

Lesson 13

Reading Material

7 Sequence and structure: Conversation Analysis

In this chapter I outline the approach to spoken discourse known as Conversation Analysis (commonly abbreviated to CA). I use the word 'outline' advisedly: CA has produced a voluminous literature, and there is far more to say about it than can be accommodated in a single chapter. Here I set out to explain what makes the CA approach distinctive, but I also focus particularly on those areas of its work which have been influential across the whole spectrum of approaches to spoken discourse.

In spite of its name, Conversation Analysis does not deal only with conversation: the approach has also been applied to talk in professional and workplace settings (Drew and Heritage 1992), to political speeches (Atkinson 1984) and to media genres such as radio phone-in programmes (Hutchby 1996). Many practitioners label their object of study not 'conversation' but 'talk-in-interaction'. This label is apt, since although CA can be used to analyse many different kinds of data, it was developed to analyse *talk* (rather than written text) and more specifically the kind of talk that is thoroughly *interactive* (as opposed to monologic sequences, such as narratives¹). The fact that talking is prototypically a joint enterprise involving more than one person, and that people normally take turns at talk, is central to the CA approach, which is concerned above all to describe sequential patterns (regularities in what follows what) which are observable in the data being analysed. Claims about what is going on in any particular bit of talk are often supported by 'going to the next turn', that is, appealing to the evidence of what happens in the following bit of talk. Speech act theorists may define an utterance as, say, a 'question' in terms of whether it meets the felicity conditions for the relevant illocutionary act (see Chapter 6); practitioners of CA would be more interested in whether it is followed by an answer. It is in answering, CA practitioners point out, that the participants display their understanding of the previous contribution as a question, and their 'orientation' to the general principle that in properly ordered talk, questions will be followed by answers.

CA is a markedly 'data centred' form of discourse analysis: in 'pure' versions (it will become clearer later what I mean by 'pure' here), the analyst is not supposed to appeal to any evidence that comes from outside the talk itself. This makes doing CA a rather different proposition from doing, say, ethnography of speaking. As we saw in Chapter 5, the ethnographic approach relates verbal behaviour to the setting in which talk occurs, who the participants are, what languages or language varieties their repertoire includes, and what they believe about all kinds of matters, ranging from child development to the mysteries of salvation. Such 'external' considerations

are taken to be potentially relevant in analysis whether or not anyone alludes to them directly in the piece of talk being analysed. In the extract from a Pentecostal meeting reproduced in Chapter 5, no one explicitly says that contributions are prompted by the Holy Spirit. The investigator Christine Callender discovered the significance of this belief by talking to community members outside the meeting, and she used that knowledge to illuminate certain features of the talk that occurred inside the meeting. CA, by contrast, does not require the analyst to gain detailed knowledge of participants' identities, their daily routines or their beliefs: instead their talk itself is treated as containing everything relevant for analysis.

Some CA practitioners (who adhere to what I am calling the 'pure' version of the approach) apply the principle that one should not look beyond the data very strictly. For them it is not just unnecessary but illegitimate for an analyst to make use of information that the participants themselves have not chosen to 'make relevant' in their ongoing interaction. This position reflects the origins of CA in a particular school of sociological thought, 'ethnomethodology'. Ethnomethodology developed in opposition to other sociological approaches (for instance structural-functionalism and marxism) which explain society – and the habitual behaviour of its members – in terms of abstractions such as 'the class system'. Ethnomethodologists reject this view: for them social order is created not by abstract structures, but by the concrete actions of people going about their everyday business. It follows that sociologists should concentrate on studying people's actions on their own terms, rather than trying to fit them into an abstract theoretical framework which may have no relevance for the actors themselves (in which case it can hardly constitute a satisfactory explanation of their actions).

Ethnomethodology raises complex theoretical and philosophical issues which I will not pursue here (but see the suggestions for further reading at the end of the chapter). In CA more specifically, the 'purist' stance which shows the influence of ethnomethodology has been the subject of heated debate, much of which revolves around questions of *power* in discourse. As Ian Hutchby notes, the position of some highly respected CA practitioners is that 'analysts should only have recourse to concepts such as power once they are forced to do so by virtue of some otherwise unexplainable interactional phenomenon' (Hutchby 1996: 482). Others, including people who use the techniques of CA as well as people who prefer other approaches, take a different view. For instance, some feminist practitioners of CA have argued that men tend to dominate conversations involving both sexes (Fishman 1983; Zimmerman and West 1975). From a feminist perspective this is a manifestation in discourse of the more general social fact of gender hierarchy: it is therefore relevant to refer to that social fact in an analysis of discourse data. To the CA 'purist', however, gender and gender hierarchy are only relevant to the analysis of a piece of data if the participants *make* it relevant in some way. If it is not explicitly an issue for the participants in their talk, the analyst has no warrant for making it an issue in the analysis of that talk.

In Chapter 11 I return to questions about power and inequality. In this chapter, though, I concentrate on the apparatus CA provides for describing spoken discourse empirically. Though CA's insistence on focusing on the data and nothing

but the data is considered by some to have its drawbacks, it is also widely acknowledged to have certain advantages: if you cannot make reference to what is *not* in the data, you are impelled to pay very close attention to the fine details of what *is* there. CA is a 'microanalytic' approach, which takes apparently mundane and unremarkable spoken interactions and finds intricate patterning in the way they are organized. Just as putting a snowflake under a microscope reveals structure and complexity which are not visible to the naked eye, so putting talk under the CA microscope defamiliarizes what we normally take for granted, and reveals the unsuspected complexity of our everyday verbal behaviour.

