

2 | Joint activities

Language use arises in joint activities. You call up your sister for an address, or talk to a friend about what to take on a picnic, or discuss news with a colleague. If you were later asked “What did you do?” you wouldn’t describe your acts of speaking. You would describe the joint activities you took part in. “I got an address from my sister.” “My friend and I decided what to bring on a picnic.” “A colleague and I traded gossip.” In each case, you take the joint activity to be primary, and the language you used along the way to be secondary, a means to an end. To account for the language used, we need to understand the joint activities.

A discourse is one type of joint activity – one in which language plays an especially prominent role. Originally the term *discourse* meant conversation or dialogue – literally, a running back and forth – but nowadays it includes lectures, interviews, interrogations, plays, novels, essays, personal letters, and much much more. But if discourses are a type of joint activity, we will never understand how they work until we understand more generally how joint activities work. This, then, is another reason for investigating joint activities.

And just as language use arises in joint activities, these are impossible without using language. Two or more people cannot carry out a joint activity without communicating, and that requires language use in its broadest sense. Yet whenever people use language, they are taking joint actions. Language use and joint activity are inseparable. The conclusion, once again, is that we cannot understand one without the other. We must take what I will call an *action approach* to language use, which has distinct advantages over the more traditional *product approach*.

In this chapter, I take up joint activities and how they work. Although I will focus on those in which language dominates, it is the joint

activities *per se* that are of interest. In most the language is merely an emergent product.

Joint activities

What is a joint activity? The approach I will take was inspired by Stephen Levinson's (1979, 1992) notion of *activity type* (1992, p. 69):

I take the notion of an activity type to refer to a fuzzy category whose focal members are goal-defined, socially constituted, bounded, events with constraints on participants, setting, and so on, but above all on the kinds of allowable contributions. Paradigm examples would be teaching, a job interview, a jural interrogation, a football game, a task in a workshop, a dinner party, and so on.

The notion of activity type, Levinson argued, is preferable to such related notions as "speech event," "episode," "form of life," and Wittgenstein's "language games" "because it refers to any culturally recognized activity, whether or not that activity is coextensive with a period of speech or indeed whether any talk takes place in it at all." For Levinson, an activity type can be either a time-bounded event ("a football game") or an ongoing process ("teaching"). I will call the first *an activity* and the second *activity*. And activity types may have a single participant ("a task in a workshop") or more than one participant ("a dinner party"). My interest is in activities with more than one participant, which are fundamentally different from autonomous activities. I will call them joint activities.

Activity types—hence joint activity types—vary on many dimensions. One is *scriptedness*, the "gradient formed by two polar types, the totally prepackaged activity, on the one hand (e.g., a Roman Mass) and the largely unscripted event on the other (e.g., a chance meeting on the street)" (p. 69). A second dimension is *formality*, ranging from "a highly formal activity on the one hand and a very informal one on the other" (p. 69). A third dimension is *verbalness*, "the degree to which speech is an integral part of each activity" (p. 70):

On the one hand, we have activities constituted entirely by talk (a telephone conversation, a lecture, for example); on the other, activities where talk is nonoccurring, or if it does occur is incidental (a game of football, for instance). Somewhere in between ... we have the placing of bets, or a Bingo session, or a visit to the grocers.

I would add two more dimensions. One is *cooperativeness*. It ranges from cooperative activities like buying groceries to adversarial, or competitive, activities like playing tennis or cross-examining witnesses in court. The other dimension is *governance*. Quartet playing, chess playing, party planning, and making acquaintance are more or less egalitarian, the participants having roughly equal roles (“A and B did something”). Lecturing to a class, interviewing an applicant, and buying a car are at the autocratic end, with one participant playing a dominant role (“A did something to or for B”). These dimensions are summarized here:

| Dimension of variation | From | To |
|-----------------------------|----------------------|----------------|
| Scripted vs. unscripted | marriage ceremony | chance meeting |
| Formal vs. informal | city council meeting | gossip session |
| Verbal vs. nonverbal | telephone call | football game |
| Cooperative vs. competitive | business transaction | tennis match |
| Egalitarian vs. autocratic | making acquaintance | class lecture |

Because there are still other dimensions of variation, the number of potential activity types is vast.

At first glance, activity types appear to have arbitrary, ad hoc, unprincipled properties, but Levinson argued that that isn’t so. These properties “can be seen to follow from a few basic principles, in particular rational organization around a dominant goal” (p. 71), although Levinson didn’t say what these principles are. The challenge is to discover these principles – at least for joint activities, especially discourses. We will start on the principles in this chapter, but they will occupy us in various ways for the rest of the book.

EXAMPLE OF A JOINT ACTIVITY

Most joint activities don’t come scripted like a marriage ceremony. They emerge in time as two or more people try extemporaneously to accomplish certain ends. Take a conversation I recorded in a California drug store as I was buying a couple of items from a clerk I will call Stone.¹

Clark walks up to a counter and places two items next to the cash register.

Stone is behind the counter marking off items on an inventory.

Clark, looking at Stone, catches her eye.

¹ The transcript here is based on a surreptitious audio recording fleshed out later by my recollection of what occurred. It is incomplete and inaccurate in many respects. Unfortunately, it is difficult to get video records of such exchanges.

Stone, meeting Clark's eyes: "I'll be right there."

Clark: "Okay."

Stone continues marking off items for fifteen seconds, puts the inventory aside, turns toward Clark, and manifestly begins to look for the items Clark is purchasing.

Clark, noting her search, points at the two items on the counter between them: "These two things over here."

Stone nods, takes the items, examines the prices on them, and rings them up on the cash register.

Stone: "Twelve seventy-seven."

Clark: "Twelve seventy-seven."

Clark takes out his wallet, extracts a twenty-dollar bill, hands it to Stone, then rummages in his coin purse for coins.

Clark: "Let's see that's two pennies I've got two pennies."

Clark hands Stone two pennies.

Stone: "Yeah."

Stone then enters \$20.02 in the register, which computes the change.

Stone (handing change to Clark): "Seven twenty-five is your change."

Clark: "Right."

Clark puts the money in his wallet while Stone puts the items and receipt in a bag.

She hands the bag to Clark, they break eye contact, and he turns and walks away.

What Stone and I did was transact a piece of business. I wanted to buy two items, so she and I exchanged the items for money. Our transaction began when she turned to me and looked for my purchases, and it ended when we broke eye contact and I left. The transaction got advanced through a series of joint actions – my catching her eye, our utterances to each other, her picking up the items, my handing her money. How does such a rich structure emerge?

PARTICIPANTS AND ROLES

In any joint activity, some people are understood to be taking part, and others not. When a string quartet plays a Haydn quartet, the four musicians are the participants. Other musicians may be standing around, even trying to play along, but unless they are recognized as ratified participants, they stand outside the playing of the quartet proper as bystanders, as nonparticipants. People get ratified as participants as the joint activity gets initiated and carried out. Also, one joint activity can be embedded within another. When the quartet plays the Haydn piece in a concert, the quartet playing (one joint activity) is embedded within the

concert performance (another joint activity), and in the concert performance, both musicians and audience are now participants. Likewise, Stone and I were the participants in our drugstore transaction, and the nearby customers were nonparticipants. Our transaction was embedded in a larger joint activity that required Stone and me to coordinate with other customers vying for her services.

