2
Changing Ideas about Reference

Herbert H. Clark and Adrian Bangerter

1 Introduction

How do people refer to things? At first, the answer seems simple: they produce the right expression in the right situation. According to John Searle (1969), for example, to refer to a dog, speakers must produce a referring expression (such as *the dog she had with her*) with the intention that it pick out or identify the dog for their addressees. But is the answer really this simple? Accounts of how people refer have changed again and again since about 1960, often dramatically. But how have they changed, and why? In this chapter, we offer a selective, largely personal history of these changes as they have played out in the experimental study of reference. Our goal is not a complete history – an impossible ambition – but a better understanding of what reference really is.

Language use isn't easy to study. It has been investigated largely by three methods – intuition, experiment and observation. With intuitions, you imagine examples of language used in this or that situation and ask yourself whether they are grammatical or ungrammatical, natural or unnatural, appropriate or inappropriate. This was Searle's method. With experiments, you invite people into the laboratory, induce them to produce, comprehend or judge samples of language, and measure their reactions. With observations, you note what people say or write as they go about their daily business. We will name these methods by their characteristic locations: *armchair, laboratory* and *field*.

Each of these methods has its plusses and minuses. Almost every analysis of language use begins in the armchair. There you imagine a wide range of utterances and situations and draw your conclusions. You are limited only by what you can imagine, but that turns out to be quite a limitation. It is impossible to imagine the hidden processes behind planning and word retrieval, and it is difficult mentally to simulate the opportunistic back and forth processes of social interaction. And armchair judgments are known to suffer from bias, unreliability, and narrowness (Schütze, 1996). The laboratory,

in contrast, is especially useful for isolating and measuring hidden processes, as inferred from reaction times, eye movements or brain activation. But when you bring language into the laboratory, you are forced to strip it of its everyday features – often in unknowable ways. The field is the best place to see how ordinary people, unencumbered by theoretical preconceptions and laboratory wiring, actually use language, but it, too, has its dangers. There you are forced to choose what quarry to track, where to track it and what to record, and these lead to their own biases. And in the field it is hard to infer what causes what.

To return to the original question, how, then, do people refer to things? Since 1960, ideas have changed not merely about the *process* by which people refer, but about the very *conception* of what reference is. Both changes, we suggest, came about when scientists got out of their armchairs and went into the laboratory and the field. Language use, we argue, cannot be studied without all three methods. You cannot even begin without armchair observations. You cannot easily draw causal claims outside the laboratory. And yet you cannot really know what language use is, in all its richness, without venturing out into the field.

2 Referring as a cooperative process

Early on, definite reference was viewed as a rather simple act: the uttering of a referring expression adequate to pick out the intended referent uniquely. In an influential paper by David Olson in 1970, the idea was this: 'Language is merely the specification of an intended referent relative to a set of alternatives' (p. 272). So speakers consider a *set of alternatives* – an array of things that they could potentially refer to – and try to find an expression that will uniquely specify the intended referent in that array. To do this, they select a description that is just sufficient for the purpose. Suppose the set of alternatives is this (in an unknown arrangement):

And you want to refer to the figure on the left. You cannot use 'the figure' or 'the circle' or 'the small figure', because these don't specify the left-most figure uniquely. You have to use 'the small circle'. For Olson, reference is the relation between a referring expression and an element of an array.

Olson's account was built on a number of unstated assumptions. Among them were these:

- Referring is an *autonomous act*. It consists of planning and producing a referring expression, which speakers do on their own.
- Referring is a *one-step process*. It consists of the planning and uttering of a referring expression and nothing more.

- Referring is *addressee-blind*. It depends on the context – the set of alternatives in the situation – but doesn't otherwise depend on beliefs about the addressees.
- Referring is *ahistorical*. It doesn't take account of past relations between speakers and their addressees.
- The referent belongs to a *specifiable set of alternatives*.

Olson's account was hardly unique. A variety of these assumptions were common to most accounts of reference in 1970, including Searle's, and some are still taken for granted. The problem is that they are each suspect – indeed, we will argue, wrong. Some have been challenged by field observations, others in the laboratory, and still others in the armchair. We begin with challenges to Olson's assumptions about addressees.

2.1 Reference and cooperation

Are speakers blind to their addressees? Unbeknownst to Olson, Paul Grice had already argued in 1967 that speakers and addressees cooperate in their use of language (Grice, 1975, 1978; see also Sperber and Wilson, 1986). In his account, they adhere to a *cooperative principle* and therefore try to follow four maxims – to be informative, but not too informative; to have evidence for what they say; to be relevant; and to be orderly in how they speak. Indeed, according to Grice's maxim of quantity ('Make your contribution as informative as is required; do not make your contribution more informative than is required'), speakers should design definite descriptions much as Olson predicted. But in Grice's proposal speakers do that because they are trying to cooperate with addressees.

By Grice's maxim of manner, speakers should also be orderly in their use of language. One way to be orderly, according to Clark and Haviland (1974, 1977), is to follow conventional practices in the design of utterances. One of these practices specifies how to express *given* and *new information*. According to Halliday (1967) originally, speakers obligatorily mark utterances for what he called *information focus*. They distinguish between '(1) information the speaker considers given – information he believes the listener already knows and accepts as true; and (2) information the speaker considers new – information he believes the listener does not yet know' (Clark and Haviland, 1977).[1] When June tells David, 'It was George who bought Julia's car', she takes it as given that someone bought Julia's car, and she takes it as new that the buyer was George.

---

[1] The Given–New distinction here is closely related to 'old' and 'new' information as introduced by Chafe (1970), and to 'presupposition' and 'focus' as used by Akmajian (1973), Chomsky (1971) and Jackendoff (1972). See also Prince (1981).

Speakers and listeners should therefore adhere to the tacit contract expressed here as a directive to the speaker (Clark and Haviland,1977, p. 9):

> *Given–New contract.*   Try to construct the given and the new information of each utterance in context: (a) so that the listener is able to compute from memory the unique antecedent that was intended for the given information; and (b) so that he will not already have the new information attached to that antecedent.

June, in saying 'It was George who bought Julia's car', must have assumed that David was able to compute, from memory, the unique antecedent to the given information, the event of someone buying Julia's car, and that he didn't already believe, or was unable to compute, that the buyer was George. The contract in turn enables listeners to rely on a *Given–New strategy*. By that strategy, David would divide June's utterance into 'X bought Julia's car' (given information) and 'X was George' (new information), search memory for a unique event of X buying Julia's car, and replace X with an index to George.

