# Locating and Conducting Discourse Analytic Research

*Stephanie Taylor*

## Introduction

The aim of this opening chapter, in conjunction with the final chapter of the book, Chapter Eight, is to provide a new researcher with a broad overview of discourse analytic research. In order to locate discourse analysis within the academic field, I will first venture a possible definition and discuss what *kind* of research it is. This will involve a brief discussion of language and a review of some of the premises underlying social research in general and discourse analysis in particular. I will also consider the role of the discourse analytic researcher. In the second half of the chapter I offer a brief guide to conducting a discourse analytic project, with sections on formulating research questions; data, including transcription; the process of analysis, and writing up a project into a research text. In Chapter Eight, I discuss the evaluation of discourse analytic research and its potential applications.
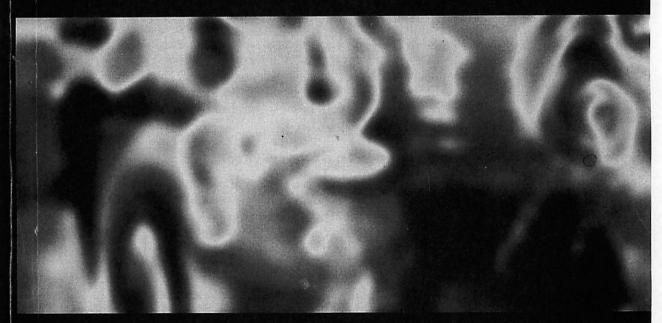
These opening and closing chapters are therefore intended as an introduction to a more complex academic area, rather than as self-contained guides to analysis. Whereas the intervening chapters, Two to Seven, consider specific forms of discourse analysis in detail, I will summarize some of the broader features of the field as a whole and discuss issues which are relevant to the work of most social researchers who would identify themselves as discourse analysts.

## 1 Locating discourse analytic research

The obvious starting point for this chapter is to define what discourse analysis *is*. You are probably already aware that the term has a wide reference (see Wetherell *et al.*, 2001). It can describe very different research activities with different kinds of data, as Chapters Two to Seven demonstrate. In order to cover this diversity, I will start with a loose definition, as follows: discourse analysis is the close study of language in use. To unpack this a little I will explore what is meant by 'language' and then present four possible approaches to discourse analysis. These are simplified descriptions and will not work as a definitive categorization system for any particular research project because the approaches are necessarily interconnected, as I will show. Discourse analysis is best understood as a field of research rather than a single practice but the four approaches provide a useful introduction.

# Discourse as Data

## A GUIDE FOR ANALYSIS

MARGARET WETHERELL • STEPHANIE TAYLOR
SIMEON J. YATES

The common starting point, then, is that discourse analysts are looking closely at **language in use** and, furthermore, they are looking for **patterns**. But what exactly is 'language'? One possible answer is illustrated in the way that most people attempt to learn a new foreign language, beginning with a few basic words and expressions. This common-sense strategy rests on a particular model of language, as a static system which can be broken down to its component parts. The student learns the parts, such as items of vocabulary, grammatical forms like plurals and tenses, and fixed expressions such as greetings, and then tries to connect them together again. In this way, hopefully, the student eventually learns to communicate with other users of the same language, by analysing what they say into its component parts, then building up appropriate messages back.

On this model, the system of language works for communication because it is a vehicle for meaning; in other words, it can be used to convey meaning from one person to another, provided that both are familiar with the elements of the language. It is as if speakers or writers encode meanings into the language and then hearers or readers decode them. Wertsch (1990, reproduced as Reading Sixteen in Wetherell *et al.*, 2001) calls this "the transmission model of communication": meaning is transmitted, or conveyed, through language, like signals through a telephone wire. However, two problems with this straightforward view of language quickly become apparent. These are especially interesting for the purposes of this chapter, because they highlight some of the concerns of the discourse analyst and distinguish certain approaches to discourse analytic research.

One problem with the model of language as a system is, of course, that the system is not static but is constantly changing. Because the elements change, the language which is in use is never the same as that in the grammar book or dictionary, however up-to-date and complete these may be. Change occurs over time (think of how quickly idiomatic expressions go out of date) and also within a single interaction. New meanings are created through the to-and-fro and the combined contributions of both (or all) parties. As a simple example, a term like 'this' or 'here' changes meaning as it refers to something different. Because these new meanings are being created, and also because the language is being used to do things, it is not sufficient to understand language as **transparent** or **reflective**. It is not a neutral information-carrying vehicle, as the transmission model of communication would imply. Rather, language is **constitutive**: it is the site where meanings are created and changed.

A small digression: the notion that language is not transparent is one of the fundamental assumptions of discourse analysis but it does create (consti-tute?) a problem for the analyst, as will be discussed further in Section 1.3. On the one hand, language is assumed to be constitutive, as I have just explained. On the other, in academia as elsewhere, no one entirely abandons the premise that talk and texts convey information about something else. So language is also assumed to be **referential**. (This is, of course, taken for granted as a feature of any academic text in which

analyses are presented or issues are discussed, including the one you are reading now!) One issue becomes the extent to which data should be treated as referential or as constitutive. For example, when a researcher interviews someone, how far can their talk be treated as a reflection of something else, such as established memories and ready-formed opinions? Should it instead be analysed as the place and process in which memories and opinions are constituted?

These questions are addressed in more detail in Chapters Four and Five (see also Section 1.3).

Another problem with the static system model relates to the uses of language. Language is an important means for *doing* things: greeting, snubbing, claiming, persuading, denying or sowing doubts (see Potter, 2001). To understand what exactly is being done, it is often necessary, again, to know what occurred at an earlier point in a situation or interaction. This applies to both the immediate interaction of a conversation and the more protracted sequence which is implied in, say, a reader's letter in response to a newspaper article, or one academic text criticizing another. In other words, to understand what is being done with language, it is necessary to consider its **situated** use, within the process of an ongoing interaction.

Both of these problems show that the model of language as a static system is over-simplified. More importantly, for the purposes of this chapter they serve to introduce the four approaches to discourse analysis. In the first approach to discourse analysis, it is precisely the variation and imperfection of language as a system which is the focus. Discourse analysts study language in use to discover how it varies and relate this variation to different social situations and environments, or different users. In contrast, the second approach to discourse analysis focuses on the *activity* of language use, rather than the language itself. Here the analyst studies language use as a process, investigating the to-and-fro of interactions (usually talk) between at least two parties and looking for patterns in what the language users (speakers) do.

A third approach to discourse analysis is rather different. The analyst looks for patterns in the language associated with a particular topic or activity, such as the family of special terms and meanings around it: a study might focus on the language associated with a particular occupation, such as social work or nursing. And a fourth possible approach to discourse analysis is to look for patterns within much larger contexts, such as those referred to as 'society' or 'culture'. Here the interest is in how language is important as part of wider processes and activities. For example, the analyst might investigate patterns in the labelling and classification of people or activities within a society. The language of categorization will be implicated with, on the one hand, the values underlying it (for example, beliefs that certain people are 'good' or 'bad') and associated philosophies or logics (such as when an activity is evaluated negatively because it is believed to have negative consequences), and on the other, the consequences and social effects of the classification. The analyst's interests will therefore

extend beyond language in use, that is, from the 'discursive' to the 'extra-discursive', probably blurring any distinction between them.

You have probably noticed that although I describe four approaches to discourse analysis, I have avoided giving a separate definition or definitions of 'discourse'. One reason is that that term is itself wide-ranging and slippery. More importantly, defining it still leaves the separate task of explaining 'discourse analysis' as a research activity. So for the remainder of this chapter I will focus on the latter, although this will sometimes involve introducing certain definitions of discourse.

My next task is to elaborate the four approaches I have outlined above. In the first approach, as I have said, the discourse analyst focuses primarily on the language itself. The patterns which are identified in the language in use may be similar to those which have conventionally interested linguists. For instance, they may be described in terms of vocabulary, structure or functions. Concepts like 'genre' or 'code'. may be employed to characterize an interrelationship between language and social situation. (See Maybin, 2001; Hodge and Kress, 1988, reproduced as Reading Twenty-one in Wetherell *et al.*, 2001.) The analyst's interest is, broadly, in regularities within an imperfect and unstable system.

In the second approach to discourse analysis, the analyst is more interested in the 'use' than the 'language' and interaction becomes the major focus. Patterns may be identified in terms of a sequence of contributions to an interaction, or a regular 'shape' like a script. This focus on use implies a particular view of the language user. She or he is not seen as a free agent using language to encode or decode a meaning in order to communicate it. To a great extent, any one person's contribution must follow on from that of the previous contribution and is inevitably shaped by what has gone before. Furthermore, meaning will be created within the interaction. The language user is therefore understood as **constrained** by the interactive context.

In the third approach to discourse analysis, the pattern within language in use which interests the analyst is the set or family of terms which are related to a particular topic or activity. This approach draws attention to how new terms enable people to talk about different things; for example, a student generally learns new vocabulary as part of the process of becoming familiar with a new field of work. This is not simply a matter of attaching different labels to already existing objects. As I have already described, there is also a sense in which language is constitutive; that is, it creates what it refers to. Of course this is not a once-and-for-all situation. Think about how meanings are created and eroded as part of ongoing social change. Technological change creates new *things* to be talked about and also new activities. For instance, the internet has given us alternative meanings for 'surfing' and 'browsing'. A language pattern of this kind (sometimes referred to as '*a* discourse' – note the article – or an 'interpretative repertoire': see Chapter Five) is therefore specific to particular circumstances. This approach would understand language as situated, but within a particular social and cultural context rather than a particular interaction (as in the second approach). In this respect, it blurs into the fourth approach to discourse analysis.

The aim of the analyst following the fourth approach is, broadly, to identify patterns of language and related practices and to show how these constitute aspects of society and the people within it. Ultimately such an analysis draws attention to the social nature and historical origins of the world 'out there' which is generally taken for granted. Controversy is basic to this form of discourse analysis because it involves the study of power and resistance, contests and struggles. The basic assumption here is that the language available to people enables and constrains not only their expression of certain ideas but also what they do.

To take a simple but well-known example, this assumption underlies the promotion of anti-racist and anti-sexist language. This is part of the equal opportunities policies of many organizations (such as universities) but is often disparaged as 'political correctness' (or 'pc'). In its simplest form, the aim is to reduce discriminatory practices in an institution by discouraging the use of discriminatory language. The proponents of this policy have criticized established images, for example in children's books, and taken-for-granted terms and expressions which contain negative stereotypes. Its opponents usually claim that such material is innocuous. Its proponents use a rationale which corresponds loosely to the notion that language is constitutive and that it blurs into practices. They argue that the way in which something, or, more likely, someone, is talked about does make a difference to the larger workings of society. It is through language, for example, that certain things or people are either categorized together or separated out as different, and through language that value is attributed or denied. This is particularly evident when language is used to classify and categorize for official purposes. For example, Rose (1985) describes how, as schooling became widespread in Britain, children began to be classified as good and bad learners. The requirements of the new education system created new categories of children, such as the 'feebleminded' child who had no obvious physical problem yet did not learn or progress through the school system.

This approach has given rise to historical studies, known as **genealogies**, which trace back the development of discourses (these are discussed more fully in Chapter Seven, which presents an example of a genealogical study). A famous example is Foucault's (1961) study of the development of the concept of madness. Rose's work, described above, also belongs in this tradition. A study of these practices explores their wider implications, such as the identities which they make available and the constraints which they set up. In general, a study based on either the third or fourth approach to discourse analysis is often on a larger scale than a study of interaction and may involve public contexts. It could be conducted, for example, through the study of official documents (see Chapter Seven: also, Helleiner, 1998; Stenson and Watt, 1999).

This fourth approach to discourse analysis draws attention to the all-enveloping nature of discourse as a fluid, shifting medium in which meaning is created and contested. The language user is not a detached communicator, sending out and receiving information, but is always **located,** immersed in this medium and struggling to take her or his own social and cultural positioning into account. Even more than with the

second approach, this fourth approach to discourse analysis understands the language user not as a free agent but as one who is heavily constrained in her or his choice of language and action, even if these are not fully determined. And of course the discourse analyst is not outside these struggles and constraints but is one such user within them.

Although I have presented these four approaches to discourse analysis as distinct, it will already be apparent that they are implicated in one another and shade together. The outline of four separate approaches works as a useful but oversimplified introduction to the field of discourse analytic work. It does not make clear the controversies within the field, and the criticisms which different practitioners would level at each other's work. As a way into these, I will now approach discourse analysis from a different direction and ask: what kind of research is this and what kind of knowledge does it produce?

## 1.1  What kind of research?

So far I have offered a loose definition of discourse analysis (as the search for patterns within language in use) and discussed what it involves, elaborating the definition to cover research activities with four different foci. In this section I want to step back a little and consider a broader question: what *kind* of research is discourse analysis?

One way to answer this would be to locate discourse analysis within disciplines. Unfortunately, as anyone who has consulted a university library catalogue will be aware, the enormous range of academic work which is referred to as discourse analysis spans many disciplines. For example, as the focus on 'language' would suggest, discourse analysis is used in linguistics and sociolinguistics, and also in literary studies (though this last area is not covered in this book). The focus on 'use' or on society gives it a place in different fields of social research including sociolinguistics, sociology and social psychology as well as more 'applied' areas, such as social policy and education. More recently, it has also been used by students of politics and human geography.