INVESTIGATING THE OBVIOUS: THE ORGANIZATION OF TURN-TAKING

Several years before I encountered CA as a student, I came across a reference to it in the British satirical weekly magazine *Private Eye*, which has for many years run a feature called 'Pseud's Corner'. Items for inclusion in Pseud's Corner are sent in by readers, and are usually examples of artists or intellectuals stating the obvious in unnecessarily pretentious language. On this occasion, a reader had sent in an observation attributed to the pioneer of CA, Harvey Sacks, about turn-taking in conversation. The observation was that 'one speaker speaks at a time, and speaker change recurs'. What readers of *Private Eye* were being invited to ridicule was the presentation of such a banal observation as the fruit of some weighty intellectual endeavour. The implication was that the proverbial person in the street could have come up with this particular insight on the basis of no research whatever, except that ordinary people would have more sense than to find the issue worth thinking about. But in a sense that is exactly the reason why Sacks *did* find it worth thinking about. Things that appear obvious – so obvious, indeed, that in everyday life no one would ever remark on them – may turn out on closer inspection to be less obvious than they seem.

Here the observation is that conversation requires speakers to take turns, and this requirement is managed in a particular way. At any given moment, the turn that is in progress will typically 'belong' to a single speaker ('one speaker speaks at a time'). Participants in conversation will not usually all talk at once, and conversely there will not usually be stretches of time in which no one talks at all. That is not to say that simultaneous speech and silence never occur in conversations: they do. But when they occur they are often treated by participants as problems which need to be 'repaired' – in other words, as something other than the normal and desirable state of affairs. In the case of simultaneous speech, what typically happens is that one speaker wins the floor (the right to talk and be attended to by co-participants) while the other(s) fall silent. In the case of a silence that becomes long enough to feel awkward, what usually happens is that someone breaks into it and claims the floor.

In addition to noticing the 'one speaker at a time [no more and no fewer]' requirement, Sacks pointed out that 'speaker change recurs'. In other words, the floor is not given to any one speaker for the duration. Nor, in ordinary conversation, is

it parcelled out in advance among the participants ('OK, you speak first, then I'll speak, then Joanne can have a turn'). Rather the floor is constantly negotiated and renegotiated as a conversation goes along. In fact this continual negotiation is a general feature of conversational organization. CA holds that talk is 'locally managed', meaning that its patterns and structures result from what people do as they go along rather than from their being compelled to follow a course of action that has been determined in advance.

The sceptical reader may still be asking, with the editors of *Private Eye*, 'OK, but so what?' There does not seem to be much of a puzzle about *why* the regularities noted by Sacks exist. If everyone spoke at once, or no one spoke at all, or the same person held the floor continuously throughout, what we would have would not be conversation, it would be chaos or silence or monologue. But the intriguing question is not '*why?*', it is '*how?*' How do conversationalists manage the organization of the floor in such a way as to produce the pattern that 'one speaker speaks at a time and speaker change recurs'? Certainly it might seem logical to suppose that conversationalists would 'naturally' strive towards a pattern which enables conversation to proceed in an orderly, rather than chaotic, manner. But even if that is so (a qualification I make because later I will suggest that speakers or communities may vary in their ideas about what kinds of talk *are* orderly/chaotic), on its own it does not explain *how* they achieve the desired outcome. Nor can participants in conversation necessarily explain this themselves. Asking them questions like 'how do you know when it's your turn to speak?' is unlikely to elicit an illuminating response, because ordinary conversationalists take their ability to converse for granted – they do not have to think consciously about what they are doing when they interact. Making explicit what ordinary conversationalists take for granted is precisely what CA sets out to do.

So what, in fact, are conversationalists doing to produce the regularities noted by Harvey Sacks? Sacks, along with two colleagues, Emmanuel Schegloff and Gail Jefferson, proposed a model of conversationalists' behaviour which they presented under the heading of 'A simplest systematics for the organization of turn-taking for conversation' (Sacks et al. 1974). The model has two main elements. First, it says that speakers are aware that a turn consists of one or more (but not fewer) 'turn constructional units'. (Sacks et al. define these units as grammatical entities, like a complete clause or sentence; but it might also be suggested, for reasons explained in Chapter 2 above, that the units of spoken discourse are delineated primarily by prosody (intonation, stress, pausing) rather than grammar.²) People who are listening to someone else's speech can use their knowledge of the possible unit types to project the end-point of the turn currently in progress. Being able to do this is important, because the end of a turn constructional unit is potentially a 'turn transition relevance place', a point at which speaker change may occur (though as we will see, it does not have to). Projecting the end of a turn involves attending to a combination of things, including the content of what is said, the prosodic and grammatical structure of the speech, and aspects of nonverbal behaviour such as the direction of the speaker's gaze. For example, if someone utters the words, 'and that's what I did today', with falling intonation and a pause at the end, listeners are likely to take that as a possible turn transition relevance place. By contrast, if someone utters the words 'and she

says to me?' (the ? denotes rising intonation), both the content and the prosody provide a clue that the clause boundary in this case does not mark the end of the turn: the speaker intends to continue with a report of what 'she' went on to say, and is entitled to continue holding the floor at least until that report has been completed.

The second element of the simplest systematics model is a mechanism for allocating turns to particular participants in a conversation. When a turn transition relevance place is reached, what ensues is not a random free-for-all, with everyone present having an equal chance of getting the floor next. Instead, Sacks et al. suggest, there is an ordered set of rules for the allocation of the next turn, which can be shown in summary like this:

- 1 Current speaker selects next speaker
or if this mechanism does not operate, then . . .
- 2 Next speaker self-selects
or if this mechanism does not operate, then . . .
- 3 Current speaker may (but does not have to) continue

The turn-allocation mechanism that takes precedence over the alternatives is for the current speaker to select the next speaker. If a speaker ends her turn by shifting her gaze to the person on her left and saying, 'and that's just like what happened to you last year, isn't it?' then the next turn is marked as belonging to whoever she has just addressed. If that person remains silent, it is an accountable silence – not just anybody's silence but specifically the silence of the person selected to speak next. There are many ways for a current speaker to select a next speaker, including asking them a question, naming them (as often happens in the classroom – 'Nahil, what do you think?' – and in broadcast talk – 'and now over to you, Bill') or aligning your body or gaze so you are seen to be addressing a specific co-conversationalist at the end of your own turn.