The Given–New strategy leads to a class of Gricean implicatures called *bridging inferences* (Clark and Haviland, 1974, 1977; Clark, 1975, 1977; Haviland and Clark, 1974; see Matsui, 2000). Suppose June describes a scene for David with one of these two sequences:

(1)   I went for a walk this afternoon. The walk was invigorating.
(2)   I went for a walk this afternoon. The park was beautiful.

When she says 'The walk was invigorating', David can find a unique antecedent for 'the walk', namely the walk mentioned in the first sentence. But when she says 'The park was beautiful', he has to do more. He must assume that she intended him to draw an implicature, or *bridging inference*, namely:

(2′)   *Bridging inference*:   June's walk took her through a park.

According to this proposal, a great many definite descriptions require bridging inferences.

Bridging inferences should take extra time or effort. In the original test of this prediction (Haviland and Clark, 1974), people were brought into a laboratory and given sequences of two sentences, to read one sentence at a time:

(3)   Horace got some beer out of the car. The beer was warm.
(4)   Horace got some picnic supplies out of the car. The beer was warm.

Inferring the referent for 'the beer' in 'The beer was warm' requires a bridging inference in (4), namely 'There was beer among the picnic supplies'. No such bridging inference is required in (3). Indeed, it took people about 200 msec

longer to read and understand 'The beer was warm' in (4) than in (3). In a later study (Garrod and Sanford, 1982), people were given sequences like these:

(5) Keith took his car to London. The car kept overheating.
(6) Keith drove to London. The car kept overheating.

To identify the referent of 'the car' in 'The car kept overheating' requires a bridging inference in (6) ('What Keith drove was a car'), but not in (5); however, it took no longer to read and understand 'The car kept overheating' in (6) than in (5). Apparently, the bridging inference took little time to create, or the proposition was already part of the discourse model (see, e.g., Brewer and Treyens, 1981), so the bridging itself added no *measurable* time to the reading process. Replace *car* in (5) and (6) with *motorcycle*, and (6) would have taken more time. Over the years it has been found that the more readily inferable the bridging inference is in context, the less extra time is needed to infer it (see Matsui, 2000).

What is remarkable about the work on bridging is how little of it has come from the laboratory. The original model was built on field observations – from books, newspapers and magazines (Clark and Haviland, 1977), and from spontaneous speech (Halliday, 1967). And the predictions of that model – we have mentioned only some of them – were sharpened by armchair judgments about a vast range of examples analogous to (1) through (6) (e.g., Clark, 1977; Clark and Haviland, 1977), and that tradition has continued (Matsui, 2000). In short, the model was refined with evidence from the armchair, laboratory *and* field. No one method would have been enough.

2.2 Reference and common ground

Given information was originally characterized as 'information the speaker believes the listener already knows and accepts as true'. But this cannot be right, as bridging inferences themselves show. When June referred to 'the park' in (2), she did *not* need to believe David already knew she had walked in a park. All she needed to believe was that David could *infer* she had walked in a park – from what he already knew plus her definite reference. But can speakers work merely from what they believe their addressees know? Not in general. Speakers and addressees must rely on *shared* information – but shared in a particular way. The needed concept is what has come to be called *common ground*.[2]

Suppose June points at a cup of coffee on a nearby table and tells David, 'The cup of coffee is for you'. How can she refer to that as 'the cup of

---

[2] For early formulations of common ground, see Karttunen (1977), Karttunen and Peters (1975), but especially Stalnaker (1978).

coffee'? According to armchair examples in Clark and Marshall (1981), she needs to find a basis for the following belief: She believes that, with this reference, (*i*) she and David will come to believe, or assume, both that there is a cup of coffee nearby and that she and David believe, or assume, *i*. This, technically, is a *reflexive* belief – her and David's belief *i includes* that very belief *i*.[3] In general, for A to have information that a proposition *p* is common ground for A and B, A needs information that:

(i)   A and B have information that *p* and that *i*.

'Having information' may include knowing, believing, assuming, seeing, hearing, even feeling, so common ground can range from mutual knowledge to mutual supposition, and its basis can range from abstract inferences to immediate perception.[4] The idea is that to make a definite reference is to presuppose that the referent can be readily and uniquely inferred from the current common ground of speaker and addressees. To make an indefinite reference is to presuppose that the referent *cannot* be so inferred.

Common ground, according to Clark and Marshall (see also Clark, 1996), divides into two broad types, *communal* and *personal* common ground:

Communal common ground.   Suppose June and David are strangers, but in talking to each other at a party, they discover they are both on the faculty at the University of Illinois. This way they mutually establish that they are both members of at least these communities – United States, Illinois, Champaign, American academics, University of Illinois employees. The two of them can therefore take as *communal* common ground all the information they believe is common to members of these communities. On the basis of residence, June can refer, for example, to 'South Prospect', 'the psychology department', 'the stadium', 'the road to Chicago', confident that David will be able to identify the referents uniquely.

Personal common ground.   Once June and David begin interacting, June may infer that certain items are *co-present* to the two of them. That includes items

---

[3] Reflexive beliefs, as formulated here, are analogous to the reflexive intentions on which Grice formulated the notion of speaker's meaning (Grice, 1957, 1968). See Harman (1977).

[4] As an aside, we note that common ground was characterized as reflexive by Lewis (1969), but as an *infinite regress* of beliefs by Schiffer (1972). The point of Clark and Marshall's paper was to argue for Lewis's characterization and against Schiffer's – at least, as a psychological model of common ground. All too many investigators have cavilled about common ground based on Schiffer's, not Lewis's characterization (see Clark, 1996, ch. 4).

they have talked about – say, her walk that afternoon. It also includes items that are perceptually co-present – say, the cup of coffee on the table. On the basis of these joint personal experiences, she can take the walk and cup to be part of their *personal* common ground and refer to them as *the walk* and *the cup of coffee*.

Do people rely on common ground in referring? Early examples from the armchair suggested yes (see Clark and Marshall, 1981; Hawkins,1978). But it soon became clear that the question is too simple. Common ground plays a role in every aspect of language use, from word meanings to politeness formulae. Here we consider its role in convention and coordination; later, we see how it forced a change in the conception of reference itself.

2.3 Reference and coordination

Even more fundamental than Grice's notion of cooperation is the notion of *coordination*: For June and David to cooperate on a task, they have to coordinate their individual actions (see Clark, 1996, ch. 2). In 1960, Thomas Schelling had already laid out several principles of coordination, and in 1969, David Lewis showed how these applied to language use. Lewis argued, in particular, that principles of coordination are needed to account for the linguistic notion of convention. Take the word *dog*. Speakers and their addressees have a recurrent problem of coordination: how are the speakers to refer to a domesticated canine such that their addressees will understand them as intended? They need what Lewis called a *coordination device*. The solution that evolved historically is the convention that the word-form /*dôg*/ mean 'domesticated canine.' What makes it a convention, according to Lewis, is that it is a regularity in behaviour, partly arbitrary, that is common ground in the community of English speakers as a coordination device for a recurrent coordination problem.

