

Cambridge University Press
978-0-521-56745-9 - Using Language
Herbert H. Clark
Excerpt
[More information](#)

PART I
Introduction

1 | Language use

Language is used for doing things. People use it in everyday conversation for transacting business, planning meals and vacations, debating politics, gossiping. Teachers use it for instructing students, preachers for preaching to parishioners, and comedians for amusing audiences. Lawyers, judges, juries, and witnesses use it in carrying out trials, diplomats in negotiating treaties, and actors in performing Shakespeare. Novelists, reporters, and scientists rely on the written word to entertain, inform, and persuade. All these are instances of *language use* – activities in which people do things with language. And language use is what this book is about.

The thesis of the book is this: Language use is really a form of *joint action*. A joint action is one that is carried out by an ensemble of people acting in coordination with each other. As simple examples, think of two people waltzing, paddling a canoe, playing a piano duet, or making love. When Fred Astaire and Ginger Rogers waltz, they each move around the ballroom in a special way. But waltzing is different from the sum of their individual actions – imagine Astaire and Rogers doing the same steps but in separate rooms or at separate times. Waltzing is the joint action that emerges as Astaire and Rogers do their individual steps in coordination, as a couple. Doing things with language is likewise different from the sum of a speaker speaking and a listener listening. It is the joint action that emerges when speakers and listeners – or writers and readers – perform their individual actions in coordination, as ensembles.

Language use, therefore, embodies both individual and social processes. Speakers and listeners, writers and readers, must carry out actions as individuals if they are to succeed in their use of language. But they must also work together as participants in the social units I have

4 | INTRODUCTION

called ensembles. Astaire and Rogers perform both individual actions, moving their bodies, arms, and legs, and joint actions, coordinating these movements, as they create the waltz. In some quarters, language use has been studied as if it were entirely an individual process, as if it lay wholly within the cognitive sciences – cognitive psychology, linguistics, computer science, philosophy. In other quarters, it has been studied as if it were entirely a social process, as if it lay wholly within the social sciences – social psychology, sociology, sociolinguistics, anthropology. I suggest that it belongs to both. We cannot hope to understand language use without viewing it as joint actions built on individual actions. The challenge is to explain how all these actions work.

The goal of this chapter is to make a preliminary case for the thesis. To do this, I will take a tour through the settings of language use, the people who play roles in these settings, and the way joint actions emerge from individual actions. It will take the rest of the book to fill out the picture and develop principles to account for how language use is a joint action.

Settings of language use

Over the years, when I have asked people for instances of language use, they have offered such examples as “conversation,” “reading a novel,” “policemen interrogating a suspect,” “putting on a play,” “talking to oneself,” and dozens more. These answers are remarkable for their range. To get a sense of that range, let us look at the answers classified by scene and medium. The scene is where the language use takes place.¹ The medium is whether the language use is spoken or signed or gestural, or written or printed, or mixed. I will use *setting* for the scene and medium combined and divide the media simply into *spoken* and *written* forms.

SPOKEN SETTINGS

The spoken setting mentioned most often is conversation – either face-to-face or on the telephone. Conversations may be devoted to gossip, business transactions, or scientific matters, but they are all characterized by the free exchange of turns among the two or more participants. I will call these *personal settings*. In monologues, in contrast, one person speaks with little or no opportunity for interruption or turns by members of the audience. Monologues come in many varieties

¹ See Hymes (1974, pp. 55–56) for a related use of setting and scene.

too, as when a professor lectures to a class, a preacher gives a sermon, or a student relates a recent experience to an entire class. These people speak for themselves, uttering words they formulated themselves for the audience before them, and the audience isn't expected to interrupt. These I will call *nonpersonal settings*.

In *institutional settings*, the participants engage in speech exchanges that resemble ordinary conversation, but are limited by institutional rules. As examples, think of a politician holding a news conference, a lawyer interrogating a witness in court, a mayor chairing a city council meeting, or a professor directing a seminar discussion. In these settings, what is said is more or less spontaneous even though turns at speaking are allocated by a leader, or are restricted in other ways. In *prescriptive settings*, in contrast, there may be exchanges, but the words actually spoken are completely, or largely, fixed beforehand. Think of the members of a church or synagogue reciting responsive readings from a prayer book, or a bride and groom reciting vows in a marriage ceremony, or a basketball referee calling foul. Prescriptive settings can be viewed as a subset of institutional settings.