People who take part in a joint activity aren't just participants *simpliciter*. They have roles in that activity – *activity roles*. In the quartet playing, one musician is first violin, another second violin, a third viola, and a fourth cello. These roles help shape what they each do and are understood as doing. In the concert performance, other people are members of the audience, a role that defines still other activities. People in joint activities get ratified not merely as participants, but as participants in particular roles.

So it was in the drugstore. Stone was the server, and I was the customer, roles that helped shape what we did and how we interpreted each other. When she said "Twelve seventy-seven," I took her to mean I was to pay her \$12.77. It was the server's job at that point to tell me what I needed to pay. But when I said "Twelve seventy-seven," she didn't take me to mean that she was to pay me \$12.77. That would have violated my role as customer. The norm for what people do in transactions is partly defined by their roles. If Stone had later attended a lecture I gave, she and I would instead be member of the audience and lecturer, roles that dictate different actions and interpretations. In other joint activities, the roles might be police officer and citizen, attorney and witness, teacher and student, supervisor and worker, narrator and audience, priest and congregation. Indeed, the roles may change from one subactivity to the next, or emerge only as the nature of the joint activity becomes clear.

The participants in joint activities also have *personal identities*. Stone and I each had our own identities, beliefs, feelings, and desires, which also helped shape what we did.

GOALS

People participate in joint activities to achieve certain dominant goals. In many activities, one person initiates the joint activity with a dominant goal in mind, and the others join him or her in order to achieve it. I initiated the business transaction with Stone to buy two items, and she joined me to complete it. Joint activities can usually be summarized by describing the dominant goal achieved:

Guide A led tourists B, C, and D through the Eiffel Tower.
 Musicians A, B, C, and D played a Haydn string quartet.
 A on white and B on black played a game of chess.
 Professor A lectured students B, C, D, et al. on labor law.
 Caller A got a telephone number from telephone operator B.
 Police officer A interrogated witness B about a crime.
 Guests A and B at a party got acquainted.
 Customer A bought items from server B.

Each description specifies the participants (A, B, et al.), their roles (e.g., guide, tourist), and the goal achieved. Add the words “managed to” to each description – for example, A *managed to* lead B, C, and D through the Eiffel Tower – and we have the same joint activity. That shows that the activity was goal directed, something the participants jointly intended to do. In some joint activities – like a gossip session – the dominant goal may be vague (e.g., “catch up on news”), or it may evolve in the course of the activity.

Although the participants may share the dominant goal, there is usually a division of labor among them. In the tour of the Eiffel Tower, the guide assumes one set of responsibilities, and the tourists another. The two have the same end goal – that the guide lead the tourists through the Eiffel Tower – but differ in what they do in fulfilling the goal. The participants’ actions and responsibilities depend on the role they inherited from the activity they are engaged in – e.g., as helper or helped, or guide or guided, or interviewer or interviewee – even in egalitarian activities. In the string quartet, the first violin’s responsibilities and actions are very different from the cello’s. These are no less joint activities because of the division of labor.

In most joint activities, the participants pursue many goals at once (Brown and Levinson, 1987; Goffman, 1974; Hobbs and Evans, 1980). Their dominant goal, as I have called it, is a *domain goal* – getting their business transacted, the chess game played, the lecture completed, the witness interrogated. But the participants also have *procedural goals*, such as doing all this quickly and efficiently, making clear moves, attending to what is being done. They also have *interpersonal goals*, such as maintaining contact with the other participants, impressing them, being polite, maintaining self-respect. They may also have *private agendas* such as deceiving the others, getting rid of them, or working the situation for personal advantage. These goals are not all alike in their influence on joint activities. They divide at least into public and private goals.

In any joint activity, certain goals become a matter of public record, what the participants are “on record” as doing in the activity. In my terminology, information is *public* in a joint activity if it is openly recognized by all the participants. Some public goals get established explicitly, as when Jack asks Kate to play chess. Others become public without being explicitly agreed to. When I approached Stone in the drugstore, the two of us took it for granted that I was there to buy something. For a goal to be a joint one, as we will see, it must be public, and it is the joint goals that define the joint activity the participants are engaged in. For Jack and Kate to play a game of chess, it must be public that they are doing so. The same goes for me buying the two items from Stone.

Other goals are *private*. Although many of these are innocent enough, others would be self-defeating if they became public. If Duncan is trying to impress Ann with his knowledge of classical music, it wouldn’t do to make that goal public. The same goes for being polite and maintaining face and for deceiving or getting rid of others. In competitive activities like chess or tennis, success hinges on keeping private goals private – even deceiving the other participants about them.² People’s private goals are sometimes in direct conflict with their public goals, making their adherence to the public goals a sham.

So joint activities are influenced by at least two types of goals. Public goals are there for all the participants to see, but private goals are hidden from view. Public and private goals have different consequences.

COORDINATION OF ACTION

Every joint activity requires coordination among its participants. If the four musicians are to play the Haydn quartet, they must coordinate. They must play the same edition in the same key, start together, stick to the same tempo, and finish together. In the drugstore, Stone and I engaged in the same business transaction, and we started, proceeded, and finished together.

How do people manage to coordinate? One way is with conventional procedures. In playing a string quartet, the four musicians exploit conventional procedures for handling their instruments, reading music, setting tempo and loudness, starting and stopping, and much more. In the drugstore, Stone and I used conventional English expressions and

² When there are more than two participants, as in team sports, there can also be coalitions with private agendas.

applied conventional procedures for specifying the price, exchanging money, bagging the items, and taking them away. But, as we will discover, people also coordinate by means of nonconventional procedures – both in and out of language use. How people coordinate is one of the fundamental issues of language use.

SECTIONS AND BOUNDARIES

Most joint activities get realized as sequences of smaller actions, many of which are themselves joint actions. For the four musicians to play the Haydn quartet, they must play the first, second, and third movements in that order. And to play the first movement, they must play the first section, beginning with the first phrase, beginning with the first measure, beginning with the first note. Playing the quartet divides into sections, or *phases*, each of which divides into subsections or subphases, and so on. What emerges is a hierarchy of joint actions.

So it goes for many unscripted joint activities. When Stone and I transacted our business, the items I wanted had to be identified, their prices identified, the money paid, and the items taken away – and in that order. We might count four main sections, each with its own goal. The identification of the items itself required several subsections, each with its own goal: I showed Stone the items and she identified them as what I wanted to buy. And so on. What emerged, again, was a hierarchy of joint actions. It differed from the Haydn quartet in that its sections and subsections weren't fixed beforehand, but were negotiated as we went along.

One reason joint activities are complicated is two or more people must come to mutually believe that they are participating in the same joint activity. In the drugstore, it wasn't enough for *me* to believe I was in a business transaction with Stone. I had to believe *she* believed the same thing at the same time. Being in the same transaction is like believing we are in the same room at the same time. If I am in the room without Stone, or she without me, neither of us believes we are in the same room at the same time. But once we are both in the room – I can see her, and I can see that she sees me seeing her – we are now doing business together.

