

1 What is discourse and why analyse it?

In 1996 a spokesperson for British Telecom (BT), the UK's largest phone company, launched a campaign to improve the nation's communication skills, explaining that 'since life is in many ways a series of conversations, it makes sense to be as good as we possibly can at something we tend to take for granted' (quoted in the *Guardian*, 30 December 1996). Analysts of spoken discourse do not usually share BT's goal of making people 'better' at talking: they begin from the assumption that people are, with few exceptions, highly skilled users of spoken language. But most would probably agree that 'life is in many ways a series of conversations', and that talking is 'something we tend to take for granted'. When linguists and other social scientists analyse spoken discourse, their aim is to make explicit what normally gets taken for granted; it is also to show what talking accomplishes in people's lives and in society at large.

The reference to 'linguists and other social scientists' in the last paragraph is meant to underline the important point that working with spoken discourse is an interdisciplinary enterprise: among those who may be engaged in it are anthropologists, linguists, philosophers, psychologists, sociologists, students of the media or education or the law. A commonly used academic label for what these various people are doing, and one which I will use myself throughout this book, is *discourse analysis*. But while it is useful in many contexts to have this generic label available, it does need to be remembered that 'discourse analysis' is an umbrella term, allowing for considerable variation in subject matter and approach. For instance, I should make clear straight away that discourse analysis is not exclusively concerned with *spoken* discourse: in principle it can deal with socially situated language-use in any channel or medium. Discourse analysts may work with written data, or data from sign languages of the deaf, and some analysts work with textual graphics and images as well (e.g. Kress and van Leeuwen 1996).

There is also a distinction to be made between analysing discourse as an end in itself and analysing it as a means to some other end. Some discourse analysts – including many of those whose disciplinary affiliation is to linguistics – are primarily concerned to describe the complex structures and mechanisms of socially situated language-use. They ask questions like 'how does turn-taking work in conversation?', or 'does the form of a question affect the form of the answer?' or 'why do people misunderstand one another?'. They study talk because they want to know about talk. But some linguists, and many researchers who are not linguists, are more interested in the idea that 'life is in many ways a series of conversations', which implies that

people's talk can be a source of evidence about other aspects of their lives. Though they may not be studying discourse as an end in itself, many sociologists and social psychologists, or researchers in education, in cultural studies and media studies, adopt methods which produce discourse data. Interviews, focus group discussions and ethnographic studies using participant observation all involve verbal interaction between a researcher and research subjects, and/or between research subjects themselves. At least some of the analysis carried out by researchers who choose these methods will involve listening to talk, transcribing it, and reflecting on its meaning and significance.

As we will see in more detail later on, there are many different varieties of discourse analysis, and there is a certain amount of argument about their relative merits. My aim in this book is to be as inclusive as possible: though I write as a linguist myself, I have tried to design my account of what discourse analysis is, and does, to be helpful to readers across a range of academic disciplines. That means trying to cover the spectrum of approaches they are likely to encounter, making clear what the similarities and differences are, and bringing out the distinctive contributions made by different approaches.

As the title *Working with Spoken Discourse* suggests, this book will not be concerned with *all* forms of discourse, but only with *spoken* discourse. In the Introduction I gave one reason for this choice: talk forms the data many social researchers turn to discourse analysis to help them interpret, and they are often unsure how to approach this data in a suitably 'analytic' way. Arguably, this is less true of written discourse. Anyone who has been educated in a highly literate society will have developed, not only the ability to read and write, but also some ability to think analytically *about* written texts. For instance, many school students have had some experience of learning how to do 'close reading' of literary texts: they have had their attention drawn to the structure of a poem or to the existence of competing interpretations of its meaning. By contrast it is much less likely that they have ever been taught to approach ordinary talk – or any kind of spoken language – in the same systematic way. Similarly, most people acquire in the course of their schooling an extensive metalanguage ('language about language') with which to describe the structures of writing: terms like *letter*, *comma*, *sentence* and *paragraph* belong to this metalanguage. They rarely possess a parallel metalinguistic apparatus for discussing the structures of spoken language: as we will see in more detail in Chapter 3, most people do not realize the extent of the differences between writing and speech.