If the current speaker does not select a next speaker, the second option is for someone other than the current speaker to select themselves, by starting to speak. When this option is activated, there is scope for simultaneous speech to occur because more than one speaker simultaneously self-selects; but as noted earlier, the normal pattern is for this situation to resolve itself, with only one of the self-selectors continuing to hold the floor. If, on the other hand, no one self-selects, it is open to the current speaker to continue. This is the third option offered by the simplest systematics model, and it ranks below the alternatives, i.e. it is only permitted if neither of the higher-ranked options has been chosen. At the next turn transition relevance place, the same options apply in the same order all over again.

The simplest systematics is, indeed, simple, and elegant in its simplicity, but how well does it account for what is actually observed in real conversations? Reading the account I have just given, you may well have thought of objections based on your

own knowledge and experience of talk. For example, turn transitions are not always as smooth as my account might seem to suggest. I mentioned only one reason why we sometimes get stretches of talk in which two or more people speak simultaneously: the possibility of more than one speaker self-selecting at a turn transition relevance place where the current speaker has not selected a next speaker. But what about the fact that people often begin speaking before the previous speaker has finished rather than waiting for the end of the turn? Are they breaking the rules?

Some cases of simultaneous speech are classed as 'overlaps': they result from the new speaker's failure to project the end of the last speaker's turn with complete accuracy. The new speaker comes in at the point where s/he *thinks* the last speaker will finish, but either the timing is slightly 'off' or else the new speaker is mistaken in thinking the last speaker is close to completion of a turn. Overlaps of this kind are common, but they are typically short, and they start close to a potential turn transition relevance place. They are not instances of rule-breaking so much as unintended errors in the application of the mechanisms described by Sacks and colleagues. In other cases, however, a new speaker may start to speak at a point in the last speaker's utterance that cannot possibly be a turn transition relevance place. This is not overlap but interruption, and for some analysts (e.g. Zimmerman and West 1975) it is not just a violation of the turn-taking system but a hostile act designed to deny the current speaker their legitimate right to the floor.

For other analysts, however, the idea that simultaneous speech is either overlap (an honest 'mistake') or interruption (a hostile 'violation') is oversimplified, and does not account for some quite common instances of simultaneous speech. To illustrate the problem, let us revisit a piece of data I used in Chapter 3, involving two women friends (they are university students, and are talking about two other women students they both know).

- Ali: and she didn't she didn't like Katie she didn't ge[t on with Katie at all]
 Beth: [no she didn't get on with] Katie
 [brackets mark sequences spoken simultaneously]

In this extract, Beth begins her turn at a point where Ali is right in the middle of an informational, a grammatical and (although to keep things simple I haven't shown it in this transcript), a prosodic unit. The word *get* in the phrase *didn't get on with* is unlikely to be taken to mark the end of any kind of unit, which makes it implausible to suggest that Beth is starting her turn on the assumption that Ali is about to finish hers. (A more likely point for that to happen would have been somewhere around the first utterance of *Katie*.) Is Beth, then, *interrupting* Ali? Two considerations make this unlikely, especially if interruption is understood as a hostile act. First, Beth does not deprive Ali of the floor: Ali does not fall silent (whereas in their study of gender and interruptions, Zimmerman and West (1975) found that the interventions they classed as interruptions often had the effect of making the interrupted party stop speaking). Ali continues speaking simultaneously with Beth until she finishes the unit

she has started (that it is indeed complete at this point is signalled by a marked slowing down with emphatic stress on *at all*). Second, Beth's 'interruption' does not mark disagreement with what Ali has said, nor lack of interest in what she has said, nor any desire to change the direction of the conversation. On the contrary, it echoes both Ali's sentiments and Ali's actual words. At the point Beth starts speaking, just after Ali has begun producing the word *get*, she appears to have heard enough to guess accurately what Ali is in the middle of saying – 'she didn't get on with Katie' – and she interjects in order to say exactly the same thing. The function of Beth's contribution, then, is to *support* Ali in holding the floor rather than to take the floor away from her.

Jennifer Coates (1996) argues that simultaneous speech with this supportive function is particularly common in the talk of women friends; I have found examples of it in the talk of young men friends too (Cameron 1997). Carole Edelsky (1981) provides evidence that supportive simultaneous speech can also occur in institutional contexts. In analysis of talk at meetings, Edelsky distinguished two types of floor, which she labelled 'F1' and 'F2'. F1 conformed to the 'one speaker speaks at a time' model, whereas F2 involved more simultaneous speech (and according to Edelsky, enabled women to share the floor more equally with men than was the case with F1). It appears, then, that in both friends' talk and some kinds of institutional talk it is possible for the floor to be organized on a principle other than 'one speaker speaks at a time'. But the simplest systematics model assumes that 'one at a time' is both normal and fundamental: there is no obvious place in the model for simultaneous speech which is neither an error nor a violation, but merely a normal feature of certain kinds of talk.