Joint activities therefore have boundaries. We can identify three stages of participants A and B with respect to joint activity J:

1. *Entry.* A and B go from not being in J to being in J.
2. *Body.* A and B are in J.
3. *Exit.* A and B go from being in J to not being in J.

In the drugstore, I believed at one point that Stone and I were just entering a business transaction, and at another point that we were just exiting from it. Stone had corresponding beliefs, though they may not have coincided exactly with mine. For me, and for her, the entry and exit defined the boundaries of our transaction.

Entries and exits have to be engineered for each joint action separately. That makes entries and exits especially important features of joint activities. In the drugstore, I tried to engineer Stone's and my entry into the business transaction by standing at the counter and catching her eye, but she put me off ("I'll be right there" "Okay") to go on with her inventory. Only once she began looking for my items did I think we had entered the transaction proper. Our exit was simpler. After paying, I took the bag she offered, turned, and walked away, assuming that she would take this as completing our transaction.

Since the sections of a joint activity are themselves joint activities or joint actions, they too each have an entry, body, and exit. In the drugstore, specifying the items I wanted to buy was a brief joint action: We entered it when I showed Stone the items ("These two things over here"), and we exited from it when she acknowledged them by picking them up. Then came identification of the price to be paid, the payment, and the transfer of goods, three other joint actions.

Joint activities don't always emerge as neatly as these examples suggest. Two joint activities can be simultaneous, as when Jack and Kate gossip while playing cards, canoeing, or cleaning the yard together. A single joint activity can also be intermittent, as when Jack and Kate, in the car, talk, lapse into silence, talk, and lapse into silence in cycles. Joint activities may also divide, as when a single conversation among four people breaks into two conversations (see, e.g., Sacks, Schegloff, and Jefferson, 1974). And they may expand and contract as new participants enter and old participants leave. The challenge is how to describe these dynamics.

Let us draw these observations together into some general claims about joint activities:

| | |
|-----------------------|---|
| Participants | A joint activity is carried out by two or more participants. |
| Activity roles | The participants in a joint activity assume public roles that help determine their division of labor. |
| Public goals | The participants in a joint activity try to establish and achieve joint public goals. |
| Private goals | The participants in a joint activity may try individually to achieve private goals. |

| | |
|--------------------|---|
| Hierarchies | A joint activity ordinarily emerges as a hierarchy of joint actions or joint activities. |
| Procedures | The participants in a joint activity may exploit both conventional and nonconventional procedures. |
| Boundaries | A successful joint activity has an entry and exit jointly engineered by the participants. |
| Dynamics | Joint activities may be simultaneous or intermittent, and may expand, contract, or divide in their personnel. |

These claims, of course, need fuller justification, and that will come as we proceed. For now we will take them as a place to start.

Advancement in joint activities

Joint activities advance one increment at a time. My transaction with Stone, for example, emerged in steps. To know what to charge me, she needed to know what I wanted to buy, so we established that first. To know what to pay, I needed to know what she was charging me, so we established that next. To know how much change to give me, she had to know how much money I was giving her, so we established that next. And so on. Each joint action added incrementally to reaching our public goals. What we did in pursuit of those goals depended on what we had done so far.

If joint activities are cumulative, what accumulates? I will argue that it is the common ground of the participants about that activity – the knowledge, beliefs, and suppositions they believe they share about the activity (Chapter 4). Although accumulation of common ground has been studied in discourse (Clark and Haviland, 1974, 1977; Clark and Marshall, 1978, 1981; Gazdar, 1979; Lewis, 1979; Stalnaker, 1978), it occurs in all joint activities.

ACCUMULATION IN DISCOURSE

When people take part in conversations, they bring with them certain prior knowledge, beliefs, assumptions, and other information. Part of this information Robert Stalnaker (1978) called their common ground:

Roughly speaking, the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of the conversation...Presuppositions are what is taken by the speaker to be the *common ground* of the participants in the conversation, what is treated as their *common knowledge* or *mutual knowledge*. (p. 320, Stalnaker's emphases)

The participants each have their own presuppositions about the conversation, but as Stalnaker argued, “It is part of the concept of presupposition that the speaker assumes that the members of his audience presuppose everything that he presupposes” (p. 321). They may, of course, be mistaken, but they realize this and have systematic strategies for resolving such discrepancies (see Chapters 7, 8, 9).

The common ground of the participants about their conversation changes as the conversation proceeds. As David Lewis (1979) put it:

Presuppositions can be created or destroyed in the course of a conversation. This change is rule-governed, at least up to a point. The presuppositions at time t' depend, in a way about which at least some general principles can be laid down, on the presuppositions at an earlier time t and on the course of the conversation (and nearby events) between t and t' . (p. 339)

These changes lead to increments to common ground. We can say that the common ground of the participants about the conversation *accumulates* in the course of that conversation.³

Assertions are prototypical linguistic actions for incrementing common ground. As Stalnaker argued, “the essential effect of an assertion is to change the presuppositions in the conversation by adding the content of what is asserted to what is presupposed. This effect is avoided only if the assertion itself is rejected.” At one point in the drugstore transaction, Stone and I presupposed that I didn’t know the total price of the items I was buying. She tried to change that presupposition by asserting that the price was \$12.77, “Twelve seventy-seven,” which I ratified by repeating “Twelve seventy-seven.” With the assertion completed, we added to our common ground the presupposition that the price was \$12.77. Other communicative acts – promises, questions, apologies, requests, declarations – increment common ground in other ways.

Accumulation of common ground occurs in all joint activities. To see how, let us begin with a rather formal joint activity, a game of chess. We can then return to the messier, spontaneous business transaction between Stone and me.

³ The common ground may also get restructured as new information accumulates.

STATES AND TRACES

In 1859 Paul Morphy and Adolph Anderssen, two master chess players of their era, met for a game of chess. They began with the chess board in its initial state SA_0 (SA stands for “state of the activity”). Playing white, Morphy made the first move, M_1 , displacing his king’s pawn (the pawn in front of his king) by two squares (to the square called “king 4”). In doing that, he changed the state of the activity from SA_0 to SA_1 . What Morphy did was *increment* the state of the activity, which I will write this way: $SA_0 + M_1 = SA_1$. Next it was Anderssen’s turn, and he displaced his queen’s pawn to the square called “queen 4.” He made the move, M_2 , not as an increment to the initial state of the game SA_0 , but as an increment to the state as it was after Morphy’s move, SA_1 . The result was SA_2 .

Morphy and Anderssen’s game was cumulative in this special sense. Each move, M_i , added an increment to the just prior state, SA_{i-1} , to produce a new state, SA_i . In symbols: $SA_{i-1} + M_i = SA_i$. The game accumulated this way:

| Time | Move | State of activity |
|------|-----------|-------------------|
| 0 | Open game | SA_0 |
| 1 | M_1 | SA_1 |
| 2 | M_2 | SA_2 |
| 3 | M_3 | SA_3 |
| ... | ... | ... |
| n | M_n | SA_n |

Put another way, each state SA_i was the cumulative result of the first i moves of the game, M_1 through M_i , whether made by Morphy or Anderssen. So the official course of the game for the first n moves can be represented by either of these two sequences:

States of the activity: $SA_1, SA_2, SA_3, \dots, SA_n$

Trace of the activity: $M_1, M_2, M_3, \dots, M_n$

The states of the game are represented by successive configurations of the chess board, and the trace, by moves that relate each two successive configurations. Given SA_0 , if you know one sequence, you can figure out the other.