Conventions, of course, are essential to most forms of reference. When people refer with proper names – *Washington, Bonaparte, van Gogh* – they count on addressees picking out the right referents by means of conventions associated with the names. And when English speakers refer to walks, parks, beer, cars, stadiums and cups of coffee, they do so in part by means of the conventions associated with the nouns *walk, park, beer, car, stadium, cup* and *coffee*. Coordinating with words requires conventions, and deploying conventions requires common ground.

Coordination *à la* Lewis and Schelling, however, goes beyond convention, for there are other coordination devices as well. What these devices have in common is that people coordinate with them by means of *salience against current common ground*. It is this formulation of coordination that is needed for demonstrative reference.

The point was made clear in a series of experiments reported by Clark, Schreuder and Buttrick (1983). In one experiment, a student named Sam

approached people on the Stanford University campus, handed them a photograph of a garden with four types of flowers in it, and asked:

(7)   How would you describe the colour of this flower?

According to a model like Olson's, and even Grice's, the referent of *this flower* is underdetermined because the description *flower* is not informative enough to distinguish among the four flowers. If so, people should find the question impossible to answer. But by a Lewis-style model of coordination, listeners expect speakers to coordinate with them so that they can readily pick out the referent by considering the referring expression (here 'this flower') against their current common ground. By the coordination model, people should find Sam's question acceptable if they can pick out such a referent.

People in this experiment had been handed one of two pictures. In the first, one flower was a bit more salient perceptually than the other three; in the second, no flower was particularly salient. For the first picture, people immediately described the salient flower 55 per cent of the time, requesting clarification the rest of the time ('Which one?' or 'This one?'). But for the picture with no salient flower, people described a flower only 15 per cent of the time, requesting clarification the rest of the time. People took the perceptual salience of a flower against common ground as the means by which they were supposed to coordinate on identifying the referent.

In the coordination model, people don't coordinate by perceptual salience *per se*, but by salience against common ground more broadly. In a second experiment (Clark et al., 1983), Sam handed people on campus a photograph of Ronald Reagan and David Stockman equally salient standing side by side. At that time, Reagan was president of the United States, and Stockman was the director of the Office of Management and Budget. Sam then asked one of two questions:

(8)   You know who this man is, don't you?
(9)   Do you have any idea at all who this man is?

By many models, people should either: (a) pick out the same man for 'this man' in both questions; or (b) refuse to pick out either man. By the coordination model, people should take Sam's presuppositions into account and find one man more salient than the other against their current common ground. With (8), Sam presupposed 'this man' was familiar, but with (9), he presupposed 'this man' was *not* familiar.

The results were clear. With (8), people answered 'Reagan' 80 per cent of the time, requesting clarification the rest of the time. They never answered 'Stockman'. But with (9), people answered 'Stockman' 20 per cent of the time, requesting clarification the rest of the time. They never answered 'Reagan'. Afterwards, all of the participants were able to identify both men in the photograph, and all rated Reagan as the more familiar for Stanford students.

For these addressees, then, the current common ground included not only the photograph, but the public familiarity of the two men, and Sam's presuppositions. They selected the referent by judging the most salient possibility against that common ground.

The picture so far is this. Speakers ordinarily perform their acts of referring for particular addressees. They design referring expressions in the expectation that their addressees will be able to pick out the referents as the most salient possibility against their current common ground. Listeners, in turn, rely on the same logic to identify the referents. This is no more, really, than an elaboration of cooperation and coordination as characterized by Grice and Lewis.[5] Once again, it took evidence from armchair, laboratory and field to create this picture. Indeed, we see with Sam's questions on the Stanford University campus the beginning of a combined method – the field experiment, or laboratory in the field.

3 Referring as an interactive process

Coordination and cooperation are profoundly social. June and David try to keep close track of what each other is doing, thinking, looking at, and when they act, they try do so by anticipating what the other will do. Yet in the models reviewed so far, referring is treated not merely as an *individual* act – something speakers do themselves – but as an *autonomous* act – something speakers do *by* themselves. Referring is a one-shot process that is complete once the speaker has uttered 'the beer', 'Washington' or 'this man'. As Searle (1969) put it: 'The unit of linguistic communication is not . . . the symbol, word, or sentence, but rather the production or issuance of the symbol, word, or sentence, in the performance of a speech act' (p. 16). For him, linguistic communication is like writing a postcard and dropping it in the mailbox. It doesn't matter whether the addressee receives, reads or understands it. But this feels wrong. How can the concepts of coordination and cooperation be so social, and the act of referring be so non-social?

3.1 Language and joint action

The answer was already present, though not entirely explicit, in Lewis's 1969 characterization of coordination. Suppose that Michael, an accomplished pianist, plays Bach's 'Two-part Invention Number 3' *as a solo*. Next, June and David, also pianists, play the same two-part invention, but *as a duet*, June playing the right-hand, and David the left.[6] We make recordings of both

---

[5] Note that speakers don't design their references for just any hearer. They can prevent bystanders and overhearers from understanding, as when June tells David knowingly, 'You know who has just arrived' (Clark and Carlson, 1982; Clark and Schaefer, 1987, 1992). This, too, is evidence that speakers aren't blind to their audience.

[6] This example is based on a childhood experience of Clark, who would team up with a girl named Jane to play Bach fugues that he couldn't play by himself.

performances. If June and David are successful, we shouldn't be able to tell the duet from the solo. But how did June and David manage, with their two individual actions, to sound just like Michael, with his single individual action?

Duets are performed by two people acting in concert. Let us represent what June and David are doing mid-performance this way:

(a)   The pair June-and-David are playing a two-part invention that consists of b and c.
(b)   The individual June is playing the right-hand part *as a part of (a)*.
(c)   The individual David is playing the left-hand part *as part of (a)*.

The duet is represented in (a), and June's and David's individual acts in (b) and (c). The schema shows how June cannot intend to play her part *as part of the duet* without believing that David is trying to play his part *as part of the same duet*, and vice versa. June's and David's acts are individual acts, but ones that cannot be carried out without the other. These are called *participatory acts* (Clark and Schaefer, 1989). Yes, June is playing the left-hand part herself, which makes it an individual act, but she is doing it as part of the duet, which makes it a participatory act. The duet itself is a *joint action*. Although Michael's performance is also an individual act, it is an autonomous one.

