

5 | Meaning and understanding

All speech, written or spoken, is a dead language,
until it finds a willing and prepared hearer.

Robert Louis Stevenson,
Reflections and Remarks on Human Life

When people take part in joint activities – business transactions, chess games, piano duets – they perform a variety of joint actions. They say things to each other, hand things to each other, nod at each other, gaze at each other, and through these advance their joint activities. Many of these joint actions, or their parts, are *communicative acts* through which they get others to understand what they mean. What sort of acts are these, and how do they work?

The traditional view is that communicative acts are performed by a speaker autonomously. In the drugstore, when Stone said “I’ll be right there,” she was making a promise on her own. Although she directed it at me, I had no real part in it. A promise expresses a commitment to do something in the future, and speakers express such commitments on their own. In that tradition, the focus is on speakers. There is no mention, no hint, that addressees have any role.

Paradoxically, the traditional view carries the seeds of its own destruction. The very notion of meaning – speaker’s meaning – requires addressees to join speakers in a special way, and so do other notions of speech acts. We will discover, on closer examination, that communicative acts are inherently joint acts, and that they are just one level of an entire ladder of joint actions. To begin, let us turn to what is at the heart of all communicative acts: meaning.

Meaning

In 1957, in a ten-page paper entitled “Meaning,” Herbert Paul Grice¹ presented a theory of meaning that revolutionized the study of language

¹ Grice, for some reason, went by Paul rather than Herbert.

use. He began by distinguishing the meaning of certain natural events, which I will call *natural signs* or *symptoms*, from the meaning of certain deliberate human acts, which I am calling *signals*. Compare these two statements:

1. Those spots mean that Margaret has the measles.
2. The doctor's hand wave means that Margaret has the measles.

The spots described in 1 are a natural sign – a symptom or direct evidence – that Margaret has the measles. If I tell a friend, “Those spots mean that Margaret has the measles,” I am committed to the belief that Margaret has the measles. But the hand wave in 2 (say, through a glass barrier in a hospital!) means what it does in part because of the doctor’s intentions toward me, which are to tell me that Margaret has the measles. Unlike the spots, the gesture bears no natural connection to measles. And if I tell a friend, “The doctor’s hand wave means that Margaret has the measles,” I am not committed to the belief that Margaret has the measles – the doctor could be wrong. Grice called these two kinds of meaning *natural meaning* and *non-natural meaning*. In my terminology, symptoms have natural meaning, and signals have non-natural meaning.

Language use depends on both natural signs and signals. Take natural signs. The sounds I hear mean that the radio is on. The shape of the object my friend is holding means that it is a book. The pitch of a caller’s voice means that he is a man. A speaker’s involuntary hesitation in uttering a word means that he probably had difficulty thinking of, choosing, or pronouncing it in time. Most things have a natural meaning, and these can be important for language use because they are all natural signs that this or that is true. What distinguishes language use is that it always involves non-natural meaning as well.

SPEAKER’S MEANING AND SIGNAL MEANING

Non-natural meaning itself, according to Grice, divides into two types: *speaker’s* (or *utterer’s*) *meaning*, and what I will call *signal meaning*. Consider these descriptions:

3. By uttering “I surrender,” Sam meant that he was surrendering.
4. By waving a white flag, Sam meant that he was surrendering.
5. By uttering “I am hungry,” Elizabeth meant that she was in need of food at that moment.
6. By pointing at her mouth and an empty plate, Elizabeth meant that she was in need of food at that moment.

All four examples describe what a speaker meant. They each fit a standard frame for speaker's meaning:

Speaker's meaning. By presenting s to A , S meant for A that p .

In this frame, S denotes the agent of the action, like a speaker or letter writer; A denotes a certain audience; and s denotes a deliberate human action, a signal, like waving a flag or uttering a sentence. The following descriptions, in contrast, describe what a signal means or meant:

7. The sentence *I am hungry* can mean that the speaker, whoever that is, is in need of food at the moment he or she utters the sentence.
8. The word *hungry* can mean "in need of food."
9. Elizabeth's utterance, "*I am hungry*," meant that she was in need of food at the moment she produced the utterance.
10. Elizabeth's gesture at her mouth and empty plate meant that she was in need of food at that moment.

All these examples fit a standard frame for signal meaning:

Signal meaning. s means or meant " y ," or that p .

As before, s denotes a deliberate human action, a signal; in addition, p denotes a proposition, and y a paraphrase.

It is odd to have to explain the difference between speaker's meaning and signal meaning. In German, they are called *Gemeintes* and *Bedeutung*, in Dutch, *bedoeling* and *betekenis*, and in French, *intention* and *signification*. For theorists working in German, Dutch, and French, they are as different as apples and oranges. Yet for theorists working in English, they are a chronic source of confusion because they have the same name – *meaning*.² In language use, it is essential to keep them straight.

Signal meaning comes in several varieties. Example 7 describes an instance of sentence meaning, one way in which the sentence *I am hungry* can be used on a particular occasion. (The same sentence can be used in other ways too.) Example 8 describes an instance of word meaning, or one way in which the word *hungry* can be used on a particular occasion. And example 9 describes an instance of utterance meaning, what Elizabeth's act of uttering the sentence meant on that particular occasion. In 9, one of the conventional meanings of the

² It is almost enough to make one believe in Benjamin Lee Whorf's linguistic determinism.

sentence Elizabeth uttered bears a relation to the meaning of her utterance, but that relation could have been very indirect, even absent. The doctor's hand wave in 2, for example, may have been a signal she and I decided on for that occasion alone. Other times it might mean nothing or something entirely different.

These distinctions are important. Words and sentences are *types* of signals, linguistic units abstracted away from any occasion on which they might be used, stripped of all relation to particular speakers, listeners, times, and places. To describe them is to describe the conventions for their use within speech communities (see Chapters 3 and 4). But utterances are the actions of producing words, sentences, and other things on particular occasions by particular speakers for particular purposes. The study of language structure is primarily about the conventions that govern words, sentences, and their meanings. But in conversations, books, and newspapers, we deal with *utterances* of words, sentences, and other things, and that requires a different approach.

Non-natural meaning isn't confined to uses of conventional languages like English, Japanese, or Dakota, nor did Grice ever intend it to be. Signals can be both "linguistic" (belonging to a conventional language), as in 3, 5, 7, 8, and 9, and "non-linguistic," as in 2, 4, 6, and 10. In the frame for speaker's meaning, the speaker is "presenting *s*" and not merely "uttering *s*." Ordinary language use depends on both. In conversation, people not only issue words, but also pause, gesture with their hands, head, eyes, and shoulders, and present other non-linguistic signals (Chapter 6). They use these in combination to say what they mean. So when I use the terms *utterances*, *speakers*, and *speaker's meaning*, I normally intend *signals*, *signalers*, and *signaler's meaning*.

