although he doesn't know it yet, we are buying a new bicycle for Harry*, the referents for *he* and *it* cannot be identified until *Harry* has been uttered. It would be premature of the addressee to ask, *Who doesn't know what?* after

the first clause. The natural place to ask such questions is immediately after the utterance is complete.

In this view the two partners assume a unit of conversation we have called the *contribution*. It consists, minimally, of the utterance of one sentence on the topic of conversation, where the sentence can be full or elliptical, or even a quasi-sentence like *Coffee, please.* But to become a contribution, the utterance has to be mutually accepted before the initiation of the next contribution, and that process may require repairs, expansions, and replacements of all or part of the initial presentation. Indeed, as Schegloff, Jefferson, and Sacks (1977) have shown, speakers are usually allowed to present utterances without being interrupted. The place their partners initiate most repairs and expansions and offer most replacements is immediately after the presentation and before the next contribution is initiated. It is for this reason that allowing a new contribution to proceed is tantamount to a mutual acceptance of the old one.

4.2 Modes of Language Use

Conversation, though fundamental, isn't the only site of language use. There are novels, newspapers, and letters—literary uses—as well as radio and television broadcasts, sermons, tape-recorded messages, large lectures, and many others. In these circumstances the participants may not have full access to one another and hence cannot adhere to the principle of mutual responsibility as it has evolved for conversation. The principle may get weakened or modified in various ways. Precisely how it is weakened or modified defines a family of *language modes*. In this paper we have described on such mode, the collaborative mode, but there are many others. We will mention just a few.

In many circumstances, as in literary forms, lectures, and radio broadcasts, writers and speakers are distant from their addressees in place, time, or both. They might be assumed to adhere to a weakened version of mutual responsibility:

> *Principle of distant responsibility*
> The speaker or writer tries to make sure, roughly by the initiation of each new contribution, that the addressees should have been able to understand his meaning in the last utterance to a criterion sufficient for current purposes.

How people adhere to the principle should depend on whether they are speaking or writing, and whether the product is extemporaneous or planned.

In spontaneous speech without concurrent listeners, speakers still monitor what they say (Levelt 1983) and can therefore change course in the process of making a reference. If so, they should still (1) initiate the process

with elementary, episodic, provisional, and dummy noun phrases and (2) repair, expand, and even replace their initial noun phrases. It is just that they do all this without feedback from listeners. In a study by Levelt (1983), people were asked to describe complex spatial networks into a tape recorder. As expected, they produced large numbers of what we have called repairs, expansions, and replacements. This is typical of such mono-logues (see also Goffman 1981; Maclay and Osgood 1959). On the other hand, people don't shorten their repeated references as much when speaking into a tape recorder (Krauss and Weinheimer 1966).

Writers with time to plan, edit, and rewrite, however, should satisfy their responsibilities to readers by eliminating everything but elementary pro-posals, and many writers do. Others retain a sprinkling of provisional noun phrases, repairs, expansions, and replacements apparently to affect a spon-taneous style or for other rhetorical effects. So here are two noncollabora-tive modes, one spontaneous and one planned, both the result of adhering to the principle of distant responsibility.

There may be several collaborative modes. In a study by Cohen (1985), pairs of people were recorded as they (1) spoke over a telephone hookup or (2) typed messages that were stimultaneously displayed on both their own and their partner's computer terminals. The task was for one partner to instruct the other in how to assemble a water pump. In both environments the two partners used methods we have argued are part of the acceptance process. But, as Cohen demonstrated, the partners with spoken access used much finer-grained methods than those in the keyboard condition. Instruc-tors on the telephone, for example, were more likely to ask their partners explicitly to identify a referent before they went on. On the keyboard, the two partners couldn't go as quickly, use nuances of intonation, or interrupt each other with such precise timing, so they apparently adapted their collaborative techniques to fit the limitations. How people adapt to such constraints in general is an open question.

Social factors also govern the collaborative mode. An army private being dressed down by a commanding officer is simply not allowed to interrupt or offer the feedback usually found among equals. Yet the officer can interrupt the private, request confirmations, offer replacements, and do much else. In a study by Ragan (1983), interviewers of job applicants initiated many side sequences, whereas the applicants never did. Further, applicants were much more likely than interviewers to qualify statements, revealing uncertainty about the adequacy of their presentations, and to seek acceptance with *you know*. So the form that collaboration takes is also adapted to certain social constraints. How the participants make these adaptations has yet to be established.

Participants in a conversation, we have argued, are mutually responsible for establishing what the speaker meant. Definite reference is only one part

Copyrighted Material

of that process. They must collaborate, in one way or another, on most or perhaps all other parts of speaker's meaning as well. Collaboration may take one form for word denotation, another for demonstrative reference, a third for assertions, and so on, yet there should be commonalities. The techniques documented for definite reference are likely useful for other parts of the speaker's meaning too.