Face-to-face conversation, the primary form of language use, is also a *joint activity*. For June and David to talk to each other, they have to coordinate their individual actions at at least these four levels (Clark, 1996):

Level 1   June produces vocalizations and gestures for David; David tries to attend to them.
Level 2   June presents words and phrases for David; David tries to identify them.
Level 3   June means something for David; David tries to understand what June means.
Level 4   June proposes something for David; David tries to consider her proposals.

June and David coordinate not only on the *content* of these four levels of action, but on the *timing*: They synchronize what they are doing. In face-to-face conversation, speaking and listening are participatory acts that require as-much coordination as playing the two parts of the Bach two-part invention.

If language is a joint activity, then we need to treat referring as part of a joint activity. An obvious solution is to treat it as a participatory act by the speaker – as one half of a joint act by the speaker and addressees. Let us review evidence for such a change.

3.2 Referring and interaction

Referring had been shown to be a social act long before Olson's proposals. In 1964 and 1966, Robert Krauss and Sidney Weinheimer described a referential communication task in which participants were given an ordered

array of four figures and asked to describe them so that a partner, hearing those descriptions, could put the figures into the same order. The figures were abstract difficult-to-describe drawings; one, for example, pictured a figure that was first described by one participant as 'an upside-down martini glass in a wire stand'. The participants repeated this task with the same few figures used many times. Olson's account would predict that the martini figure should be referred to each time as 'the upside-down martini glass in a wire stand', because that was the description originally needed to distinguish it from its alternatives. In fact, the descriptions got shorter on successive references, going from 'the upside-down martini glass in a wire stand' to 'the inverted martini glass', to 'the martini glass', and finally to 'the martini'. In a model like Olson's, referring shouldn't reflect previous references, except in pronouns. But as Krauss and Weinheimer demonstrated, it does. Referring is a historical process.

Even more remarkable was Krauss and Weinheimer's (1966) finding on feedback. In the same task, some participants were paired with active partners who provided on-going feedback such as 'uh-huh', 'which one?' and 'got it'. Others were paired with a tape-recorder into which they spoke to future partners. When there was feedback, referring expressions became shorter over successive references, as in the martini example; the average length dropped from ten to two words. But when there was no feedback, there was no shortening.[7] A model like Olson's has no explanation for such a phenomenon.

Krauss and Weinheimer's experiments raised two important issues. What role does feedback play in acts of referring? And how do speakers appeal to previous references in the current act of referring?

3.3 Referring and collaboration

Common ground isn't a homogeneous body of well-established propositions. When June and David are talking, some elements are firmly established at the moment one of them speaks, but others are in doubt. Most propositions haven't even been assessed for whether or not they are in common ground. Also, June and David must reassess their common ground with each new utterance – indeed, each new bit of utterance – and with each new joint perceptual experience. So people in conversation are faced with an ongoing practical problem: how are they to assess and establish what is common ground as they race along in their conversation?

Establishing common ground is a central problem in acts of referring, and people have evolved strategies for doing that. In one documented example

---

[7] Later studies (e.g., Bavelas, Coates and Johnson, 2000, 2002; Clark and Krych, 2004; Fox Tree 1999) showed that references designed without feedback are also inferior in quality; some are even impossible to grasp.

(from Sacks and Schegloff, 1979), Ann wanted to refer to a couple named Ford. She thought her friend Betty might recognize the couple by that name, but she wasn't sure. How should she design her reference? Ann's solution was ingenious:

(10)   *Ann*:   . . . well I was the only one other than than the uhm tch *Fords?*,
                 uh Mrs Holmes Ford? You know uh
                 ⌈the the cellist?
       Betty: ⌊Oh yes. She's she's the cellist.
       Ann: Yes. Well she and her husband were there.

Ann produced *the Fords* with a so-called *try marker*, a rising intonation (indicated by the question mark) followed by a slight pause (Sacks and Schegloff, 1979). In doing that, she implied that Betty might *not* recognize the couple by that name, and she was requesting Betty to say yes – 'uh-huh' or a nod – if she *did*. When Betty didn't say yes immediately, Ann tried *Mrs Holmes Ford*? And when Betty didn't say yes to that, she tried *the cellist*? This time Betty said 'Oh yes', even interrupting Ann to keep her from trying yet another expression, and confirmed her understanding with 'She's she's the cellist'. Only then did Ann go on.

Most models of definite reference assume that speakers refer as if they were writing to distant readers. Searle's and Olson's accounts are two examples. These *literary models*, as Clark and Wilkes-Gibbs (1986) called them, make several tacit assumptions: (1) References are expressed with standard noun phrases (e.g., proper names, definite descriptions, pronouns). (2) Speakers use these noun phrases intending their addressees to be able to identify the referents uniquely against their common ground. (3) Speakers discharge their intentions simply by the issuing of such noun phrases (see Searle, 1969). And (4) the course of the process is controlled by the speakers alone.

Ann's reference to the Fords is at odds with all four assumptions: (1) The noun phrase *the Fords*? with its added try marker is not standard. (2) In using it, Ann did not necessarily expect Betty to be able to identify the referent uniquely; that is why she added the try marker. (3) Ann did not discharge her intention to get Betty to identify the referent simply by producing the noun phrase. On the contrary, (4) she used *the Fords*? to initiate a process that required Betty's collaboration. She recognized that the process was controlled not just by her but by Betty too. If Betty had nodded immediately after *the Fords*?, Ann would have stopped there. As it was, she had to try three noun phrases (*the Fords?, Mrs Holmes Ford*? and *the cellist*?) before going on. Ann's act of referring was not an autonomous act, but a *participatory* act. It required Betty's coordinated participation in a duet-like joint action. Indeed, her act was opportunistic in a way that she could not have done alone.

3.4 Referring and grounding

The collaborative nature of referring was documented in an experiment by Clark and Wilkes-Gibbs (1986). The experiment was a referential communication task, but with several changes from Krauss and Weinheimer's method. Instead of one naïve participant and a confederate of the experimenter, there were two naïve participants – a *director* and a *matcher*. Instead of arrays of four abstract line drawings, there were 12 hard-to-describe Tangram figures – abstract geometric shapes that vaguely depicted people. And there were six trials on which the director got the matcher to arrange his 12 figures in the same order as the director's.

Although the new experiment confirmed many features of the original task, it raised new issues. As in the original task, directors needed fewer words per figure as they referred repeatedly to each figure – from 42 words on the first trial to ten on the sixth. But along with the drop in words was a drop in the number of *turns* per figure – from five on the first trial to two on the sixth. Why? Because of the way directors and matchers collaborated with each other on each reference.