Speaker's meaning and signal meaning, though different, are obviously connected. Speakers mean something only by using signals, and signals mean something only because they are used by speakers to mean something. Still, speaker's meaning is logically prior in several respects. Many signals have no *conventional* meaning. What these mean gets fixed only by what speakers meant in using them on particular occasions. One noon at a lunch with friends, I reminded my wife of an impending dentist appointment by taking an obvious look at my watch. But looking at one's watch doesn't conventionally or usually mean "you are due at the dentist's soon." Here what the speaker is inferred to mean helps us determine what the signal means, and not just vice versa.

The principle is general. We cannot talk about a signal having meaning without assuming an agent or speaker behind it. The doctor's hand wave from the other side of the glass meant that Margaret had the measles only because I assumed the doctor waved with those intentions in mind. If I discovered that the doctor was waving at someone behind me, or that her gesture wasn't intended to be the one we had agreed on, I wouldn't take her to mean that Margaret had the measles. The same is true of a word like *hungry*. It can mean "in need of food" only because of a convention that it can mean that – there is a community of people who can mean "in need of food" by uttering it in the right situations (Chapter 3). So to say "*s* meant (or means) something" is tantamount to saying "somebody meant (or can mean) something by *s*": the agent, the somebody, must be included. In this sense, speaker's meaning is primary, and signal meaning secondary.

WHAT IS SPEAKER'S MEANING?

It was Grice's insight that speaker's meaning has to do with getting other people to do things, but only by certain means.³ Suppose Sam took you to the window to let you see the rain outside. He got you to believe that it was raining out, and you recognized his intention to get you to believe that. Still, you wouldn't say, "By presenting this scene to me, Sam meant that it was raining out." For speaker's meaning, Grice argued, your recognition of Sam's intention must serve as part of your reason for thinking that it's raining out. If, instead, Sam had said simply, "It's raining out," his intentions would have been essential. If you had thought he was practicing a line from a play, or reading from a novel, or offering an example of a present progressive verb, you wouldn't have taken him as meaning it was raining out. You thought that was what he meant in part because you recognized his intention that you think that.

Precisely how to formulate speaker's meaning has been debated ever since Grice's first proposal. Here is a formulation that is faithful to Grice's original idea, but has been amended in several ways:⁴

³ See Grice (1957, 1968), Schiffer (1972), Strawson (1964), and Searle (1969), among others.

⁴ This formulation is based on some but not all arguments in Strawson (1964), Grice (1968, 1982), Searle (1969), Bach and Harnish (1979), Harman (1977), Sperber and Wilson (1986), Récanati (1986), and Thomason (1990).

Speaker's meaning (reflexive)

In presenting s to audience A , a speaker S means for A that p if and only if:

- (i) S intends in presenting s to A that A recognize that p in part by recognizing that i .

Speaker's meaning is a type of intention. When I say to you "Please sit down," my intention is for you to recognize that I want you to sit down. But as part of your reason for thinking this, you must recognize my very intention in presenting what I did. So speaker's meaning is a *reflexive* intention: intention i contains a reference to i itself. Grice called it an *m-intention* (for "meaning intention"), which he took to be essential to all non-natural meaning.

SIGNALING AND RECOGNIZING

Grice's m-intention—the heart of speaker's meaning—is a curious type of intention: It is one the speaker cannot discharge without the audience's participation. When I say "Please sit down" and mean you are to sit down, I rely on you doing your part by recognizing what I mean. In Grice's formulation, my intention depends directly on your recognition of that intention. I can discharge my intention to shake a stick, an autonomous action, without anyone else's actions. But I cannot discharge my intention to do my part of our hand shake, a joint action, without you doing your part. Here my individual act is a participatory act, which I perform as part of a joint act that requires you to do your part too (Chapters 1 and 3). The same is true of signaling and recognizing. The principle I wish to defend is this:

Signal recognition principle. Signaling and recognizing in communicative acts are participatory acts.

The joint act of one person signaling another and the second recognizing what the first meant I will call a *communicative act*.

To see how signaling and recognizing work, let us examine them from the inside, as actions in progress. Recall that when Ann and Ben play a flute-piano duet, we have a joint action r and their individual participatory actions:

Ensemble A-and-B are playing a flute–piano duet r in situation w if and only if:

0. the duet r includes 1 and 2;
1. A is playing the flute part as part of r ;
2. B is playing the piano part as part of r .

So for Ann to be playing her flute part as part of the duet, she must be playing it believing Ben to be playing his piano part as part of the same duet. If half way through the duet she thinks Ben is no longer doing his part – he has stopped playing because his music blew away – she will no longer consider them to be playing the duet – to be performing participatory actions as parts of that duet. If she does play on, she will consider herself to be playing alone. The point is crucial. Ann's and Ben's participatory actions are interlinked: Ann cannot consider herself to be playing her part as part of the duet without assuming Ben is playing his part as part of the same duet, and vice versa.

So it goes with the participatory acts of signaling and recognizing. Suppose Ann presents signal s to Ben (e.g., she utters “Please sit down”) meaning that p (e.g., that he is to sit down). Again we have a joint act r and participatory acts (1) and (2):

Speaker's meaning (joint)

In presenting s to A, speaker S means for A that p if and only if:

0. the communicative act r includes 1 and 2;
1. S presents s to A intending that p as part of r ;
2. A recognizes that p as part of r .

When Ann utters “Please sit down” as part of r , she expects Ben to do his part. Ben must recognize what she means in part by seeing that she is uttering “Please sit down” with the intention in 1. As in the duet, Ann's and Ben's actions are linked: Ann cannot consider herself to be asking Ben to sit down without assuming that Ben is intending to recognize these intentions, and vice versa.

Consider Ann's and Ben's actions half way through her utterance. If she thinks Ben is no longer doing his part, she will no longer consider them to be communicating; she will no longer consider herself to be asking Ben to sit down. Suppose Ann assumes Ben knows Dutch and says “Gaje even...” when Ben interrupts with “What?” before she can finish “zitten alsjeblieft.” Although she begins her utterance intending Ben to recognize that she wants him to sit down, she is forced to abort that intention undischarged when she realizes Ben isn't doing his part. And although Ben may realize she has been trying to signal him, he realizes that she isn't succeeding (see Chapters 8 and 9).

The two-part representation just given brings out several basic properties of speaker's meaning. It divides communicative acts into their two natural parts – signaling and recognizing. Part 1 specifies the speaker's

actions and responsibilities, and part 2, the hearer's actions and responsibilities. It also shows how the two actions are linked – how A's intentions depend on B's recognition, and vice versa. Finally, instead of putting all the onus on speakers, it treats speakers and addressees as partners. The idea, in short, is to treat signaling and recognizing for what they are – two parts of a joint act.

SIGNALING AS A COORDINATION DEVICE

Signals aren't important merely because they mean things. They are important because they are used in discourse to accomplish the participants' goals. When the server in the drugstore said "I'll be right there," she meant that she would be ready to serve me soon. But she was using the signal to coordinate her and my actions at that point in our transaction. Viewed in isolation, a signal is an act by which a speaker means something. Viewed within joint activities, it is an act by which the participants coordinate the next step in their ongoing activity. Signals are coordination devices.