Another answer to the question 'what kind of research is discourse analysis?' would be in terms of the classic dichotomy between qualitative and quantitative research methods. Certainly, researchers often become interested in discourse analysis because they want to find a form of qualitative analysis for interviews or documentary material (see Chapters Six and Seven). However, discourse analysis can also use quantitative methods (some examples are discussed in Chapter Three). Moreover, the qualitative/ quantitative distinction is extremely broad.

In other words, attempts to locate discourse analysis by categorizing it as qualitative rather than quantitative are not much more useful than locating it within a particular discipline. To answer the question 'What kind of research is this?', the discourse analyst will have to return to the broad premises underlying social research and locate her or his own project with reference to these and to certain established debates. The next two sections will discuss discourse analytic research in these terms (for a much fuller discussion of these debates, see Wetherell, 2001a).

## 1.2  Locating social research in general

### *Epistemological issues*

I have asked, 'what is discourse analysis?' and then 'what kind of research is it?' Another question would be about what we can use discourse analysis to find out. In other words, 'what kind of knowledge does discourse analysis produce?' This is a question about **epistemology**, meaning that it is concerned with the status of knowledge.

I have said that discourse analytic researchers are looking for patterns in language in use and I have suggested that these might be patterns within language or patterns of activity. In addition to identifying such patterns, the researcher will be making some kind of epistemological claim about them. There is a range of possibilities here but I will discuss them in terms of two broad traditions.

The first tradition is associated with the physical sciences and embraces **positivism** and **postpositivism**. These are themselves broad and varied traditions (see Smith, 1998, for a fuller discussion) but they share several assumptions. One is that through the use of appropriate methods, which have become well established, the researcher can obtain knowledge of the world and its workings, particularly of the causal relationships which operate within it. Identifying such relationships enables the researcher to apply the research to real-world problems by making accurate predictions, and possibly interventions. The knowledge obtained through the research is generalizable to other contexts because it is universal.

Another claim made in these traditions is that research produces knowledge which is value-free and objective, unaffected by any personal bias or world-view of the researcher. Good research is considered to produce neutral information and contribute to a cumulative process which aspires towards finding universal truths. The whole or final truth about the world may not be attainable, but successive researchers attempt to approach it, testing hypotheses and taking a fallibilistic approach in which previous findings are treated as provisional and open to further testing (see Seale, 1999). Each researcher's procedures and findings are, ideally, published, scrutinized and rigorously evaluated in terms of validity, reliability and replicability (see Chapter Eight; also Sapsford, 1999: 107 and 139). So here is one set of related claims, that research produces knowledge which is universal, in that it holds across different situations and different times, and is value-free. In ordinary terms, this is knowledge with the status of truth: it is enduring and it is separate from the opinions and values of the researcher.

A contrasting tradition (or more precisely, again, a composite of different traditions) is more strongly associated with the social than the physical sciences and with work around critical theory, postmodernism and poststructuralism. Underlying it are quite different epistemological claims. Researchers in this tradition do not usually aim or claim to capture the truth of reality but to offer an interpretation or version which is inevitably partial. This is not simply because they have more limited ambitions but is a consequence of various assumptions or premises! The first of these is that the complexity and also the dynamic nature of the social world mean that a

researcher can seldom make confident predictions about it. There are too many factors operating in any situation and the relationships which operated in the past will not necessarily be those that prevail in the future. As a result, the researcher's aim is to investigate meaning and significance, rather than to predict and control (Banister *et al.*, 1994: 3).

A second premise is that no neutral single truth is possible in the social sciences because these involve the study of other people who have their own viewpoints. Any account of a social phenomenon or situation inevitably reflects the observer/researcher's partial understanding and special interest. To claim it as pure knowledge or truth would therefore be to deny the diversity of viewpoints and experiences of the other people who are involved (Said, 1978). A third and more complex premise is that no single truth is possible because reality is neither single nor regular: there are multiple realities and therefore multiple truths (this is not just an epistemological position, about the status of knowledge, but an **ontological** one, about the nature of the world itself). Furthermore, truth claims cannot be checked because accounts of the world are not simply reflections or records of what already exists. They themselves constitute and change what they purport to describe (this point relates to the discussion of language as constitutive in Section 1.1).

The epistemological claims being made in this second broad tradition are that the knowledge obtained by research is **partial**, **situated** (i.e. specific to particular situations and periods rather than universally applicable) and **relative** (i.e. related to the researcher's world view and value system). Such claims are particularly associated with qualitative research, including forms of discourse analysis such as those presented by Edley in Chapter Five and Carabine in Chapter Seven. Some other forms of discourse analysis, most notably conversation analysis (see Chapter Two) are closer to the first epistemological tradition.

To understand what kind of research discourse analysis is, it is not enough to study what the researcher *does* (like following a recipe!). We also need to refer back to these epistemological debates and their wider implications. It has been suggested that these debates have created for researchers a "double crisis of representation and legitimation" (Denzin and Lincoln, 1998: 21). The crisis of **representation** is that the researcher cannot claim to offer 'objective' knowledge of reality and of the world out there, but only a biased 'subjective' account. Indeed, since objectivity is impossible, the terms 'objective' and subjective' cease to be applicable. The most extreme interpretation of this position would be that research cannot tell us about the world but only about the world-view of the researcher! The crisis of **legitimation** is that there are no well-established procedures for evaluating the knowledge obtained. The main reason for this is that there is no assumption that the researcher's findings or claims can be checked against objective reality. Reality remains inaccessible so any attempt to verify results, for example, by duplicating the research, simply produces another unreliable version. It is as if the researcher is forever trapped inside a building in which all the windows contain distorting glass. It is impossible to go outside to get an open view of the world. The view from one window can only be checked against the view from

another, which is also limited and distorted by the glass, but in a slightly different way.

I will return to the implications of this double crisis in Chapter Eight. It is worth noting, however, that the first issue, the problem of representation, is of particular interest to discourse analysts because representation is also the topic or focus of their research. We can imagine a potential (perhaps irritating!) spiral in which such a researcher conducts an analysis, then analyses her or his own representation of the analysis, then analyses the representation of the analysis of the representation of the original analysis, and so on! (For an example of a piece of work in this spirit, see Ashmore, 1989.)

The second issue, the problem of legitimation, might suggest that research can never be assessed or evaluated by anyone but the original researcher. This would create a situation in academia which Seale and Silverman have described as "methodological anarchy" (1997: 380). Aside from other problems, it would lead to professional difficulties for researchers whose future employment depends on others' assessment of the quality of their work! Fortunately, such anarchy does not, in practice, prevail, for reasons which will be discussed later in this chapter and also in Chapter Eight. The double crisis therefore needs to be understood by prospective researchers but it should not be exaggerated: a wide variety of research, including discourse analytic research, continues to be conducted, written up (for example, in academic journals), evaluated and widely debated.

### *Generalization*

Another related point to consider is the claim that a researcher is making about **generalization**. I have said that discourse analysis involves the search for patterns within language in use. How general are these patterns and what form do they take? The answers to these questions again relate back to the analyst's premises or assumptions.

I have suggested that an analysis which focuses on the language itself is likely to approach language as a system, albeit one which is imperfect and fairly fluid. Consequently, the claims made are likely to be generalized. The analyst identifies features which occur across a range of contexts, or in a particular category of context (such as a type of interaction). The component elements might be defined structurally. Another possibility (not the only alternative) would be to analyse them in terms of functions. The analyst might identify them as part of an investigation of their wider effects, such as how they assist in the creation or reinforcement of groups and difference. In this form of analysis the body of data is sometimes very large (often referred to as a **corpus**) and the analysis might involve electronic sorting methods, such as concordancing (this is explained in Chapter Three), and possibly also some quantitative analysis. This broad description applies to much discourse analytic work in sociolinguistics and critical discourse analysis (see Chapters Three and Six).

If the discourse analysis focuses on interaction, it is still possible to generalize but the *basis* for doing so will be different. As I explained in Section 1.1, the analyst is likely to approach every interaction as unique,

unfolding in a way which cannot be predicted and involving the development of new and context-specific meanings. So how can it be possible to generalize, to claim that the *particular* features which are identified in one interaction have a wider existence and relevance? One possibility is for the analyst to identify patterns or features which are common to many different interactions, such as sequences of 'turns' taken by speakers. This is demonstrated in the analysis by Wooffitt in Chapter Two. Like other conversation analysts, Wooffitt does not argue that such a pattern is inevitable or automatic; rather, it is an example of how speakers co-ordinate their talk/action by drawing on common knowledge. Such knowledge is sometimes referred to as 'members' methods' because it is shared by the members of a society. A somewhat different approach would be to generalize about interactions on the basis of the roles which speakers take up (such as doctor and patient). Commonalities could again be explained in terms of shared social knowledge (by arguing, for example, that people know how they 'should' speak in these roles and generally comply) or commonly held ideas and beliefs, perhaps discussed in terms of culture, or, at a deeper level, as related to cognitive processing, or enduring category memberships, such as gender.

A somewhat different approach is to consider interaction on a larger scale. It can be argued that *all* language use is interactive, addressed to others and responding to what has gone before. The interaction may address a party who is not present, as in the case of a letter. It may be directed towards someone who is imagined rather than known, such as a hypothetical political opponent or potential critic. Billig (1987) has described this form of language use as **rhetorical**. The analyst interested in rhetorical work may aim to identify the positions and arguments being addressed or countered, as a general feature of talk or other language use around a certain issue (this is one aspect of the analysis presented by Edley in Chapter Five).

An alternative to generalizing is to study something which has a *particular* significance in itself. This is the approach taken in studies using material from people or events which seem uniquely important. Examples include political or public figures like Lech Walesa (see Jaworski and Galasinski, 1998) or Diana, Princess of Wales (see Wetherell, 2001b); or historic events, like the Iran-Contra hearings in the USA (Lynch and Bogen, 1996). Similarly, a study could focus on features which exist only in one context, making the claim that they are not ephemeral, even if not completely stable, and also that their ramifications are fairly widespread. The claim would not be that the feature recurs but that it is significant and persistent. An example would be a study of racial categories employed within a *particular* society. Similarly, a researcher might look at practices which are unique to one society or historical period but (probably) relevant to a large area of that society. Historical (genealogical) analysis adds to the explanatory power of such analyses (see Chapter Seven). In these cases, the researcher would not be attempting to generalize up from component elements or particular instances but rather to describe some aspect of a whole.

To locate the research undertaken by a particular discourse analyst and understand what kind of research it is, it is necessary to look at the position she or he takes on these issues. The issues described in this section are common to much social research. The following section discusses others which are more specific to discourse analytic research.

## 1.3 Issues in discourse analytic research

I have said that discourse analysis is the close study of language in use. One debate around it involves the status of language as **topic** or **resource** (see also Chapter Five). The issue here is whether the analyst is studying talk or language *itself* or using the language as a resource for studying something else. Take the example of a psychologist who wants to study an emotion, like anger, and sets out to do so by collecting examples of people talking about their anger. How should this talk be analysed? Treating it as a resource, the psychologist might try to generalize about the causes of people's anger, how long it lasts, the physical experiences associated with it, and so on. However, an alternative approach would be to view this 'talk about anger' as the topic of the study. In this case, the aim would be to find patterns within the talk, for example, in how references to anger are located within interactions and how other speakers respond (see Edwards, 1997, reproduced as Reading Seventeen in Wetherell *et al.*, 2001, for a detailed example of this approach to the study of emotion).

Of course, it is a general feature of discourse analysis that language is not treated as information about something else but is somehow problematized. That is, the analyst approaches language as a topic. Nonetheless, the topic/resource question does arise, particularly in the fourth approach to discourse analysis which I have outlined in which the focus is on patterns across wider social or cultural contexts (see Section 1.1). For example, du Gay (1996: 148), investigating "the powerful discourse of enterprise", quotes workers in retail organizations who talk about the uniforms they have to wear. His interest is in their feelings about wearing a uniform and the conflicts they have with the company around how the uniform *should* be worn, as issues which relate to conformity and resistance. In this case, du Gay is using the talk about these issues as a resource in his analysis of identities and practices within the workplace, as part of the larger study of the discourse of enterprise.

Another debate concerns whether the analyst should investigate **process** or **content**. I have suggested that some discourse analysts are concerned with an ongoing, probably spoken, interaction, and with how speakers talk and what they do through talk. Other analysts may focus more on the content than the process of interaction. In this case, language use may be analysed as a completed whole, as if it were a finished performance. In practice, this amounts to a focus on, in the one case, the connections between consecutive utterances, and in the other, recurring elements in the body of talk (whether words, images, ideas or whatever) which are considered out of sequence. The focus, in turn, affects the way data are presented in the writing up of the project, as either a complete interactional sequence or as isolated extracts or a summary of features. This distinction between talk as an ongoing process and the content of talk is important in

ethnomethodology and conversation analysis, described below (see Chapter Two and also ten Have, 1999 for a more detailed account).

