The Language of
News Media

ALLAN BELL

The Language of News Media

Language in Society

The Language of News Media

ALLAN BELL

BLACKWELL
Oxford UK & Cambridge MA

Contents

Figures and Tables

Editor's Preface

A series editor should be careful to avoid exaggeration when introducing volumes in the series, and I am conscious that Allan Bell is not the first author to write for the Language and Society series for whom I have made the claim that he or she is uniquely qualified to write on their given topic. It is difficult, however, to avoid observing that Allan Bell is the only scholar I have ever come across who is both an experienced and practising journalist and an academic sociolinguist with an international reputation. As a New Zealand-based journalist, Allan Bell is a writer who is familiar with both the electronic media and print journalism in many parts of the (particularly English-speaking) world; and, within academic sociolinguistics and linguistic variation theory, he is very well known indeed as a scholar who has produced solidly empirically-based work of very considerable theoretical importance. In particular, his 1984 media-based theoretical paper 'Language style as audience design' is widely regarded as a sociolinguistic classic.

The use of language in the presentation of news, and elsewhere in the media, represents a form of interaction between language and society which affects us all. His surely unique combination of knowledge and expertise has now enabled Allan Bell to produce a book which provides original and exciting insights into this area which will be of great importance for students and researchers in communications and media studies, as well as in sociolinguistics.

Peter Trudgill

*For the Memory of
my Mother and Father*

Introduction and Acknowledgements

It is my belief — and one of the themes of this book — that audiences are an important influence on media content. The book is addressed to an audience of all those who are interested in how media work, how language works, and particularly how the two interact. Its contents should be accessible to people who have a lay interest in these issues as well as to people who study such matters. I hope it will also be of interest to journalists and others in the news media to learn something about the nature of the linguistic work they do.

The book's disciplinary background is in sociolinguistics and discourse analysis, but it draws on a wide range of mass communications theory and research. The book should serve to introduce the study of mass media to students of sociolinguistics. It is less appropriate as an introduction to linguistics for communications researchers, although I have generally avoided technical linguistic terminology.

What is media discourse like? What can it tell us about media? What can it tell us about language? These are important questions for both sociolinguistics and mass communication research. They are also important for society at large. It is my belief that true, responsive communication among people and between peoples is worth striving for. The language of news media is prominent and pervasive in society, and it is worth understanding how that language works, how it affects our perceptions of others and ourselves, how it is produced, how it is shaped by values. This book addresses central issues in the nature of media language and discourse, its production and its reception. It does not, however, enter into the debate over what effect media may have on the use of language in society, for instance whether media language is debasing everyday language. 'The media and language' is the subject for another look.

To explain the background to this book: I have been both studying

and making media language for nearly 20 years. I began research on media language in 1972, and several years later moved to work in journalism. For a decade I alternated employment as an editor and journalist with semi-employment as a researcher in sociolinguistics and mass communications. In that time I worked in monthly magazines, a weekly newspaper and a daily news service, covering especially environmental, scientific and agricultural issues.

Daily journalism leaves no time or mental space for reflection let alone research, but I am now able to combine rather than alternate the two strands of practice and theory. For this my thanks go to the Head Office of the New Zealand Department of Scientific and Industrial Research, where I work half-time as a journalist and media consultant, and particularly to Kevin Sloan for allowing me several months away to write full-time and so complete this book. My other working life is spent as a freelance researcher, and I am grateful to the Department of Linguistics at Victoria University of Wellington, which has adopted me as an honorary research fellow.

Participant observation is a method which has produced insights into the nature of news and its production. Much of what I write here is the fruit of a converse approach — observant participation. I have drawn many examples and observations from my own experience and news stories I have written, or from stories by journalists who worked with or for me.

The content of this book has been, as I say, a long time brewing. Some of the work (on style in news language) originates in doctoral study of 15 years ago, some (especially on discourse analysis of news stories, and news comprehension) I have come to quite recently. Calling up all one's intellectual debts over such a period is difficult. But first credit goes to my principal research colleague, Janet Holmes of Victoria University's Department of Linguistics. She has been a continuing co-worker and encourager in a number of projects, and has commented helpfully on most of this manuscript.

I thank other individuals who have contributed insights and encouragement on aspects of my work: William Labov, Walt Wolfram, Ralph Fasold, Joy Kreeft Peyton, Nikolas Coupland, Howard Giles and Gerhard Leitner. And I remember the late Werner Droescher, who started me off in linguistics at the University of Auckland, and the late Colin Bowley, who saw me through a doctorate there.