Viewing signals as coordination devices gives us yet another perspective on speaker's meaning and audience's understanding. In the cumulative model of joint activities, participants use utterances and other signals to increment their current common ground. A signal is then the speaker's way of introducing into the discourse a shared basis for the piece of common ground to be added. Recall that a shared basis b for common ground has three properties (Chapter 4):

p is common ground for members of community C if and only if:

1. every member of C has information that b holds;
2. b indicates to every member of C that every member of C has information that b holds;
3. b indicates to members of C that p .

A signal that is recognized satisfies the same three properties:

1. S and A have information that S presented s to A;
2. s indicates to S and A that S and A have information that S presented s to A;
3. s indicates to S and A that S means for A that p .

So when the server uttered "I'll be right there," she was providing a shared basis for the next step in our transaction, a shared basis for incrementing our common ground.

Signaling is the prototypical coordination device in joint activities.

If the drugstore server wants to coordinate her actions with mine, her usual strategy is to present a signal and get me to recognize what she meant by it. Her signal serves as a shared basis for a mutual belief that we can then add to our common ground. In that way it carries the discourse forward to the next step.

Speech acts

Speakers get their addressees to recognize what they mean, in Grice's scheme, by taking actions toward them – by signaling them. What sorts of actions are these? One of the first to take this question seriously was John Austin. His 1957 William James Lectures were called "How to do things with words" (Austin, 1962), but they were really about how to do things with utterances. In them he proposed a general theory of speech acts – acts that people perform in speaking – in which he distinguished among many things people do with utterances. Some details of his argument have been eclipsed by work since then, but many of his basic insights remain.

Certain actions we take, Austin argued, are designed to get our audience to do things on the basis of their understanding of what we mean. Suppose I speak to my son, and he responds, as follows:

I request of him "Please pass the horseradish." He says "Okay," and passes it.
 I ask him "What are you doing?" He answers "Getting ready to leave."
 I tell him "That book is terrific." He believes me and starts reading it.
 I warn him "Bruno is coming." He believes me and gets frightened.

My son complies with my request, answers my question, comes to believe what I assert, follows my advice, and gets scared, all based on his understanding of what I meant. In Austin's terminology, these are *perlocutionary effects*, or *perlocutions*, of my actions, and my acts in getting him to do them are *perlocutionary acts* (see Davis, 1979). Some perlocutionary effects are intended, and others aren't. If I unintentionally make my son laugh by asking him for the horseradish, his laugh is still a perlocution.

Perlocutions aren't part of understanding itself. My son could have understood my request for the horseradish, but refused to comply. He could have understood my assertion about the book, but not believed me. All he needed for understanding was to recognize my meaning. The act of getting the audience to recognize the speaker's meaning Austin called an *illocutionary act* and the recognition itself came to be called an *illocutionary*

effect (Searle, 1969). My request, question, assertion, and warning are illocutionary acts, and my son's understanding of them are illocutionary effects.

TYPES OF ILLCUTIONARY ACTS

Illocutionary acts come in many types. They include telling, asserting, requesting, ordering, asking, promising, apologizing, thanking, firing, and baptizing – there are over 150 such illocutionary verbs in English (Verschueren, 1980). Is there any order behind these acts? John Searle (1975c) argued there is. The primary way they differ is in what he called their *illocutionary point* – their publicly intended perlocutionary effect. For some illocutionary acts, the point is to get listeners to do things; for others, it is to commit the speaker to doing things; and so on. Searle used this notion to divide illocutionary acts into five main categories, the last of which I have divided into two:⁵

1. *Assertives*. The point of an assertive is to get the audience to form, or to attend to, the belief that the speaker is committed to a certain belief. When Sam told you, “It’s raining out,” he was trying to get you to think he believed it was raining out. The prototypical assertive is the assertion, but the category also includes diagnoses, predictions, notifications, confessions, denials, disputations, retorts, conjectures, suppositions, and many others.

2. *Directives*. The point of a directive is to get the audience to do things. When I asked my son, “Please pass the horseradish,” I was trying to get him to pass me the horseradish. Directives fall into two major classes: requests for action (as with most commands and suggestions), and requests for information (as with most questions). With my question, “What are you doing?” I was asking my son for information. Directives vary in how forceful they are – from mild hints to stern commands – and in other ways too.

3. *Commissives*. The point of a commissive is to commit the speaker to a future action. The prototype is the promise, as when George says to Jane, “I’ll get some coffee,” committing himself to Jane to getting some coffee. One subtype is the conditional promise, or offer, as when George says to Jane, “Can I get you some coffee?” committing himself to getting her coffee if she wants it.

4. *Expressives*. The point of expressives like thanking, apologizing, congratulating, and greeting is to express certain feelings toward the

⁵ The emendations come from Bach and Harnish (1979) and Hancher (1979).

audience. When Verona says to Wilfred, “Sorry I’m late,” she takes for granted that she came late and tries to get Wilfred to believe she regrets it.

Next come illocutionary acts Searle called *declarations*. These rely on codified conventions of institutions such as the law, the church, and organized games. Within these institutions, speakers can do certain things by virtue of a privilege the institution grants them because of their role as judge, priest, referee, or whatever. Declarations divide into two main subcategories.

5a. *Effectives*. The point of an effective is to change an institutional state of affairs. In industry, a boss may fire, promote, or appoint someone. In court, a judge may indict, pardon, or sentence someone. A policeman may arrest someone. In football, a referee may start the game and call time outs. In church, a minister may baptize, marry, or bless someone. In each case, the speaker has the institutional power to change things merely by saying, “You’re fired,” “You are hereby sentenced to three years in jail,” or “Time out” in the appropriate circumstances.

5b. *Verdictives*. With verdictives, the point is to determine what is to be the case within the institution. In baseball, umpires have to judge whether a ball that has been pitched has passed through the strike zone – whether it has crossed the plate between the batter’s shoulder and knees. The umpire may try to be accurate, but when he says “Strike,” his verdict is law from then on regardless of whether the ball actually passed through the strike zone. As far as the game is concerned, the ball did pass through the strike zone, and the pitch was a strike. Verdictives also occur when a jury finds a prisoner innocent or guilty, when the presiding officer in a meeting rules a motion out of order, and when a journal editor accepts or rejects a paper for publication.

Ilocutionary act	Ilocutionary point
assertives	to get the addressee to form or attend to a belief
directives	to get the addressee to do something
commissives	to commit the speaker to doing something
expressives	to express a feeling toward the addressee
effectives	to change an institutional state of affairs
verdictives	to determine what is the case in an institution

Searle’s scheme, as summarized here, has many problems. One is that it doesn’t generate all potential illocutionary acts. We can invent new rituals, new games, new social customs, each with its own special illocutionary acts, *ad infinitum*, but the scheme has no principles to say what is

allowed, and what isn't. Another problem is that every illocutionary act is assumed to belong to one and only one category. But consider a general's order to a sergeant. Under military regulations, that order changes an institutional state of affairs just as surely as a judge's sentencing does—the sergeant could be court-martialed for not obeying—and that makes it an effective. Yet it is also surely a directive. The same goes for other illocutionary acts (see Hancher, 1979; Wunderlich, 1977). Despite its problems, the scheme is useful as a gross classification and for its widely accepted nomenclature. I shall use it for both.