A third major dispute is between '**etic**' and '**emic**' analyses. This has been described as the difference between "using an imposed frame of reference" (etic) and "working within the conceptual framework of those studied" (emic) (Silverman, 1993: 24). At the widest level, the issue here is who determines what language, especially talk, is 'about'. Should more weight be given to the interpretation of an outsider (the analyst) or an insider (the participant in the interaction)? An emic approach is particularly associated with conversation analysis, also known as **talk-in-interaction** (this is presented more fully in Chapter Two). Conversation analysis assumes that within an interaction, such as a conversation, the speakers jointly create meaning. The talk is about what all parties make it about as the interaction proceeds, in the way they 'orient' to the previous utterances. Talk therefore becomes 'about' what speakers make it about, employing their shared members' knowledge of how to do things and how to move forward, step-by-step. In an example quoted by Schegloff (1997), a dinner-table request to pass the butter, made by a woman to a man, becomes 'about' gender as he teasingly answers 'Ladies first' but helps himself. Conversation analysts have, therefore, criticized researchers who approach a body of talk with preconceptions of what it contains (see ten Have, 1999: 102). This also has implications for the collection of data, suggesting that a researcher who sets up interviews on a certain topic and subsequently analyses them as data on that topic may be imposing an incorrect interpretation on the interactions which took place.

## 1.4 The discourse analytic researcher

The previous sections have outlined theoretical issues which are common to all social research and also some which are specific to discourse analytic research. I have said that in order to locate an example of discourse analysis and understand what kind of research it is, we need to consider these issues. In addition, they have implications for the **role** of the discourse analytic researcher. A major question here is how far the researcher can be separate from the research. This section outlines the concept of **reflexivity** which suggests that separation is impossible, and then considers the implications of the identity of the researcher for data collection and analysis. Finally, it relates the role of the researcher to ethical issues in social research.

I suggested in Section 1.2 that in one research tradition associated with positivism and postpositivism, the researcher aims to be neutral, conducting the research efficiently but exerting no bias on the processes of data collection and analysis. This neutrality is essential to one of the conventional criteria for evaluating such research, **replicability** (see Chapter Eight). According to this criterion, a different researcher (or researchers) should be able repeat a research project and obtain the same or similar results. It is based on the ideal of a detached researcher who does not influence the data or the participants, that is, a researcher who does not cause **reactivity**.

However most researchers, particularly those in the second, more social scientific tradition outlined in Section 1.2, would consider that such neutrality is impossible because the researcher and the research cannot be meaningfully separated. The argument here is that a basic feature of social research is its reflexivity, namely, the way that the researcher acts on the world and the world acts on the researcher, in a loop. If this is accepted, the researcher moves from the 'service' role of a faceless technician – which is implied in the first tradition – to a central and visible position. Detachment is impossible so the researcher's influence must be taken into account and even utilized (Hammersley and Atkinson, 1995: 19). Doing this requires the researcher to be self-aware. It involves the imagined act of stepping back to observe oneself as an actor within a particular context. The researcher attempts to understand how her or his own presence and actions influence the situation. This includes considering the relevance of the researcher's identity to the research, an important area for discourse analytic research, particularly for studies which involve interviewing.

The **identity of the researcher** becomes relevant to discourse analytic research in several ways. First, it influences the selection of the topic or research area. The researcher is likely to conduct a project which chimes with her or his personal interests, sympathies and political beliefs. This is usual in all research but is perhaps particularly true when projects are relatively small and involve only one or two researchers, as in most discourse analytic projects. The researcher's special interests and, possibly, personal links to the topic are not in themselves a sufficient basis for research, but they are a probable starting point for the project. They are not seen negatively as bias but as a position to be acknowledged. (For an interesting example of a researcher setting out her own relationship to the research topic and the participants, see the first chapter of Frankenberg, 1993.)

The researcher's identity is also relevant to data collection. For instance, it can affect an interview in several ways. The gender of the interviewer may act as a constraint on certain topics or areas of talk. Sometimes different interviewers are used in a project so that male participants are interviewed by a male researcher, female participants by a female researcher. Nonetheless, it is important not to exaggerate the common ground provided by gender. A participant may feel ill at ease with an interviewer who appears older, younger, more confident, or richer, or because of numerous differences other than gender, many of which may be conveyed in a first impression by the interviewer's appearance and accent. (Conversely, in some circumstances it may be easier for a participant to talk to someone who is different than someone who appears similar to themselves.) It can be argued, therefore, that the interviewer should not try to approach participants as an insider who shares their situation or interests, but simply as an outsider. This may also be a more honest acknowledgement of the power differences between them. (The issue of power differences is discussed in more detail in the next sub-section on ethics. Also see Croghan and Miell, 1998, for an account of a project in which the differences between interviewer and participants are noted and

the interview is acknowledged to be an unequal and apparently high-risk situation for the participants, women who are clients of welfare agencies.)

In an interview situation, taking account of reflexivity includes considering how the interviewer's questions influence the answers given. The questions may raise topics and problems which the participants would not otherwise have considered and, alternatively, discourage other topics as unsuitable. The general style of the interview is relevant here. Atkinson and Silverman (1997: 309) suggest that we live in an interview society. One consequence of this may well be that participants bring 'wrong' expectations to the interview. A researcher needs to be aware that the interviewer's manner, the questions asked and even the setting of the interview may invite different kinds of talk, such as the confession of personal feelings associated with a therapy interview or the expert knowledge given by interviewees on the television news. Participants may feel uncomfortable because the interview seems to ask them for a kind of talk which they do not want or feel able to provide (see Shakespeare, 1998: 41–59 for a discussion of different kinds of interviews and the talk which they may seem to invite or expect).

It can be argued that the identity of the researcher also influences interpretation and analysis, through the knowledge and general world view which she or he brings to the data. At the most basic level, the researcher needs to understand the language and references used by the interview participants or the writers of documents. (Some discourse analytic projects do translate material e.g. de Cillia *et al.*, 1999, but this is usually for the writing up of the project; the analysis is generally conducted on the untranslated material.) Even when the researcher and participants speak the same language, there may be barriers to understanding. For instance, if the project analyses material from another locality or time, the researcher may be unfamiliar with local idioms which are used and also with references to people and events (see Chapter Seven).

This practical point shades into the more theoretical issue of the boundary between content and relevant **context**. Some researchers, especially those working in the conversation analysis tradition, consider that contextual information is only relevant as speakers orient to it in their talk, that is, when it ceases to be background. Analysts may use material which they had no role in collecting and know only as audio-recordings of interactions which they did not witness. In this case, clearly, the analyst cannot depend on local knowledge or a similar background or experience to that of the participants. In contrast, other analysts include background information, especially about interview participants (see Mehan, 1996, reproduced as Reading Twenty-five in Wetherell *et al.*, 2001). Certain analysts would consider it important to have the double role of interviewer/ analyst, in order to bring the experience of the original interaction to the interpretation (see Edley in Chapter Five, for example). They would discuss their feelings during the interaction and their own connection to the topic and participants (e.g. Hollway, 1984, reproduced as Reading Twenty in Wetherell *et al.*, 2001).

Again, at the analytic level, taking account of reflexivity can change the interpretation of the data. In an analysis of talk, for instance, rather than

looking 'through' what is said to the topic, as if it were transparent, the analyst would be looking *at* the talk itself, as a form of interaction. To give a more specific example, imagine the talk includes discussion of racial conflict. If the talk is analysed as information about racial conflict, this assumes that the talk is transparent and the speaker is a fairly truthful, neutral source of information, including information about her or his own views and attitudes. (An analysis which takes this approach unproblematically would not usually be considered to be discourse analysis.) Alternatively, the analyst can focus on what is happening in the interview where the participant is interacting with the interviewer and doing work to present herself or himself in a certain way. The focus of the analysis is now racism as a practice that is maintained and reproduced through talk. The analyst will look, for example, at how people talk about racial conflict, what they respond to, what rhetorical moves they make. (For an example of such an analysis, see Wetherell and Potter, 1992, Chapter 7.)

Altogether, a strong argument can be made for the claim that the identity of the researcher influences the project, especially the collection of data. However, as with the collection of data through interviews, it is important to acknowledge the limits to the closeness and understanding which can be achieved between analyst and participants. I have outlined arguments that language constitutes reality and that there is no single social truth to be discovered by the researcher. The same arguments would suggest that discourse analytic projects cannot reveal the 'true' inner states of language users. As a general rule, this means that the analysis should be confined to the discourse rather than to the people who produced the talk or documents. The researcher should not aim, for example, to reveal the intentions and meaning or beliefs of speakers or writers, or to see through their words to some underlying meaning, or to uncover attitudes or beliefs of which the speakers themselves are unaware. These are all large and dangerous claims for discourse analysts to make.

The extent of the researcher's influence is inevitably difficult to assess. Reflexivity extends to the writing up of the research (see Atkinson, 1990: 6–7). At the simplest level, this is shown by the researcher adopting a policy of openness with the aim of showing her or his place within the research process. The aim is to position her or himself within the project, as part of the social world in which the research is being conducted. In practical terms this means including some self-description and accounts of her or his own relation to the topic, participants and data. Other influences and constraints can similarly be acknowledged and discussed in the written account of the project (see Section 2.5). Self-awareness is incorporated in the writing up of the research, with the researcher assessing and qualifying claims as they are made, rather than presenting them as statements of truth.

The final point to note here is that if the identity of the researcher is relevant to a project, there may be cases in which that identity is problematic. In other words, a researcher may need to accept that she or he is not the appropriate person to conduct a particular project. This can be a difficult and controversial issue. De la Rey (1997) describes a South

African conference at which black women complained about being made the object of study of white women academics. In the prevailing political and social circumstances, sympathy and goodwill could not make such projects acceptable to the category of people that they focused on and were probably intended to help. This suggests that a researcher's identity is relevant not only to the way that a project is conducted but also, at an earlier stage, to the design of the project and even to the decision to conduct it.

## Ethics

Reflecting on the project should also involve attention to the **ethics** of the research. Any researcher has ethical obligations, but these are highlighted when the researcher acknowledges her or his own presence within the research process and also abandons the claim to be discovering truth. Ethical concerns are always relevant because of the power relations between a researcher/analyst and the participants in a project. Throughout this chapter, reference has been made to 'participants' rather than 'subjects' in order to emphasize that the researcher has no rights over the other people who contribute to a research project. (This follows the recommendations of the British Psychological Society.) However, it would be an exaggeration to suggest that the researcher and participants meet as equals. In general, the researcher has more power than the participant and must be careful not to abuse it. Power can come from holding the status associated with being an academic and, supposedly, an expert. This status is particularly strong for researchers in certain disciplines; for example, it is important not to exploit the often unwarranted reverence which many lay people have for a 'psychologist'. (Mehan, 1996, reproduced as Reading Twenty-five in Wetherell *et al.*, 2001, is an account of a project in which this reverence is central to the analysis.) The researcher also has power because, as the person setting up the project, she or he has more information about it. Participants who agree to be interviewed, for example, may not anticipate the impact on themselves or others if their words are later published. Although most academic publications are not widely read outside the field, a few may be (newspapers often base articles on 'interesting' new research presented at academic conferences, for example). And, of course, non-academic participants may well have friends and relatives working within academia. For all these reasons, the responsibility is on the researcher to anticipate negative outcomes and take steps to protect participants.

Many projects will be designed to conform to the requirements of a university ethical committee or to the guidelines or codes of ethical practice issued by professional associations, such as the British Psychological Society or British Sociological Association. Researchers need to be generally guided by the obligation not to harm or distress participants. It should not be assumed that the end justifies the means: the interest or 'relevance' of a project cannot warrant dispensing with ethical requirements. As a first step, participants should be guaranteed anonymity: real names are not published and it may also be necessary to remove other information from transcripts (occupation, relationship to a public figure, and so on) which might

identify a participant. (Notice, though, that there are limits to the confidentiality which can be guaranteed, for example, with relation to certain legal issues.) Particular problems arise when participants are drawn from a small community or narrow category of people whose defining feature makes them interesting to a researcher but also readily identifiable to others. This could be the case for people in an unusual occupation, especially if they are subject to media attention (e.g. professional sportspeople). It could also apply to people in a particular category, such as a profession, who are a minority within that category, perhaps because of race or sexual orientation (e.g. black lawyers; gay and lesbian police officers). In difficult cases, if confidentiality cannot be guaranteed to the participants, the proposed research project may have to be abandoned.

A second important ethical requirement is that the researcher should obtain **informed consent** from participants regarding their involvement in the project and also the use of the data they provide. In practice, making sure that the consent is informed, that is, that participants fully understand the implications of their involvement, can be difficult. There are limits to how much detail (for example, about the theoretical background and aims) most participants will want and also to how much the researcher may be able to provide (for example, about eventual findings and their publication). But there is still an obligation on the researcher to inform the participants as fully as is practicable and to obtain consent in advance. Also, it should be noted that obtaining such consent does not absolve the researcher from other obligations. (As a practical point, note that signed consent forms should not be included in any place, such as an appendix to a thesis or dissertation, where they may be accessible to others and so identify a participant.)

Thirdly, the researcher must observe the participants' legal rights. These can be complex, especially as they relate to protected documents, such as legal or medical records, or to participants who are legally minors. For example, a researcher planning to interview children about their lives would probably need to establish in advance whether full confidentiality could be promised if participants disclosed illegal behaviours, such as drug use and under-age sexual relationships, or situations which endangered them, such as abusive domestic relationships. More generally, there can be legal issues around the ownership and re-use of data, such as audio-recordings. Despite the fact that banks of qualitative data have been established to enable researchers to re-use recordings and transcripts collected in previous projects, the copyright of an audio-recorded interview is held by the interviewee. A researcher therefore has a legal obligation to obtain permission from a participant to use the material (see Bornat, 1994: 172, for a detailed account of the legal requirements and a sample consent form; also see ten Have, 1999, for a further discussion of confidentiality and consent issues).