As well as presenting my own work, this book incorporates a wide range of other people's research and findings on media language. For this I am indebted to a score of scholars, whose published work I have drawn on freely but most of whom I do not (or did not) know personally: Oliver Boyd-Barrett, Harald Burger, Jack Cappon, Howard

Davis, Teun van Dijk, Mark Fishman, Johan Galtung, Herbert Gans, the Glasgow University Media Group, Erving Goffman, Barrie Gunter, Andreas Jucker, Hannes Kniffka, Denis McQuail, Marie Holmboe Ruge, Philip Schlesinger, Michael Schudson, Gaye Tuchman, Paul Walton and Ruth Wodak. They have all produced insightful work, and I hope my presentation does them justice.

To those hosts who have over the years taken in this 'freelance academic nomad' (I owe the title to Walt Wolfram), I am grateful: the Center for Applied Linguistics, Washington, DC, where I was a Visiting Research Associate in 1981; University of Reading, where I was Leverhulme Visiting Fellow in 1982; Victoria University's Stout Research Centre, for its hospitality on a couple of occasions; and Linguistics Departments at Georgetown University, Washington, DC, University of Pennsylvania, University College, London, and the University of Stuttgart.

Several chapters of this book report findings from a project on media coverage and public understanding of the climate change issue in New Zealand. I acknowledge funding contributed to this project by the Department of Scientific and Industrial Research and Ministry for the Environment. I am also indebted to Jenny Neale, Andrew Matthews and Peter Clare for their professional inputs to the project.

And an appropriately brief 'par' — 'graf to the Americans — to thank those journalists and editors with whom I have worked. I have learnt most from those I had the most trouble with.

Other journalists have also taken time out to talk to me as a researcher about their work, especially Radio New Zealand's staff, and Lindsay Clark, formerly of Wellington's *Dominion* newspaper. Philip Carpenter has been a most patient publisher awaiting a work which has had the gestation period of several elephants. And Ann Bone's editing skills have helped to clarify the end product. I am indebted to Peter Trudgill, both for his general contribution to my work and in his role as editor of the Language in Society series.

The book is dedicated to my mother, who died just before it was finished, and my father, who died 30 years earlier. Finally, I thank Susan Jordan — companion for precisely as long as I have been working on media language — for her constant encouragement and support, and occasional insight and research assistance. She has been ably seconded by Sonny and Thorcas.

One point of writing style in this book: in using generic pronouns, my policy is roughly to alternate *she* and *he*, giving precedence to *she* in most contexts.

Allan Bell, Wellington

1

Media and Language

People in Western countries probably hear more language from the media than they do directly from the lips of their fellow humans in conversation. Society is pervaded by media language. Even in a nation as small as New Zealand, the media pour out daily almost two million words of that primary media genre, news, through some 35 newspapers, newscasts carried by a hundred radio stations and three television networks. In larger countries, the production multiplies. The American blockbuster Sunday newspapers print close to a million words each. The production of media language is huge, although only a fraction of all the face-to-face talk individuals produce. But media language is heard not just by one or two people but by mass audiences. It is the few talking to the many. Media are dominating presenters of language in our society at large.

Within the media, news is the primary language genre. It fills pages of the daily newspaper and hours of radio and television time. Even in broadcasting, where it occupies a small minority of airtime, news is seen by both media organizations and audiences as the focus of media content. Also common to all three daily media is the other dominant genre, advertising, which bulks larger than the news in many daily papers. Some of our data and examples in this book will be drawn from advertising, but most will come from the news since this is the most researched and arguably the most central (cf. McQuail 1987) genre.

News was not always so dominant. The year 1930 was early days for radio. The youthful British Broadcasting Corporation sometimes found there was a shortage of news deemed worthy to be broadcast. If this happened, no attempt was made to fill the gap. The announcer just said: 'There is no news tonight'. At that time, the BBC had a total

news staff of four. It carried news only after six o'clock at night, by agreement with news agencies and the press, who feared for their monopoly. It was allowed to broadcast no more than 400 eye-witness accounts of events per year (quoted in Schlesinger 1987: 20). In this later generation, the declaration that there is 'no news tonight' comes as a shock, a challenge to convention, even to the shape of reality itself. Now there is always news – unless a strike makes us do without.

In the news are carried the stories and images of our day. News is determined by values, and the kind of language in which that news is told reflects and expresses those values. Audiences feel that the way in which language is used must affect the content of what we receive from the media. We will touch on some but not all of the questions which concern people about the media and their language. One question which we will not address here is this: whether, in a world saturated in media language, the way the media use language is changing language itself. In this book we examine the characteristics of news language not its effect on other language.