ILLOCUTIONARY ACTS AND THEIR RECOGNITION

How do speakers get their addressees to recognize the illocutionary act they are performing? The classical answer is that they do so by their choice of sentence modality (e.g., Vanderveken, 1990). In English, there are five modalities:

Modality	Examples
Declarative	That book is awful. It is raining out.
Yes/no interrogative	Is it raining out?
WH-interrogative	What are you doing?
Imperative	Pass the horseradish.
Exclamatory	What a beautiful day! Is it ever hot out!

To assert something, you choose a declarative; to ask a question, an interrogative; to make a request or command, an imperative; and for an exclamation, an exclamatory. Your partners, by noting your choice of modality, can immediately recognize the illocutionary act you are performing.

This view is inadequate from the very start (Levinson, 1983). With only five modalities, we should be able to distinguish only five types of illocutionary acts, but we easily distinguish scores. The imperative, for example, can be used for at least these illocutions (Sadock and Zwicky, 1985; Sperber and Wilson, 1986, p. 250):

Illocutionary act	Example
Commands	To the rear, march
Requests	Please pass the horseradish
Promises	Mow the lawn and I'll pay you a dollar
Threats	Stop or I'll shoot
Warnings	Watch out!
Offers	Have some cake
Well wishing	Have a good trip
Advice	For a dry martini, mix six parts gin with one part vermouth
Curses	Go to hell
Exclamations	Well, look at you!
Exhortations	Fly American Airlines

Worse, these illocutionary acts range over four of Searle's five main categories. There is much the same variation for declarative and interrogative modalities.

A more sophisticated view makes use of what Levinson (1983) has called *illocutionary force identifying devices*, or *ifids*. The idea is that speakers use conventional devices in addition to sentence modality for specifying the illocutionary act they are performing. They might mark an utterance as a request with *please*, *Why not?*, or *I'd appreciate it very much if*; as a promise or offer with *I'll* or *Let me*; and so on. But it is easy to see that ifids, while informative, cannot do the job alone. Many utterances do not have enough ifids to pin down the illocutionary act the speaker is performing. The bare *Sit here* can be used as a request, command, advisory, threat, promise, exhortation, or offer, and it has no ifids to tell us which.

RECOGNITION AND UPTAKE

There is something missing in this picture of speech acts. At the center are speakers and what they do, but if there are any listeners, they are nowhere to be seen. It is as if the official portrait of a wedding included a groom but no bride. The terms *speech acts*, *illocutionary acts*, and *perlocutionary acts* describe what speakers do, but there are no comparable terms for what listeners do – as if their actions were irrelevant. Searle (1969) even argued: “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 letter and dropping it in the mail. It doesn't matter whether anybody receives, reads, or

understands it. This view is, of course, absurd.⁶ There can be no communication without listeners taking actions too – without them understanding what speakers mean.

Austin recognized the problem, but his suggestions were ignored by most who followed (e.g., Bach and Harnish, 1979; Searle, 1969, 1975). Suppose, Austin said, that he has promised a friend to return some money by uttering “I’ll pay you back tomorrow.”

*It is obviously necessary that to have promised I must normally (A) have been heard by someone, perhaps the promisee; (B) have been understood by him as promising. If one or another of these conditions isn’t satisfied, doubts arise as to whether I have really promised, and it might be held that my act was only attempted or was void.*⁷ (Austin, 1962, p. 22)

That is, promises require recognition by the addressees, who hear and understand what is being promised. This is nothing less than a coordinated action by the addressees. In my terminology, that makes a promise and its recognition – two participatory actions. They are the two parts of a joint action or communicative act.⁸

Austin noted a similar problem for perlocutionary acts. To complete certain illocutionary acts, he argued, the speaker has to secure their acceptance. His examples included betting, marrying, giving, and appointing:

My attempt to make a bet by saying “I bet you sixpence” is abortive unless you say “I take you on” or words to that effect; my attempt to marry by saying “I will” is abortive if the woman says “I will not.” (p. 36)

Likewise, a person cannot give or bequeath something to others, or appoint them to some position, without their acceptance, either “expressed or implied.” Illocutionary acts like these are cooperative or bilateral instead of unilateral (Hancher, 1979). Without your acceptance, I may have tried to give you something, or appoint you to some position, but I will have failed.

Diehard unilateralists might deny that betting, giving, bequeathing, appointing, and their kind are illocutionary acts at all – even though they are on everyone’s list of illocutions. These acts, they could argue, are

⁶ See also Streeck (1980).

⁷ Later, Austin asked, rhetorically: “One of the things that cause particular difficulty is the question whether when two parties are involved ‘*consensus ad idem*’ is necessary. Is it essential for me to secure *correct understanding* as well as everything else?” (p. 36).

⁸ Later I will adopt the term *uptake* but with a more restricted meaning than Austin’s.

really *pairs* of illocutionary acts. A bet consists of (1) a proposal by the bettor (“I bet you sixpence”), and (2) its uptake by the bettee (“I take you on”). What these examples show, they might continue, is that the pair of illocutionary acts is achieved jointly. They show nothing about the proposal and uptake as separate acts. What is needed to complete an illocutionary act, however, is not its uptake, but its recognition. For Austin to make a promise to his friend, he “must normally have been heard [and] have been understood by him as promising.” This requirement isn’t hard to satisfy, but it takes the friend’s coordinated actions. He and his friend have to work jointly to establish, to a reasonable criterion, that his friend has understood him as intended (see Chapter 8).

These paths lead to a new outlook on speaker’s meaning, illocutionary acts, and perlocutionary acts. Speaker’s meaning is a type of intention that can be discharged only through joint actions. Illocutionary acts, as Austin himself realized, can be accomplished only as parts of joint actions, and the same is true of perlocutionary acts. The issue is how to bring the long neglected addressee back into the picture.

SOCIAL PRACTICES

Whatever its status, Searle’s classification of illocutionary acts illustrates one point over and over again: Illocutionary acts have their origins in social practices. Acts such as arresting, overruling, and calling time out – the effectives and verdictives – belong to highly codified social activities and wouldn’t exist without the social institutions in which they are formalized. All the other acts belong to well-developed social activities as well. It is just that these activities are informal and not codified. Directives arise when one person wants another person to do something and has some authority to oblige the other to do it. The authority may not come from a formal institution, but it does come from accepted social practices – as in ordering food in a restaurant, asking a librarian for a book, or asking a bank customer for identification. Assertives, commissives, and expressives work in similar ways. Illocutionary acts arise in joint activities (see also Cohen and Levesque, 1990).

The stronger assumption, which Austin seems to have held, is that illocutionary acts cannot be defined without reference to the joint activities of which they are parts. We cannot specify what constitutes a marriage vow, christening, bequest, or bet – Austin’s primary examples – without saying how they are performed within the appropriate ceremonies. Although Austin was later faulted for stressing these

institutional and conventional features, his assumption still seems fundamentally correct. The problem lies in our understanding of joint activities. It is easy to specify how sentencing, indicting, and dismissing are created within well-codified court procedures. It is more difficult to specify how offers, greetings, and questions are created within uncodified social practices.