This and the previous section have suggested some of the ways in which a social researcher, such as a discourse analyst, can influence the research which she or he is conducting. The notion of reflexivity has been introduced and the ethical responsibilities which the researcher has towards participants in the project have been outlined.

## 2    Conducting discourse analytic research

In the first part of this chapter I offered a brief definition of discourse
analysis and outlined some (simplified) approaches to this form of research.
I will now move on to the main concern of the book, conducting discourse
analytic research. In contrast to the detailed presentations in Chapters Two
to Seven, the remainder of this chapter is a more general guide to doing
research, with sections on formulating research questions, choosing data,
transcribing recorded talk, the process of analysis and writing up research.

### 2.1 The research question

Formulating research questions involves, as the term implies, narrowing the
broad focus and wider theory of a project down to one or more questions
which the research can be designed to answer. Although it might seem
logical that the researcher will decide on a research question at the start of
the research process, in practice she or he will probably need to conduct
quite extensive background reading and library searches before formulating
it. (Carabine gives a good account of this process in Chapter Seven.) A new
researcher should also look for other studies in the broad area of the
research topic. This can be initially disconcerting. Despite the apparently
infinite range of possible research, it is usual to find that an original and
exciting idea for a project has already been used by several other people!
However, this does not mean that the proposal should be abandoned.
Although it would be undesirable to duplicate another study exactly, these
apparent duplicates will almost certainly differ from the new project in a
number of important respects, possibly including the disciplinary area. With
further work, a distinctive research question can be developed.

It is also possible for a new project to formulate its research questions in
terms of a previous project, for example, by aiming to investigate the same
topic or features in a different context. Seale (1999: 80) suggests that even
those social researchers who do not accept the principle of replicability
should design new studies to build cumulatively on previous findings. Of
course, the accumulation of data cannot be equivalent to that associated
with truth claims (see the discussion of comparison, below), but there is
still a strong case for formulating a research question to refer back to
previous research, either by building on it or challenging it.

The research question may be written as a question or alternatively as a
**hypothesis**, that is, a statement which may or may not be supported by
the findings of the research. In the positivist and postpositivist tradition, this
would be related to a truth claim: the researcher's aim would be to
establish or verify information about the world. The testing of a hypothesis
is part of a **fallibilistic** approach, where the aim of establishing
(provisional) truths is pursued negatively, by testing and challenging a
point to see if it stands up (see Smith, 1998: 344, on falsification). However
researchers outside this tradition may still formulate the research question
as a hypothesis, with the difference that the final claims made in relation to
it will be more provisional and, usually, confined to a more limited and
specific context.

It is worth noting that extra care needs to be taken if the research question is formulated in terms of a comparison. Initial thinking about a problem will often lead a researcher to formulate a question which compares findings from different situations or participants. Another possible comparison is of the past with the present, usually to demonstrate change. However, a comparative design can raise special practical and theoretical problems. Practically, it can double the size of the project, which may make it unfeasible, especially for a researcher working alone. In the case of comparisons over time, there is often no available data for the 'past' which is to be compared with the 'present' or 'recent' situation. (Carabine, in Chapter Seven, discusses this problem.)

In addition, comparative designs raise the problem of whether like is being compared with like. The researcher may be drawing on several conflicting premises and traditions in the project design. First, there is the experimental tradition of positivist research in which researchers compare two situations which differ in only one feature, in order to measure the importance of that feature. However, even in this tradition, such minimal difference is only assumed to be attainable in a highly controlled and artificial situation. Second, and relatedly, there is the assumption that any difference resides in the phenomena under study rather than in the researcher: this is a return to the notion of the neutral researcher with unimpeded access to the meaningful world. A researcher who has rejected this assumption will have to present a reasoned justification for comparing different kinds of data or data from different situations. (Note that some discourse analysts would even challenge whether it is possible to make comparisons between two different interviews!)

A discourse analytic researcher planning a comparative project will therefore need to relate the project design to the assumed connections between language, meaning, context and interpretation. It is difficult to avoid theoretical confusion at an early stage of the project, so the researcher should not settle too quickly on the final form of the research question (or questions). The aim should be to refine and reword it a number of times, as the project design is developed. This will probably involve becoming more specific, incorporating terms from the theoretical background to the research, and referring specifically to the data for the project. It is possible in some cases that the final form of the question will not be decided on until the data have been collected.

The final wording of the research question needs to be very precise and will probably only emerge gradually. The researcher probably begins with rather general interests, for example in terms of topic area (such as race or gender) and material to be analysed (for example, from particular documents or interview participants). These interests are not in themselves research questions but will need to be refined. Silverman (1993: 2) draws a useful distinction between a problem that is discussed in the world around us, like homelessness, and a researchable topic. He points out that many new researchers make the mistake of attempting to research problems or answer questions which are impossibly large and wide-ranging. This is not, of course, a bad place to begin thinking: the mistake is to stop at this point

instead of continuing to explore the possibilities and complexities before settling on the final form of the research question.

## 2.2 Discourse analytic data

It is likely that some prospective analysts, especially students, may have become interested in discourse analysis, and possibly turned directly to this section, because they already have access to a body of material, perhaps documents or recorded interviews, which they hope will provide the basis for some interesting research. This is a common situation and can lead to a successful project. However, it is important to consider exactly what constitute data, and not to approach the analysis with the assumption that whatever material is to hand somehow contains revelations, like gold within the dross, which can be extracted by a conscientious researcher working without preconceptions or any particular focus. Discourse analysis is not a neutral, technical form of processing but always involves theoretical backgrounding and decision making (see, for example, Potter, 1996: 130).

The above paragraph refers to a prospective researcher who already has a body of *material* rather than data. It is easy to assume that the data for discourse analysis already exist, and that new data are constantly being generated. In the world outside academia, newspapers are being printed and discarded, officialdom is producing documents, people are talking in a million contexts. Surely, it seems, all that needs to be done is to collect up the data and begin the analysis! However material only becomes data through certain considered processes, including selection. What count as data will depend on the researcher's theoretical assumptions, about discourse and also about the broad topic of the research.

I have said that one of the processes by which material becomes data is **selection**. There are several different criteria for selecting a **sample**. Most quantitative research, particularly surveys, uses a sample which is large enough to be representative of a population as a whole (that is, to include all the features which might be of interest) and which also permits generalization, based on assumptions about the frequency and regularity of features or phenomena. It is part of the special usefulness of quantitative work that large amounts (quantities!) of data can be analysed and summarized. Chapter Three presents some examples of quantitative discourse analyses, using computers.

In contrast, the analysis of qualitative data, including qualitative discourse data, is relatively inefficient and labour-intensive. It is often difficult to put the data into a succinct form for either analysis or presentation. The researcher is therefore likely to use a much smaller sample. This may, nonetheless, be designed to be as broad and inclusive as possible. For an interview study, participants may be selected to provide a balanced sample in terms of main population categories, such as gender. On the other hand, they may be selected *because* they belong to a particular limited category. For example, Wetherell and Potter (1992) sought participants who shared the same class and ethnicity. For a study of whiteness as a racial category, Frankenberg (1993) chose to interview white, female participants. These could be seen as examples of a sample selected to represent a 'specimen perspective' (ten Have, 1999: 50). The selection is not made to represent a

population as a whole but a particular category within it: the aim is therefore to find participants who are 'typical' rather than exceptional.

A similar argument could be made for the selection of documents, and in both cases the claim that the selection are specimens could be supported by reference to statistics. Alternatively, the analyst might select documents because they are not broadly representative but highly specific, claiming that they are worthy of analysis because, for example, they are associated with powerful or well-known people. For example, Edwards and Potter (1992: Chapter 3) analyse newspaper articles and extracts from *Hansard* referring to the then-Chancellor of the Exchequer, Nigel Lawson; Stubbs (1996) analyses writings of the founder of the Boy Scout and Girl Guide movements, Baden-Powell; Jaworski and Galasinski (1998) analyse the talk of the former President of Poland, Lech Walesa. In these cases, there is an implied or stated argument that the origin or context of the material relates it to wider social practices. This can be similar to the argument of relevance as a criterion for evaluation (see Chapter Eight).

A somewhat different basis for selection could be justified in analyses which assume that language use reflects the knowledge or skills shared by members of the same culture. ('Culture' may be rather loosely defined here: it does not necessarily mean a distinctive national culture or any neatly bounded grouping.) For example, conversation analysts (see Chapter Two) assume that speakers hold in common a tacit knowledge of certain rules for conversational interaction and can recognize departures from these, such as when a conversational action does not receive an expected response. Discursive psychologists (see Chapters Four and Five) may assume that shared knowledge includes alternative understandings and constructions, for example, of the connection between mind and body (Chapter Four) or of what constitutes a 'good' couple relationship (Chapter Five). Following this, an analyst might select a relatively small sample of data with the justification that the patterns revealed in it indicate knowledge which is shared by other members of the culture.

Selection also extends to which **features** of material are relevant to the analysis and are, therefore, part of the data. For example, a study which focuses on language and assumes it is a (loose) system (as in the first approach to discourse analysis described in Section 1.1) might investigate patterns in its component parts, perhaps using computerized searching techniques such as concordancing (see Figure 2 in Chapter Three for an example). Following linguistics and sociolinguistics, Chapters Three and Six both select such features. (Another example is in Chapter 4 of Stubbs, 1996; a study which analyses language as part of an investigation of sexism.) The analyst might also be interested in features other than the words of the documents, such as the general appearance and layout. Again, Chapter Six provides an example of an analysis which considers these features. (Layout is also a feature in the analyses of local government documents by Stenson and Watt, 1999, and the analysis of a newspaper page by Kress, 1994.) In such cases, the data probably consist of original or photocopied documents rather than, say, a word-processed copy of the text.

In defining data, the discourse analyst must distinguish between the data themselves and **context** or **background information**. Some social researchers, such as ethnographers (see Hammersley and Atkinson, 1995, for an overview), do not make such a distinction: they attempt a form of total analysis in which it can be said that everything the researcher observes forms part of the data. Discourse analyses may be conducted as part of a larger ethnographic study (for example, du Gay, 1996; Back, 1996). Alternatively, some discourse analysts collect detailed background information to inform their analyses (see Mehan, 1996, reproduced as Reading Twenty-five in Wetherell *et al.*, 2001). Background information may also serve to define the data. For example, if the analyst's aim is to identify discourses within a certain class of official documents (see Chapters Six and Seven; see also Jagger, 1997, which analyses government talk and texts), those documents will only acquire their status as a result of their use within administrative systems, which may need to be established by the researcher.

Some analysts consider that the talk constitutes the data and no other information is needed. For a researcher interested in discourse as interaction, such as a conversation analyst (see Chapter Two), the focus is on what happens within the interaction. Background information is largely irrelevant and may even distort the interpretation. For example, it can be argued that including in the data the information that, say, one speaker is female and one is male amounts to a claim that gender is relevant to the interaction, when perhaps it is not. Is this necessary background information which should be provided, for example, about an interaction between a counsellor and a client or a doctor and a patient? (See below and Section 2.3 on transcription.) Similarly, there can be arguments for or against the relevance of other membership categories, such as age group, race, occupation and class. It could also be argued that the context and/or the physical location of the interaction are relevant to the data collection and analysis because they establish certain expectations for the kind of talk that is appropriate (see Shakespeare, 1998: 43–59). What counts as part of the data will depend, finally, on the particular project and the theory underpinning it.

Another process by which material becomes data is through **transcription**. This is discussed in more detail in the next section (see also Chapter Two) but some broad points can be made here. First, many (though not all) discourse analysts who study audio-recorded talk would not regard the talk itself as data without a process of further selection, through transcription. Furthermore, that basic material may provide different kinds of data, again depending on the focus of the analysis. For example, some social psychologists have investigated recurring patterns of word use, imagery and ideas within talk which have been defined as 'interpretative repertoires' (see Chapter Five). The exact definition is not necessary here (see Potter and Wetherell, 1987; Wetherell, 1998) but the approach is broadly similar to the third approach to discourse analysis outlined in Section 1.1. For a researcher interested in repertoires, the data are usually a body of interview material from different speakers. The researcher probably works from transcripts which record the words spoken

by consecutive speakers but little further detail (see Gill, 1993; Marshall and Wetherell, 1989).

In contrast, for a conversation analyst who is interested in talk as interaction (like Wooffitt in Chapter Two), the data include not only the features of talk which are common to printed language (basically, the words) but also other aspects of the spoken interaction, such as the sequential organization of utterances from different speakers, including interruptions and pauses. Consequently, if the researcher works from transcripts, these will need to be much more detailed. The researcher will probably use a smaller quantity of data and will analyse each interaction separately. The analysis may require that the talk is 'naturally occurring' rather than collected through research interviews.