The above reference to metalanguage reminds us that issues of terminology and definition tend to loom large in all academic enterprises; discourse analysis is no exception. I have already used several different terms – *conversation*, *talk* and [spoken] *discourse* – for what might appear to be much the same thing. Do all these terms, however, really mean the same thing? The question would be unlikely to arise in the context of an ordinary, non-academic exchange, but an important part of being 'analytical' is being able to reflect on and ask questions about the conceptual frameworks and vocabulary we take for granted in everyday life. So at this point I want to look more closely at some of the key terms that are relevant to discourse analysis, beginning with the apparently straightforward term *conversation*.

CONVERSATION, TALK, DISCOURSE

ACTIVITY

(Note: this activity is more interesting to do in a group, and it is especially interesting if the group includes speakers of more than one language.)

As quickly as possible, list all the words you can think of that are used to describe different kinds of talk in each of the languages/varieties you know. Now examine your list more closely.

- How would you define each of the terms you have listed (for the benefit of someone learning the language, for example)? Is each one distinct from all the rest or is there overlap? Is there any disagreement in the group about the definition of certain terms?
- What dimensions of contrast (e.g. formal v. informal, serious v. non-serious) seem to be important in distinguishing different kinds of talk?
- If different languages/varieties are represented in your list, do they all make similar distinctions?
- How many of the kinds of talk you have listed could you also describe as ‘conversation’? If that term applies better to some cases than others, why do you think that is? If it is inapplicable in some cases, what makes it inapplicable?

In ordinary usage, *conversation* usually refers to spoken rather than written language. I say ‘usually’ because recently an interesting exception has become noticeable: people who regularly interact with others via the internet, for instance in ‘chat rooms’, sometimes refer to what they are doing as ‘talking’ or to their exchanges as ‘conversations’, though the medium is actually written language. The term ‘chat room’ makes an explicit parallel with a certain kind of informal conversation, namely *chat*. If we are being analytical, these usages might prompt some questions. Are *interactivity* (the fact that on-line exchange can involve a relatively rapid ‘back and forth’ between participants) and *spontaneity* (the fact that contributions to chat-room exchanges are typically composed without much planning or editing) more salient characteristics of what we call ‘conversation’ than the channel or medium of interaction? Is the actual language people produce in chat rooms more similar to face-to-face speech than other kinds of written language?

We (that is, English speakers, though the same thing is true for speakers of many other languages too) have quite a large vocabulary for distinguishing different kinds of talk. We can describe interactions in terms of their tone, level of formality and subject-matter using terms like *argument*, *blether*, *chat*, *discussion*, *gossip*. We can describe spoken language events in terms of their setting, context or purpose using terms like *interview*, *debate* and *seminar*. Is *conversation* just a generic term that subsumes all the others, or does it cover only some of the possibilities? Is a seminar a ‘conversation’? Is the talk I have with my doctor when I visit her surgery ‘conversation’ in the same way the talk I have with her if I run into her at the supermarket is

‘conversation’? The activity on p. 9 is intended to encourage you to think about your own understanding of what ‘conversation’ is, and more generally what different kinds of talk are recognized by language-users in the community or communities you belong to.

One point the activity might illustrate is that *conversation* in English has both ‘generic’ and more ‘specific’ uses. It is generic in the sense that we can use it to describe a relatively broad range of different kinds of spoken interaction. More specifically, though, it seems most ‘natural’ to apply it to interaction which is characterized by informality, spontaneity and egalitarian relationships between the participants (if your boss asks you to come and have a ‘conversation’ about your punctuality, you tend to suspect euphemism, or irony). Certainly, for me as an English-speaker it seems more natural to use the word *conversation* in connection with ‘chat’ or ‘gossip’ than for a seminar or a medical consultation. Each of these has features of conversation, but intuitively I feel it is not the prototypical case.