The most striking finding was how often the partners made references with techniques not described in grammars of English – techniques that required contributions from both participants. Speakers initiated the process of referring with a variety of standard and non-standard noun phrases. Most of the following types are from the Tangram task, although a few are from other conversations:

1. **Elementary noun phrases**, such as *the guy leaning against the tree*.
2. **Self-repaired noun phrases**, such as *the guy reading with, holding his book to the left*.
3. **Other-repaired noun phrases**, as when B repairs A's 'Monday' in this example (Schegloff, Jefferson, and Sacks, 1977, p. 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 to*mor*row.
   A: Yah.
4. **Episodic noun phrases**. These are noun phrases produced in two or more intonation units, such as *the person ice skating, with two arms*.
5. **Instalment noun phrases**. In these, the speaker produces a first instalment of a noun phrase, waits for an acknowledgement, and then produces the next instalment, and so on, as here:
   Director: And the next one is *the one with the tail to the right*,
   Matcher: Okay.
   Director: *With the square connected to it*.
6. **Expanded noun phrases**. These are noun phrases expanded at the instigation of the addressee, as in the italicized expansion instigated by the matcher's 'uhhh':
   Director: Okay, the next one is *the rabbit*.

Matcher: Uhhh –
Director: *That's asleep, you know*. It looks like it's got ears and a head
           pointing down?
Matcher: Okay.

7. **Trial noun phrases**, or noun phrases with try markers, such as *the
   person ice skating that has two arms*? (or *the Fords*?). With these, speakers
   request a yes or no answer, which then leads to two different courses of
   action.

8. **Holder noun phrases**. These are conventional expressions used as
   stand-ins until the speaker or addressee can come up with a better noun
   phrase, as in 'it may take a hell of a long time to come, . if he puts it
   into the diplomatic bag, as u:m – *what's his name*,. Mickey Cohn did',
   (1.1.83, Svartvik and Quirk, 1980). Other examples: *the whatchamacallit,
   what's-its-name*, and *you-know-who* (see Enfield, 2003).

9. **Invited noun phrases**. With these, speakers invite a completion from
   their addressees (see Lerner, 1987; Wilkes-Gibbs, 1986), as here:
   Director: And number twelve is, uh,
   Matcher: *Chair*
   Director: With a chair, right.

Examples like these challenge the notion of referring as an autonomous act.
In every case, speakers require coordinated actions from their addressees. In
cases (3), (5), (6), (7), (8) and (9), there is an explicit request for help. In the
rest, there is an implicit one.

What do all these noun phrases have in common? It is clearly not that
they constitute complete references in and of themselves. When speakers
produce *the rabbit* in (6) or *the whatchamacallit* in (8) or *uh* in (9), they recog-
nize that these are hardly adequate to complete the references. Rather, they
use these noun phrases to *initiate* a process with two intertwined goals:

Identification.   Speakers are trying to get their addressees to identify or
pick out a particular figure under a particular description.

Grounding.   Speakers and their addressees are trying to establish the
mutual belief that the addressees have identified the referent well enough
for current purposes.

Identification is the traditional goal of referring – it is central to Olson's and
Searle's models. What is new here is *grounding*.

Grounding is a very general notion. *To ground a thing is to establish it as
part of common ground well enough for current purposes* (Clark, 1996; Clark and
Brennan, 1991; Clark and Wilkes-Gibbs, 1986). In conversation, speakers
and addressees ordinarily try to ground everything that gets said (Clark and
Schaefer, 1989), and that includes the identification of referents (Clark and

Wilkes-Gibbs, 1986; Schober and Clark, 1989). The grounding process is neatly illustrated in Ann's reference to the Fords. It consisted roughly of these steps:

Step 1	Ann says *the Fords*?	Betty doesn't respond, implying failure of identification
Step 2	Ann says *Mrs. Holmes Ford*?	Betty still doesn't respond, implying failure of identification
Step 3	Ann says *the cellist*?	Betty says *oh yes*, confirming the identification with *she's the cellist*
Step 4	Ann accepts Betty's evidence by going on	Betty accepts Ann's judgment

Each type of noun phrase we listed initiates a different grounding process, one with different presuppositions and different prospects.

With examples like these, it is no longer possible to maintain that referring in conversation is an autonomous, one-shot act by the speaker. Rather, it is a participatory act that requires coordinated actions from the addressee. We would never have discovered this in an armchair. We needed conversation from the field – or at least from field-like conditions in the laboratory.

3.5 Reference and conceptual pacts

In history books, the same individual might be referred to as *Napoleon, the loser at Waterloo*, and *the man who instituted modern French civil law*. Although these pick out the same individual, they do so under different descriptions – *qua* person named Napoleon, *qua* loser at Waterloo, and *qua* man who instituted modern French civil law. Each description reflects a different conceptualization, a different perspective. Indeed, every reference (so far as we can tell) is intended to pick out: (i) an individual; (ii) under a description. Descriptions play an important role in referring (a role absent from Searle's account, implicit in Olson's, and undeveloped in others). Let us see how.

One of the Tangram figures was referred to by various people as 'the rice bag', 'the whale', 'the complacent one', and 'the stretched-out stop sign' (Schober and Clark, 1989), and another one as 'the ice skater' and 'the ballerina' (Clark and Wilkes-Gibbs, 1986). Reaching these perspectives took coordination. It wasn't enough for a director to see a figure as, say, an ice skater. He or she had to get the matcher to see it that way too. Consider this first reference to a Tangram figure:

(11) *Director*: Okay, the next one looks, is the one with the person standing on one leg with the tail.
    Matcher: Okay.

> *Director*: Looks like an ice skater.
> *Matcher*: Okay.

First, the director got the matcher to identify the figure as 'one with the person standing on one leg with the tail'. But then he offered a more memorable perspective ('ice skater'), which the matcher agreed to ('okay').[8]

What the director and matcher established in this exchange was a *conceptual pact*, a temporary agreement about how they were to conceptualize the referent (Brennan and Clark, 1996). Participants establish conceptual pacts each time they achieve a reference. Typically, speakers offer a conceptualization that is tentative or provisional, using hedges such as *sort of, kind of* and *looks like*. Once their addressees agree to the conceptualization, speakers drop the hedges and treat it as a conceptual pact that they can appeal to later. The director in (11), for example, referred to the figure from then on as *the ice skater*. To ignore the pact and call it *the ballerina* would implicate that the first perspective needed to be replaced by the second. Once two people have established a conceptual pact, it is cooperative to continue using it – unless there are reasons to change.