The description of language, usually talk, as '**naturally occurring**' has various possible meanings (see Chapter Four). In the most idealized form, it would probably refer to informal conversation which would have occurred even if it was not being observed or recorded, and which was unaffected by the presence of the observer and/or recording equipment. However, there are both technical and ethical problems with the covert recording necessary to obtain material of this kind (see Section 1.5 on ethics and informed consent). The closest satisfactory approximation is probably material which is obtained when the researcher has permission to record and the participants have become sufficiently accustomed to the presence of the recording equipment to act as if it was not operating (e.g. Goodwin, 1990). Of course, the researcher still has ethical obligations to the participants which may make some of the material unusable.

This approach assumes that naturally occurring talk is talk which is informal and occurs outside the context of situations with a declared purpose and particular venue. It can be argued that talk also occurs naturally in more structured situations such as courtrooms (Drew, quoted in Chapter Four), medical consultations (Coupland and Coupland, 1998), telephone calls to service providers (Sacks, quoted in Chapter Two), and even counselling sessions (Silverman, 2000; Edwards, 1997: 154). The naturalness here would not necessarily refer to whether the speakers were relaxed or unselfconscious, but to the talk being uninfluenced by the presence of the observer/recorder. (Again, the same ethical issues would apply.) One advantage for the researcher of using talk from these more structured situations might be that it is clearer in some ways what the talk is 'about'. However, this point is problematic. Medical consultation is not necessarily 'about' the medical problem under discussion or even about doctor/patient relations. But if the researcher is interested in interaction (as in my second approach to discourse analysis), talk from these situations might be taken as examples of a certain kind of interaction and could be analysed to show, for example, common features and patterns (e.g. Coupland and Coupland, 1998).

These various kinds of supposedly natural talk are usually contrasted with more conventional research interviews in which the researcher/ interviewer attempts to initiate talk which is 'about' something by conducting interviews specifically for the purpose of the research, usually working with a prepared list of questions or discussion topics. Discourse

analysts who are interested in interaction may challenge the use of research interviews on two grounds. The first is that the interview is unnatural because the interviewer controls the interaction and influences the talk. The second is that the researcher incorrectly assumes that the talk is about the official topic of the interview, imposing her or his own interpretation on the talk (see the discussion of emic/etic in Section 1.3). Some researchers have attempted to avoid the first of these problems by setting up group discussions, usually in small groups, rather than one-to-one interviews (see Edley and Wetherell, 1997; Augoustinos *et al.*, 1999). Unfortunately, this can raise new problems of 'about'-ness: perhaps the interaction is now about group relations rather than the apparent topic of discussion? It can also raise new ethical problems related to confidentiality and respect for participants, especially if the topic of the research is a sensitive one. The researcher has an obligation to respect participants and avoid causing them distress, but one participant may feel no such obligation to another in the group and be careless of her or his feelings.

As this discussion has shown, discourse data vary widely. For every project, the researcher must establish the justification for the data being used, even if this is done cryptically through reference to a previous study. To conclude this section on discourse data, I will give some brief descriptions of published examples. In the first, as part of a study of nationalism, Billig (1995) analyses the main British newspapers on a single day. The data include the content of the news stories, the order and form of their presentation (for example, which pages they appear on), and also the language used. There is also some analysis at the level of the component parts of language, specifically of the deistic use of words like 'we' and 'our' (deixis means, loosely, that a word has no meaning except as it refers to another word: see Billig, 1995: 106). The nature of the data for this project follows from the ways in which both nationalism and discourse are theorized. Part of the researcher's task is to establish a connection between the broader topic or argument of the study and what are used as data, that is, between nationalism and particular features of newspapers. It is important to note here that this link is *not* created by treating language as transparent or reflective, and therefore as 'true' descriptions of phenomena. In this case, it would not be an adequate form of analysis to search the newspapers for all the references to 'nationalism' (although such a search might be a useful first step in coding material and choosing pieces for further analysis: see Section 2.4; also Potter and Wetherell, 1987: 167).

A second published example by Gill (1993) analyses interview transcripts in a study of equal opportunities for women in broadcasting, particularly radio. The interviews are conducted specifically for this study (in contrast to naturally occurring language, discussed above) and the researcher chooses to interview not women but men. The justification for this derives from a particular view of employment and of the factors which influence whether or not women are employed. A third study by Helleiner (1998) analyses documents rather than interviews. The data used here are parliamentary speeches and official reports on travellers and traveller children. In this study, part of the researcher's task is, again, to relate the data to the topic of concern. In both studies, the arguments which do this are subtle. This

theorizing of the relationship between the general topic, the definition of discourse, and the data to be analysed cannot be taken for granted but needs to be established for each study.

This summary of issues around data and data collection, particularly as these relate to interviews, may seem frustrating, especially to a new researcher, since it raises many problems without offering firm solutions. It is always important to remember that no research is perfect, so every study must justify itself coherently and acknowledge (probably rather briefly!) the criticisms which can be made of it. This is particularly true about the selection of data. This will always be open to criticism and researchers working from very different premises may ultimately find each other's work and findings unacceptable. A project which is praised in one research tradition may be rejected by academics working in another discipline or theoretical area.

## 2.3 Transcription

An important aspect of data collection and selection in research involving talk is **transcription**. This is the process which, perhaps bizarrely, turns that talk into a document, a **transcript**. Doing transcription is a time-consuming process (estimates of how long it takes to transcribe an hour of recorded material range from about four hours for the simplest transcription of an audio-recording to perhaps more than twenty hours for detailed transcription of video). The best way to learn about it is probably to try doing it (see Chapter Two for a suggested activity). However, the following activity is intended to introduce some of the many issues around transcription.

---

## Activity 1

Below are examples of transcript from five published discourse analytic projects. Look over them and try to answer the following questions for each transcript. The questions in Part A are concerned more with the practical process of transcription; those in Part B with the theoretical approach underlying it. You will probably want to read the discussion of Questions 1–4 before proceeding to Questions 5–6.

### Part A

1  How many speakers are there?

2  How does the transcript differ from standard written text?

3  In addition to the words spoken, what details does the transcript include?

4  What extra information do you need to understand the transcript?

### Part B

5  In what respects is the transcript a construction rather than a neutral record of talk?

6  What does the transcript show about the form of discourse analysis being used?

## Extract 1

705:  *nie zauważyłem że hydra komunistyczna odrasta że pajęczyny*
706:  *tworzy że wlazło mi w całą gospodarkę i to to musi powodować*
707:  *niezadowolenie.*
705:  'I didn't notice that the communist hydra had been growing back
      that it spins webs
706:  that it has gone [me] into the entire economy and that must cause
707:  discontent.'

235:  *dzisiaj przymus wymaga przeciwstawić się (..) komunie która głowę*
236:  *podnosi*
235:  'today there is an obligation to oppose the commune which
236:  raises its head.'

<div align="right">Jaworski and Galinski, 1998, 'The last Romantic hero: Lech Walesa's<br>image-building in TV presidential debates', <em>Text</em>, vol.18, pp.524–44</div>

## Extract 2

Tony:        W't's 'e g'nna <u>do</u> go down en pick it <u>up</u> lat<u>er</u>? er
             something like (      ) [well that's <u>aw</u>]:ful
Marsha:                             [H i s <u>friend</u> ]
Marsha:      Yeh h[is <u>friend</u> <u>Stee-</u>      ]
Tony:             [That really makes] me ma:d,

             (0.2)

Marsha:      ʹhhh <u>Oh</u> it's disgusti[ng ez a <u>matter</u> a'f ]a:ct.
Tony:                            [P o o r J o e y,]
Marsha:      I- <u>I</u>, I <u>told</u> my <u>ki</u>:ds. who <u>do</u> this: down et the <u>Drug</u>
             Coalition ah <u>want</u> th'<u>to:p</u> back.h ʹhhhhhhhhh ((1.0))
             SEND OUT the <u>WO</u>:RD.hhh hnh
             (0.2)
Tony:        Yeah.
Marsha:      ʹhhh <u>Bu</u>:t u-hu:ghh his <u>friend</u> <u>Steve</u> en <u>Brian</u> er driving
             up. <u>Right</u> after:: (0.2) <u>school</u> is out. En then <u>hi</u>'ll
             drive do:wn <u>here</u> with the:m.

<div align="right">Schegloff, 1997, 'Whose text, whose context?', <em>Discourse and Society</em>,<br>vol.8, pp.165–87</div>

## Extract 3

Shell:       I'm quite, I'm certainly in favour of a bit of Māoritanga. It is
             something uniquely New Zealand, and I guess I'm very
             conservation minded (yes) and in the same way as I don't like
             seeing a species go out of existence I don't like seeing (yes) a
             culture and a language (yes) and everything else fade out.

Williamson:  I think it's important they hang on to their culture (yeah) because
             if I try to think about it, the Pākehā New Zealander hasn't got a
             culture (yeah). I, as far as I know he hasn't got one (yeah) unless
             it's rugby, racing and beer, that would be his lot! (yes) But the
             Maoris have definitely got something, you know, some definite
             things that they do and (yeah). No. I say hang onto their culture.

<div align="right">Wetherell and Potter, 1992, <em>Mapping the Language of Racism</em></div>

## Extract 4

```
1   D:  But er despite all of those things, in the majority
2       of people the disease does come back
3       (0.8)
4   D:  even from the beginning.
5   P:  Yes
6   D:  And: (0.4) if it does come back we can try
7       other drugs which may control it for a little while
8   P:  mm um
9   D:  but generally all that you can try and do is control
10      the symptoms.
11  P:  Yes mm.
12  D:  Uhm, the first time gives us the best chance
13      for a longer (0.5) survival hopefully long term
14  P:  hhm
15  D:  but the odds are generally against that.
16  P:  Yes um (0.4)
17  D:  But if we do nothing for these sorts of diseases it
18      kills you within a couple of months.
19  P:  Yes
```

Seale and Silverman, 1997, ensuring rigour in qualitative research',
*European Journal of Health*, vol.7, pp.379–84.

## Extract 5

*(nearing the end of the consultation)*

```
1   Doctor:     but er (.) I think things are OK and I think they'll
                improve
    Patient:    yes
    Doctor:     OK?
5   Patient:    well what the hell do I expect (.) I must be mad!
    Doctor:     (amused, reassuringly) I don't think you're mad
    Daughter:   chuckles
    Patient:    (laughs) I'm just a crabby old =

    Daughter:   = a thirty year old er brain inside an eighty year old
10              [(laughs)
    Doctor:     [that's right (.) it's frustrating isn't it?
    Patient:    (frustrated) I can't [get on
    Daughter:                        [oh but she's better than she says
    Doctor:     she is much better
15  Daughter:   yes
    Doctor:     OK
    Patient:    yes (warmly) thanks [ever so much
    Doctor:                         [that's alright (2.0) look after
                yourself
```

Coupland and Coupland, 1998, 'Reshaping lives, c9onstitutive identity
work in geriatric medical consultations, *Text* vol.18, pp.159–89.

## Discussion of Questions 1–4

### 1   How many speakers are there?

In most transcripts the speaker is indicated on the left, like a playscript. So there are two speakers in Extract 2 (Tony and Marsha), two in Extract 4 (D and P) and three in Extract 5 (Doctor, Patient and Daughter). However in Extract 3 a different convention is used: the short comments of a different person, presumably the interviewer, are transcribed in brackets in the body of the talk of the main speakers. And in Extract 1 no interviewer or other person is included, although there was presumably at least one present (the title of the article indicates that the talk is from a television debate). It may be that no one else spoke during the stretches of talk which are transcribed here.

Other people may also have been present during the talk in other transcripts (for example, perhaps a researcher was in the consultation room in Extract 5). This raises the issue of whether the presence of other people is relevant: do they only count as part of an interaction and participants in the research if they speak?

### 2   How does the transcript differ from standard written text?

In these examples, Extract 1 is closest to standard written text, followed by Extract 3. Although a transcript is talk which has been in some way 'written down', it does not necessarily use standard sentences with punctuation such as commas and full stops, for the good reason that most of the time we do not speak in such sentences. (Notice, however, that elements of standard punctuation, like full stops and colons, may be used in a transcript to indicate other features of talk, as discussed in the answer to Question 3, below.) Extract 1 probably uses these conventions because the talk has been translated, and even here the sentences have been modified slightly, using layout and square brackets. Extract 3 organizes the talk into sentences but includes some of the irregularities typical of ordinary unscripted talk (for example, "I'm quite, I'm certainly" and "I, as far as I know he hasn't got one").

Notice that Extract 5 uses some conventions from a particular kind of writing, a dramatic script. Notes in brackets give information about the context ('nearing the end of the consultation'), sounds made ('laughs' and, not in brackets, 'chuckles') and also a speaker's feelings and speaking style ('amused, reassuringly', 'warmly') but whereas in a script, this information would be include as *instructions* to the director and actors, here it is included as description (see the discussion of Question 5 below).

As you are probably aware, it can be disconcerting to read a transcript for the first time, especially a transcript of your own talk, precisely because of the differences between spoken and written language. This may need to be discussed with participants if transcripts are to be returned to them for checking (see Chapter Eight): they may feel that the transcript gives a poor impression of them, as incoherent or poorly educated!

### 3   In addition to the words spoken, what details does the transcript include?

The most obvious extra detail is included in Extract 1 which presents the talk twice, in different languages, with the words in the original language (Polish) translated line by line.