Mass communication has several characteristics which distinguish it from face-to-face communication and offer advantages to the linguist: multiple originators, a mass simultaneous audience, a fragmented audience, absence of feedback, and general accessibility to the public. We shall see that these characteristics have a profound effect on the shape of media language, on how it is produced, on audiences' ability to understand media content, and on communicators' ability to make themselves understood.

This book deals mainly with language as it is used in the mainline, daily news media – press, television and radio. A wider definition of printed media could call in magazines, books, posters, record covers, bumper stickers, T-shirts – each getting further from the core media. We could include records, cassettes, videos, films. We might cover newer media such as teletext – but it tends to reproduce press-style content and style in broadcast format. But I will confine myself to those media which have a mass audience, and a continuous or daily production cycle. The massness of the core media is characterized by their general availability to all people within a given geographical area. Anyone with a radio or television set or spare change to spend can receive the mainline mass media.

1 WHY STUDY MEDIA LANGUAGE?

First, because it is there. The uses to which language is put in the mass media are intrinsically interesting to us as language users and receivers. The linguistic means advertisers use to try and persuade us, the distinctive manner in which DJs speak, the way news stories are told: these are all interesting uses of language in their own right. How the media use language often seems larger than life, and research which just describes such uses has its own interest.

Secondly, as we have noted, media generate a lot of the language that is heard in society. This is reflected in frequent public comment about how the media use language. Criticism of the media's language use, and the presumed bad effects which those usages are having on everyday speech, are a commonplace of public debate – ironically, conducted in the media's own columns.

A third reason for looking at media language is that language is an essential part of the content of what the media purvey to us. That is, language is a tool and expression of media messages. Both the general public and researchers – to say nothing of the communicators themselves – concern themselves with the content of what is transmitted by the media and with the way in which language carries that content.

Fourth, media language offers the linguist advantages over face-to-face communication. In collecting data from ordinary conversation, one of the biggest problems faced by sociolinguistic researchers is Labov's 'Observer's Paradox' (1972a: 209): that we want to observe and record speakers talking the way they do when they are not being observed and recorded. In the media, this is a non-problem since media language is already intended for mass public consumption. The radio broadcaster is already doing all the necessary monitoring in order to cater to her public, and the fact that someone is recording her makes no change in her awareness of the way she speaks.

Another advantage is availability. Media language is easier to collect than conversation. It is also there in large quantities. The average newspaper may provide you with 100,000 or more words of text. The problem is not so much getting enough language to analyse but deciding how to restrict yourself to a manageable amount. Finally, media offer the potential for good quality recording of spoken language. Direct-line recording off radio or television means that recordings can be almost as good in quality as the originals, with none of the problems of interference, traffic, background television or children's noise which bedevil face-to-face recording.

In sum, assessing the range of research with which I am familiar, these seem to be the main reasons why researchers have studied media language:

- accessibility of media as a source of data for some language feature they want to study
- interest in some aspect of media language in its own right, such as headline language
- interest in the way the media use some language feature also found in ordinary speech
- taking advantage of how the media communication situation manipulates language in a revealing way, for instance in news copy editing
- interest in media's role in affecting language in wider society
- interest in what language reveals about the media's structure and values
- interest in what media language reveals as a mirror of the wider society and culture
- interest in how media language affects attitudes and opinions in society through the way it presents people and issues.

Some of the data and examples I use in this book are drawn from my own work, both as researcher and journalist, the rest are from other researchers. Most are from New Zealand and the United Kingdom, with a small proportion from Europe and the United States. Besides being the geographical areas I am most familiar with, these are also the ones where most research has been done and published. There is an imbalance here, with Asia, Africa and South America under-represented. The countries of the North dominate research on media as well as the production of media content, particularly through material in and on the English language. Writing from a country (Aotearoa/New Zealand) which lies socially and politically between the North and the South – a power in the South Pacific but an ex-colony of Europe – I regret the lack of alternative examples which means having to reproduce Western media dominance.

As well as being concerned with what makes media language tick in its own right, this book focuses on the light it can cast on two related questions: what do the patterns of media language tell us about language, and what do the patterns of media language tell us about news and media? That is, I am concerned to address central questions in linguistics and sociolinguistics, and central questions in mass communication research. Media language can tell us things both about media and about language.