Conceptual pacts lead to several predictions about referring (Brennan and Clark, 1996; Van Der Wege, 2000). In one referential communication task, a first group of participants were given a set of 12 figures that included three shoes, whereas a second group were given a set that included just one of the three shoes. As expected, the shoe common to both sets was called, for example, *the dress shoe* by the first group, but *the shoe* by the second group. Speakers were no more informative than they needed to be – conforming to Grice's Maxim of Quantity. After a few trials with the same sets, the first group was given a set of 12 figures that had only the one shoe in it. On Olson's or a Gricean account, they should immediately switch to calling it simply *the shoe*. But they didn't. Most continued calling it *the dress shoe*, even though that was more informative than necessary. Why? Once the two partners had established the conceptual pact 'dress shoe', it was cooperative to continue using that pact. Indeed, the more firmly they had established that pact in the initial trials, the longer they continued using *the dress shoe* before simplifying it to *the shoe*. More than that, speakers tended *not* to retain *dress shoe* for the lone shoe when they got a new partner (see also Wilkes-Gibbs and Clark, 1992).

The picture of referring has changed once again, and this time into a truly social act. Speakers perform acts of referring *in collaboration with* their addressees. Referring is a participatory act: like playing in a duet, it requires the co-participation of the addressee. In conversation, therefore, speakers go

---

[8] One pair was unable to reach a single perspective, so the director referred to the figure thereafter as 'your monk and my machine gun'.

beyond standard noun phrases. They use both standard and non-standard forms to initiate the grounding of that reference – establishing the mutual belief that their addressees have understood it as intended. Speakers and addressees come to agree not only on which individual is being referred to, but on how that individual is to be conceptualized. Once again, these conclusions come from the judicious combination of evidence from armchair, laboratory and field.

4 Referring with language and gestures

At first, referring was treated as an autonomous act based mostly on literary or armchair examples. But once investigators began studying conversations on audiotape that conception changed, and referring was treated as a participatory act – one half of a joint act by speakers and addressees coordinating with each together. Even in this conception, referring was treated as a *linguistic* act, one achieved exclusively through language. But once investigators began studying conversations on *video*tape, that conception changed yet again. Let us use the term signal as an act by which speakers mean something for their addressees *à la* Grice (1957). When people were videotaped in conversation, they were found to exploit a range of signals that were not linguistic at all. Referring was seen to be a multi-method process, one that normally requires more than one *method* of signalling.

Work from the field shows that speakers routinely *anchor* their references to the material world – to actual people, artifacts, rooms, buildings, landscapes, events, processes (Clark, 2003; Goodwin, 2003; Hindmarsh and Heath, 2000). When June refers to a nearby building, 405 North Mathews Avenue, as 'that building', it isn't enough for David to understand that she is referring to a particular building in their current common ground. He is to understand that she is referring to that huge building *over there* – the one she is now pointing at – and she expects him to look at it to confirm. Recall that the act of referring establishes two things: (i) an individual as the referent; (ii) a conceptualization or perspective on that individual. Schematically,

referring = indicating + describing

Describing may be doable through language alone, but indicating requires something more – the locating of an individual in relation to the speaker and addressee's here and now.

4.1 Methods of indicating

Indicating is a method of signaling built on C. S. Peirce's notion of *index* (see Buchler, 1940). Pointing is a prototypical example. When June points at 405 North Mathews Avenue, she intends David to construe her pointing as an index to that building. Unlike symbols (such as the words *building* and

work), indexes work by means of *intrinsic connections* between them and their objects. To quote Peirce, an index designates its object 'because it is in dynamical (including spatial) connection both with the individual object, on the one hand, and with the senses or memory of the person for whom it serves as a sign, on the other hand' (Buchler, 1940, p. 109). There is an intrinsic connection between June's pointing gesture (the timing and direction of her finger and hand motions) and 405 North Mathews Avenue. Indexes often work by directing our attention. According to Peirce, 'A rap at the door is an index. Anything which focuses the attention is an index' (Buchler, 1940, p. 108).

In indicating, then, speakers get their addressees to *focus attention* on individual objects. In discourse there is always a focus of attention, and it is often exploited, for example, in referring with pronouns (Brennan, 1995; Grosz and Sidner, 1986). Indicating is analogous but focuses *visual* or *auditory* attention on *physical* objects, events and states. Speakers can indicate in two main ways (Clark, 2003). One is by directing the addressees' attention to an object. This is what June does in pointing at 405 North Mathews Avenue for David. This technique is called *directing-to*. Another way is by placing an object within the addressees' current or future locus of attention. June, at a bookstore, places a book on the checkout counter to indicate that book to the clerk as one she wants to buy. This technique is called *placing-for.* Directing-to and placing-for are exploited in acts of referring both when required and when circumstances permit.

Focusing attention creates even more difficulties for Olson's model. By focusing addressees' attention on a subset of potential referents, speakers don't have to specify the referents in relation to *all* alternatives; they can get away with an otherwise ambiguous description (Beun and Cremers, 1998; Clark, Schreuder and Buttrick, 1983). But then referring is not a one-step process anymore: indicating (focusing attention) and describing are at least two components. We will return to this question later.

4.2 Indicating within acts of referring

Deictic references, which contain words like *this, that, here, there, I* and *you*, are often incomplete without visible or audible acts of indicating. Such acts are common in field observations, as in this example (Schegloff, 1984, p. 280):

(12) *Frank*: why:nchu put that t the end uh the ta:ble there [pointing]

Although Frank refers to the dish 'that' in front of Marge without a gesture, he refers to the end of the table 'there' by pointing at it. Without the gesture, he couldn't get Marge to recognize where precisely on the table he meant. So in referring to that location, he used a *composite signal* – a combination of describing (the use of *there* to mean 'in that place') and indicating (his pointing gesture).

The problem is that many indicative gestures are performed *without* affiliated noun phrases. Here is an example from the field (Schegloff, 1984, p. 284):

(13) *Linda*: en I'm getting a sun tan [pointing at her two cheeks in turn]

As Linda says this, she points first at her left cheek and then at her right, asserting that she is getting a tan on her cheeks. She might be paraphrased as saying 'I'm getting a sun tan **on my face**', but she doesn't say 'on my face'. Examples like this – and they are legion – bring out the radical change required for multi-method communication. Linda and her addressee are to combine Linda's linguistic signal ('en I'm getting a sun tan') with her gestural signal (her pointing to her cheeks) to get an interpretation something like 'I'm getting a sun tan on my face'. We don't know of formal systems of pragmatics in which this is possible.