Extracts 2, 4 and 5 attempt to differing degrees to record *how* the words were spoken, for example, by showing emphasis with italics, block capitals and underlining. Extract 3 does this to a lesser extent using spelling ('yeah' instead of 'yes', for example) and standard punctuation (an exclamation mark) and some phonetic conventions (the marks over some vowels, for example, in Māoritanga). All of the examples, except Extract 4, use spelling to indicate the contracted forms (didn't, it's, I'm) which are more commonly used, following complex conventions, in spoken than written English. Extract 5, as mentioned, also has additional notes similar to those used in a script to describe how something was said. Extract 2 and, to a lesser extent, Extract 1 use less obvious notations, like colons to indicate that a sound is drawn out. Notice that very detailed transcripts, like Extract 2, may borrow some conventions from phonetics, but they do not draw fully on any phonetic system, such as those used in dictionaries to indicate pronunciation; in other words, the attempt to indicate how participants speak is always partial and therefore selective.

Sounds other than words are indicated through the script-like directions in Extract 5 and spelled-out forms (oh, er, um, hhm) in Extracts 2, 4 and 5. In addition to the sounds made, Extracts 2, 4 and 5 indicate pauses and overlapping of talk. These are additional details about the sequence of the talk, which is recorded in all the transcripts using the left-to-right and top-down conventions of written English but notice that this is interrupted in two examples. In Extract 1, the talk in the original language (Polish) is followed directly by the English translation. In Extract 3, the speaker, Shell's, words do not follow from Williamson's but are taken from a different interview. In Extract 3, therefore, the top-down presentation indicates not sequence but *similarity*, a point I will return to.

There is other information which could be included in a transcript. The list of transcription symbols provided in Chapter Two indicates some possibilities, such as pace and intonation (the rising or falling tone noticeable, for example, in questions). There could be more information about the context or situation: for example, nothing in Extract 2 indicates that this is part of a telephone conversation. Some transcripts include details of actions, body language (such as smiles or grimaces) and gaze (see Swann and Graddol, 1994, for an example of the latter using notation and Suchman, 1997, for an example using words). West (1996) quotes from a transcript which distinguishes the different 'voices' used by a doctor who variously addresses a child patient and the parent and makes 'notes' into an audio-cassette recorder during a consultation.

These points raise the question of who the transcript is made *for*. A researcher working alone might include extra information in a transcript to assist her or his own analysis (this is similar to the 'metadata' described in Chapter Three). A professional transcriber or a researcher working as part of a team might add detail about context or speakers to help a different

analyst understand what was happening at a certain point in the talk. (I recently returned to an audio-recording to help me understand its transcript and, after some time listening and re-listening, realized that the stilted quality of the interview, extended pauses and, on the recording, occasional poor sound quality and strange background noises were all the consequence of the participant having been interviewed while he was eating lunch!) On the other hand, a transcript included within a publication (like the five examples above) might have extra detail added or taken away in order to improve the readability. An example is the line numbering included in Extracts 1, 4 and 5. Similarly, arrows or other margin symbols can be used to direct the reader's attention to key points.

In addition to what is included in the transcript itself, other information about the talk is likely to be given in the text and elsewhere. Notice how much information we are told about the talk in the titles of the publications containing Extracts 1 and 5, for example. Footnotes could have been used to explain local terms and references, such as Māoritanga, Pākehā and Maori in Extract 3, if the analyst had judged that this was necessary. Perhaps it would also be useful to know that 'rugby, racing and beer' is a well-established expression with particular associations. And so on: the decisions about what to include and explain blur the distinction between transcription and interpretation.

### 4   What extra information do you need to understand the transcript?

The above discussion indicates some of the information which will have to be supplied to enable a reader to understand a transcript: a list of transcription symbols, if these are not obvious (see Chapter Two, Section 4, for a list of the best known symbols, devised by Gail Jefferson); supplementary information in the transcript itself or the accompanying text. Some social researchers may present transcripts without an accompanying analysis as a way of giving a voice to participants, but this is a different research aim. Discourse analysis requires that even the most readable transcript is presented only as part of a larger research text.

### Discussion of Questions 5–6

### 5   In what respects is the transcript a construction rather than a neutral record of talk?

I have already said that a transcript is a construction in so far as it cannot be a total record of talk. Each of these examples therefore constructs the talk through what is included or excluded.

As a brief summary, none of the extracts includes a record of actions and body language, constructing the talk as more meaningful. They exclude details about location which other analysts might consider relevant. (Compare, for example, Suchman, 1997, and Goodwin, 1995, analysing talk as part of workplace interactions). These transcripts also variously omit: other participants in the talk (in Extract 1); information about how the words are spoken (probably all to a greater or less extent, with Extract 2 least); details of the sequence, including pauses and overlap (Extracts 1, 3 and perhaps 4); contextual information, such as the date, occasion and

previous talk (though some of this is including in the accompanying texts for particular extracts); information about the speakers, such as gender and age, and so on. All the features discussed in answer to Question 3 are omitted from one or other of the transcripts. All of this selection works to construct certain features as meaningful.

The analyst's interpretation can also be reinforced by the inclusion of extra information in a transcript. For instance, all of these extracts, except Extract 1, give additional information within the transcript about the speakers and their relationships and roles. As a general point, notice that names, even when pseudonymous, can convey information about, for example, gender and, more subtly, class and ethnic background. The use of first names only may imply that the participants have an intimate relationship, family names only that they are speaking in a more formal role (see Billig, 1999). The identification of speakers as Doctor and Patient, or D and P as in Extract 4, not only reports these roles but suggests they are salient: in other words, the suggestion is that the identity of 'doctor' is more important here than the speaker's gender or any other identity. The role, therefore, constructs the talk as a certain kind of occasion and could be criticized for imposing this interpretation. We could ask: are the participants orienting to this?

In addition, more subtly, the transcript is a construction in that it *is* a transcript i.e. it is written rather than spoken; it selects a beginning and end point and it confers a special significance on *this* talk by discussing it in an academic research text. In other words, it is a selection from a much larger body of available material.

## 6   What does the transcript show about the form of discourse analysis being used?

As discussed above, a transcript emphasizes certain features of talk simply by recording them. By seeing what is included or excluded, we can draw some conclusions about the analyst's focus.

Transcripts which record in detail how people speak, or record sounds which are not words, suggest that these features are meaningful. In contrast, a transcript which presents talk as similar to writing, punctuating it as complete sentences, for example, suggests that the analyst is concerned with the meaning which resides in those features which talk shares with writing i.e. words. Whereas Extract 3 records the exact words used by speakers, Extract 1 suggests that meaning can be expressed in words, but the same meaning can be expressed in different ways, in slightly different words, or even in a different language. (Alternatively, it could be argued that Jaworski and Galasinski, from whom Extract 1 is taken, do not subscribe to the latter position but consider translation an inferior substitute, and this is the reason that the reader is not given just the translation but also the original talk, in Polish.)

A transcript which identifies participants in terms of roles (doctor and patient) is likely to focus on patterns in language in use associated with those roles and generalize accordingly. Transcripts which present the talk of different speakers in sequence suggest that the analyst is interested in meaning which is created through interaction, that is, jointly constructed.

Again, the focus could be mainly on the words used. This might require an extended sequence of talk (see Chapter Five). Alternatively, depending on the detail which is included, the focus could be on the conversational 'moves', such as interrupting, speaking over someone else, or failing to answer (Extracts 4 and 5 but particularly Extract 2) as part of social practices.

Transcripts which present parallel examples, such as Extracts 1 and 3, suggest that these contain patterns. From Extract 1, we can guess that the analysts are interested in patterns within the talk of a single speaker, whereas Extract 3 presents a pattern across the talk of different speakers, perhaps emphasizing the use of shared resources (see Chapter Five).

A final, and controversial, point to note is that an epistemological argument may be subtly conveyed by the sheer *quantity* of detail included in a transcript. On the one hand, an extremely detailed transcript may suggest that the analysis is derived directly from the data, downplaying the role and influence of the analyst, as in the positivist and postpositivist tradition outlined in Section 1.2. Relatedly, by making the transcript relatively difficult to understand, the detail may suggest that the analyst is a detached and objective technician rather than the involved interpreter implied by reflexivity (see Billig, 1999). On the other hand, an easily read transcript which uses standard writing conventions can work as a claim for the common-sense, uncontroversial nature of the interpretation presented.

### Concluding comments

A transcript is a written-down version of what has been said, and in some cases done. The simplest form is probably a record of the words spoken. It can, in rare cases, be made from 'life' (West, 1996, quotes some examples). More usually, the transcript is made from an audio or video-recording which can be replayed and paused to allow the inclusion of more detail than an observer or participant could record, or even notice, in 'real time'. One advantage of the transcript is, therefore, the detail it contains. Another is its relative convenience and portability. Most analysts find they can work more quickly and efficiently from a transcript than by repeatedly listening to or watching recordings. However, it should be noted that working from recordings *is* a viable alternative, especially for preliminary analysis, such as coding (see Section 2.4 and Potter and Wetherell, 1987: 173). Furthermore, some researchers consider that transcripts should always be used in conjunction with the original recordings on which they are based, as a point of good practice. Others, more prosaically, may choose to work directly from recordings because the production of transcripts is, inevitably, extremely time-consuming. The ideal might seem to be to pay someone else to do the transcribing, but in practice this is expensive and often unsatisfactory.

Because, ultimately, transcription is a matter of selection and is therefore itself part of the analysis, not a separate stage (see Ochs, 1979: 44; also Riessman, 1993: 56–60), the researcher may need to refine transcripts made by someone else. (In Chapter Two, Wooffitt compares a transcript produced

by a professional transcriber with the revised version he has produced for analysis.) A further use of transcripts is for the presentation of data to others. This has implications for the evaluation of research (see Chapter Eight). However, some researchers analyse from the recordings and then later transcribe short sections for presentation.

Transcripts involve selection, and precisely what is selected will depend on the theory and aims of the research project, as discussed above. It is important to emphasize, therefore, that there is no one way to do transcription, any more than there is one theoretical approach to discourse analysis. Correspondingly, there is not one set of transcription symbols. Although the best known are those devised by Gail Jefferson (a list of these is included in Chapter Two) most researchers using these will probably draw on a selection only. Other researchers will develop their own systems, perhaps based on standard written English. They may also use symbols to mark particular features which are important to their own projects. For a new analyst, the priority will be to decide which features of talk are significant and justify the decision, probably with reference to other published research.

This might suggest that transcripts are only usable for the particular project for which they are produced. However, in practice, researchers may conduct new analyses of transcripts produced for earlier projects. (Some research councils and other bodies compile archives of material for this purpose.) Researchers may also analyse transcripts produced for other purposes, such as the British parliamentary record, *Hansard* (see, for example, Activity 1 in Chapter Seven; also Edwards and Potter, 1992: 116). In these cases, there may be the problem that the transcript lacks details which the researcher would have wanted.

This might suggest that the ideal transcript is one which includes *everything*. However, as Activity 1 in this chapter has shown, this ideal is not attainable. First, there is the theoretical point which has been repeated throughout this chapter, that language is not assumed to be transparent or reflective. Logically, therefore, a transcript as a form of language cannot neutrally reflect the talk or interaction which it purports to record. It is itself a construction. Second, there are two practical objections to very full transcripts. As has already been mentioned, producing these involves unnecessary expenditure of time and money. In addition, excessive detail makes a transcript very large and extremely difficult to read and to work with.

The conclusion must be, therefore, that it is neither practical nor desirable to produce highly detailed transcripts as a substitute for identifying the focus of the study. Obviously it is reasonable and desirable to expect that the extended process of analysis will identify features which went unnoticed during transcription, itself an extremely useful first stage, especially if carried out by the analyst, but it is not feasible to treat transcription as a substitute for thinking and making decisions about the material. Approaches based in conversation analysis, also known as talk-in-interaction, do require the extremely detailed transcripts, but producing these may not be the first step in the analysis (see Chapter Two; also ten Have, 1999: 95).

There are other issues with transcription than those related to the level of detail. Ochs (1979: 46) points out that the convention by which consecutive utterances are placed below each other, like a play script, encourages the assumption that a particular utterance is a response to the one immediately preceding it. She suggests that this may not be the case when an interaction involves several people, or a child: "Young children frequently 'tune out' the utterances of their partner, because they are otherwise absorbed or because their attention span has been exhausted, or because they are bored, confused, or uncooperative". This comment may seem equally applicable to adults!

Decisions about the detail and the forms of notation used in the transcript are ultimately based in the theoretical approach. They can be challenged and disputed. An analyst may distort the meaning of the original talk by cutting an utterance short or extending it, in the same way that a news editor might change the import of a politician's comments by editing a film in a certain way. More subtly, an interpretation may be affected by the transcription (see the discussion in Chapter Eight of Seale and Silverman, 1997: 381–2, which includes Extract 4 above).

On the other hand, the elaborate notation of details which are not relevant to the analysis can make a transcript difficult to read. It can also reduce the amount of material which can practically be analysed and also reproduced in the writing up of the analysis (see Section 2.5). If the analyst's aim is to identify sets of related terms, such as discourses or interpretative repertoires (as in the third simplified form of discourse analysis described in Section 1.1), then she or he probably does not require the same level of detail as an analyst, like a conversation analyst, who is studying interaction (see Chapter Two).