In this book I want to consider many kinds of spoken interaction, and to examine the similarities and differences among them. Therefore, when I discuss spoken interaction in a generic sense I will use two other words in preference to *conversation*: one is *talk* and the other is *discourse*. Just to complicate matters, one of these terms is more generic, or at least more inclusive, than the other. *Talk*, to state the obvious, refers only to spoken language-use, whereas *discourse*, as I have noted already, can refer to language-use in any channel or medium. But that is not the only difference between the two terms. *Discourse* is evidently a more ‘technical’ term than *talk*. And like a lot of technical terms, it is also ‘contested’ – that is, it has generated a lot of debate among scholars about what it means and how it should be used. In fact, the term *discourse* is notorious for the arguments surrounding it and the confusion it can cause. A major source of potential confusion is that the meaning of the term tends to vary quite significantly depending on the academic discipline and the theoretical preferences of the person who uses it. The range of meanings *discourse* can have in academic discussion is an issue that needs to be clarified sooner rather than later. So before I go any further: what is *discourse*?

LANGUAGE ‘ABOVE THE SENTENCE’ AND LANGUAGE ‘IN USE’

The most straightforward definition of *discourse* is the one often found in textbooks for students of linguistics: ‘language above the sentence’. Of course, that is not at all straightforward unless you understand some basic assumptions in linguistics, so let me spell them out.

Linguists treat language as a ‘system of systems’, with each system having its own characteristic forms of structure or organization. For instance, the sound system of a language (its phonology) does not have the same kinds of units, or the same rules for combining them, as the grammatical system of that language. As your units get larger (e.g. words are larger than sounds and sentences are larger than words), you metaphorically move ‘up’ from one level of organization to the next.

If discourse analysis deals with ‘language *above* the sentence’, this means it looks for patterns (structure, organization) in units which are larger, more extended, than one sentence.

One of the earliest discourse analysts, the linguist Zellig Harris (1952), posed the question: how do we tell whether a sequence of sentences is a *text*¹ – that is, the sentences relate to one another and collectively form some larger whole – as opposed to just a random collection of unrelated bits? The answer to that question, Harris thought, would make clear what kind of structure exists ‘above the sentence’. Texts would have this structure, whereas random collections of sentences would not.

Plainly, language-users do routinely interpret sequences longer than a sentence as texts in which the parts combine to form a larger whole. Consider the following example, a simple text produced by a child, which is discussed in a famous article by the conversation analyst Harvey Sacks (1972).

The baby cried.
The mommy picked it up.

One obvious instance of ‘structure above the sentence’ in this example is the pronoun *it*, which is *anaphoric* (referring back). It comes in the second sentence but it refers to something mentioned in the first: ‘the baby’. A reader or hearer automatically takes it that the ‘it’ which the mommy picked up must be the baby, and not some previously unmentioned object like a rattle or a banknote. The pronoun is a *cohesive* device, tying the two sentences together, and cohesion is a property of texts.

But there is more to say about what makes this sequence work as a text. For instance, it is natural to read it as a *narrative*, in which the sequence of events in the text mirrors the sequence of events in the reality being reported: the baby cried *and then* the mommy picked it up. Indeed, in this case we are likely to infer not merely sequence but causality: the mommy picked up the baby *because* it cried. The way we process the text as a narrative implies that we are following a general procedure for dealing with structure ‘above the sentence’: where A and B are sentences, we assume that A followed by B means ‘A *and then* B’ or ‘A *and consequently* B’.

But there is a problem with Zellig Harris’s proposal about distinguishing texts like the child’s story from random collections of sentences which are not texts: we seem to have a strong tendency to apply the principles just described to any sequence we are confronted with, however bizarre. Michael Stubbs (1983: 93) quotes a radio announcer who once said:

Later, an item about vasectomy and the results of the do-it-yourself competition.

Why does this raise a smile? The default assumption is that the parts of the announcement on either side of the conjunction *and* relate to one another in the same way as the two parts of the baby/mommy sequence, and so we reason that the announcer must be referring to a do-it-yourself vasectomy competition. However, I am confident that most of us immediately go on to reject this interpretation. We recognize that the announcer did not intend the structural relationship that is implicit in the organization of his discourse. But this recognition does not come from pondering the details of the announcer's language, and deciding that, contrary to our initial assumption, his utterance has no structure. Rather, we realize that the scenario we have conjured up by applying the usual procedure is implausible and ridiculous: no one would organize, or enter, a competition in which men performed vasectomies on themselves. In other words, we take account not simply of the linguistic properties of the announcement, but also (and in this case more significantly) of what we know about the world.