These are some of the specific issues which researchers have addressed through their study of media language:

- What are the ideologies behind different television reports on a single event, particularly through the way the news actors are labelled (Davis and Walton 1983a)?
- What are the discourse structures of one story reported in 250 newspapers in a hundred countries (van Dijk 1988a)?
- What language styles are used on radio stations in New Zealand, and why do they differ from each other (Bell 1982a)?
- Is British television news coverage of industrial issues biased (Glasgow University Media Group 1980)?
- How acceptable is the standard of spoken English on BBC radio (Burchfield, Donoghue and Timothy 1979)?
- What linguistic resources does a presenter use to create a relationship with the audience (Coupland 1985, Montgomery 1988)?
- How do copy editors edit international news (Bell 1984a)?
- How well do people understand the same news story rewritten in different ways (Lutz and Wodak 1987)?
- How did broadcast language develop from the early days of radio (Leitner 1983c)?
- Has British news language become more American over the past 100 years (Bell 1985)?
- How do radio announcers manage to produce a stream of fault-free talk (Goffman 1981)?
- What is the conversational structure of news interviews on BBC radio (Jucker 1986)?
- How do newsworkers structure headlines and lead paragraphs (Kniffka 1980)?

2 MEDIA LANGUAGE RESEARCH AND THE DISCIPLINES

The study of media language has much to offer to the different disciplines on whose territory it touches, including linguistics, sociolinguistics, discourse analysis, semiotics, communication studies, sociology and social psychology. The principal questions within these disciplines can be illuminated by the study of media language, and that study can itself be illuminated by framing research questions from a variety of disciplinary traditions.

Sociolinguists have tended to regard media language as a second-class means of getting access to the real thing – conversational data (for example Labov 1972a: 211). I would suggest there are rather

better reasons for turning to the media for language data. Mass communication provides a situation where data on certain (socio)linguistic issues can be found more readily, or phenomena observed more clearly, than in face-to-face communication. For instance, the way in which broadcasters shift their styles when they move from one station to another is crucial evidence in the issue of what causes style shift (Bell 1984b, and chapter 6 below). It indicates that style shift cannot be explained in terms of a mechanistic view of the amount of attention paid to speech (Labov 1972a).

In sociolinguistics there is also a phenomenon known as initiative or metaphorical style (Blom and Gumperz 1972, Bell 1984b). This occurs when a person adopts a style, or even a language, which is not just a response to the situation in which they find themselves but which actually redefines the situation as something different. This occurs, for example, when a speaker switches from local dialect to a prestigious national dialect in order to win an argument with a friend or family member. But data on such switches is very hard to collect, since by definition they occur rarely and in unexpected circumstances.

However, it turns out that broadcast advertisements provide a situation which matches closely the conditions that produce such style shifts (chapter 7 below): the need to influence the audience through language, and a short time to do it in. And advertisements are far more readily recorded. So broadcasting provides something of a laboratory instance of initiative shift, which simulates the conditions which produce it in face-to-face communication, and enables us to record it in adequate quantities from a readily available data source. In addition such a study provides a fascinating view of how language works in advertising and touches on concerns of the general public: what are the means which advertisers use to get at us?

To linguists, media language can provide data relevant to questions of theoretical importance. For instance, news stories are the common narratives of our time. Their discourse structure casts light on the way in which stories in general are told and structured (chapter 8). News production processes mean we can gain access to language on the production line as it is composed and edited by journalists and newsworkers. It illuminates the ways people compose and amend written discourse (chapter 4). The media can provide data for diachronic linguistics – the study of language change (chapter 7). In the print media we now have a precisely dated, large and consistent archive stretching back for several hundred years since the first recognizably modern newspapers were produced in Europe in the seventeenth century.

To the social psychologist, concerned with attitudes to language or the means by which individuals and groups interact with each other, the media offer a public data source, and their own field of study. The media carry many evidences of interaction between groups and the part language plays in this. They also carry explicit comment on language varieties and their acceptability in society. And they reveal in their own language usage the ways in which people can mould their language in order to cater to their different audiences (chapter 6). Concerns with intergroup relations and images which one group holds of another are well focused in the words media use to describe people and groups: are they 'terrorists' or 'freedom fighters' (chapter 9)?

To the sociologist of language, media are one of the main language-forming institutions in society, along with education and government. How does the use of language in the media affect groups in a society? Can the media contribute to a society's language planning goals? In the sociology of organizations, the issue of organizational control is neatly encapsulated in the ways in which media attempt to control how staff use language (chapter 3). And in the micro-sociology pioneered and practised so insightfully by Erving Goffman, we find a long study devoted to the language of radio presenters and what it can tell us about their self-presentation (Goffman 1981).