Stone’s and my business transaction was cumulative in much the same way. Our transaction began with an initial state SA_0 . In approaching the counter, I made the first move, M_1 , catching her eye, adding an increment to SA_0 to produce SA_1 . She made the second move, M_2 , saying

"I'll be right there." Her move incremented not the initial state SA_0 , but the state of activity SA_1 after I had caught her eye. As in chess, each move M_i added an increment to the just prior state SA_{i-1} to produce a new state SA_i . In symbols: $SA_{i-1} + M_i = SA_i$.

The current state of Stone's and my transaction was incremented not just by what we did, but by each event we jointly recognized as advancing our joint activity. When Stone entered the two prices on the cash register, it rang a bell when it had the total. Once Stone and I jointly heard the bell, we mutually knew the total was available, so she could assume I would understand what she meant by "Twelve seventy-seven."⁴ So our business transaction accumulated this way:

| Time | Joint event | State of activity |
|------|------------------|-------------------|
| 0 | Open transaction | SA_0 |
| 1 | E_1 | SA_1 |
| 2 | E_2 | SA_2 |
| 3 | E_3 | SA_3 |
| ... | ... | ... |
| n | E_n | SA_n |

As with the chess game, we can represent the official course of the business transaction through the first n moves either (1) by the states of joint activity SA_1 through SA_n or (2) by a trace of the joint activity E_1 through E_n .

OFFICIAL AND UNOFFICIAL

In most joint activities, the states and events that become public – mutually known to the participants – divide into those that are officially part of the activity and those that aren't. The division is clear in chess. Chess moves M_i are official parts of the game because they are added to the official trace, alter the official board, and advance the game. In fact, a trace, or record, of the moves is all we formally need to know about a game; it is what is reproduced in chess books and studied by chess aficionados. Other actions may become public even though they play no official role in the game. When Morphy adjusted a piece on the board, moaned over a blunder, sipped water, or took a long time on a move, these were added to Morphy's and Anderssen's common ground, but as information outside the game proper. Chess players are careful

⁴ In craps, the count on the dice on any throw is outside the players' control, yet it is a joint event that advances the game.

to distinguish the official from the nonofficial at every point in their game.

Stone and I made much the same distinctions in the drugstore. We viewed some of the public events around us as part of the transaction proper, and others as outside it – as when she marked off her inventory, and I dropped several coins. Keeping track of which public states and events are official and which aren’t is essential to the orderly advancement of any joint activity.

TIME

Like any idealization, this model of joint activities is incomplete in many ways. One way is time. It assumes that events occur in sequence, and otherwise time doesn’t matter. But when actions are simultaneous or continuous, their interpretation can be affected by the passage of time.

Many actions in joint activities aren’t sequential, as the model suggests, but simultaneous or overlapping in time. In the drugstore, while I was saying “These two things over here,” Stone was following my gesture with her eyes, and I took her as having understood me when her eyes lit on the two items I was pointing at. If she hadn’t followed my gesture, or if her eyes had lit on the wrong objects, I would have continued “Over here” or “No, over here.” Indeed, simultaneity is the rule for one of the commonest linguistic actions in discourse – acknowledgments like North American “uh huh” and “yeah” and British “m.” These often overlap with what they acknowledge or they wouldn’t work as intended (see Chapter 8).

Also, many actions in joint activities aren’t discrete, bounded events, as the model suggests, but continuous actions in real time. Suppose in the drugstore I had wanted two items on a shelf behind Stone, and I had said, pointing, “Those two things [pause] right [pause] over a little more [pause] there,” timing “there” to coincide with her hand reaching the correct items. In referring to the two items, I exploited the continuity of her attention and hand movement and the continuity of my utterance. Demonstrative references with “this,” “that,” “here,” “now,” and “just” often depend crucially on the continuity of such actions (see Chapters 6, 8, 11).

Often, what is important to actions is timing and not just sequence as the model has it. Take the exchange:

| | |
|------------------------------|----------------------|
| Stone, meeting Clark's eyes: | I'll be right there. |
| Clark: | Okay. |

If I had produced “Okay” with no delay, I would have meant “I accept without conditions,” but if I had let one second pass first, I would have shown impatience and meant “I accept only reluctantly.” One of our goals must be to account for simultaneity, continuity, and the timing of events – to bring time into models of joint activity.

These observations give us a first picture of how joint activities work. The official course of a joint activity starts in an initial state SA_0 and the activity advances with each joint event E_i that adds to the previous state SA_{i-1} to form the current state SA_i . Other states and events become public but are not officially part of the joint activity proper.

Representations of common ground

What accumulates in a joint activity, I have argued, is the common ground of the participants. For most activities, the common ground at any moment divides into three parts:

1. *Initial common ground.* This is the set of background facts, assumptions, and beliefs the participants presupposed when they entered the joint activity.
2. *Current state of the joint activity.* This is what the participants presuppose to be the state of the activity at the moment.
3. *Public events so far.* These are the events the participants presuppose have occurred in public leading up to the current state.

It is worthwhile looking at these divisions, because they will help us understand a number of phenomena in language use. It is easier, however, to identify these divisions in chess than in other joint activities, so let us take up chess before returning to Stone’s and my transaction in the drugstore.

INITIAL COMMON GROUND

When Morphy and Anderssen entered their chess game, they each presupposed a vast amount of common ground. They presupposed the rules for chess – how the pieces move and capture, how the two players take turns, who begins, what constitutes a check, mate, and draw. They presupposed how to interpret the chess board and the pieces on it – who is attacking whom, what are the possible next moves. They presupposed the etiquette for chess playing – who sits where, when to keep silent, where to keep score. They presupposed a great deal about the strategies, tactics, and effective procedures in chess – e.g., opening gambits, deceptive moves, end game tactics. Having played each other before, they presupposed something about each other’s personal strategies,

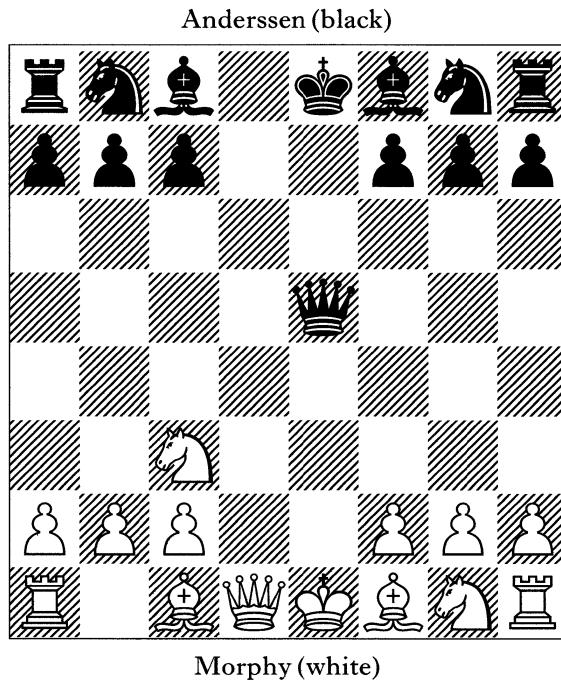
strengths, weaknesses, practices, habits. And so on. All this formed the starting context of their game – their *initial common ground* CG_0 .