Real-world knowledge is also relevant to the interpretation of the baby/mommy text. There is no purely structural reason why we have to take it that the mommy who picked the baby up is the baby's own mother, since the child does not specify that by using a possessive pronoun, referring only to 'the mommy'. Nevertheless I would bet that most readers did make that assumption. The text follows a familiar script whereby babies cry and are picked up by their mothers to stop them crying. It is imaginable that a crying baby might be picked up by a total stranger who was, however, the mother of some other baby; but that would not be many people's first guess.

I am suggesting, then, that we make sense of discourse partly by making guesses based on knowledge about the world. If that is accepted, then arguably the definition of discourse as 'language above the sentence', and of discourse analysis as a search for structure at a level higher than sentence structure, is not adequate. That definition suggests that single sentences and texts have a similar kind of organization: the difference is one of *scale*. But is that really the case? To be grammatical, a sentence must contain certain constituents in a certain order: it is conformity to structural rules that makes the difference between grammatical sentences and 'word salad' (like **stood boy the on up chair a* – the asterisk is a linguist's convention for denoting ungrammatical sequences). But our ability to decide whether and how discourse makes sense appears to involve much more than quasi-grammatical generalizations about what can go with or follow what.

It might also be asked whether the characteristic features picked out by discourse analysts have much to do with the size of the units being analysed – the fact that they are larger than a sentence.² Henry Widdowson (1995) has pointed out that a 'text' can in fact be smaller than a sentence. He observes for instance that the legend LADIES on the door of a public lavatory is a text, as is the letter P which is used in Britain to indicate a space for parking cars. A single word or letter cannot have 'structure above the sentence'. So what makes these examples texts? Widdowson's answer is that in the contexts he is concerned with, each of them is intended to convey a complete message. Of course, what we take that message to be does depend on the context, and once again, its interpretation relies on real-world knowledge that is not

contained in the text itself. Looking up the word *ladies* in a dictionary would not, on its own, make clear what message it conveys when written on a door. (Someone who spoke English but was unfamiliar with the concept and etiquette of public lavatories might think it meant ‘there are ladies behind this door’.) A great deal of general knowledge and contextual information has to be brought to bear on even the most banal texts we encounter if those texts are to serve their communicative purpose. A distinctive feature of discourse analysis, as opposed to the study of syntax (sentence structure), is its overt concern with what and how language communicates when it is used purposefully in particular instances and contexts, and how the phenomena we find in ‘real language’ (implicitly contrasted to the idealized, made-up example sentences most often discussed by analysts of syntax) can be explained with reference to the communicative purposes of the text or the interaction. From this standpoint a better definition of *discourse* than ‘language above the sentence’ might be ‘language in use’: language used to do something and mean something, language produced and interpreted in a real-world context.

Deborah Schiffrin (1994) suggests that the two definitions of *discourse* I have just outlined correspond, roughly speaking, to two important currents or tendencies in twentieth-century linguistics. One is *formalism* or *structuralism*: an interest in the abstract form and structure of language. The other is *functionalism*: an interest in what language is used to do. But Schiffrin goes on to point out that treating this as an absolute distinction would be an oversimplification. Because its meaning is so dependent on context, discourse is not amenable to ‘pure’ formalist analysis. Conversely, functionalists have always been concerned with form as well as function. They are interested in how the two are connected, suggesting that language has a certain kind of formal organization because of the purposes it is designed to serve.

Most discourse analysts who locate themselves within the academic discipline of linguistics are concerned with both form and function, though the balance between these concerns may vary. But not all discourse analysts are linguists, and not all would define their goals in terms of improving our understanding of language as such. Many social scientists (including, in fact, some linguists) are more interested in discourse as a source of evidence or insight about social life and social relations. Their questions are not like Zellig Harris’s, primarily about the way language works. Rather they use discourse analysis as a qualitative research method for investigating social phenomena: sexual harassment, attitudes to the monarchy and youth subcultures are among the topics that have been investigated in this way.³ But investigators doing this kind of work often adopt a definition of the term *discourse* which differs from the ones we have examined so far. That alternative definition is now sufficiently influential across disciplines to merit more detailed consideration.