Cooperation

If Austin is right, to understand what speakers mean, we must look at the joint activity or social practice they are engaged in. Grice argued much the same point in his own 1967 William James Lectures ten years after Austin's. He put it this way (Grice, 1975, p. 45):

Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction. The purpose or direction may be fixed from the start (e.g., by an initial proposal of a question for discussion), or it may evolve during the exchange; it may be fairly definite, or it may be so indefinite as to leave very considerable latitude to the participants (as in a casual conversation). But at each stage, some possible conversational moves would be excluded as conversationally unsuitable.

The participants of a conversation, Grice argued, therefore expect each other to adhere to the *cooperative principle*, which he expressed as an exhortation to speakers:

Cooperative principle. Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

In Grice's view, people take it for granted that "contributions" to conversations are to be interpreted against the "accepted purpose or direction of the talk exchange." One might pursue Grice's insight in many ways. He chose to apply it to the problem of what people mean by their utterances.

SAYING AND IMPLICATING

To see what speakers mean, Grice argued, we generally go beyond what they actually say. He asked us to imagine A standing next to an obviously immobilized car and striking up a conversation with passerby B:

- A: I am out of petrol.
 B: There is a garage round the corner.

All B has said is that there is a garage, a gas station, around the corner. Yet that isn't all A takes him as doing. A can suppose B was trying to offer information relevant to the situation at hand – that A is stranded and has just remarked that he is out of gasoline. So B must also mean, in Grice's words, "that the garage is, or at least may be open, etc."⁹ This he called an *implicatum* but is more often called an *implicature*. So in Grice's scheme, speaker's meaning divides into two parts: *saying* and *implicating*.

What is the difference? What is said (in Grice's special sense) is what speakers mean mostly through the conventional content of the sentences they utter – indeed, through only that part that affects the truth of their utterances. In uttering "There is a garage round the corner," B is saying only that there is a garage around the corner. The rest of what B meant is implicated. Some implicatures are conventional and, therefore, part of the sentence meaning. The ones I shall be concerned with Grice called *conversational implicatures*. One example is B's implicature that the garage may be open and selling petrol. A recognizes it not because of any conventional link with what B said. Rather, as Grice put it, A "works it out."

For Grice, conversational implicatures have three main properties (but see Nunberg, 1981; Sadock, 1978). (1) They are *non-conventional*. They are not conventionally associated with the words or sentence uttered. "There is a garage round the corner" doesn't conventionally mean that the garage is open. Yet (2) they are *calculable*. Speakers intend addressees to be able to work them out. A is to work out that B means that he believes the garage may be open. Conversational implicatures are those parts of what speakers mean that addressees recognize only by "working them out." Finally, (3) they are *defeasible* – the speaker can cancel them, rendering them null and void. B could have said "There's a garage round the corner, but I doubt if it's open," canceling the implicature A would otherwise work out.

FOUR MAXIMS

How are implicatures to be worked out? Since Grice argued that every utterance "contributes" to the "accepted purpose or direction of the talk exchange," we might have expected him to develop the notions of

⁹ Grice's "etc." is usually ignored, but it is important. He seems to be suggesting that we may not be able to enumerate A's implicatures explicitly – that unlike what it said what is implicated may be vague or lacking in clear limits.

“contribution” and “accepted purpose” and show how implicatures follow, but he didn’t. Instead, he offered four rules of thumb, four *maxims*, that he argued enable listeners to work out implicatures. Paradoxically, he expressed the maxims as exhortations to speakers (Grice, 1975, pp. 45–46):

- | | |
|-------------------|--|
| Maxim of quantity | <ol style="list-style-type: none"> 1. Make your contribution as informative as is required (for the current purposes of the exchange). 2. Do not make your contribution more informative than is required. |
| Maxim of quality | <ol style="list-style-type: none"> 1. Do not say what you believe to be false. 2. Do not say that for which you lack evidence. |
| Maxim of relation | Be relevant. |
| Maxim of manner | <ol style="list-style-type: none"> 1. Avoid obscurity of expression. 2. Avoid ambiguity. 3. Be brief (avoid unnecessary prolixity). 4. Be orderly. |

Once listeners take for granted that speakers adhere to these maxims and to the cooperative principle itself, they can work out what the speakers are implicating.

Speakers create implicatures in two main ways. The first is by direct appeal to the maxims. Take this invented exchange:

Burton: How many children do you have?

Connie: I have two children.

All Connie has said is that she has two children, which would be literally true even if she had three or four or twelve. Yet, by the maxim of quantity, Burton can assume she has been as informative as she needs to be for the current purposes of this exchange. And because he was asking for the *total* number of children, she must be giving him the total. Contrast that exchange with this one:

Burton: Do you have two quarters I could borrow for the pay phone?

Connie: Yes, I have two quarters.

Here, Burton is trying to find out not how many quarters Connie has in total, but merely whether she has two quarters he could borrow. She may have three, four, or twelve quarters, but she is being “as informative as is required for the current purposes of the exchange” by saying that she has two quarters. In these contrasting circumstances “I have two children” implicates “and no more than two children,” whereas “I have two quarters”

does *not* implicate “and no more than two quarters.” These are meanings Connie expects Burton to work out. The other maxims apply directly in similar ways.

The second method of creating implicatures is by blatantly violating, or *flouting*, a maxim. In the following example, Kate is describing a visit to a women’s college (1.3.560):

Kate: and . um then, . a bell rang, -- and - millions of feet, . ran, . along corridors,
you know, and then they . it all died away, it was like like sound effects
from the Goon Show

When Kate claimed “millions” of feet ran along the corridors, she was blatantly violating the maxim of quality, “Do not say what you believe to be false.” The violation was so blatant that she could expect her audience to reason: “Kate flouted the maxim, yet was otherwise cooperative. She must therefore not have meant ‘millions’ literally, but as hyperbole. It only *seemed* as if there were millions of feet.” Flouting maxims also leads to understatement, metaphor, irony, sarcasm, and other tropes.

Both methods of implicating have serious difficulties. Flouting maxims, for example, is really a type of joint pretense in which speakers and addressees create a new layer of joint activity. Kate and her audience jointly pretend that she heard “millions” of feet run along the corridors. Pretending to say something is not the same type of action as actually saying something, so hyperbole and other such tropes require a different explanation (see Chapter 12). Another difficulty for both methods lies in the notion of saying itself.

PROBLEMS WITH SAYING

In Grice’s scheme, implicatures are based on what is said. But what is saying? According to Grice – though he was vague on this point – it is the literal meaning of the sentence uttered with its ambiguities resolved and its referents specified. Take B’s “There is a garage round the corner.” In British English, *garage* is ambiguous between “parking structure” and “service station,” so to know what B said we must choose between them. We must also identify the time referred to in *is* and the object referred to with *the corner*. Behind Grice’s scheme are three assumptions:

- Assumption 1** What is said is logically prior to what is implicated.
- Assumption 2** The way listeners determine what is said is different in principle from the way they “work out” what is implicated.
- Assumption 3** What is said is well defined for every type of utterance.