References

Asher, S. R. (1979). Referential communication. In G. J. Whitehurst and B. J. Zimmerman, eds., *The functions of language and cognition.* New York: Academic Press.

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

Brown, R. (1958). How shall a thing be called? *Psychological Review* 65, 14–21.

Brown, R., and E. H. Lenneberg (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology* 49, 454–462.

Clark, H. H. (1985). Language use and language users. In G. Lindzey and E. Aronson, eds., *The handbook of social psychology,* vol. 2. 3rd ed. New York: Harper and Row.

Clark, H. H., and T. B. Carlson (1981). Context for comprehension. In J. Long and A. Baddeley, eds., *Attention and performance IX.* Hillsdale, NJ: L. Erlbaum Associates.

Clark, H. H., and C. R. Marshall (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, and I. A. Sag, eds., *Elements of discourse understanding.* Cambridge: Cambridge University Press.

Cohen, P. R. (1985). The pragmatics of referring and the modality of communication. *Computational Linguistics* 10, 97–146.

Donnellan, K. S. (1978). Speaker reference, descriptions and anaphora. In P. Cole, ed., *Syntax and semantics 9: Pragmatics.* New York: Academic Press.

Duncan, S. D. (1973). Toward a grammer for dyadic conversation. *Semiotica* 9, 29–47.

Elffers, J. (1976). *Tangram: The ancient Chinese shapes game.* New York: McGraw-Hill.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form.* New York: Academic Press.

Glucksberg, S., R. M. Krauss, and E. T. Higgins (1975). The development of referential communication skills. In F. E. Horowitz, ed., *Review of child development research,* vol. 4. Chicago: University of Chicago Press.

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

Goffman, E. (1981). *Forms of talk.* Philadelphia, PA: University of Pennsylvania Press.

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

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

Grice, H. P. (1968). Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language* 4, 1–18.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan, eds., *Syntax and semantics 3: Speech acts.* New York: Academic Press. Also in D. Davidson and G. Harman, eds. (1975). *The logic of grammar.* Encino, CA: Dickenson.

Grice, H. P. (1978). Further notes on logic and conversation. In P. Cole, ed., *Syntax and semantics 9: Pragmatics.* New York: Academic Press.

Jefferson, G. (1972). Side sequences. In D. Sudnow, ed., *Studies in social interaction.* New York: Free Press.

Jefferson, G. (1973). A case of precision timing in ordinary conversation. *Semiotica* 9, 47–96.

Kamp, J. A. W. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof, eds., *Formal methods in the study of language,* part 1. Amsterdam: Mathematical Centre Tracts.

Krauss, R. M., and S. Glucksberg (1969). The development of communication: Competence as a function of age. *Child Development* 40, 255–256.

Krauss, R. M., and S. Glucksberg (1977). Social and nonsocial speech. *Scientific American* 236, 100–105.

Krauss, R. M., and S. Weinheimer (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science* 1, 113–114.

Krauss, R. M., and S. Weinheimer (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology* 4, 343–346.

Krauss, R. M., and S. Weinheimer (1967). Effect of referent similarity and communication mode on verbal encoding. *Journal of Verbal Learning and Verbal Behavior* 6, 359–363.

Kripke, S. (1977). Speaker's reference and semantic reference. In P. A. French, T. E. Uehling, Jr., and H. K. Wettstein, eds., *Contemporary perspectives in the philosophy of language.* Minneapolis, MN: University of Minnesota Press.

Laver, J. D. M. (1973). The detection and correction of slips of the tongue. In V. A. Fromkin, ed., *Speech errors as linguistic evidence.* The Hague: Mouton.

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

Maclay, H., and C. E. Osgood (1959). Hesitation phenomena in spontaneous English speech. *Word* 15, 19–44.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review* 77, 257–273.

Ragan, S. L. (1983). Alignment and conversational coherence. In R. T. Craig and K. Tracy, eds., *Conversational coherence: Form, structure, and strategy.* Beverly Hills, CA: Sage Publications.

Sacks, H., and E. A. Schegloff (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas, ed., *Everyday language: Studies in ethnomethodology.* New York: Irvington Publishers.

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

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist* 70, 1075–1095.

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. (1981). Discourse as an interactional achievement: Some uses of un-huh and other things that come between sentences. In D. Tannen, ed., *Analyzing discourse: Text and talk.* Georgetown University Roundtable on Languages and Linguistics 1981. Washington, DC: Georgetown University Press.

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

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

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

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language.* Cambridge: Cambridge University Press.

Stalnaker, R. C. (1978). Assertion. In P. Cole, ed., *Syntax and semantics 9: Pragmatics.* New York: Academic Press.

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

Yngve, V. H. (1970). On getting a word in edgewise. In *Papers from the Sixth Regional Meeting,* Chicago Linguistic Society, Chicago, IL.

Zipf, G. K. (1935). *The psychology of language.* Boston: Houghton-Mifflin.