Stone and I presupposed much the same type of information as the initial common ground CG_0 for our business transaction. We presupposed the standard procedures in American culture for exchanging money for goods – particularly in drugstores of this type. (Other businesses, even other types of drugstores, work differently.) These procedures have sometimes been called scripts or frames, and they specify what happens in standard situations of this cultural type (see, e.g., Minsky, 1975; Schank and Abelson, 1975). Among other things, they specify the roles and responsibilities of server and customer, actions for establishing the price, actions for exchanging money for goods, and actions for releasing the goods to the customer. Stone and I also presupposed that we both spoke English, that I had enough money, and other such things.

The point is this. People entering a joint activity presuppose a great deal about carrying out that activity. That information is represented in chess as rules, regulations, and etiquette. The analogous information is no less important in the drugstore even though it is represented in uncodified scripts or frames.

CURRENT STATE OF THE ACTIVITY

One part of the initial common ground of the participants in a joint activity is the initial state of the activity SA_0 . For Morphy and Anderssen's chess game, the initial state was represented by the chess board in its starting configuration. After ten moves, the current state was represented as shown in the illustration.

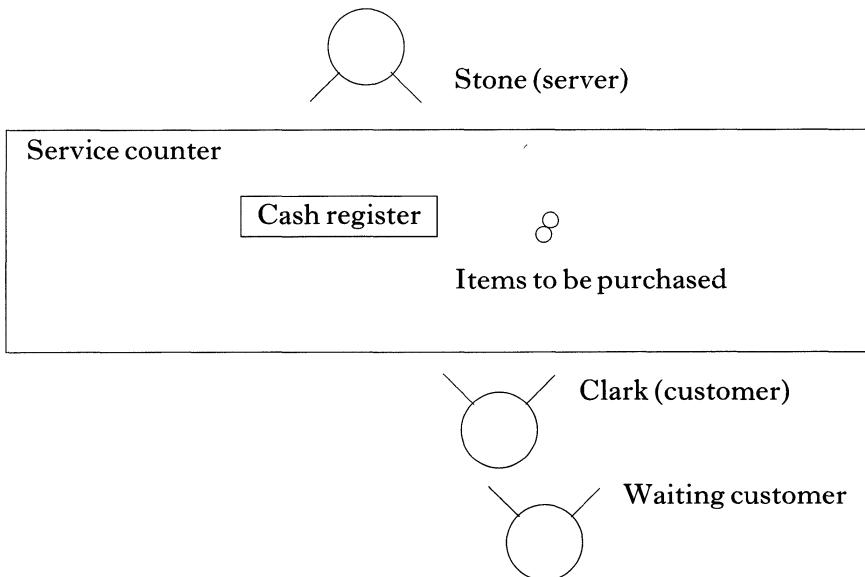


What is remarkable about chess – indeed about most games – is that the current state of the activity is represented in quite a concrete form. The chess board and its pieces are an *external representation* of the current state.⁵

External representations are particularly useful. Take Morphy and Anderssen's chess board after their tenth move. Morphy and Anderssen could see at a glance where each piece resided at that point and, by elimination, which pieces had already been captured. They could see that the black queen was threatening the white king – a “check” – and that there were several pieces Morphy could move to defend it. The chess board is the representation in which most rules of chess are stated – how each piece moves and captures, what constitutes checks and mates. So for Morphy and Anderssen, the chess board and chess pieces weren't mere patches of color and lumps of wood, but were elements of a scene they interpreted according to a highly developed understanding of the game.

Stone and I had an external representation of the current state of our business transaction too, and it was the scene around us. We entered the transaction with the scene in an initial configuration, but by the time I

⁵ I am indebted to Stuart Card for discussions about external representations.



had said “These two things over here,” the scene looked something like the bird’s-eye-view illustration. At a glance, Stone and I could see that she was the server and I the current customer, that the items I was referring to were those on the counter, that she had yet to ring them up on the cash register, and so on. The scene helped us keep track of where we were in our transaction. So for Stone and me, the scene wasn’t simply a set of brute objects. It was a scene we interpreted according to a highly developed understanding of how things work in such transactions.

External representations are more important to joint activities than is usually supposed. The chess board and drugstore scene illustrate some of their properties:

1. *Physical model.* The chess board and drugstore scene are physical models: They can be viewed, touched, and manipulated. Football fields, tennis courts, craps tables, courtrooms, classrooms, churches, and their contents are particularly useful because they too can be viewed, touched, manipulated.

2. *Markers.* External representations contain markers that denote elements of the joint activity. The squares on the chess board are markers for physical locations, and the chess pieces are markers for imaginary objects—such as kings and bishops—that can move, capture, and be captured. In the drugstore, the money and the receipt were markers for elements in the business transaction. Markers range from cards in card games and positions in queues to altars in churches and witness stands in courtrooms.

3. *Locational interpretation.* The markers are interpreted in part by their spatial location with respect to other markers. Anderssen's queen is interpreted as on his king 4 and in position to capture Morphy's king.

4. *Manipulability.* Some markers can be moved or altered, and the participants interpret these changes by the locations and forms that result. Move a wooden marker on a chess board, and you move the queen and change who she is threatening and is threatened by. Hand a twenty-dollar bill to a server, and you change who's in possession of the money.

5. *Simultaneous and parallel accessibility.* External representations are ordinarily accessible to all participants at the same time and in parallel. Morphy and Anderssen could study any part of their board simultaneously, and Stone and I could check out any part of the scene simultaneously.

It is hard to exaggerate the value of these representations. First, they are highly reliable representations of the current state of the activity. The chess board shows precisely where Anderssen's queen is, and because the board is simultaneously accessible to both players, they can both assume it to be part of their common ground. It is hard to dispute the position of a piece. This reliability is especially important in adversarial and business activities. Tennis, basketball, and football have scoreboards, gin rummy has counters, and business transactions have money, cash registers, and receipts, all to prevent disputes about the current state of the activity. And second, external representations are a particularly effective memory aid and medium for imagining moves. The chess board surely helped Morphy and Anderssen recall what they had just done and imagine what they should do next. The drugstore scene helped Stone and me recall our past actions and anticipate our future ones.

External representations are especially valuable as a medium for the actions themselves. Morphy and Anderssen played chess in part by displacing the pieces on the chess board. Stone and I transacted our business in part by manipulating the money, goods, cash register, receipts, and paper bags in the drugstore scene. Many joint activities would seem impossible with such representations.

PUBLIC EVENTS SO FAR

People also keep track of a third division of common ground: the public events since the beginning of their joint activity. In chess, the principal public events are the players' moves, as represented in a record of the game so far. Here, for example, is the official record of Morphy and Anderssen's first ten moves (in so-called descriptive notation):

| White (Morphy) | Black (Anderssen) |
|-----------------------|--------------------------|
| 1. P – K4 | P – Q4 |
| 2. P × P | Q × P |
| 3. Kt – QB3 | Q – QR4 |
| 4. P – Q4 | P – K4 |
| 5. P × P | Q × Pch |

Each line denotes a move by white (e.g., “P – K4” or “Pawn to king 4”) then one by black (e.g., “P – Q4” or “Pawn to queen 4”). Simple moves are denoted by “–”, captures by “X,” and checks by “ch.” The record will eventually end with a mate, draw, or resignation. As master players, Morphy and Anderssen presumably took much of this record as part of their common ground.