A transcript therefore constructs a certain version of the talk or interaction which is to be analysed. This does not, of course, mean that it is false or misleading, but simply that it is not neutral. It selects out the features which the analyst has decided are relevant, that is, what the analyst counts as data.

## 2.4 The process of analysis

The title of this section is perhaps misleading in that it suggests that analysis is a distinct and separate activity. However, as I have already emphasized, the process of analysis includes transcription and to some extent also the definition and selection of data. Describing analysis in the abstract is necessarily inadequate, which is why the central chapters of the book, Chapters Two to Seven, work through examples from specific projects. However there are some useful general points which can be made for researchers new to discourse analysis.

The first of these is that most of the discourse analytic approaches presented in this book are qualitative. The nature of the analysis is therefore relatively open-ended and also circular, or **iterative**. The researcher is looking for patterns in the data but is not entirely sure what these will look like or what their significance will be. She or he must therefore approach the data with a certain blind faith, with a confidence that there is something there but no certainty about what. Conducting

analysis involves going over data again and again, whether listening to recordings or reading transcripts or documents, noting features of interest but not settling on these. It involves working through the data over quite a long period, returning to them a number of times. Data analysis is not accomplished in one or two sessions. (These points are also made strongly in Chapter Seven, including in the useful guide to conducting one form of analysis, a Foucauldian genealogical discourse analysis.)

As possible patterns emerge, it is useful to note them but continue searching. Eventually there will be a range of possibilities to explore further. It will almost certainly be necessary to focus on some at the expense of others, leaving unfinished avenues for later exploration. Discourse data are 'rich', which means that it is probably impossible to reach a point where the data are exhausted, with nothing more to find in them because the analysis is complete.

Basic to this process is some kind of sorting and categorizing to identify patterns. The term '**coding**' has conventionally been used for the classification of research data into categories (Seale, 1999: 102–5). The common feature of software packages for data analysis is that they facilitate coding, for example, making it easy to underline or otherwise mark up sections of data (see Seale in Silverman, 2000, for a discussion of the use of computers to analyse qualitative data). Potter and Wetherell (1987) recommend broadly coding transcript as a starting point for discourse analysis in an approach similar to that presented by Edley in Chapter Five. This can be done using word-processor cut-and-paste functions. Notice, however, that they assume the use of broad and overlapping categories, whereas other approaches to data analysis have tended to set up exclusive coding categories, like those used in the analysis of survey data (Fielding, 1993). The principal difference, however, between discourse analyses and other data analyses is not this initial process of analysis but the analytic concepts involved. As has been discussed, these derive from how the research is located theoretically. Analytic concepts are given by the theoretical tradition, the research questions, and so on. The discourse analyst searches for patterns in language in use, building on and referring back to the assumptions she or he is making about the nature of language, interaction and society and the interrelationships between them. It is this theoretical underpinning rather than any sorting process which distinguishes discourse analyses.

Finally, with regard to writing up the analysis (see Section 2.5 below), notice that the final presentation of the analysis is not a record of the process but a summary of selected findings. Findings are presented for a reader so the most interesting or complete patterns are selected from others which are perhaps less complete. Writing up condenses a large amount of work. It is a common mistake on the part of researchers to allow too little time for data analysis compared to collection or the preliminary analytic task of transcription. The following section on writing up presents a fairly standard format for a research text in which the relative sizes of the sections do not in any way correspond to the time spent on any aspect of the research process.

## 2.5  Writing up the research

Several of the previous sections have emphasized that discourse analytic studies vary considerably, for example, in their data, foci and wider theoretical background. The design and conduct of a research project have been presented in terms of a series of decisions which confront the researcher, rather than through a straightforward model which can be followed. It would be an exaggeration to say that at the point of writing up, the decision making ceases. However, the positivist and postpositivist tradition has passed on a useful model for writing up research which continues to be relevant for many other researchers as well. Before introducing this, I should make a couple of points. It is, as I say, a loose model and, of course, has to be modified according to the discipline which the researcher is working in and also the purpose of the writing. For instance, a thesis is different from a journal article in several ways, apart from length, and a book chapter may be different again. Also, this model of the **research text** has been strongly challenged as a result of the double crisis of representation and legitimation discussed in Section 1.1 and some writers have rejected it completely. Nonetheless, it can provide a useful guide and starting point for producing a final research text, especially for a new researcher.

As a loose guide, the research text is divided into six parts. In practice, the number of these is not fixed, as I explain, but the text should be planned and written as several discrete sections. Students sometimes approach academic writing with a romantic image of a writer who, fired with inspiration, begins at the top of a blank page or screen and creates the text in a single creative flow, from the opening sentence to the ringing conclusion! This is over-optimistic and almost certainly inefficient. If the text is not planned, key points will probably be left out. Some sections, such as the account of how the project was conducted, can be prepared, at least in note form, while the research is being carried out. Others, like the introduction and conclusion, probably need to be written out of sequence. And, of course, the finished research text will not be read like a work of fiction, from start through to finish. Potential readers may skim the beginning and end first, and possibly the section headings and references, to see how the work connects to their own research interests. After that they may read the text as a single document. All of these points affect the process and final product of the writing up.

The **title** and the first **introductory section** of the research text present the research question or questions, the main claims which are being made, and an overview of the content of the whole text. The introduction should also be closely related to the closing section or paragraphs of the text, so that by looking at them both, as described above, an impatient prospective reader can obtain a fair understanding of what the text says. If the research is being written up for publication in an academic journal, there will also need to be an introductory **abstract**. This is an even more condensed summary of the text, usually in 100–200 words. One way to organize an abstract is to write a single-sentence summary for each section of the research text. The abstract will need further amendment and polishing before it is complete, but this is usually an effective way to begin.

The introduction will either flow into, or else be followed by as a separate section, the main **theoretical section** of the text. Following the conventions of good academic research practice (see Chapter Eight on evaluation), the broad task of this section is to locate the project in relation to an established tradition of academic work. The links may be positive or negative, in that the project may be presented as following on from or challenging previous research. This section explains where the current project is situated and also attempts to anticipate criticisms, justifying the approach being taken and possibly arguing against alternatives, particularly those that are best known or most recently published. For a discourse analytic researcher, there is likely to be a double task in locating the project theoretically. First, the text needs to present a position on the established theories and research-based publications related to the broad topic of the research: the most relevant will probably be those from the researcher's own discipline. The researcher may present the new project as following one particular theoretical tradition rather than another, or may argue for a modified or even new theoretical position. Second, this section presents the theory of discourse underpinning the research, perhaps through references to published research or perhaps as a more abstract discussion.

The next section of the research text is often referred to as the '**methods**' section (or chapter, in the case of dissertation or book). Conventionally, this consists of a full and detailed account of how the research was carried out. For a researcher who accepts the principle of replicability (see Chapter Eight), the aim is to enable a future researcher to replicate the project and, hopefully, confirm the findings. Other researchers, including most discourse analysts, may not expect that the research will, or even can, be replicated, but still use the section to describe the setting up of the project, and the collection and analysis of the data. The difference is that a researcher who accepts the concept of reflexivity (see Section 1.4) will not necessarily edit out problems and false starts: the aim is not to 'smooth' or idealize the research process. The researcher uses this account to acknowledge her or his own relationship to the research and discuss the constraints and limitations which operated. In addition, by describing exactly what was done, the researcher enables readers to assess the research. The detailed account can therefore work as a form of evaluation (see Chapter Eight). Finally, this openness has an ethical dimension: the researcher avoids concealment and accepts responsibility for the research and its outcomes (see Section 1.5).

Although the account of the research may assist others to evaluate it, this is not in itself sufficient. The evaluation of qualitative research is a complex and disputed area (see Chapter Eight). The methods section is therefore likely to include a discussion of the criteria which the researcher claims are relevant to the particular project, and arguments to support her or his claims that the research is of value.

The fourth section of the research text presents the **data analysis** and the **findings**. This is often a difficult section for the discourse analyst to structure because of the quantity of data used in the project. I will outline three possible ways to organize the section, each of which raises its own problems. The first is to present the data in full and work through the

analysis to show the reader exactly how the data were interpreted and the conclusions were reached. This is the convention which is generally followed in conversation analysis for example (see Chapter Two). It operates as a form of reliability (see Chapter Eight), in that the analysis is open to scrutiny and criticism. The main problem, of course, is that only a limited quantity of data can be analysed. This is a practical difficulty, in that the analyst may have to leave out most of the material collected in larger-scale projects. It also has theoretical implications: either the claims which are made must be confined to the data presented, or there must be a theoretical justification for generalizing from that data to other cases (see Section 1.2).

A second structure for the analysis section is used in other traditions of social research which also use large quantities of qualitative data, such as ethnographic research (see Hammersley and Atkinson, 1995). The analysis and interpretation are conducted 'off-stage'. The analysis section presents only a summary of the data, perhaps with illustrative examples, and then explains the findings and conclusions and justifies them through argument. The process of analysis is not demonstrated completely because the reader is not presented with all of the data. An example of a discourse analytic research text which uses this structure is Brown (1999). The data for his study are twenty-two 'self-help' books, which obviously cannot be included in full. The research text presents summaries of the narratives, devices and the concept of self found in these, illustrated with brief quotations. The author's interpretation is written as a summary, without the data on which it is based. The general advantage of this second way of organizing the analysis section is clear: claims can be based on much larger quantities of data. A disadvantage is that the process of analysis is less open. Also, it can be difficult to find appropriate examples of data to illustrate the general claims being made: a feature which appears across a large sample may not be visible in a short extract. Finally, because discourse data tend to be rich, an extract offered as an example of a particular feature may be open to further analysis which distracts from the point it is intended to illustrate.

It should be noted, however, that the distinction between these two structures may blur. Even when the data for the analysis are presented in full, there has already been a process of selection and decision making, acknowledged or unacknowledged, about *which* data to use out of all that were gathered in the project. Interestingly, ten Have (1999: 24) analyses key papers in the conversation analytic tradition and shows how the authors do not limit the discussion and analysis to the data in the research text. He cites instances in which a claim is supported by reference to a larger body (or corpus) of material which is not presented to the reader. Reference is made to background or contextual information. In addition, reference is made to what might be described as general knowledge or common sense: "knowledge that any competent member (i.e. of society) is assumed to have on the basis of his or her own experience".

A third possible structure for an analysis section is only relevant if the theoretical approach permits language to be analysed out of the original context. One example of such a study is presented in Stubbs (1996). The focus of Stubbs' research is the means which are available within the

English language for indicating a speaker's or writer's point of view and attitude to a point being discussed (this is known as 'modality'). The data are examples from several very large corpora (a corpus, plural 'corpora', is a body of language data: the term is discussed more fully in Chapters Three and Eight). Particular short extracts, most of about four to ten words in length, are analysed in detail. In a few cases, a limited amount of background information is provided e.g. "A BBC radio news reader reported an explosion in a water-processing plant which had killed sixteen people" (Stubbs, 1996: 197). Claims are supported by references to the frequency with which a feature occurred, though this is not given in numbers e.g. "In my corpus data, the commonest surface form is ... " (p.212). The text is mainly written as a series of general statements, a style which is discussed in more detail below.

There may be other decisions around how to present data, such as how to make transcripts more readable (as was discussed in the previous section). At the stage of writing up the analysis, the original transcripts may be edited down and simplified or, alternatively, supplemented to make them more accessible to readers (for example, by indicating key utterances with an arrow, or with a label in an extra column on the right). In some instances the data may be translated (e.g. de Cillia *et al.*, 1999, see also ten Have, 1999: 93).

The fifth section of the research text is the **discussion** and/or **conclusion**. The divisions between sections will largely follow the conventions which operate in a particular discipline but also depend on how much data are included. One format would be for each piece of data to be analysed and the analysis to be discussed, in sequence. Another would be for all the data to be analysed and then the significance of the analysis discussed in a separate section (probably titled 'Discussion'). The conclusion may be very short, consisting mainly of a brief summary of the principal claims which refers back to the introduction. (The introduction and conclusion may well be written at the same time, to ensure that they 'tie together' well.) Alternatively, the conclusion may be a long section which contains the main discussion of the analysis. In either case, following again from the positivist and postpositivist tradition, the conclusion often addresses imagined researchers who would aim to replicate or build on the research as part of a cumulative tradition. It indicates directions for future research, perhaps to extend or clarify the findings of the current project or to fill gaps in a field. It may also discuss the applications of the findings of the project (see Chapter Eight).

The final two sections of the research text are the **references** and the **appendices**. These are, of course, differently formatted and do not continue the 'flow' of the account of the main text. However, I describe them here as sections to indicate their importance. The list of references contains all the works which are cited in the other sections. Some research texts require a fuller **bibliography** which also includes key works which have informed the project, even if these aren't referred to directly in the text. It therefore gives some indication of the traditions in which the analyst is working (though, of course, it will also include publications which are challenged and argued against, especially in the theoretical section). In a

discourse analytic project, the appendices often contain extra data or information about data collection. For example, they may present samples of official documents if these are the data for the analysis. They may include extracts from transcripts which are too large to be included in the main text. However, it is not usually practical (for reasons of length), or desirable (for reasons of confidentiality) to include full interview transcripts as appendices. For an interview-based study, particularly for a larger research text such as a thesis, the appendices may include the question schedule (the list of questions or topic areas followed) and an example of the consent form used (note that this should not be an example which contains the name or signature of any participant). Appendices may also be used to present summarized information about participants (age, occupation etc.) in table form, if this is relevant to the selection of the data.