The question this raises is whether Sacks and colleagues make assumptions about talk-in-general which are not, in fact, universally valid. In CA, analysts look not merely for regular sequential patterns in data, but for evidence that participants themselves are *orienting* to the existence of those patterns. Thus if the analyst's claim is that 'one speaker speaks at a time', one would expect participants in talk to display their orientation to that pattern by treating instances of simultaneous speech as problems requiring repair. This is what is happening when, say, a speaker who inadvertently overlaps another apologizes and stops speaking to let the other finish, or when three speakers who simultaneously self-select sort themselves out in such a way that only one continues. But in something like Carole Edelsky's 'F2' sequences, or in the conversation between Ali and Beth reproduced above, there is no display of orientation to the 'one speaker speaks at a time' pattern, and this is what motivates speculation that some other system of floor organization may be operative.

Some readers of my account of the simplest systematics – especially, perhaps, readers who do not live in an exclusively English-speaking society – may have wondered about the issue of *cultural* differences in relation to the organization of conversational floors. Ethnographers have reported on various societies which do not appear to fit the prototype assumed by Sacks et al., ranging from the Antiguan village in which Reisman (1974) observed 'contrapuntal conversation' (i.e. at least two speakers habitually speak at a time) to the Nordic and native American speech communities which tolerate extended silence in conversation (see Tannen and

Saville-Troike 1985). Anyone who has been exposed to a variety of cultural practices will be able to produce anecdotes about other differences. When I was teaching in Sweden, for example, an Argentinian student explained how hard it had been for her to learn Swedish turn-taking conventions. In Argentina, she said, the rule had been roughly 'grab the floor and then talk until someone interrupts you'; in Sweden the rule appeared to be something more like 'wait for the other to finish before you start speaking'. As a result, the Argentinian had often found herself in the midst of what struck her as an interminable monologue, desperately wishing for someone else to claim the floor from her. Meanwhile, as she later realized, her Swedish interlocutors were waiting for her to show some sign of finishing so they could take a turn. Both conventions basically conform to the 'one speaker at a time' model; the difference is subtler, to do with how and when the turn-allocating mechanism of self-selection may be used.

The existence of variations like these underlines the need for analysts of spoken discourse to be cautious about making quasi-universal claims. But that is not a criticism of the CA approach as such. Rather it restates what is actually a fundamental principle of CA – that analysts must be attentive, and analysis faithful, to the fine-grained detail of the particular data being analysed, and to the patterns that appear to be salient for participants themselves.

'ONE THING AFTER ANOTHER': ADJACENCY PAIRS

The turn-taking system provides a basic framework for the organization of talk-in-interaction, since it allows for the floor to be alternated systematically between speakers – in other words, for participants to *interact* rather than simply acting individually in an uncoordinated manner. But if the orderly exchange of speaking turns is a necessary part of spoken interaction, it is not on its own sufficient to guarantee that there is genuine and meaningful interaction, as we can see from the following example (quoted by Labov 1972a: 299):

Therapist: what's your name?

Patient: well, let's say you might have thought you had something from before, but you haven't got it anymore

Therapist: I'm going to call you Dean

In this exchange, the parties seem to have no problem coordinating the orderly exchange of turns. What makes the exchange odd is rather the apparent lack of co-ordination in the *content* of their turns. The therapist asks a question, but although the patient responds as expected in the sense that he unhesitatingly takes the floor, he fails to provide anything like a meaningful answer to the question he has been asked. A Gricean pragmaticist might say that the patient is flouting the maxim of relation. A practitioner of CA would not disagree with the assessment of the remark

as irrelevant, but s/he would consider it important to clarify that the patient's response is irrelevant, specifically, *as a response to the immediately preceding contribution*. CA places great emphasis on the idea that conversation is 'one thing after another': it is an activity that unfolds in time, and what I say now must inevitably constrain what you can meaningfully say next. In the above example, the first turn is a question requesting information. The second speaker is thus constrained to provide the information requested – not in a minute or an hour or a year, but in the next turn.

Here it may be objected that questions are not always followed immediately by answers; sometimes, indeed, they may be followed by further questions (there is a parlour game based on this possibility, in which the players must speak entirely in questions – for instance 'how are you?', 'why do you ask?', 'what's the matter, are you paranoid?', 'what makes you think I'm paranoid?' etc.). But if one question is followed by another question, or indeed by an irrelevant remark like the patient's in the 'what's your name?' example, normally that only defers and does not cancel the expectation that the original question must be answered. The strength of that expectation is underlined, in fact, by what happens in cases where it is *not* met.

In a study called 'On questions following questions in service encounters', Merritt (1976) looked at exchanges like the following:

- 1 Customer: pint of Guinness please
 - 2 Bartender: how old are you?
 - 3 Customer: twenty two
 - 4 Bartender: OK coming up
- [constructed example]

Merritt points out that the implication of responding to a request (turn 1) with a question (turn 2) rather than simply by supplying what has been requested is that the server's ability or willingness to provide the desired service depends on the answer to the question. In the above example, turn 2 conveys to the customer, not that the server is ignoring his request in favour of idle questions about his date of birth, but that the provision or non-provision of beer is *conditional* on the answer to the question 'how old are you?'. The question is taken as relevant by the customer, not only because of his background knowledge that alcohol can only be purchased by persons over a certain age, but also and importantly because of its position in the sequence of turns: it occurs immediately after a request for service, but before the server has fulfilled his obligation to provide that service. Hence what happens in turn 3 is that the customer fulfils his own obligation to provide an answer to the question posed in turn 2; the server then responds in turn 4 to the original request made in turn 1.

In the 'what's your name?' sequence, by contrast, when the therapist does not get a straight answer to his first question he uses his next turn to effectively answer it himself. That suggests he has some reason to believe the patient is not going to honour his obligation to provide an answer to the question. (In fact, he knows that the patient suffers from schizophrenia and for that reason often does not conform to 'normal'

expectations in spoken interaction. He also knows the patient's name. Even so, he apparently is not content to let the question simply remain unanswered.)