There are major problems with all three assumptions.

According to assumption 1, listeners have to know what is said in order to work out what is implicated. Even Grice's own example shows the problem with this assumption. To determine what B said, A had to decide whether *garage* meant "parking structure" or "service station." But he could only determine that it meant "service station" by first working out what B was implicating, namely, that B's remark was relevant to A's being out of petrol. Suppose the exchange had gone this way:

- A: I think I am parked in an illegal parking zone.
 B: There is a garage round the corner.

This time A would work out a different implicature and choose "parking structure" instead. That is, the only way A could determine what B was *saying* was by working out what B must be *implicating*, and this violates assumption 1. The very notion of literal meaning is problematic, which also undermines assumption 1 (Gibbs, 1989, 1994; Searle, 1978, 1980).

According to assumption 2, listeners determine what is said according to one set of principles or procedures, and they "work out" (or *calculate*) what is implicated according to another. But listeners often have to calculate parts of what is said. Consider the novel word meanings in these remarks from a friend:

- The photographer asked me to do a **Napoleon** for the camera.
 Diane's approach to life is very **San Francisco**.
 Never ask two **China trips** to the same party.

I cannot determine what my friend has said (in Grice's sense) because the literal meanings of *Napoleon*, *San Francisco*, and *China trip* don't fit these sentences (see Chapter 3). When I decide that "do a Napoleon" means "tuck my righthand under my coat flap," I apply the same principles or procedures that I apply in working out implicatures (Clark and Clark, 1979; Clark, 1983; Clark and Gerrig, 1983; Nunberg, 1979; Sag, 1981). But if I have to "work out" what is said for *Napoleon*, *San Francisco*, and *China trip*, that violates assumption 2.

Indirect reference is another problem, and is illustrated in Grice's own example. When A tells B "I am out of petrol," he uses *I* to refer to himself and, thereby, indirectly to his car. After all, it isn't A but A's car that is out of petrol.¹⁰ How does B determine that? Knowing the conventional mean-

¹⁰ For similar examples, consider "I am parked up the street," or "I am the blue Volvo over there," or "Could you please fill me up with gasoline."

ing of *I* isn't enough, because in other situations A could have meant "my lawn mower" or "my service station" or "the can for my Molotov cocktails." What A is saying with *I* is something B must "work out" as he would any implicature, and the same goes for all indirect references (see Chapter 4). Indirect references are another violation of assumption 2.

Finally, according to assumption 3, what is said must be well defined for every type of utterance. If it weren't, we would have no basis for working out implicatures. But counter-examples are plentiful. The first type are *phrasal utterances*. When you tell a bartender, "Two pints of Guinness," are you saying (in Grice's sense) "I'd like" or "I'll have" or "Get me" or "Would you get me" or "I'd like you to get me two pints of Guinness"? There is no way in principle of selecting among these candidates. Whatever you are doing, you don't appear to be *saying* that you are ordering beer, and yet you cannot be implicating it either because you cannot cancel the order – it makes no sense to say "Two pints of Guinness, but I'm not ordering two pints of Guinness." Saying simply isn't well defined for phrasal utterances.¹¹

Another type of counter-example are utterances like "hello," "well," and "ah" (see Chapter 6). Traditionally, these are said to have not literal meanings but conventional uses. The dictionary defines *hello* as "an informal expression used to greet another," *well* as "used to express surprise," and *ah* as "used to express various emotions, such as surprise, delight, pain, satisfaction, or dislike." So when a friend tells you, "Helen is coming today," and you utter a delighted "Ah," what are you saying? Because literal meaning isn't defined for "ah," it is impossible to specify either what is said or what is implicated.

The same goes for nonlinguistic signals (Chapter 6). In conversation,

¹¹ Wittgenstein (1958), in *Philosophical Investigations*, describes a communication system between builder A and his assistant B (see also Chapter 10).

A is building with building-stones: there are blocks, pillars, slabs and beams. B has to pass the stones, and that in the order in which A needs them. For this purpose they use a language consisting of the words "block," "pillar," "slab," "beam." A calls them out; – B brings the stone which he has learnt to bring at such-and-such a call. (p. 3)

About these phrasal utterances, Wittgenstein remarks:

But what about this: is the call "Slab!" in example (2) a sentence or a word? – If a word, surely it has not the same meaning as the like-sounding word of our ordinary language, for in (2) it is a call. But if a sentence, it is surely not the elliptical sentence: "Slab!" of our language... But why should I not on the contrary have called the sentence "Bring me a slab" a lengthening of the sentence "Slab!"?... And why should I translate the call "Slab!" into a different expression in order to say what someone means by it? (p. 9)

speakers use their hands, body, face, eyes, and voice to make a wide range of both indicative, or deictic, gestures (e.g., pointing) and iconic gestures (e.g., smirking). These signals are essential to what speakers mean, and yet Grice's notion of what is said doesn't apply to them at all.

To sum up, the cooperative principle has offered an influential account of many phenomena, and Grice's insights have been widely adopted. Most attention has been focused on the maxims – how they should be formulated and applied. Some investigators have offered their own versions of the maxims (e.g., Horn, 1984; Kasher, 1977; Leech, 1983; Levinson, 1987); Dan Sperber and Deirdre Wilson (1986) have even reduced them all to the maxim of relevance. This effort seems misdirected, because Grice's rules of thumb can never be more than just that – rules of thumb. Although Grice recognized that speakers and addressees must cooperate, the maxims were exhortations to speakers, not addressees, and coordination became a sequence of two autonomous actions, the first by speakers and the second by addressees. For a proper understanding of speaker's meaning, we must return to three notions the maxims are based on: (1) "the accepted purpose or direction of the talk exchange," and (2) how people "contribute" to that accepted purpose or direction by means of (3) signals, both linguistic and nonlinguistic. These are just the topics I will consider in the next three chapters.¹²

LEVELS OF COMMUNICATIVE ACTS

There are many speech acts besides illocutionary and perlocutionary acts. According to Austin (1962), when I say to you "Please sit down," I am performing these acts among others:

<i>Phonetic act</i>	I am producing the noises that constitute "Please sit down."
<i>Phatic act</i>	I am uttering the words <i>please</i> , <i>sit</i> , and <i>down</i> .
<i>Rhetic act</i>	I am using the words <i>please</i> , <i>sit</i> , and <i>down</i> with a certain sense and reference.
<i>Locutionary act</i>	I am saying to you "Please sit down."
<i>Illocutionary act</i>	I am asking you to sit down.
<i>Perlocutionary act</i>	I am trying to get you to sit down.

Some of these acts differ in level of action – producing noises is at a lower level than asking you to sit down – and others differ in function. There is no

¹² In this book, however, I will not take up many of the particular linguistic phenomena that have been accounted for by direct appeal to the maxims, though I will take up many phenomena that have been accounted for as flouting of the maxims (Chapter 12).

mention of acts by addressees. So from our perspective, the list is incomplete and lacking in organization. The ideal scheme would have levels and include both speakers and addressees. I will propose just such a scheme.