Indicative gestures can also be performed by addressees. In one experiment (Clark and Krych, 2004), one participant (called the director) was required to tell another participant (called the builder) how to put Lego blocks together to form a small, seven-block abstract sculpture. When the directors could watch their partners work, the builders would often ground the directors' references with forms of directing-to and placing-for. The following is an exchange from the middle of a sculpture (with overlapping actions marked):

(12) *Director*: and now get
⌈[.75 sec] a-uh eight piece green, [1.50 sec]
Builder: ⌊[begins reaching into a collection of blocks, picks one out, then **exhibits** it]
Director: and join the two so that it's all
⌈symmetric,
Builder: ⌊[**poises** block above a location]
Director: yeah right in the center.

In (12), once the builder finds what he thinks is the right block, he holds it out for the director to inspect, and the director confirms its correctness by going on. That is, the builder *exhibits* the block to the director, a form of directing-to. Once the builder believes he knows where the block goes, he then *poises* it just above the location, to which the director says 'yeah'. Poising is also a form of directing-to.

As examples like these show once again, referring is a participatory act: the course of an utterance – its phrasing and timing – is determined not by the speaker alone, but by the speaker in collaboration with the addressees. In (12), the director delays going on for fully 1.5 seconds while the builder retrieves and exhibits the right block. In another example, the director says,

'Put it at the end of the red that's o- the other end'. She alters the course for her utterance (at 'o- the other end') when the builder points at the wrong location. In another example, a director says, 'And put it on the right hand half of the **yes** [0.30 sec] of the green rectangle'. She interrupts her reference to say 'yes' to confirm the location where the builder has poised the block. In the building of Lego models, examples like this were common.

Referring, indeed, is far more accurate and efficient when it is bilateral and multimethod. In the Lego experiment (Clark and Krych, 2004), half the directors were able to see the builders' workspace – their hands and blocks – and the other half were not. Partners took less than half the time to build their models when the builders' workspace was visible than when it was not. Language alone was no substitute for language and gesture together. Still other directors were asked to tape-record their instructions for future builders, so they were forced to design references unilaterally. The result was disastrous. Builders made over ten times as many errors when they worked from the tape-recorded instructions as when they collaborated with the directors live.

4.3 Indicating and describing in acts of referring

Indicating works, fundamentally, by *locating* referents for addressees. Recall Sam handing people a photograph of Reagan and Stockman and asking 'You know who this man is, don't you?' To refer to Reagan, Sam needed to get his addressees to locate the right *indicatum* – the image of Reagan in the photograph. That image, of course, was merely an index to the actual referent of 'this man' – to Reagan himself. Most demonstrative references work by means of two indexes: (i) the pointing locates a proximal indicatum (e.g., Reagan's image); and (ii) the indicatum is an index to the intended referent (e.g., Reagan himself).

But what does it take to locate an indicatum? In Clark and Marshall (1981), it was argued that some methods of locating things should be preferred to others. It requires fewer assumptions to point at a book and say 'that book' than to describe its location 'the book third from the left on the top shelf'. Indeed, if David asks June, 'In which building do you work?', it would be obtuse of her to answer '405 North Mathews Avenue' when she could simply point and say 'that building'. June's aim is not simply to name the building's location, but to bring the building into her and David's *joint focus of attention*.

Pointing itself can be more or less precise in establishing a joint focus of attention. Let us distinguish *close pointing* from *distant pointing* (Wilkins, 1999). With close pointing, the things pointed at are more or less within arms' reach; they can often be touched. With distant pointing, the things pointed at are out of arms' reach. *A priori*, close pointing should be more precise than distant pointing. If so, speakers should need to supplement distant pointing with other locative descriptions. They shouldn't need to do that with close pointing.

Pointing and locative descriptions do indeed trade off in referring at different distances, as shown in an experiment by Bangerter (2004). In that experiment, two participants sat next to each other facing a large board with 20 pictures of people placed at random locations on it. The director had a sheet with the people's names, and his or her job was to get the matcher to write down the right name for the right picture on an answer sheet. The two of them could talk as much as they liked. The board of pictures was placed at five distances away from the participants, from arm's length to a metre away.

Several things happened as the targets got further away. The directors pointed less often, from 82 per cent to 23 per cent of the time. They also used fewer deictic expressions (*this, that, here* or *there*), from 57 per cent to nearly none (4 per cent). Remarkably, as they used fewer deictic expressions, they included more descriptions of the pictures' locations (e.g., 'she's on the very right hand side' or 'the person right below the person laughing'). Indeed, the locative descriptions compensated precisely for the absence of deictic expressions, as they rose from 43 per cent to nearly all (97 per cent). In a control condition, partners could see the board but not each other, and thus could not point. They identified almost every picture by describing its location and features. But the location descriptions preceded the feature descriptions 91 per cent of the time, suggesting that speakers were trying to focus their addressees' attention before describing the pictures.

In short, the two partners started the referring process by bringing the location of the picture into their joint focus of attention and only then described its features. They used pointing to do this whenever possible. When pointing was ambiguous or impossible, they compensated by describing the location.

5 Conclusions

Since about 1960, theoretical conceptions of reference have gone through remarkable changes. It was originally treated as an autonomous, one-shot act by the speaker to enable a listener to identify the intended referent. It was viewed as addressee-blind or at least addressee-myopic. But with the arguments of Grice and Lewis, it came to be viewed as a cooperative, or coordinated, act that required speakers to consider their addressees. Furthermore, if face-to-face conversation is the primary setting for language use – it was the only setting for most of the history of both humans and their languages – then referring cannot be an autonomous act. It must be a participatory act – one half of a joint act by speakers and addressees working together. According to this view, speakers may initiate the process of referring, but they count on the active participation of their addressees. The act of referring was no longer viewed as unilateral, but as bilateral, its course determined by the actions of both speakers and addressees. Still later,

referring came to be seen as multi-method. It required indicating as well as describing, and indicating ordinarily requires locating referents in a joint focus of attention.

It took all three methods of language analysis – armchair, laboratory and field – to change these theoretical conceptions of reference. No one method did it alone. There are lessons to be learned, therefore, from the history we have presented. One is that the final arbiter of a theory or model of referring must be whether or not it can account for the acts that arise in everyday language. One can imagine many of these uses from an armchair, but many other uses can only be discovered in the field. And one can examine these acts under the microscope of laboratory experiments, in which people are constrained to behave in certain ways. But the evidence about what people can do with constraints applied can only be interpreted against what people can do, and do do, in the field. Language has evolved as a natural phenomenon. Just as we can never know the true behaviour of bears, penguins or dolphins by studying them in a zoo, we can never know the true nature of language without studying it in the wild.

References

Akmajian, A. (1973). The role of focus in the interpretation of anaphoric expressions. In S. R. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*: 215–26. New York: Holt, Rinehart and Winston.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science* 15(6): 415–19.

Bavelas, J. B., Coates, L., and Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology* 79(6): 941–52.