The point that is illustrated by the examples given above is that spoken interaction is often structured around pairs of adjacent utterances (utterances which occur one after the other) in which the second utterance is not just related to the first but functionally dependent on it. If the first utterance is a question, the next utterance will usually be heard either as an answer to the question or, as in the Guinness example, as a move that has to be made in order to put the speaker in a position to answer the question subsequently. The sequence 'Question-Answer' is one example of what CA labels an 'adjacency pair', meaning a pair of utterances in which the second is functionally dependent on the first. Merritt's study of questions following questions shows that the two parts of an adjacency pair may become separated (and thus no longer adjacent). This separation is normally brief, however, and the second adjacency pair which is inserted between the two parts of the first one is understood to be necessary for the completion of the original transaction (as opposed to being some sort of digression from it).

Question-Answer is not the only type of adjacency pair identified by analysts. Another type is Greeting-Greeting. If someone produces an utterance that is hearable as a greeting, the person they are addressing is constrained to produce a greeting in return. They have a choice about what they say (or do: a nod, smile or raised eyebrow may be sufficient acknowledgement). But if the first part of the greeting-greeting sequence is not followed by a second part, this will be perceived as leaving an accountably empty slot – or in plainer English, as a snub. The fact that we perceive the non-return of greetings as rude, unfriendly or socially inept behaviour provides supporting evidence for the claim that a greeting is not a free-standing turn but part of a larger unit, an adjacency pair. If there is no second part we feel the return greeting is not merely absent but missing.

Some first turns present the producer of the second turn with a choice. For instance, invitations, offers, suggestions or proposals may be either accepted or declined. A turn which solicits the addressee's opinion on some proposition made by the speaker may be met with either agreement or disagreement. Analysts have noted an interesting pattern in the way the alternative possibilities are typically handled by conversationalists. Consider the following example, in which three friends are negotiating what they will eat for dinner.

Transcription conventions: ? = rising intonation; (.) = pause

- 1 Daphne: I was thinking we could have fish?
- 2 Julia: fine
- 3 Anita: well actually (.) I've stopped eating fish now because of you know
- 4 the damage it does to the ocean

In this sequence Daphne, who will be the cook on the occasion under discussion, makes the others a proposal regarding the menu: she suggests having fish. Julia accepts

this proposal; Anita by contrast declines it. Whereas Julia's acceptance is brief and unelaborated, Anita's turn is considerably longer. It begins with *well*, a 'discourse marker' (Schiffrin 1987) which is commonly used to indicate that the speaker is about to say something which may conflict with the assumptions or wishes of a previous speaker. Anita also pauses briefly before she gets to the actual message, a polite rejection of Daphne's proposal that they should have fish for dinner, 'I've stopped eating fish now'. Finally, Anita provides an account of why she no longer eats fish: she has been persuaded to stop by ecological arguments about the damage intensive commercial fishing does to the ocean environment.

The difference between Julia's reply and Anita's illustrates the workings of what CA calls a 'preference system'. The 'preferred' response to a proposal is acceptance, and it can be identified as the preferred option because it is typically performed without hesitation or elaboration. It is possible to accept proposals more elaborately, of course – Julia could have said something like 'ooh, fish! I love fish!' – but the point is, acceptance *can* be done without elaboration. Refusal by contrast is almost never done in the same way. Refusal is the 'dispreferred' response, and is identified as such because it is typically performed as Anita performs it here: she is hesitant about coming to the point and she elaborates her refusal by giving a reason for it. A similar pairing is agreement (preferred) and disagreement (dispreferred).

In discussions of the concept 'preference' in CA, analysts have often been at some pains to emphasize that the 'preferred'/'dispreferred' distinction is made purely on the basis of formal patterns: preferred responses are prompt and short while dispreferred ones are hesitant and elaborate. The word 'preference' is not intended, these analysts warn, to have its ordinary English meaning of something that is considered more desirable than something else. People's attitudes regarding what is or is not desirable are the province of psychology, and irrelevant to the concerns of CA. CA has simply noted the existence of a regular pattern in talk: acceptances and agreements are typically done in one way, while refusals and disagreements are done in a different way.

But one might ask whether this line of argument is either convincing or enlightening. There are, after all, some fairly obvious reasons why people might generally 'prefer', in the ordinary sense of the word, to accept invitations, offers and proposals rather than refusing them, or to agree with others' opinions rather than disagreeing. Readers who recall the discussion of politeness in Chapter 6 will probably have noticed that the 'preferred' option is also the less threatening to the addressee's face. In the 'fish' example, for instance, Julia can make her acceptance without mitigation, because acceptance does not inconvenience Daphne or hurt her feelings. Anita on the other hand is doing something that both imposes on Daphne (she must think of something else to have for dinner) and hints at disagreement with or even disapproval of Daphne (Anita thinks commercial fishing and the consumption of its products are morally unjustified practices, whereas Daphne evidently does not share those views). It is not surprising that Anita chooses to mitigate the threat to Daphne's face by using politeness strategies such as hesitation. The particular patterns described by CA, then, could be seen as outcomes of the general politeness principles set out by researchers like Brown and Levinson (1987). People make choices about how to

perform certain acts in the light of their assessment of how face-threatening those acts are. If this account is accepted, then the underlying difference between a 'preferred' and a 'dispreferred' move is that the latter is more face-threatening than the former and is therefore typically performed with more mitigation (for an extended argument to that effect, see Owen 1983).

At the beginning of this chapter I noted that Conversation Analysis has not confined itself to examining the kind of informal face-to-face talk that is most readily brought to mind by the term *conversation*. In the following sections I will consider some of the contributions CA has made to our knowledge and understanding of particular kinds of talk: talk on the telephone, and talk in institutional contexts.