The person speaking isn't always the one whose intentions are being expressed. The clearest examples are in *fictional settings*: John Gielgud plays Hamlet in a performance of *Hamlet*; Vivien Leigh plays Scarlett O'Hara in *Gone with the Wind*; Frank Sinatra sings a love song in front of a live audience; Paul Robeson sings the title role in the opera *Otello*; or a television pitchman makes a sales pitch to a television audience. The speakers are each vocalizing words prepared by someone else – Shakespeare, Cole Porter, the news department – and are openly pretending to be speakers expressing intentions that aren't necessarily their own.

Related to fictional settings are the *mediated settings* in which there are intermediaries between the person whose intentions are being expressed and the target of those intentions. I dictate a letter for Ed to my secretary Annie; a telephone company recording tells me of the time or weather; a television news reader reads the evening news; a lawyer reads Baker's last will and testament at a hearing; a recording is triggered in a building announcing a fire and describing how to find the fire escape; and a UN interpreter translates a diplomat's French simultaneously into English. When I dictate a letter to my secretary Annie and say "I'll see you Saturday," the person I expect to see on Saturday isn't Annie but the addressee of my letter Ed.

Finally, there are *private settings* in which people speak for them-

6 | INTRODUCTION

selves without actually addressing anyone else. I might exclaim silently to myself, or talk to myself about solving a mathematics problem, or rehearse what I am about to say in a seminar, or curse at another driver who cannot hear me. What I say isn't intended to be recognized by other people – at least in the way other forms of speaking are.² It is only of use to myself.

WRITTEN SETTINGS

When printing, writing, and literacy were introduced, people adapted spoken language to the printed medium, so it is no surprise that written uses have many of the characteristics of spoken ones. The written settings most like conversations are the personal settings, when people write to others they are personally acquainted with, as when I write my sister a letter, or write a colleague a message on the computer. In computer settings where the writing and reading on two terminals are simultaneous, the experience can resemble conversation even more closely.

Many written messages, however, are directed not at individuals known to the writer, but at a type of individual, such as “the reader of the *New York Times*” or “the reader of *Science*.” These are *nonpersonal settings*. So a newspaper reporter writes a news story for readers of the *New York Times*, or an essayist writes on Scottish castles for readers of *Country Life*, or a physicist writes a textbook on electricity and magnetism for university undergraduates, or a car owner writes to the service department of Ford Motor Company. The reporter may know a few of the *New York Times*' readers, yet he or she is directing the news story at its general readership. Fiction, too, is usually directed at types of individuals, often defined very broadly, as when Henry James wrote *The Turn of the Screw*, and Edgar Allan Poe wrote “The Masque of the Red Death,” and William Shakespeare wrote *Hamlet*. In written fiction, the author is writing for an audience, but as with spoken fiction, the intentions expressed are not his own.

Written settings, like spoken ones, can introduce intermediaries between the person whose intentions are being expressed and the intended audience. These again are mediated settings. Usually, the person actually writing the words is doing so in place of the person who appears to be doing the writing or speaking. Examples: The Brothers Grimm

² See the discussion of “response cries” (Goffman, 1978) in Chapter 11.

write down the folktale “Aschenputtel”; a translator translates *Hamlet* into French; a ghost writer writes Charlie Chaplin’s autobiography; a speech writer writes a speech for the President; my secretary types the letter to Ed from my dictation; and the manuscript editor for this book edits my writing. The President’s speech writers, for example, write as if they were the President, who later reads the words as if they were his or her own. We make the pretense that the speech writers weren’t even involved in the process. Recorders, translators, ghost writers, secretaries, and manuscript editors, in their different ways, do much the same thing.

In some written settings, the words are selected through an institutional procedure. An advertising firm composes an advertisement for a magazine; a drug company composes the warning label for an aspirin bottle; a food company labels a package as baking soda; the US Senate legislates the wording of a new tax law; and the California legislature decides on the wording of state road signs. Although one person may have composed the words, it is the institution – the ad agency, drug company, or legislature – that is ultimately responsible, approving the wording as faithful to the institution’s collective intentions.