As well as offering a model for the structure of a research text, the positivist and postpositivist tradition originally established a certain writing style. In the same way that the researcher was not acknowledged as a presence in the research process, she or he was not referred to in the research text. Most obviously, the writing style avoided the use of the first person (references to 'I' or 'me') or expressions used in talk to indicate uncertainty or opinion (such as 'might', 'perhaps', 'seemed'). Also, as discussed above, the text, including the methods section, was written in grammatical forms which generalized the research procedures and findings and so suggested their universal nature. (An example from an early psychological study presents the findings as a 'law': "Of several responses to the same situation, those which are accompanied or closely followed by satisfaction to the animal will ... be more likely to recur", Thorndike, 1911: 244.) These features became conventions of academic writing, although they have broken down more recently in most social science disciplines. Particularly in social research such as discourse analysis, it is now more usual for the writer to 'appear' in a text through the use of 'I'. Alternatively, there may be references to 'we', though this is more ambiguous in tone: it may seem to refer to the research team, but it can also evoke the royal use of 'we' by a single person, and so return to the formality and the authority claims of the older style. Another change of style is that accounts of research are commonly written in the past tense as descriptions of particular events and situations rather than generalized procedures.

The conventional text model and writing style have been described because they offer a useful guide for new researchers. However, as the description also aims to show, there are theoretical associations and issues of power and authority attached to them. By using 'I' the writer introduces the voice of the author/researcher and hence brings her or him into the text. This can work against the impersonal authority of the conventional 'truth' text described above. Ideally, it can draw attention to the contribution of others to the research. Beckett (1996) is an example of an ethnographic/oral history study which attempts to give equal weight to the participants' and the researcher/analyst's voices. However, in other hands, this introduction of the researcher's voice can, less desirably, elevate the researcher to a position associated with a different kind of authority, the authority of charisma and celebrity associated with journalism and

autobiographical writing. There is no neat answer to this problem but it is an example of the ongoing debates about the writing up of research. Some researchers have abandoned the conventional model in favour of more innovative, loosely structured pieces of writing, sometimes referred to as 'messy texts', which aim to move from a single voice to multiple voices. These issues also relate directly to the ultimate purpose of any research, which is discussed further in Section 3 in Chapter Eight

## Conclusion

In this opening chapter, I have covered a very large amount of ground, albeit mostly in summary form. I have indicated some of the premises of discourse analytic research and the range of approaches which it encompasses. I have discussed some of the issues which the analyst must consider and offered an introductory guide to conducting discourse analytic research. The next six chapters present elaborated and detailed examples of the work of particular researchers, as a practical guide for readers who want to conduct their own discourse analytic projects. In the final chapter of the book, Chapter Eight, I resume my more general account and complete this overview with a discussion of the evaluation of discourse analytic research and its potential applications.

## References

Ashmore, N. (1989) *The Reflexive Thesis: Wrighting Sociology of Scientific Knowledge*, Chicago, University of Chicago Press.

Atkinson, P. (1990) *The Ethnographic Imagination*, London, Routledge.

Atkinson, P. and Silverman, D. (1997) 'Kundera's *Immortality*: The interview society and the invention of the self', *Qualitative Inquiry* vol.3 pp.304–25.

Augoustinos, M., Tuffin, K. and Rapley, M. (1999) 'Genocide or failure to gel? Racism, history and nationalism in Australian talk', *Discourse and Society*, vol.10, pp.351–78.

Back, L. (1996) *New Ethnicities and Urban Culture: Racism and Multiculture in Young Lives*, London, UCL Press.

Banister, P., Burman, E., Parker, I., Taylor, M. and Tindall, C. (1994) *Qualitative Methods in Psychology: A Research Guide*, Buckingham, Open University Press.

Beckett, J. (1996) 'Against nostalgia: Place and memory in Myles Lalor's "oral History"', *Oceania*, vol.66, pp.312–22.

Billig, M. (1987/1996) *Arguing and Thinking: A Rhetorical Approach to Social Psychology*, Cambridge, Cambridge University Press.

Billig, M. (1995) *Banal Nationalism*, London, Sage.

Billig, M. (1999) 'Whose terms? Whose ordinariness? Rhetoric and ideology in conversation analysis', *Discourse and Society*, vol.10, pp.543–58.

Bornat, J. (1994) 'Recording oral history' in Drake, M. and Finnegan, R. (eds) *Sources and Methods for Family and Community Historians: A Handbook*, Cambridge, Cambridge University Press in association with the Open University.

Brown, S. (1999) 'Stress as regimen' in Willig, C., (ed.) *Applied Discourse Analysis: Social and Psychological Interventions*, Buckingham, Open University Press.

Coupland, N. and Coupland, J. (1998) 'Reshaping lives: Constitutive identity work in geriatric medical consultations', *Text*, vol.18, pp.159–89.

Croghan, R. and Miell, D. (1998) 'Strategies of resistance: 'Bad' mothers dispute the evidence', *Feminism and Psychology*, vol.8, pp.445–65.

de Cillia, R., Reisigl, M. and Wodak, R. (1999) 'The discursive construction of national identities', *Discourse and Society*, vol.10, pp.149–73.

de la Rey, C. (1997) 'On political activism and discourse analysis in South Africa' in Levett, A., Kottler, A., Burman, E. and Parker, I. (eds) *Culture, Power and Difference: Discourse Analysis in South Africa*, London, Zed Books.

Denzin, N. and Lincoln, Y. (1998) 'Introduction: Entering the field of qualitative research', in Denzin, N. and Lincoln, Y. (eds) *The Landscape of Qualitative Research*, Thousand Oaks, Sage.

Drew, P.(1990) 'Strategies in the contest between lawyers and witnesses' in Levi, J.N., and Walker, A.G. (eds) *Language in the Judicial Process*, London, Sage.

du Gay, P. (1996) *Consumption and Identity at Work*, London, Sage.

Edley, N. and Wetherell, M. (1997) 'Jockeying for position: The construction of masculine identities', *Discourse and Society*, vol.8, pp.203–17.

Edwards, D. (1997) *Discourse and Cognition*, London, Sage.

Edwards, D. and Potter, J. (1992) *Discursive Psychology*, London, Sage.

Fielding, J. (1993) 'Coding and managing data' in Gilbert, N. (ed.) *Researching Social Life*, London, Sage.

Foucault, M. (1961/1997) *Madness and Civilisation: A History of Insanity in the Age of Reason*, London, Routledge.

Frankenberg, R. (1993) *White Women, Race Matters: The Social Construction of Whiteness*, London, Routledge.

Gill, R. (1993) 'Justifying injustice: broadcasters' accounts of inequality in radio' in Burman, E. and Parker, I. (eds) *Discourse Analytic Research*, London, Routledge.

Goodwin, C. (1995) 'Seeing in depth', *Social Studies of Science*, vol.25, pp.237–74.

Goodwin, M.H. (1990) *He-Said-She-Said: Talk as Social Organisation Among Black Children*, Bloomington, Indiana University Press.

Hammersley, M. and Atkinson, P. (1995) *Ethnography: Principles in Practice*, London and New York, Routledge.

Helleiner, J. (1998) 'Contested childhood: The discourse and politics of traveller childhood in Ireland', *Childhood*, vol.5, no.3, pp.303–24.

Hodge, R. and Kress, G. (1988) *Social Semiotics*, Cambridge, Polity.

Hollway, W. (1984) 'Gender difference and the production of subjectivity' in Henriques, H. *et al.*, *Changing the Subject*, London, Methuen.

Jagger, E. (1997) 'The production of official discourse on "glue-sniffing" ', *Journal of Social Policy*, vol.26, pp.445–65.

Jaworski, A. and Galasinski, D. (1998) 'The last Romantic hero: Lech Walesa's image-building in TV presidential debates', *Text*, vol.18, pp.525–44.

Kress, G. (1994) 'Text and grammar as explanation' in Meinhof, U. and Richardson, K. (eds) *Text, Discourse and Context: Representations of Poverty in Britain*, Harlow, Longman.

Lynch, M and Bogen, D. (1996) *The Spectacle of History: Speech, Text and Memory at the Iran-Contra Hearings*, Durham: NC and London, Duke University Press.

Marshall, H. and Wetherell, M. (1989) 'Talking about career and gender identities: A discourse analysis perspective' in Skevington, S. and Baker, D. (eds) *The Social Identity of Women*, London, Sage.

Maybin, J. (2001) 'Language, struggle and voice: The Bakhtin/Voloshinov writings' in Wetherell, M., Taylor, S. and Yates, S.J., *Discourse Theory and Practice: A Reader*, London, Sage in association with The Open University.

Mehan, H. (1996) 'The construction of an LD student: a case study in the politics of representation' in Silverstein, M. and Urban, G. (eds) *Natural Histories of Discourse*, Chicago, University of Chicago Press.

Ochs, E. (1979) 'Transcription as theory' in Ochs, E. and Schiefflin, B. (eds) *Developmental Pragmatics*, New York, Academic Press.

Potter, J. (1996) 'Discourse analysis and constructionist approaches: Theoretical background' in Richardson, J. (ed.) *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*, Leicester, BPS.

Potter, J. (2001) 'Wittgenstein and Austin: Developments in linguistic philosophy' in Wetherell, M., Taylor, S. and Yates, S.J., *Discourse Theory and Practice: A Reader*, London, Sage in association with The Open University.

Potter, J. and Wetherell, M. (1987) *Discourse and Social Psychology: Beyond Attitudes and Behaviour*, London, Sage.

Riessman, C. (1993) *Narrative Analysis*, Newbury Park, Sage.

Rose, N. (1985) *The Psychological Complex: Psychology, Politics and Society in England 1869–1939*, London, Routledge and Kegan Paul.

Said, E. (1978) *Orientalism*, London, Penguin.

Sapsford, R. (1999) *Survey Research*, London, Sage.

Schegloff, E. (1997) 'Whose text? Whose context? *Discourse and Society*, vol.8, pp.165–87.

Seale, C. (1999) *The Quality of Qualitative Research*, London, Sage.

Seale, C. and Silverman, D. (1997) 'Ensuring rigour in qualitative research', *European Journal of Public Health*, vol.7, pp.379–84.

Shakespeare, P. (1998) *Aspects of Confused Speech: A Study of Verbal Interaction Between Confused and Normal Speakers*, Mahwah: NJ, Lawrence Erlbaum Associates.

Silverman, D. (1993) *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*, London, Sage.

Silverman, D. (2000) *Doing Qualitative Research: A Practical Handbook*, London, Sage.

Smith, M. (1998) *Social Science in Question*, London, Sage.

Stenson, K. and Watt, P. (1999) 'Governmentality and "the death of the social"?: A discourse analysis of local government texts in South-east England', *Urban Studies*, vol.36, no.1, pp.189–201.

Stubbs, M. (1996) *Text and Corpus Analysis: Computer Assisted Studies of Language and Culture*, Oxford, Blackwell.

Suchman, L. (1997) 'Centers of co-ordination: A case and some themes' in Resnick, L., Saljo, R., Pontecorvo, C. and Burge, B. (eds) *Discourse, Tools, and Reasoning: Essays on Situated Cognition*, NATO ASI Series F: Computer and Systems Sciences, vol.160, Berlin, Springer

Swann, J. and Graddol, D. (1994) 'Gender inequalities in classroom talk' in Graddol, D., Maybin, J. and Stierer, B., (eds) *Researching Language and Literacy in Social Context: A Reader*, Clevedon, Multilingual Matters in association with The Open University.

ten Have, P. (1999) *Doing Conversation Analysis: A Practical Guide*, London, Sage.

Thorndike, E.L. (1911) *Animal intelligence*, New York, Macmillan: cited in Roth, I. (1990) *Introduction to Psychology*, Hove, Lawrence Erlbaum and Associates Ltd in association with The Open University.

Wertsch, J. (1990) 'The multivoicedness of meaning' in Wertsch, J., *Voices of the Mind*, London, Harvester Wheatsheaf.

West, C. (1996) 'Ethnography and orthography: a (modest) methodological proposal', *Journal of Contemporary Ethnography*, vol.25, October, pp.327–52

Wetherell, M. (1998) 'Positioning and interpretative repertoires: Conversation analysis and post-structuralism in dialogue', *Discourse and Society*, vol.9, pp.387–412.

Wetherell, M. (2001a) 'Debates in discourse research' in Wetherell, M., Taylor, S. and Yates, S.J., *Discourse Theory and Practice: A Reader*, London, Sage in association with The Open University.

Wetherell, M. (2001b) 'Themes in discourse research: The case of Diana' in Wetherell, M., Taylor, S. and Yates, S.J., *Discourse Theory and Practice: A Reader*, London, Sage in association with The Open University.

Wetherell, M. and Potter, J. (1992) *Mapping the Language of Racism: Discourse and the Legitimation of Exploitation*, Hemel Hempstead, Harvester Wheatsheaf.

Wetherell, M., Taylor, S. and Yates, S.J. (eds) (2001) *Discourse Theory and Practice: A Reader*, London, Sage in association with The Open University.