POWER, KNOWLEDGE AND PRACTICE: DISCOURSE(S) AND THE CONSTRUCTION OF SOCIAL REALITY

I mentioned above that discourse analysis is an increasingly popular qualitative research method in social science. The word *qualitative* in this context contrasts with

quantitative. Discourse analysis is an alternative to using standardized instruments like questionnaires, which yield statistical data. A researcher who wants to find out using quantitative methods what people do in their leisure time might ask a sample of subjects to fill in the same questionnaire and then produce a statistical digest of their responses (e.g. '50% of women under 35 reported shopping was their main leisure activity'). A researcher who decides to use discourse analysis as a method would be more likely to spend time talking in depth to a sample of the people s/he is interested in, encouraging them to explore the subject in their own way and in their own words. The researcher would record subjects' talk, transcribe it and analyse it – not in order to make statistical generalizations, but in order to point out recurring themes in the way people talk about leisure activities.

This method is sometimes criticized on the grounds that it will not tell us accurately what people *really* do in their leisure time. The criticism is not without validity, but a discourse analyst might point out that it could equally be made of the questionnaire/statistical analysis approach. When people answer a researcher's questions, whether in a face-to-face interview or by completing a written form, they are constructing a certain representation of themselves for the researcher's benefit: they may be telling the researcher what they think s/he wants to hear or what they would like her/him to believe. This is another version of the process of self-construction that goes on in ordinary talk, which is always produced with an eye to the situation and the person(s) to whom it is addressed. Arguably it is an unavoidable element of all communicative acts: people simply do not answer questions, in any situation, without first making some assessment of who is asking and why. (As a simple illustration of this point, recall the last time a doctor or nurse asked you how many units of alcohol you consume in a week. Was your answer affected by your assumptions about what use the nurse or doctor planned to make of the information? Was your answer accurate and truthful? Would you have given just the same answer to your mother, your best friend, or a prospective employer? What is your attitude to the questions I am asking now – for instance, if you don't drink alcohol, are you offended by my apparent assumption that all my readers do? Would that affect your response to me, if this were a conversation?)

Researchers who favour discourse analysis over supposedly more 'objective' methods argue that paying attention, not merely to *what* people say but to *how* they say it, gives additional insight into the way people understand things. It is less about collecting facts than about studying interpretive processes. Such researchers may also argue that analysing 'real' talk does a better job than standardized instruments of capturing the messiness of real life, and to that extent could be seen as more rather than less 'accurate'. Giving people a multiple-choice questionnaire obliges them to choose one option from a set constructed by someone else: they check box A, and that makes them look as if they are committed to A while rejecting B and C. Yet when people talk it often becomes clear that matters are more complex than that: they don't dismiss B and C out of hand, and they have their doubts about A. Standardized instruments produce an impression of certainty and consistency which is, arguably, misleading. Another advantage that might be claimed for discourse analysis is that it generates data by getting people to engage, or observing them while they engage, in

an activity – talking – which is normal and familiar to them, rather than asking them to undertake an unusual or artificial task. Life may or may not be ‘in many ways a series of conversations’, but it is in no way a series of box-checking exercises.

Social researchers who do discourse analysis often want to make the point that even when we talk ‘in our own words’, these words may not actually be ‘ours’ at all, in the sense that they are not original or unique to any one individual. As one analyst, Jay Lemke, has put this point:

We speak with the voices of our communities, and to the extent that we have individual voices, we fashion them out of the social voices already available to us, appropriating the words of others to speak a word of our own. (1995: 24–5)

Within any community there is a finite range of things it is conventional or intelligible to say about any given concern. When people talk about shopping, or drugs, or the royal family, what they say will be drawn from the community’s repertoire of things it is possible to say rather than representing some unique perspective on the topic. This is not to suggest that people never say anything novel or unexpected, or that they do not have ideas of their own. But language-using is an *intersubjective* rather than purely subjective process: a ‘voice’ that is wholly individual runs the risk of being incomprehensible. Hence Lemke’s point that individuals’ ways of talking are formed using resources that are shared with others in their communities. Discourse analysis can be seen as a method for investigating the ‘social voices’ available to the people whose talk analysts collect.