Written language is used in private settings as well. I can write in my diary, scribble a reminder to myself, take notes on a lecture, make a grocery list, or work out a mathematics proof on paper. As in the spoken settings, I am writing solely to myself for later use.

What follows are examples of the major types of spoken and written settings, but these types are hardly exhaustive. Humans are creative. For each new technology – writing systems, printing, telegraph, telephones, radio, audio recording, television, video recording, telephone answering machines, interactive computers, and voice recognizers – people have developed new settings. With no end to new technologies, there is no end to the settings they might create. Our interest must be in the principles by which these new forms are created.

8 | INTRODUCTION

| | Spoken settings | Written settings |
|---------------|---|--|
| Personal | A converses face to face with B | A writes letter to B |
| Nonpersonal | Professor A lectures to students in class B | Reporter A writes news article for readership B |
| Institutional | Lawyer A interrogates witness B in court | Manager A writes business correspondence to client B |
| Prescriptive | Groom A makes ritual promise to bride B in front of witnesses | A signs official forms for B in front of a notary public |
| Fictional | A performs a play for audience B | Novelist A writes novel for readership B |
| Mediated | C simultaneously translates for B what A says to B | C ghostwrites a book by A for audience B |
| Private | A talks to self about plans | A writes note to self about plans |

CONVERSATION AS BASIC SETTING

Not all settings are equal. As Charles Fillmore (1981) put it, “the language of face-to-face conversation is the basic and primary use of language, all others being best described in terms of their manner of deviation from that base” (p. 152). If so, the principles of language use may divide mainly into two kinds – those for face-to-face conversation, and those that say how the secondary uses are derived from, or depend on it, or have evolved from it. Language uses are like a theme and variations in music. We look first at the theme, its melody, rhythm, and dynamics, and then try to discover how the variations are derived from it. Fillmore added, “I assume that this position is neither particularly controversial nor in need of explanation.” Still, it is worth bringing out what makes face-to-face conversation basic and other settings not.

For a language setting to be basic, it should be universal to human societies. That eliminates written settings, since entire societies, and groups within literate societies, rely solely on the spoken word. One estimate is that about a sixth of the world’s people are illiterate. And most languages as we know them evolved before the spread of literacy. We can also eliminate spoken settings that depend on such technologies as radio, telephones, television, and recordings, since these are hardly universal. Most people participate only rarely in nonpersonal, institutional, and prescriptive settings, and even then their participation is usually restricted to certain roles – audiences of lectures, parishioners,

court observers. People do often participate in fictional settings, but usually as audience. The commonest setting is face-to-face conversation.

Face-to-face conversation, moreover, is the principal setting that doesn't require special skills. Reading and writing take years of schooling, and many people never do get very good at them. Even among people who know how to write, the most that many ever do is personal letters. Simple essays, to say nothing of news stories, plays, or novels, are beyond them. It also takes instruction to learn how to act, sing, lead seminars, chair meetings, and interrogate witnesses. And most people find it difficult to lecture, tell jokes, or narrate reasonable stories without practice. Almost the only setting that needs no specialized training is talking face to face.

Face-to-face conversation is also the basic setting for children's acquisition of their first language. For the first two or three years, children in both literate and illiterate societies learn language almost solely in conversational settings. Whatever they learn from books also comes in conversational settings, as their caretakers read aloud and check on what they understand. Children may learn some language from other media, but they apparently cannot learn their first language from radio or television alone.³ In school, the language of peers is influential in the dialect acquired, and that too comes from conversational settings. Face-to-face conversation is the cradle of language use.

NONBASIC SETTINGS

What, then, makes other settings not basic? Let us start with the features of face-to-face conversation listed here (Clark and Brennan, 1991):

- | | | |
|---|-----------------------|---|
| 1 | Copresence | The participants share the same physical environment. |
| 2 | Visibility | The participants can see each other. |
| 3 | Audibility | The participants can hear each other. |
| 4 | Instantaneity | The participants perceive each other's actions at no perceptible delay. |
| 5 | Evanescence | The medium is evanescent – it fades quickly. |
| 6 | Recordlessness | The participants' actions leave no record or artifact. |
| 7 | Simultaneity | The participants can produce and receive at once and simultaneously. |

³ For evidence, see Sachs, Bard, and Johnson (1981) and Snow, Arlman-Rupp, Hassing, Jobse, Joosten, and Vorster (1976).