WHY IS IT SO HARD TO GET OFF THE PHONE?

Telephone conversation has been studied quite intensively by CA practitioners. One argument for choosing this kind of data is that the phone restricts participants in talk to the oral-aural channel: there is no visual information available, and thus an analysis focusing on spoken language is not leaving out an important dimension of the participants' own experience. Not surprisingly, analyses of telephone conversation show that it is both similar to face-to-face talk in some ways and different in others. One example of simultaneous similarity/difference concerns the way calls often start:

Phone rings in Beth's location

- 1 Beth: hello
- 2 Ann: hello
- 3 Beth: oh hello Ann what's up
- 4 Ann: nothing much just had something I wanted to ask you

The puzzle here is that on the surface it looks as if we have a sequence of three greeting turns ('hello', 'hello', 'oh hello Ann') when the normal organizational principles would lead us to expect only two, one from each of the participants. But what seems to be a deviation from the norm in face-to-face conversation actually conforms to the principle of the adjacency pair; it just adapts it to the particular circumstances of the phone call. Beth's first turn is not a greeting, but rather the second part of an adjacency pair whose first part is the ringing of the phone (strictly speaking I should have numbered this event as turn 1 in my transcript – it stands in for Ann's opening turn). Analysts of telephone talk refer to the first part as the 'summons' and the second part as the 'answer'. In support of the claim that the answer is not the first turn in an exchange of greetings, we might also note that people do not always use a greeting formula like 'hello' in this position. Many recite their telephone number, for instance; in some countries it is conventional just to say one's name.

There is, furthermore, a good reason for supposing that the person who answers the phone must be taking the second rather than the first turn in the opening sequences. The person who answers is by definition not the person who has initiated the conversation: the initiator is the person who decided to make a phone call in the first place. And since it is only the initiator who knows *why* s/he is calling (the responder may not even know *who* is calling), the initiator needs to be in the position of taking the first turn in subsequent adjacency pairs.³ In the example above the initiator, Ann, produces the first greeting at turn 2 and the responder returns the greeting in turn 3, coupling this with an acknowledgement that she recognizes the caller and a formulaic enquiry ('what's up?') that could be either about the caller's general state or about her particular reason for calling. The caller then uses her next turn, turn 4, to introduce the actual business of the call: 'just had something I wanted to ask you'.

One question that is frequently asked by phone users is why it is difficult to *finish* a telephone conversation, and why getting off the phone with someone can become such a protracted business. In fact, CA practitioners have pointed out that conversations in general do not just 'end'; rather they have to be 'closed' (Schegloff and Sacks 1973). On the phone this requirement is particularly demanding linguistically: closure has to be done entirely through talk, since other options for disengagement, involving for instance body language or physical movement, are limited or unavailable.

So how do people in practice manage phone-call 'closings'? Here is an extract from a call made by Daphne, who is trying to get five people, including Paula, together for a meeting on a Saturday. At this point Daphne has elicited the information she needs, regarding which Saturdays in the next few weeks Paula has free.

Transcription conventions: (.) brief pause, ? rising intonation, : lengthened segment

- 1 Daphne: so that's any Saturday except the seventh (.) right I'll let you know when I've heard from the others
- 2 Paula: OK
- 3 Daphne: right then
- 4 Paula: I might be in London before that though shall I give you a ring?
- 5 Daphne: yeah do that
- 6 Paula: OK I will
- 7 Daphne: good
- 8 Paula: OK::
- 9 Daphne: take care then
- 10 Paula: you too
- 11 Daphne: bye
- 12 Paula: OK bye

The most obvious feature of this extract which might lead to the impression that the end of the call is 'dragging on' is the number of turns in which the participants merely utter formulaic remarks like 'good', 'right then', 'OK', rather than exchanging

any actual information. Why does this happen? Essentially, because it is necessary for telephone interactants to make sure that there is joint agreement to end the exchange. Since it is possible on the phone to cut someone off at any time (by hanging up), and this is regarded as a highly aggravated act of impoliteness (akin to just walking off in the middle of a face-to-face encounter) phone users who do not wish to be impolite have to negotiate the breaking off of contact carefully. Before they can close, each participant has to signal she has no more to say by 'passing' on the opportunity to introduce new information (some CA practitioners refer to 'unmentioned mentionables', things that could be mentioned but so far have not been). On the phone this 'passing' is signalled by taking a turn but putting no new content into it.

At turn 3 in the sequence above, Daphne 'passes'; she has an opportunity to go on to some matter other than the business the pair have just transacted, but what she says is merely 'right then'. This is potentially a 'pre-closing', an indication that she is ready to close and an invitation to Paula to pass on her own next turn. Paula, however, does not pass; she brings up the previously unmentioned topic of her intention to visit London. Daphne is thus obliged to respond to the new topic, and Paula completes the sequence by agreeing to ring when she makes the visit. At turn 7 Daphne again indicates readiness to close by saying 'good'. This time Paula also passes: she says 'OK', with marked lengthening on the last syllable which carries the implication of finality. With the pre-closing successfully accomplished by the two women each passing up the opportunity to introduce new matter, the actual closing can begin. It takes them ten turns to get from the end of their main business (at turn 2) to the end of the call, and it is not uncommon for this portion of a call to be far more extended than it is in this example. Since it is open to either party *not* to accept the other's invitation to pass, the endings of phone calls have a 'recursive' structure – the conversation in principle goes on until the parties manage to arrive at a mutual agreement to end it.