ACTION LADDERS

Many actions come in hierarchies that I will call *action ladders*. Consider Alan calling an elevator to take him up:

Level	Action in progress from t_0 to t_1
5	A is getting an “up” elevator to come
4	A is calling an “up” elevator
3	A is activating the “up” button
2	A is depressing the “up” button
1	A is pressing the right index finger against the “up” button

Alan is taking five distinct actions, but they are cotemporal – they begin and end together. The act of pressing the finger against the “up” button, for example, is in progress over the same time interval (t_0 to t_1) that the act of activating the “up” button is in progress.

It is tempting to say that Alan is really doing only one thing. It is just that I have described it in five different ways. It is easy to show, however, that Alan *is* doing five things and they are in a causal relation going up the ladder. As we move up the ladder, Alan presses his finger against the “up” button *in order to* depress “up” button, which he does *in order to* activate the “up” button, which he does *in order to* call an “up” elevator, and so on. Or as we go down the ladder, Alan is getting an “up” elevator to come *by means of* calling an “up” elevator, which he is doing *by means of* activating the “up” button, and so on. I will call this property *upward causality*. With upward causality, the relation between any two actions in a ladder is asymmetric, irreflexive, and transitive.¹³

Upward causality leads directly to a property I will call upward completion:

Upward completion. In a ladder of actions, it is only possible to complete actions from the bottom level up through any level in the ladder.

Alan, for example, might press his finger against the “up” button without depressing it because it was stuck. If so, he would complete level 1 while failing to complete level 2. Or he might depress the button (level 2)

¹³ See Goldman (1970) for a discussion of what he calls “level-generational” acts.

without activating it (level 3) because it was defective. Or he might activate the button (level 3) without calling an “up” elevator (level 4) because the elevators were turned off that day. Or he might call an “up” elevator (level 4) without getting it to come (level 5) because some idiot had propped the doors open. Actions in such a ladder are completed from the bottom up.

Upward completion entails another property I will call downward evidence:

Downward evidence. In a ladder of actions, evidence that one level is complete is also evidence that all levels below it are complete.

When Alan sees the “up” light go on, he has good evidence that he has activated the “up” button (level 3). Because of upward completion, that same evidence is also evidence that he has succeeded in pressing his finger against the “up” button (level 1) and in depressing it (level 2). On the other hand, when Alan feels the “up” button depress under his finger (level 2), that isn’t necessarily evidence that he has activated the “up” button (level 3). That is what makes “up” buttons without lights so frustrating. Impatient button pushers have no idea when they have succeeded, so they jab at the buttons over and over and over again.

Austin referred to action ladders in his discussion of speech acts,¹⁴ but only two of his speech acts fit such a ladder, illocutionary and perlocutionary acts: I am trying to get you to sit down by asking you to sit down. Locutionary and illocutionary acts do *not* fit this scheme, as Austin was careful to point out: I am asking you to sit down not *by* saying “Please sit down,” but *in* saying that. This is why Austin used the Latin prefix *in-* in coining the term *illocutionary act*. According to Austin, phonetic, phatic, and rhetic acts are all aspects of locutionary acts. To form such a ladder, we will have to knead Austin’s scheme into a new shape.

JOINT ACTIONS

Any ladder of actions for language use must satisfy several requirements. It must represent the *joint* actions of speakers and addressees as they coordinate what they do. It must capture their actions in progress, not just at the end of a signal. Its levels must conform to upward causality, upward completion, and downward evidence. And, as Grice’s analysis

¹⁴ Austin’s example: A man shoots a donkey, which he does by firing a gun, which he does by pulling the trigger, which he does by tensing his trigger finger (Austin, 1962, p. 107).

demands, it must accommodate signals of all types – flag waving, belfry lanterns, and gestures as well as words, phrases, and sentences. The proposal here is that in ordinary conversation we have at each moment an action ladder of at least four levels, each level consisting of a joint action. I will take up these levels in the order 3, 2, 1, and 4.

Level 3. We have already met one level of joint action: signaling and recognizing. When I say to you “Please sit down” or gesture to a chair, I mean you are to sit down, and you in coordination recognize my intention. The joint action, expressed in the notation of Chapter 3, is this:

Joint[A signals to B that p , B recognizes that A means that p]

In the terminology I will use, the speaker is *signaling that p*. Signaling subsumes Austin’s locutionary acts (saying that you should sit down) and his illocutionary acts (asking you to sit down) – and it isn’t confined to linguistic signals. And I will describe addressees as *recognizing*, or *understanding*, what speakers mean by their signals, though later I will revise this notion radically (Chapter 7).

Level 2. I signal something to you, in turn, by getting you to identify my behavior as a particular signal – as an act by which I mean a specific thing for you. I do this by presenting the signal (an instance of the sentence *Please sit down*, or a gesture toward the chair) for you to identify. I cannot get you to identify the signal without your help. You and I must coordinate what I present with what you identify, and that too is a joint action:

Joint[A presents signal s to B, B identifies signal s from A]

In my terminology, the speaker *presents* a signal to the addressees, and they, in turn, *identify* the signal.

Level 1. I present a signal for you to identify, in turn, by executing a bit of behavior specifically for you to perceive – by articulating “Please sit down” in your hearing or by moving my arm within your vision. I cannot get you to perceive my behavior without your coordination. In conversation, you must be attending to and perceiving it precisely as I am executing it.¹⁵ This too results in a joint action:

¹⁵ In asynchronous settings, like writing and reading a letter, writers intend readers to attend to their marks, not simultaneously, but at a later time, and readers attend to the marks on that assumption. The delay between executing the marks and attending to them is only one of the reasons why coordination takes a different form in asynchronous settings – and has different consequences (Clark and Brennan, 1991).

Joint[A executes behavior *t* for B to perceive; B attends perceptually to behavior *t* from A]

In my terminology, speakers *execute* a behavior for addressees, who in their turn *attend* to that behavior.

Level 4. What am I doing by asking you to sit down – by performing an illocutionary act? I am proposing, suggesting, posing, or putting forward a project for us to carry out jointly – namely, that I get you to sit down. Now, getting you to sit down is another thing I can't do by myself. It is a joint action that I am projecting for the two of us to do, and that requires us to coordinate our actions. I will call this joint action a *joint project*.

Joint projects have two parts. In my terminology, the speaker *proposes* a joint project, and the addressees *take it up*. I propose that you sit down, and you take up my proposal by sitting down or by agreeing to sit down. A *proposal* is expected to be followed by its *uptake*. Recall that Austin argued that “My attempt to make a bet by saying ‘I bet you sixpence’ is abortive unless you say ‘I take you on’ or words to that effect,” and that marrying, bequeathing, and appointing also require uptake.¹⁶ That makes betting, marrying, bequeathing, and appointing joint projects. Getting you to do something and getting you to accept my beliefs are also joint projects; they are initiated by the illocutionary acts of asking and telling.