Morphy and Anderssen, however, presumably improved on this bare record to form what I will call an *annotated record*. (1) They almost certainly represented the purpose of each move as “attacking the knight,” “pinning the rook,” and “defending the queen.” (2) They probably also grouped moves into purposeful sequences. Certain opening moves, for example, are called the Ruy Lopez, *Gioco Piano*, and the Sicilian defense, and Morphy and Anderssen would have presupposed such interpretations. Master players divide games into such sequences in recalling them (Chase and Simon, 1973), and surely represent these groupings during games as well. (3) Morphy and Anderssen probably also interpreted certain moves as “blunders,” “bold moves,” “surprises.” (4) They may also have noted unusually delayed or fast moves as evidence of which moves were difficult and which were easy.

Morphy and Anderssen might also have abstracted away from the annotated trace to form what I will call an *outline record*. They might have represented the first ten moves this way: “We opened by exchanging pawns; Anderssen brought his queen out; we exchanged pawns again; and Anderssen’s queen put Morphy in check.” Details are left out, but the main thrust of what occurred is clear.

In the drugstore, Stone and I kept track of much the same type of record. At the point when Stone said “Seven twenty-five is your change,” she and I presupposed a sequence of public events so far: I had caught her attention; I had specified the items to be bought; she had rung them up; I had handed her a twenty-dollar bill and then two pennies; and she had computed the change. As in chess, we formed an annotated record. I hadn’t merely handed her a twenty-dollar bill, but had given

her money to pay for the items I was buying. And so on. How we annotated these events was determined by our expectations of what should have happened, on the frame or script for buying items in a drugstore (see Bower, Black, and Turner, 1979). Stone and I probably also abstracted away from many details to form an outline record.

DISCREPANCIES IN COMMON GROUND

As Stone and I talked, we each kept track of our own representations of common ground. She and I were, after all, individual agents with individual beliefs, judgments, and perceptions. Still, the very reason I kept track of our common ground was to have a representation I believed was identical to Stone's – at least to a certain degree of accuracy. I represented my beliefs about our common ground, and any piece of information I thought wasn't part of their common ground wasn't part of it. It is something like the two of us watching a tennis match and each keeping a log of the score, the net balls, the faults, and other such things. We expect our logs to be identical to a certain degree of accuracy.

Despite our best efforts, Stone's and my representations of our common ground were discrepant. Most of these discrepancies went undetected. Other times they might have become obvious to Stone or me or both. If I had detected a discrepancy, I would have had two main options. I could have brought it up and corrected it, or let it go and lived with the consequences. The discrepancy might be so slight that it wouldn't be worth my while to correct it. But I must keep track of every discrepancy I leave uncorrected. I must realize that Stone's representation of our common ground differs from mine in that one piece of information. For this reason it is often more efficient to correct a discrepancy immediately, and that is just what people tend to do (Chapters 8 and 9).

In brief, what the participants take to be common ground in a joint activity falls into three main parts – what they presupposed on entering the activity, the current state of the activity, and the public events that led up to the current state. Each of these parts divides further into the information that is officially part of the joint activity and the information that isn't. What is striking is how this common ground is represented. The current state of the activity, in particular, is often carried in an external representation, like a chess board or drugstore scene, that plays a central role in the course of that activity.

Discourse as a joint activity

What, then, is a discourse? I suggest it is simply a joint activity in which conventional language plays a prominent role. If it is, everything we have learned about joint activities should apply, and I will argue that it does. Can we in principle distinguish discourses from other types of joint activities? The answer, I believe, is no. All joint activities depend on signals or communicative acts – on language in its more general sense. If so, the distinction may be otiose, even misleading.

LANGUAGE IN DISCOURSE

Joint activities vary a great deal in how heavily they rely on conventional language. They lie on what I will call a *discourse continuum*, as illustrated here:

| | | |
|----------------------|---|---|
| Mostly linguistic | 1 | telephone conversations, newspaper items, radio reports, novels |
| | 2 | face-to-face conversations, tabloid items, television reports, science texts |
| | 3 | business transactions, plays, movies, coaching demonstrations, apprenticeship lessons, bridge games |
| | 4 | basketball games, tennis matches, two people moving furniture, making love |
| Mostly nonlinguistic | 5 | playing a string quartet, waltzing, playing catch |

In category 1, almost everything is done by means of conventional language. In the next category, much of what is done requires conventional language, but much also relies on gestures, pictures, video sequences, graphs, and diagrams, without which the language would be incomprehensible. In the middle category, the linguistic and nonlinguistic actions are more balanced and interdependent. In category 4, the focus is on physical actions, the conventional language being largely incidental. At the bottom end, finally, we arrive at joint activities that may take no conventional language at all.

If discourse is a distinct type of joint activity, where on this continuum do we draw the line? The items in categories 1 and 2 are clear examples of discourses, but so are most of the items in category 3. If we draw the line after category 3, how are we to treat the talk that does arise in categories 4 and 5? The yelling among players in a basketball game bears a relation to the ongoing activity even if it isn't essential to it (Levinson, 1992). Throughout the continuum, the conventional language used cannot be understood without viewing it against the joint activity it is part of.

To see this, consider one traditional approach to discourse – sometimes called text linguistics – in which discourses are treated as purely linguistic objects, as *texts*. We ordinarily think of texts as written records of what is uttered in conversations, speeches, or story tellings, but in this tradition they are more abstract. In Halliday and Hasan's (1976) approach, a text is any sequence of sentences that can be given a coherent interpretation. It doesn't matter how or why it was created. The important thing to explain is why it is well or ill formed, just as grammars try to explain why a sentence is grammatical or ungrammatical (van Dijk, 1972, 1977). In this tradition, texts are assumed to be complete in themselves, characterizable independently of surrounding events.

The notion of text, however, makes no sense across most of the discourse continuum. In all but category 1, the text of a joint activity is patently incomplete, no matter how we view language use. Take the text of Stone's and my transaction in the drugstore:

- Stone: I'll be right there.
- Clark: Okay.
- Clark: These two things over here.
- Stone: Twelve seventy-seven.
- Clark: Twelve seventy-seven.
- Clark: Let's see that's two pennies I've got two pennies.
- Stone: Yeah.
- Stone: Seven twenty-five is your change.
- Clark: Right.

To know what "I'll be right there" meant, we need to know that I had just caught her eye and was waiting to be served. To know what Stone's "Twelve seventy-seven" meant, we need to know that she had just rung up my two items on the cash register. To account for each line of text, we need to know where the participants were in the larger joint activity (see Morgan and Sellner, 1980). This holds for all texts in categories 2 through 5. Many of these lack textual coherence even though they are entirely coherent in the joint activity.