BEYOND THE 'BASICS': INSTITUTIONAL TALK

As I noted in Chapter 1, many discourse analysts are inclined to privilege what they call 'ordinary talk', meaning the kind that prototypically occurs in private or domestic settings between people whose relationship is close rather than distant. For various reasons, and in particular because children typically acquire their competence as conversationalists by talking to intimates (e.g. parents, siblings and other caregivers) in domestic settings, this kind of talk is regarded as 'basic'. Generally, CA practitioners follow this line of argument. In relation to 'institutional' talk – business meetings, service encounters, doctor–patient consultations, talk in courtrooms or in classrooms – analysis tends to be organized around the question, 'what special considerations apply that make talk in this institutional context different from ordinary talk?'

In their editorial introduction to a collection of articles on workplace interaction, Paul Drew and John Heritage summarize what they take to be the main differences between institutional and ordinary talk (Drew and Heritage 1992: 22):

- 1 Institutional interaction involves an orientation by at least one of the participants to some core goal, task or activity (or set of them) conventionally associated with

the institution in question. In short, institutional talk is normally informed by *goal orientations* of a relatively restricted conventional form.

- 2 Institutional interaction may often involve *special and particular constraints* on what one or both of the participants will treat as allowable contributions to the business at hand.
- 3 Institutional talk may be associated with *inferential frameworks* and procedures that are particular to specific institutional contexts.

To see what this means in practice, let us consider the case of interactions in a magistrate's court, which were studied by Sandra Harris (1984). The court Harris observed dealt mainly with people who had not paid previously imposed fines or maintenance orders. Defendants were questioned to determine why they had not paid and what resources they had to pay with. The purpose of this was to enable magistrates (in the English legal system, officials who preside over court proceedings which are too 'minor' to require a judge and jury) to determine how much the defendant should pay over what period. Thus the talk Harris observed in court had a very specific goal (cf. point 1 in Drew and Heritage's list), which was reflected in many features of the interaction: for example, that it took the form of a series of questions and answers on the topic of the defendant's intentions and financial resources, followed by the announcement of the magistrate's decision.

One 'special and particular constraint' which applied in the magistrates' court, and which often applies to some degree in institutional talk, though it rarely applies in ordinary conversation, was a constraint on the right to ask questions. Essentially, defendants were not permitted to ask questions, whereas they were obliged to respond to questions posed by the representatives of the institution. Harris observed occasional instances where defendants *did* ask questions, and were immediately censured for doing so (showing that their questions were not treated as 'allowable contributions to the business at hand'). She also observed that defendants were censured for answering questions in ways the magistrate deemed inappropriate. The following extract illustrates these observations (Harris 1984: 5):

*Transcription conventions: italics = emphasis, (6) = pause duration in seconds,
– = short pause, [= simultaneous speech*

Magistrate: I'm putting it to you again – are you going to *make* an offer – uh – uh – to discharge this debt (6)

Defendant: Would you in my position

Magistrate: I – I'm not here to answer questions – you answer *my* question

Defendant: One rule for one and one for another – I presume (3)

Magistrate: Can I have an answer to my question – please (6)

The question is – are you prepared to make an offer to the court – to discharge – this debt (2)

Defendant: What sort of minimal offer [would be required

Magistrate: [it's not a bargaining situation – it's a *straight* question Mr H. – can I have the answer

Here the defendant twice responds to the magistrate's questions with questions of his own. On the first occasion he is explicitly told that the magistrate is 'not here to answer questions' (which means it is inappropriate for defendants to ask any); on the second the magistrate cuts him off and complains that he is 'bargaining' rather than giving a straight answer. The magistrate asks the defendant three times in this short extract to answer the question he has been asked, demonstrating the validity of the defendant's presumption that there's 'one rule for one and one for another'.

The 'one rule for one and one for another' principle applies rather consistently in asymmetrical talk (that is, talk in which the participants do not have equal power, status, responsibility or control). Not only in courtrooms but also in other kinds of institutional encounters, such as medical consultations, interviews, radio phone-in programmes and school classrooms, there is an asymmetrical distribution of questions and answers. Typically it is the dominant party who has the right (and/or the obligation) to ask questions, while the subordinate party has a restricted right to ask questions. This is no coincidence, for the position of questioner gives its occupant considerable power to direct interaction. A question forms the first turn of an adjacency pair, and it does not merely require an answer, it also constrains what will count as a relevant or appropriate answer.

Magistrates in Harris's study often exploited this fact by putting questions in a particularly 'conductive' syntactic form, that is, a form which constrains the recipient to produce a particular answer. An example is the tag question 'that means you've got to pay thirteen pounds, doesn't it?' (Harris 1984: 17). The magistrate is not interested in the defendant's opinion on the matter; rather s/he is making an assertion ('you've got to pay thirteen pounds') and the purpose of encoding it as a question is to get the defendant to demonstrate understanding and agreement. The implication of a 'yes' answer is that the defendant recognizes the court's power to impose the payment and intends to comply with the order. A 'yes' thus serves the purposes of the institution, and the magistrate's discourse is designed to make the institution's preferred outcome the most likely outcome. Using questions obliges the defendant to respond rather than just remaining silent, but choosing conductive forms minimizes the defendant's opportunities to produce responses which are protracted, unexpected or challenging to the presuppositions of the court. Of course it remains *possible* for a defendant to say 'no' or embark on some other kind of response altogether – people are not sociolinguistic automata – but this will be treated as 'marked' and accountable behaviour. It transgresses both the norms of the context and the expectations embedded in the form of the question.