Joint projects are usually achieved by two actions in sequence. “Please sit down” is followed by your sitting down, and “I bet you sixpence” is followed by your “I take you on.” But for a ladder of actions, the paired actions by the speaker and addressee must be cotemporal – they must be in progress simultaneously. The joint action I will argue for is this: By asking you to sit down, I am *proposing* a joint project; and by understanding my request, you are *considering* taking up that proposal. The joint action is this:

Joint[A proposes joint project *w* for A and B; B considers joint project *w* for A and B]

In this scheme, proposing is different from signaling, and considering is different from recognizing. The differences are easy to see in conversations with more than two participants. In the following exchange, two British academics, Arthur and Charles, are interviewing a prospective student, Beth (3.1.174):

¹⁶ So what I am calling uptake is only one part of what Austin called uptake. It is the “taking on” part of the bet, once it has been heard and understood.

Arthur: u:m - well you are . proposing . taking on . quite something Mrs.
 Finney aren't you,
 Beth: yes, I am,

In the course of Arthur's utterance, Arthur is asking Beth a question, and she is trying to recognize what he means. Charles, the third participant in the conversation, is also trying to recognize what Arthur means. These actions are all at level 3. At level 4, Arthur is proposing that Beth answer his question, and she is considering taking up that proposal. But Arthur is *not* proposing a joint project for *Charles* to consider. Speakers propose joint projects for addressees and *not* for all participants. That is precisely what distinguishes addressees from participants (Clark & Carlson, 1982a, b).¹⁷ In short, Arthur means what he means for both Beth and Charles to recognize, but proposes what he proposes for only Beth to consider. Once Charles has understood Arthur's question, he is done. The point is subtle, but essential for distinguishing level 4 from level 3.

THREE ACTION LADDERS

Individual action ladders, like the one for calling an "up" elevator, describe the several actions that are in progress during a single slice of time. Remarkably, the four joint actions just described also form such a ladder. To see this, let us consider an utterance by a university instructor (Adam) to a student (Bart) (3.5b.552):

Adam: sit down **here** [pointing at a chair] would you¹⁸

And let us focus on the actions in progress over the time interval in which Adam is producing the word *here* and gesturing at a chair (in boldface). We can identify three distinct action ladders over this interval, one for Adam's actions, one for Bart's actions, and one for their joint actions.

The ladder of Adam's individual actions is really a reformulation of Austin's speech acts, and it looks like this:

¹⁷ Clark and Carlson (1982a, b) provide a broad range of evidence that distinguishes "participant-directed informatics" from "addressee-directed illocutionary acts." That is equivalent to the distinction here between signaling something to all participants and proposing joint projects only for addressees.

¹⁸ Gestures weren't marked in the transcripts, but it is reasonable to assume that a gesture like this accompanied the word *here*.

Level A's actions in progress

-
- | | |
|---|---|
| 4 | A is proposing to B that B sit here for A. |
| 3 | A is asking B to sit here. |
| 2 | A is presenting to B a signal composed of "here" plus pointing at the chair. |
| 1 | A is executing for B's perception the articulation of "here" and the movement of his arm. |

Within this interval, Adam is *in the process of*, or *in the middle of*, proposing, asking, presenting, and executing things. These actions *in progress* form a genuine action ladder with upward causality, upward completion, and downward evidence. The same time interval yields a ladder for Bart's actions:

Level B's actions in progress

-
- | | |
|---|---|
| 4 | B is considering A's proposal that B sit here for A. |
| 3 | B is recognizing A's request for B to sit here. |
| 2 | B is identifying A's signal as composed of "here" plus pointing at the chair. |
| 1 | B is attending to A's articulation of "here" and the movement of A's arm. |

During this interval, Bart is in the process of, or in the middle of, considering, recognizing, identifying, and attending to things, and may not have completed any of them. Bart's ladder is also an action ladder complete with upward causality, upward completion, and downward evidence.

These two ladders are linked. Adam's actions at each level are participatory actions – parts of joint actions – each linked to a participatory action by Bart. The result is a ladder of *joint* actions, which, in general, looks like this:

Level	Speaker A's actions	Addressee B's actions
4	A is <i>proposing</i> joint project w to B	B is <i>considering</i> A's proposal of w
3	A is <i>signaling</i> that p for B	B is <i>recognizing</i> that p from A
2	A is <i>presenting</i> signals to B	B is <i>identifying</i> signals from A
1	A is <i>executing</i> behavior t for B	B is <i>attending</i> to behavior t from A

At each level we find a joint action by Adam and Bart. And like the single ladder, the joint ladder has upward causality, upward completion, and downward evidence. Adam must get Bart to attend to his voice or

movement (level 1) in order to get him to identify the word and gesture he is presenting (level 2). Adam must succeed at that in order to get Bart to recognize what he means (level 3), and he must succeed at that in order to get Bart to consider the joint project he is proposing (level 4). Likewise, evidence that Adam got Bart to understand what he means (level 3) is also good evidence that he got Bart to attend to his voice and arm movement (level 1) and to identify the word and gesture (level 2). Again, causation goes upward, and evidence downward. Because there is no natural terminology for the joint actions in this ladder, I will make do with these cumbersome names:

- Level 4** Proposal and consideration
- Level 3** Signaling and recognition, or meaning and understanding
- Level 2** Presentation and identification
- Level 1** Execution and attention

With this analysis, we move from Austin's mixed collection of speech acts to a ladder of joint actions performed in the use of language. Its advantage is that it satisfies upward causality, includes what addressees do, and specifies the link between speakers' and addressees' actions. Yet this is so far only a blueprint. It will take the next several chapters to fill in the details.

Conclusions

To communicate is, according to its Latin roots, "to make common," to make known within a group of people. As we saw in Chapters 2, 3, and 4, people have to coordinate closely to make a piece of information common for them – to add it to their common ground. The same argument applies to what is traditionally called communication, and it leads to the conclusion: Communicative acts are joint acts.

Surprisingly, this conclusion is entailed by Grice's very characterization of speaker's meaning. Suppose Ann says "Please sit down" to Ben, meaning that he is to sit down. Her meaning is a type of intention that she cannot discharge without Ben doing his part in recognizing that intention. For speakers to mean something, they must act jointly with their addressees. The same holds for the various types of speech acts – locutionary acts, illocutionary acts, perlocutionary acts, and the rest – as Austin himself seemed to recognize. It is time to take the jointness of these actions seriously.

Communication with language takes actions at many levels, as Austin

also recognized. I have argued that these levels form a ladder of joint actions. An action ladder is a set of cotemporal actions ordered with upward causality, upward completion, and downward evidence. In language use, these levels are joint actions. At the bottom, Ann executes behaviors and, in coordination with her, Ben attends to them; by these joint actions, Ann presents a signal and, in coordination, Ben identifies it; by these joint actions in turn, Ann signals something to Ben and, in coordination, Ben recognizes what she means; and by these joint actions, Ann proposes a joint project and, in coordination, Ben considers her proposal. These may not be the only levels, but they are the main ones.