The fundamental issue is what to include in language use. If we include any signal – any communicative act – then language use is present across the entire discourse continuum. When Morphy moved his king's pawn, he was really communicating with Anderssen. Normally, chess players communicate by displacing chess pieces on a shared board for their opponents to see, but that isn't essential. In correspondence chess, the players send each other post cards with messages like "Pawn to king

4" and "Pawn captures pawn.," and in blindfold chess, they do much the same thing. Chess moves are a type of communicative act, and the game advances entirely by means of these acts. In this sense, a game of chess belongs in category 1 of the discourse continuum, even though it doesn't rely on a conventional language like English.

The official moves in the drugstore transaction were all communicative acts as well – though some were more than that. My catching Stone's eye was just as much a request for service as her "I'll be right there" was a promise to serve me. Other moves were composite actions. When I handed Stone a twenty-dollar bill, at one level we accomplished the joint physical act of moving the piece of paper from my hand to hers. At another level, we accomplished the joint act of changing the twenty dollars from my possession to hers – an act necessary for our transaction. To accomplish this, I had to communicate what I was doing – I might have intended Stone to hold the bill for a moment, to change it for two tens, or to check whether it was a forgery. By manifestly handing it to Stone at that state of our transaction, I was declaring that it was payment for the articles I was buying, and in taking my money, she accepted that declaration (see Chapters 5 and 6). Our transaction cannot be accounted for without including all communicative acts.

Discourses, then, are not a distinct type of joint activity, at least if we include all communicative acts, as we must in a full account of language use (Chapters 5 through 9). I will use the term when I want to emphasize the language being used.

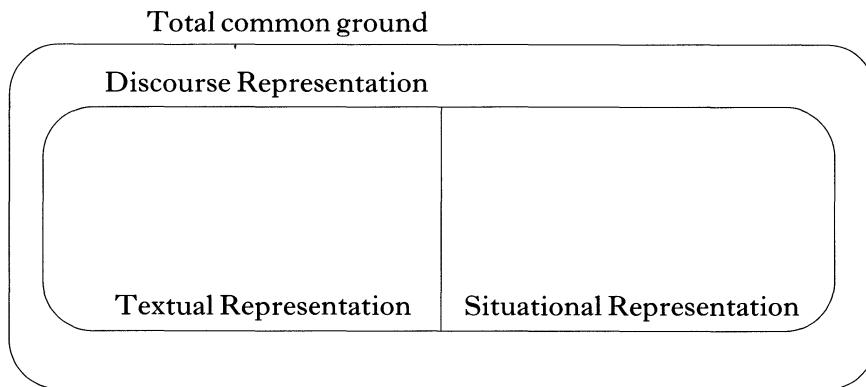
MODELS OF DISCOURSE

How do people represent the accumulating common ground in a discourse? If a discourse is a joint activity, they should represent (1) the initial common ground, (2) the current state of the activity, and (3) the public events so far. They should also distinguish between those public events that are official to the discourse and those that are not. Most theories of discourse have focused on categories (2) and (3) and then only on the official parts.⁶ I will follow this tradition as far as it goes, but add distinctions when needed.

The idea is that the participants in a discourse keep track of a *discourse representation*, which has two main parts. One part is the *textual*

⁶ For discourse models in this tradition, see van Dijk and Kintsch (1983) and Johnson-Laird (1983), among others.

representation, a representation of the language and other signals used during the discourse. The other is the *situational representation*, a representation of the situation being talked about. The picture is this:



The participants in a discourse need to keep track of the utterances and other signals they have used during the discourse, and they do this in a textual representation. Take the moment just after Stone and I had completed this exchange in the drugstore:

Stone: I'll be right there.

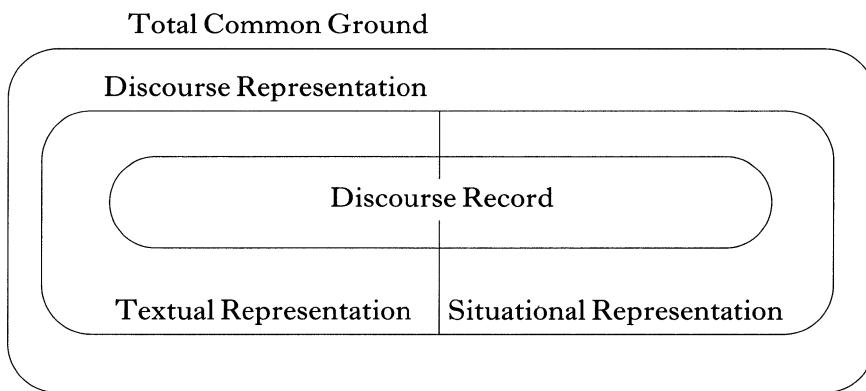
Clark: Okay.

If we understood each other completely, we each would have represented, as part of our common ground, at least these aspects of our signals: (1) the sounds produced; (2) the utterances issued; (3) the words, phrases, and sentences uttered, and their syntactic arrangement; (4) the meanings of the words, phrases, and sentences uttered; and (5) the two turns as constituting the exchange. These are things Stone and I had in our textual representations – but only for a brief time after our utterances. These include all aspects of the signals we needed for producing and understanding what we said.

The participants in a discourse also keep track of the situation they are talking about, and that they do in a situational representation. It represents such elements as: (1) the participants, time, place, and pertinent surroundings; (2) the referents of all expressions used by the participants (e.g., the referents of "I," "will be," "right," and "there"); (3) the social commitments established by what the participants said (e.g., Stone's promise and my acceptance); (4) the piece of the larger

transaction accomplished in the exchange. As we saw earlier, part of the situational representation takes the form of an external representation, like the scene in the drugstore, a type of information ignored in most situational representations. With these elements, Stone and I knew what was going on and could decide what to do next. In this view, situational representations represent what the participants have been doing, and textual representations, the communicative devices for taking those actions.

Part of the discourse representation has a privileged status, what I will call the *discourse record*. It represents the official states and events in the current joint activity. These are considered *on record*, as having advanced the joint activity. In chess, the discourse record would contain all of the chess moves, and in the drugstore, all of Stone's and my public actions that advanced our transaction. What is on record stands in contrast to those public states and events that are considered *off record* and not official parts of the joint action. With the addition of the discourse record, the picture now looks like this:



INCREMENTS TO DISCOURSE

People add to discourses mainly through communicative acts, or signals, especially linguistic utterances. Since our interest is in language use, we will want to understand how these work. As before, it is instructive to begin with chess.

A chess move is really a speech act called a declaration (see Chapter 5). When Morphy displaced the pawn on his first move in his game with Anderssen, it was as if he had said to Anderssen, "I hereby declare to you

that my king's pawn moves to my king four." In the notation developed in chess for these declarations, Morphy's move is expressed as "P – K4." Let us suppose that Morphy and Anderssen were playing correspondence chess, and that Morphy sent "P – K4" on a post card to Anderssen. In doing so, he presupposed that he and Anderssen shared a *vocabulary* in which P denoted "pawn," K "king," and – "moves to," and a *syntax* in which "P" denoted the subject, "–" the verb or relation, and "K4" the object. These are parts of a conventional language. In English translation, Morphy's statement was "Pawn moves to king 4."