Drew and Heritage's third point about institutional talk is that it may be associated with particular *inferential frameworks*. In other words, people engaged in talk that has a certain institutional goal may use their understanding of what the institution wishes to achieve to make inferences about the meaning or function of an utterance that they might not make in other contexts. Sandra Harris noted that defendants in the magistrates' court regularly interpreted questions not as straightforward requests for information but as *accusations*. Defendants anticipated that magistrates would treat their failure to pay off debts as a sign of unwillingness rather

than inability to pay, and they often framed their responses to rebut the charge of deliberate non-compliance. For instance, one defendant, asked the question 'are you supporting anyone else?' answered 'not at all – no – I live on my own sir' (Harris 1984: 15–16). This answer, which is both emphatic and markedly deferential, suggests the defendant construed the question as implying that he had taken on a financial commitment to a new partner or family while neglecting to pay maintenance to his ex-wife and three children. In fact, it is unclear whether or not the magistrate who asked the question intended that inference. Magistrates need to know a defendant's circumstances before deciding how much their weekly payment should be, and 'are you supporting anyone else' could in principle be no more than an attempt to elicit relevant information. However, some questions posed by magistrates were clearly accusatory (for example, 'would you like to explain why . . . you've chosen to ignore the order of the court please' (p.19)). It is understandable that defendants extended the same inferential framework to questions whose status as accusations was less obvious: when talking to someone who has the power to send you to prison, it is advisable to err on the side of caution and adopt a defensive stance.

ACTIVITY: ANALYSING SCRIPTED TALK

To conclude this chapter, here is an activity that brings together some of the main themes discussed above, but in relation to a slightly unusual kind of data. In the discussion of turn-taking I introduced the axiom that talk-in-interaction is 'locally managed', organized by the participants themselves as they go along rather than following some preordained path. Recently, however, increasing numbers of companies providing service to customers, especially via the telephone, have been doing something that appears to be in conflict with this fundamental principle: writing standard scripts for routine transactions, which employees are required to perform. It is interesting to consider to what extent these scripts, which are usually designed by people with no expert training in the analysis of spoken discourse, embody the organizational principles which have been found by CA to apply in non-scripted, spontaneous interaction.

Overleaf I reproduce (with slight modifications to disguise the identity of the organizations concerned) scripts which are given to employees in two call centres where people are employed to deal with customer enquiries by telephone. One centre provides directory assistance to phone users, and the other authorizes credit card transactions. Read the scripts and then consider the following questions:

ACTIVITY

- 1 In what ways does the script in example A reproduce features that are typical of unscripted spoken interaction, and in what ways (if any) does it seem to deviate from the normal organization of talk? If you notice 'deviant' features, what might account for them and what problems might you expect them to cause in practice?
- 2 How is example A affected by the fact that it is a script for a transaction conducted by telephone rather than face to face?
- 3 In example B, how would you describe the difference between the script for authorized applications and the script for unauthorized applications? Why are these scripts so different?
- 4 Do Drew and Heritage's observations on the characteristics of institutional talk (quoted above) seem relevant/illuminating in relation to these scripts? To what extent and in what ways do the scripted transactions exemplify 'asymmetrical' talk?
- 5 If you have experience of dealing with service workers who are following a script, or if you have ever done a job that required you to engage in scripted interaction yourself, how do you feel about the practice of scripting talk? Why?

Example A: Conducting a 'standard search' for a telephone number

- Agent: XYZ Directories, which name please? 1
- Caller: Jones
- Agent: Jones, thank you.
Which town please?
- Caller: Cardiff 5
- Agent: Cardiff, thank you.
Which address please?
- Caller: Number 28, Acacia Avenue
- Agent: Number 28, Acacia Avenue, thank you.
Just searching for you. (pause) 10
Sorry to keep you waiting.
Thank you. Your number is 0123 456789
- Caller: Thank you.
- Agent: You're welcome.
- Caller: Goodbye. 15
- Agent: Goodbye.

Example B: dealing with requests for credit authorization

When a transaction is authorized, say to the caller:

This application has been accepted.

If a transaction is not authorized, say to the caller:

Unfortunately this application has been declined but thank you for calling.

SUGGESTIONS FOR FURTHER READING ABOUT CA

An introduction to CA which is both detailed and accessible is Hutchby and Wooffitt's *Conversation Analysis: An Introduction* (1998). Max Atkinson's *Our Masters' Voices* (1984), a study of political speechmaking as a form of interaction (the audience doesn't get a speaking turn, but it does get the chance to clap and cheer) is a readable piece of empirical CA. Other 'classic' CA work is collected in edited volumes including Gumperz and Hymes (1972), Scheinkein (1978) and Turner (1974). On institutional talk, probably the best single source is Drew and Heritage's collection *Talk at Work* (1992). Some readers may be interested in the origins of CA and its theoretical roots in ethnomethodological sociology. One account is provided in John Heritage's *Garfinkel and Ethnomethodology* (1984). Harvey Sacks, who is generally regarded as the pioneer of the approach that became CA, died early in the field's history; he never wrote a textbook, but the lectures he gave on conversation have been a source of inspiration to many of the analysts who followed him. They circulated among insiders in unpublished form for many years, and were eventually edited for publication as *Lectures on Conversation* (Sacks 1995).

NOTES

- 1 CA does have something to say about narrative, and I pursue the subject further in Chapter 10 below.
- 2 For discussion of the relationship between grammatical and interactional structure, see Ochs et al. (1996); the specific issue of how conversationalists project turn transition relevance places is examined in detail in Ford and Thompson's contribution to that volume.
- 3 It should be noted here that technological change may raise new and interesting questions about telephone talk. For instance, Jane Davies (pers. comm.) comments that now many people have caller display on their phones (especially their mobile phones), which means they can identify certain callers as soon as the phone starts ringing, it is increasingly common for the responder's first turn to be a greeting like 'hi Jane!' rather than an answer to the initial summons. This might prompt the question whether callers who have been recognized and greeted in this manner then have to negotiate themselves back into first position to enable the call to proceed smoothly (caller display tells us who's calling, but not why); and if so, how they do it.