Bavelas, J. B., Coates, L., and Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication* 52(3): 566–80.

Beun, R.-J., and Cremers, A. H. M. (1998). Object reference in a shared domain of conversation. *Pragmatics and Cognition* 6(1–2): 121–52.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes* 10(2): 137–67.

Brennan, S. E., and Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(6): 1482–93.

Brewer, W. F., and Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology* 13(2): 207–30.

Buchler, J. (ed.) (1940). *Philosophical Writings of Peirce*. London: Routledge and Keegan Paul.

Chafe, W. L. (1970). *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

Chomsky, N. (1971). Deep structure, surface structure, and semantic interpretation. In L. A. Jakobovits and D. D. Steinberg (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Psychology, Linguistics, and Anthropology*. Cambridge: Cambridge University Press.

Clark, H. H. (1975). Bridging. In R. C. Schank and B. L. Nash-Webber (eds), *Theoretical Issues in Natural Language Processing*. New York: Association for Computing Machinery.

Clark, H. H. (1977). Bridging. In P. N. Johnson-Laird and P. C. Wason (eds), *Thinking: Readings in Cognitive Science*: 411–20. Cambridge: Cambridge University Press.

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H. (2003). Pointing and placing. In S. Kita (ed.), *Pointing: Where Language, Culture, and Cognition meet*. Hillsdale, NJ: Lawrence Erlbaum.

Clark, H. H., and Brennan, S. A. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine and S. D. Teasley (eds), *Perspective on Socially Shared Cognition*: 127–49. Washington, DC: APA Books.

Clark, H. H., and Carlson, T. B. (1982). Hearers and speech acts. *Language* 58: 332–73.

Clark, H. H., and Haviland, S. E. (1974). Psychological processes in linguistic explanation. In D. Cohen (ed.), *Explaining Linguistic Phenomena*: 91–124. Washington: Hemisphere Publication Corporation.

Clark, H. H., and Haviland, S. E. (1977). Comprehension and the Given–New contract. In R. O. Freedle (ed.), *Discourse Production and Comprehension*: 1–40. Hillsdale, NJ: Erlbaum.

Clark, H. H., and Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50(1): 62–81.

Clark, H. H., and Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber and I. A. Sag (eds), *Elements of Discourse Understanding*: 10–63. Cambridge: Cambridge University Press.

Clark, H. H., and Schaefer, E. F. (1987). Concealing one's meaning from overhearers. *Journal of Memory and Language* 26: 209–25.

Clark, H. H., and Schaefer, E. R. (1989). Contributing to discourse. *Cognitive Science* 13: 259–94.

Clark, H. H., and Schaefer, E. F. (1992). Dealing with overhearers. In H. H. Clark (ed.), *Arenas of Language Use*: 248–97. Chicago: University of Chicago Press.

Clark, H. H., Schreuder, R., and Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior* 22: 1–39.

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

Enfield, N. J. (2003). The definition of WHAT-d'you-call-it: semantics and pragmatics of recognitional deixis. *Journal of Pragmatics* 35(1): 101–17.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes* 27(1): 35–53.

Garrod, S. C., and Sanford, A. J. (1982). Bridging inferences in the extended domain of reference. In A. Baddeley and J. Long (eds), *Attention and Performance*, Vol. IX: 331–46). Hillsdale, NJ: Erlbaum.

Goodwin, C. (2003). Pointing as situated practice. In S. Kita (ed.), *Pointing: Where Language, Culture, and Cognition Meet*. Hillsdale, NJ: Lawrence Erlbaum.

Grice, H. P. (1957). Meaning. *Philosophical Review* 66: 377–88.

Grice, H. P. (1968). Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language* 4: 225–42.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (eds), *Syntax and Semantics, Vol. 3: Speech Acts*: 113–28. New York: Seminar Press.

Grice, H. P. (1978). Some further notes on logic and conversation. In P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*: 113–27. New York: Academic Press.

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

Halliday, M. A. K. (1967). Notes on transitivity and theme in English. Part 2. *Journal of Linguistics* 3: 199–244.

Harman, G. (1977). Review of 'Linguistic behavior' by Jonathan Bennett. *Language* 53: 417–24.

Haviland, S. E., and Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior* 13: 512–21.

Hawkins, J. A. (1978). *Definiteness and Indefiniteness: A Study of Reference and Grammaticality Prediction*. London: Croom Helm.

Hindmarsh, J., and Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics* 32(12): 1855–78.

Jackendoff, R. S. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA., MIT Press.

Karttunen, L. (1977). Presupposition and linguistic context. In A. Rogers, B. Wall and J. P. Murphy (eds), *Proceedings of the Texas Conference on Performatives, Presuppositions, and Implicatures*: 149–60. Arlington, VA: Center for Applied Linguistics.

Karttunen, L., and Peters, S. (1975). Conventional implicature in Montague grammar. *Berkeley Linguistics Society* 1: 266–78.

Krauss, R. M., and Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science* 1: 113–14.

Krauss, R. M., and Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology* 4: 343–6.

Lerner, G. H. (1987). *Collaborative Turn Sequences: Sentence Construction and Social Action*. Unpublished Ph.D. dissertation, University of California, Irvine.

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

Matsui, T. (2000). *Bridging and Relevance*. Amsterdam: J. Benjamins.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review* 77(4): 257–73.

Prince, E. F. (1981). Toward a taxonomy of Given–New information. In P. Cole (ed.), *Radical Pragmatics*: 223–55. New York: Academic Press.

Sacks, H., and 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*: 15–21. New York: Irvington Publishers.

Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson and J. Heritage (eds), *Structures of Social Action: Studies in Conversation Analysis*: 262–96. Cambridge: Cambridge University Press.

Schegloff, E. A., Jefferson, G., and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language* 53: 361–82.

Schelling, T. C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

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

Schober, M. F., and Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology* 21: 211–32.

Schütze, C. T. (1996). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.

Stalnaker, R. C. (1978). Assertion. In P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*: 315–32. New York: Academic Press.

Svartvik, J., and Quirk, R. (eds). (1980). *A Corpus of English Conversation*. Lund: Gleerup.

Van Der Wege, M. M. (2000). *Agreeing on Reference*. Unpublished Ph.D. dissertation, Stanford University.

Wilkes-Gibbs, D. (1986). *Collaborative Processes of Language Use in Conversation*. Unpublished Ph.D. dissertation, Stanford University.

Wilkes-Gibbs, D., and Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language* 31(2): 183–94.

Wilkins, D. P. (ed.). (1999). *Manual for the 1999 Field Season. Language and Cognition Group*. Nijmegen: Max Planck Institute for Psycholinguistics.