But these elements aren't enough. Which pawn and whose king 4 was Morphy referring to? Morphy had to presuppose that he and Anderssen shared not only the language system, but two other pieces of knowledge. One was the current state of the game, SA_i. For "P – K4," the pieces were in their starting position, and it was Morphy's move. The second was the set of potential moves {M_{i+1}} at this juncture – e.g., all of Morphy's pawns could move one or two squares forward. Given SA_i, the pawn Morphy was referring to must be one that could move to K4. Given {M_{i+1}}, the only pawn in the right position was a king's pawn, and because it was Morphy's turn, it must have been Morphy's king's pawn, and K4 must have been Morphy's king 4.

A more telling example is "P × P," or "Pawn captures pawn." Although Morphy used "P × P" for both his second and his fifth moves, he meant quite different things on the two occasions. The first time he meant "My king's pawn (on my king 4) captures your queen's pawn (on your queen 4)," and the second time, "My queen's pawn (on my queen 4) captures your king's pawn (on your king 4)." What he meant depended on the current state of the game and the ways he could advance it. It depended on how it could increment the game.

Chess expressions like "P × P" illustrate an important property of communicative acts: *efficiency of expression*. In this notation, each expression includes no more information than is necessary to select the intended move out of all possible moves {M_{i+1}} given SA_i. In Morphy's second and fifth statements "P × P", the captured piece needed to be specified only as a pawn "P" because it was uniquely identifiable by that description. In Morphy's twenty second statement "R × KBP," the captured piece was also a pawn, but it had to be specified as the king's bishop's pawn "KBP" to distinguish it from other pawns that a rook R could have captured. The efficiency of expression, then, is this: Participants in a joint activity try to express no more than they need to

express for their joint efficiency in advancing the current purposes of that activity (see Chapter 5).⁷

Efficiency of expression applies just as readily to English. Here is one of Stone's and my exchanges in the drugstore:

Stone: Twelve seventy-seven.
 Clark: Twelve seventy-seven.

Here again we find the same expression with two different uses. Stone meant "The cost of the items you want to buy is \$12.77," and I meant "I am confirming that the amount is \$12.77." What Stone meant depended on the current state of our transaction (she had just entered the prices of the two items on the cash register) and the potential next moves. She expressed no more than she needed to. She could presuppose I expected her to specify the cost of the items at that point, and that it would be expressed in dollars. And I expressed no more than I needed to to confirm the precise dollar amount. Note that our utterances weren't sentences, but nominal phrases. If Stone and I were trying to be efficient, these are just what we should expect. Phrasal utterances are common in spontaneous talk.

APPROACHES TO LANGUAGE USE

Language use has been studied in two broad traditions. The *product tradition* grew out of the linguistic study of sentences, words, and speech sounds – the products of language use. It was strongly influenced by the work on generative grammars by Noam Chomsky and his colleagues. The *action tradition*, in contrast, grew out of the philosophical and sociological investigation of intentions and social actions.⁸

In the product tradition, sentences, words, and phonetic segments are treated as linguistic types abstracted away from speakers, times, places, and circumstances in which they might have been produced. Sentences have a syntactic structure; words have a phonological and morphological structure; segments have a phonetic structure; and words

⁷ For an extreme example, see von Savigny's (1983) discussion of headlight blinking on European roads.

⁸ One line of this tradition was developed by such philosophers as Austin (1962), Grice (1957, 1968, 1975, 1978), and Searle (1969, 1975a, 1975b, 1978, 1980), and another, by such sociologists as Goffman (1967, 1971, 1974, 1976, 1978, 1981a), Sacks and Schegloff (1979), Sacks et al. (1974), Schegloff (1968, 1972, 1979, 1982) and Jefferson (1972, 1973, 1978).

have conventional meanings that combine by certain rules of combination. The structure of these items determines only their *potential* uses. To specify an *actual* use, we have to fill in what is missing from “context.” For Stone’s utterance of the sentence *I’ll be right there*, we must note the circumstances – who was speaking, when she was speaking, what she was pointing at, etc. – and fill in the person referred to by *I*, the place referred to by *there*, and the time referred to by *will be* and *right*. We must also look to the “context” to decide whether the speaker was using the sentence to make an assertion, a threat, an apology, or what. The approach is product-centered: You start with the products – the sentences, words, and phonetic segments abstracted away from the circumstances – and fill in the missing content from the actual circumstances.

The product approach has several drawbacks. As we have seen, theories of language structure cannot be extended to cover discourse structure, so the approach has to be fitted out with an entirely different type of analysis to handle discourse. Theories of language structure deliberately abstract away from speakers, listeners, times, places, and the circumstances of utterance, so when these theories are extended to language use, the participants, times, places, and circumstances tend to be relegated to secondary roles and given short shrift. In this approach, sentences, words, and phonetic segments tend to be treated as static timeless objects, whereas utterances play out in real time and, as we have seen, their continuity, simultaneity, and timing count. Also, theories of language structure are concerned solely with conventional languages like English, Japanese, or American Sign Language, so the communicative acts in discourse that are *not* part of conventional languages – eye gaze, gestures, nods, smiles, and manifest actions such as my handing Stone a twenty-dollar bill – are excluded on principle and then ignored.

Perhaps the greatest drawback of the product approach is its attitude toward “context.” In logic, the object of study is well-formed formulas, such as “*p* implies not *q*,” and the rules by which they can be used to make inferences. These theories are exclusively syntactic: They don’t depend on the referents of *p* and *q*. When logics like this were taken as models for sentence meaning and, later, language use, it was hard to shake the attitude that the referents of utterances were of only secondary interest. The stricture seemed to be that theories of language use shouldn’t appeal to “context” until they were forced to. One result is that there has been little investigation within the product tradition of the “context” that does get appealed to.

In the action tradition, the focus from the beginning has been on what

people do with language. Discourse isn't merely a linguistic structure. Speakers, listeners, times, places, and the circumstances of utterance are taken into account, and in at least part of the tradition, so are the continuity, simultaneity, and timing of utterances and other actions. Attention is paid to the gamut of communicative acts, from utterances to eye gaze. "Context" is generally given the prominence it deserves and is not treated as the refuge of last resort.

The difference between the two approaches may at first seem slight, but it is fundamental, for it leads to radically different theories of language use. Although we must appeal to results from both traditions, it is the action tradition that will set us off in the right direction.

Conclusions

When people use language, it is generally as part of a joint activity. Now, joint activities range from planning a party or transacting business to playing chess or playing in a string quartet, and they have properties all their own. They take the coordinated actions of two or more participants in particular roles. They each have an entry and an exit, and most emerge in sections and subsections. Most establish a dominant goal, and the participants advance toward that goal one increment at a time. Each of these increments adds to the common ground of the participants, changing what they take to be the current state of the activity.

The argument is that joint activities are the basic category, and what are called discourses are simply joint activities in which conventional language plays a prominent role. If we take language use to include such communicative acts as eye gaze, iconic gestures, pointing, smiles, and head nods – and we must – then all joint activities rely on language use. Chess may appear to be nonlinguistic, but every chess move is really a communicative act, and every chess game a discourse.

Joint activities advance largely through identifiable joint actions by the participants. What are joint actions, and how do they work? These are questions for Chapter 3.