Many social researchers today would argue that people’s understandings of the world are not merely expressed in their discourse but actually shaped by the ways of using language which people have available to them. Another way of putting this is to say that reality is ‘discursively constructed’, made and remade as people talk about things using the ‘discourses’ they have access to. Evidently, the word *discourse* in this formulation is not being used in the way linguists typically use it, to mean ‘language above the sentence’ or ‘language in use’. An obvious difference is that the linguist’s *discourse* has no plural, whereas social theorists often talk about *discourses*. This plural usage reflects the influence of the philosopher and cultural historian Michel Foucault, who defined *discourses* as ‘practices which systematically form the objects of which they speak’ (Foucault 1972: 49).

To see what Foucault meant, let us consider the case of ‘drugs’. The word *drugs* might seem to name a pre-existing ‘real-world’ category (of substances that affect the mind and body in certain ways). But if we think about the way the word is most often used, it becomes evident that it does not simply denote *all* substances that have certain effects: depending on context, it denotes either those which are medicinal, or those which are used non-medicinally and are also illegal. Caffeine, nicotine and alcohol are clearly mood-altering substances, but if we hear a report on ‘drugs’ we do not immediately think of coffee, cigarettes and beer. We certainly do not think of coffee-bean importers as drug traffickers or of tobacconists as drug dealers. What this suggests is that our category of ‘drugs’ (in the non-medicinal sense) has been formed through a particular set of *practices*: legislation (making some substances illegal),

policing (trying to prevent breaches of the law and to catch people who do break it), the practices of the courts (where stories are told about why people have broken the law and decisions are made about how to deal with them), of social and charitable agencies (which try to reduce the harm caused by drugs), schools (which practise 'drugs education') and the media (which report on 'the problem of drugs'). Buying, selling and using illegal substances are also practices relevant to the understanding of 'drugs' as a category, though fewer people are involved in these practices compared with the numbers exposed to education or media reporting.

With so many practices and agencies involved, not surprisingly there are multiple 'discourses' on drugs. We may be working with the same category, but we can discuss it in different ways. For instance, there is a 'law and order' discourse in which drug-use is discussed as a crime, committed by people who are 'bad'. An alternative discourse is 'medical': people who use drugs are sick, and need treatment rather than punishment. There is a 'social' discourse in which drug-taking arises from deprivation and hopelessness. In contrast to these negative ways of talking, there is also a discourse in which drug-using is defined as a recreational activity, enjoyed without ill effects by the majority of those who engage in it. Another positive discourse suggests that using drugs may help people attain greater spiritual awareness.

Each of these ways of talking about 'drugs' has a history, but in some practices (and many discussions) they are not kept distinct. Drugs education, for instance, typically aims to persuade young people that they should not use drugs (it is unhealthy, illegal and dangerous), but some programmes also discuss the idea that drug-use is pleasurable, on the grounds that the appeal of drugs must be acknowledged if young people are to take warnings about the dangers seriously. Some programmes assume that many or most people will experiment with drugs, and aim to teach them how to minimize the risks involved. So drugs education may mix, in various proportions, elements of the 'law and order', 'medical' and 'recreational' discourses. Together, the various ways of discussing drugs and the practices that go along with them form a network of concepts and beliefs that set the agenda for debate and define what we perceive as reality on this subject. This is what theorists mean when they say that reality is 'discursively constructed'.

It might be asked what the sense of the word *discourse* that I have been discussing has to do, specifically, with *language*. Recall Foucault's definition, quoted above: although he calls discourses 'practices', he goes on to say that they 'form the objects of which they *speak*'. The link between practice and speaking (or more generally, language-use) lies in Foucault's concept of 'power/knowledge'. In the modern age, Foucault points out, a great deal of power and social control is exercised not by brute physical force or even by economic coercion, but by the activities of 'experts' who are licensed to define, describe and classify things and people. Definition, description and classification are practices, but they are essentially practices carried out using language. Words can be powerful: the institutional authority to categorize people is frequently inseparable from the authority to do things to them. Thus for instance, experts define mental health and mental illness, and on the basis of their definitions, individuals can be classified as mentally ill and detained in psychiatric institutions. Experts produce definitions of good or adequate 'parenting', and parents

who do not meet the minimum standard may have their children taken away from them. Experts elaborate a concept of ‘intelligence’ and devise ways of measuring it (such as IQ tests); this may have real-world consequences for individuals’ education and employment prospects.