10 | INTRODUCTION

- | | | |
|----|---------------------------|--|
| 8 | Extemporaneity | The participants formulate and execute their actions extemporaneously, in real time. |
| 9 | Self-determination | The participants determine for themselves what actions to take when. |
| 10 | Self-expression | The participants take actions as themselves. |

If face-to-face settings are basic, people should have to apply special skills or procedures whenever any of these features are missing. The more features are missing, the more specialized the skills and procedures. That is borne out informally.

Features 1 through 4 reflect the *immediacy* of face-to-face conversation. In that setting, the participants can see and hear each other and their surroundings without interference. Telephones take away copresence and visibility, limiting and altering language use in certain ways. Conversations over video hookups lack copresence, making them different too. In lectures and other nonpersonal settings, speakers have restricted access to their addressees, and vice versa, changing how both parties proceed. In written settings, which lack all four features, language use works still differently.

Features 5 through 7 reflect the *medium*. Speech, gestures, and eye gaze are evanescent, but writing isn't, and that has far-reaching effects on the course of language use. Speech isn't ordinarily recorded, but when it is, as on a telephone answering machine, the participants proceed very differently. In contrast, writing is ordinarily relayed by means of a printed record, and that leads to dramatic differences in the way language gets used. With written records and no instantaneity, writers can revise what they write before sending it off, and readers can reread, review, and cite what they have read. Most spoken settings allow the participants to produce and receive simultaneously, but most written settings do not. Being able to speak and listen simultaneously gives people in conversation such useful strategies as interrupting, overlapping their speech, and responding "uh huh," and these are ruled out in most written settings.

Features 8 through 10 have to do with *control* – who controls what gets done and how. In face-to-face conversation, the participants are in full control. They speak for themselves, jointly determine who says what when, and formulate their utterances as they go. In other settings, the participants are restricted in what they can say when. The church, for example, determines the wording of many prayers and responses. In fictional settings, speakers and writers only *make as if* they are taking certain actions – Gielgud is only play-acting his role as Hamlet – and that

alters what they do and how they are understood. And in mediated settings, there are really two communications. Wim says “Heeft u honger?” in Dutch, which David translates for Susan as “Are you hungry?” Susan is expected to hear David’s utterance knowing it is really Wim who is asking the question. The less control participants have over the formulation, timing, and meaning of their actions, the more specialized techniques they require.

What about private settings? These are sometimes considered the basic setting for language use. We all talk to ourselves, the argument goes, so private settings are surely universal. When we do talk to ourselves, however, the principal medium is the language we have acquired from others. People who know only English use English; people with only Chinese use Chinese; and people with only American Sign Language use American Sign Language. We may develop additional ways of talking to ourselves, but these too are derived from our social ways of talking. In talking to ourselves, we are making as if we were talking to someone else. Private settings are based on conversational settings.

In brief, face-to-face conversation is the basic setting for language use. It is universal, requires no special training, and is essential in acquiring one’s first language. Other settings lack the immediacy, medium, or control of face-to-face conversation, so they require special techniques or practices. If we are ever to characterize language use in all its settings, the one setting that should take priority is face-to-face conversation. This is a point I will take for granted in the rest of the book.

Arenas of language use

Language settings are of interest only as arenas of language use – as places where people do things with language. At the center of these arenas are the roles of *speaker* and *addressee*. When Alan is addressing Barbara, he is the speaker and she the addressee. Now, Alan is speaking with the aim of getting Barbara to understand him and to act on that understanding. But he knows he cannot succeed unless she takes her own actions. She must attend to him, listen to his words, take note of his gestures, and try to understand what he means at the very moment he is speaking. Barbara knows all this herself. So Alan and Barbara don’t act independently. Not only do they take actions *with respect to each other*, but they *coordinate* these actions with each other. In the term I introduced