I have explained the sense of *discourse* that comes from the work of Foucault because this usage of the term is now quite common, and students of discourse analysis in a variety of disciplines are likely to encounter it. However, it should not be supposed that all social researchers who adopt discourse analysis as a method are committed to the ideas of Foucault, or those of any other theorist. Some sociologists and social psychologists use *discourse* rather as some linguists do, to mean ‘language in use’. There are varying views on whether and to what extent social reality is ‘discursively constructed’: you do not have to believe in the discursive construction of reality to regard what people say as a source of insight *about* reality.

As I pointed out earlier, though, any researcher who sets out to investigate some aspect of reality by studying discourse will end up with data in the form of *language*. And it is easy to underestimate the complexity of those data. As practised users of at least one language, researchers may be tempted to assume that it requires no special expertise to interpret linguistic data – that this is simply an extension of our ordinary, everyday behaviour as participants in verbal interaction. But that is at best only partly true. Being able to do something yourself is not the same as being able to analyse it from the outside. Discourse is not pure content, not just a window on someone’s mental or social world; it has to be considered *as discourse*, that is, as a form of language with certain characteristics which are dictated by the way language and communication work. It is not only linguists who can benefit by paying attention to the ‘how’ as well as the ‘what’, the form as well as the content of people’s discourse. Conversely, linguists have something to gain by attending to other social scientists’ insights into what discourse does, or what social actors do with it.

SUMMARY

This chapter has been concerned with the meaning of the term *discourse* and the goals or purposes of analysing it. The view of discourse analysis taken here and throughout this book is a ‘holistic’ one, which acknowledges that discourse analysis is several things at once. It is a method for doing social research; it is a body of empirical knowledge about how talk and text are organized; it is the home of various theories about the nature and workings of human communication, and also of theories about the construction and reproduction of social reality. It is both about language and about life.

Part I continues, though, not with these grand abstractions, but with some concrete, practical considerations. The first requirement for any kind of discourse analysis is a body of data to analyse. In the next chapter we will look at the options and problems involved in collecting spoken discourse data.

SUGGESTIONS FOR FURTHER READING ABOUT 'DISCOURSE'

A book whose purpose is to 'unpack' the complex term *discourse* is Sara Mills's *Discourse* (1997). A shorter survey of various tendencies in contemporary discourse analysis is provided by the editors' Introduction to *The Discourse Reader* (Jaworski and Coupland 1999). This volume also includes an edited extract from Michel Foucault's *The History of Sexuality*, 'The incitement to discourse', which gives something of the flavour of what Foucault and his followers mean by the term. For a more traditionally 'linguistic' perspective on discourse and discourse analysis, a good source is the second chapter of Deborah Schiffrin's textbook *Approaches to Discourse* (1994), which is titled 'Definitions of discourse'. Teun van Dijk's edited two-volume *Discourse Studies: A Multidisciplinary Introduction* (1997) is a useful reference source for both linguists and others.

NOTES

- 1 In the next few paragraphs the word *text* occurs several times, and readers may find themselves wondering if it is just a synonym for *discourse*. In fact that is a disputed question. Some writers use the two terms more or less interchangeably (this is how I am using them in this section); some refer to spoken *discourse* but written *text* (i.e. the difference is one of medium); others (e.g. Widdowson 1995) make a more theoretical distinction. Briefly, Widdowson argues that *text* is the linguistic object (e.g. the words on a page in a book, or the transcript of a conversation) whereas *discourse* is the process of interaction/interpretation that produces meaning from language. In speech discourse comes first, and produces a text; in writing text comes first, and readers produce discourse from it.
- 2 As will be discussed in more detail below (Chapter 3), the sentence is in any case essentially a unit of *written* rather than spoken language.
- 3 For readers who want to follow up any of these examples, Kitzinger and Thomas (1995) is a discourse analytic study of sexual harassment; Billig (1992) is a study of discourse about the (British) monarchy; and Widdicombe and Wooffitt (1995) is a study of youth subcultures using the method of Conversation Analysis.