In communication studies, the study of media language is crucial to understanding the messages the media construct (chapters 8 and 9). The field of semiotics and cultural analysis, and communication studies in general, has long concerned itself with the language in which such messages are framed as a clue to the underlying structures of meanings, often in a search for bias and stereotyping. Such analysis benefits from close application of linguistic analysis, and may in fact be either misleading or unconvincing if it approaches language with inadequate methods (chapter 10).

Sociolinguistics and communication studies share a historical interest in similar research questions. Early paradigms put forward by pioneers in the two fields are uncannily close. Basic issues of communications research were encapsulated by Harold Lasswell in 1948:

Who says what in which channel to whom with what effect? (Lasswell 1960: 117)

Compare this with a prescription for the sociolinguistic enterprise:

Who speaks what language to whom and when? (J. Fishman 1965: 67)

3 THEMES OF THE BOOK

This book is based in the concerns and frameworks of sociolinguistics and discourse analysis. It does not introduce students or scholars of mass communication to those disciplines. A number of good introductory texts on sociolinguistics exist for that purpose, although discourse analysis is not so well served.[1] I have, however, tried to keep technical linguistic detail to a minimum, and to explain it when it arises, so the book should be accessible to those without a background in linguistics.

On the other hand, the book does aim to introduce linguists to media research. For those linguists who wish to inform themselves more about mass communications, there are a number of introductions, overviews and collections of key papers in print.[2] While no adequate introduction to the study of news media language exists in English (although in German, see Burger 1984), there are several important collections of papers on media language: Davis and Walton (1983b), Leitner (1983b), Baetens Beardsmore (1984), van Dijk (1985a). Later chapters in this volume present much of this research.

Finally, thumbing through my own book, I want to draw out three themes which run through the volume: the importance of the processes which produce media language, the notion of the news story, and the role of the media audience. Media language is the product of multiple hands, and the processes by which it is moulded and modified are both crucial and enlightening for an understanding of the eventual news text, its form and its content. This theme occupies us in chapters 3, 4 and 10. The idea of the story is central to news. Newsworkers – journalists, copy editors, and others directly involved in production – do not write articles. They write stories – with structure, order, viewpoint and values. The story, its generation, conventions and ideology, are our subject in chapters 4, 8, 9 and 10. Lastly there is the audience, whose role in news language seems peculiarly ambivalent. While apparently passive receivers of material determined by producers, audience members bring the power of their own choices, understandings and preconceptions to media reception. The nature of the audience, their role in influencing media language styles, and their comprehension of media content are the theme of chapters 5, 6, 7 and 11.

2

Researching Media Language

Good research on media language requires some familiarity with three things:

- language and how to analyse it
- media and research on media
- the specific media under study and their output.

Many small-scale and student research projects make use of media language data. The reasons are obvious: the media provide data which are good in quality, adequate in quantity, comparatively easy to access, and not modified by an observer effect. But media language research is littered with examples of flawed methods or mistaken conclusions which arose simply because the researchers did not acquaint themselves either with the basics of media, or with linguistic analysis, or with the media they were about to study.

In this chapter I background the structures, operations and content of the media as a prelude to the discussion of media language. I do this by looking at methods of conducting research on media language, particularly the decisions required to gather an actual sample or corpus of data. The chapter thus serves primarily as an introduction – but also as a guide – to how media language research is conducted.

News media form a kind of speech community producing their own variety of language, and the first rule of studying a speech community also applies here: get to know it. Direct, personal observation of media production processes is valuable, but a lot can be learned by close acquaintance with media output. The sociologist Erving

Goffman's long, perceptive essay on radio language (1981) was the fruit of close listening to radio output.

Availability is one of the attractions of media as a linguistic data source, especially for projects with severe time limitations. Buying a newspaper or recording a radio station are easier than many other kinds of linguistic fieldwork, which may require face-to-face interviewing in often difficult circumstances (Labov 1984, Milroy 1987). But the apparently easy accessibility of media language is deceptive. The obstacles to collecting media language may be different from those of recording face to face, but they are nonetheless real obstacles.

1 UNIVERSE AND SAMPLE

Any research beyond a small-scale glance at a couple of media texts means defining the kind of media language one wants to study and how it will be collected. This involves two steps: making a clear and consistent delineation of exactly what is to be collected, and limiting the amount of data to be gathered to manageable proportions while ensuring it remains representative. The first step defines the universe of discourse or the population, and the second draws a valid and reliable sample of that universe. Strictly speaking, defining the universe precedes any sampling, but in practice there is a to-and-fro between the two steps as you become more familiar with what you are working on, especially with the volumes of data which certain decisions will produce. This book draws on data from many other researchers as well as my own. The following list shows the kind of samples they used to gather the data, ranging from recordings of many months of broadcast news, or stories from a hundred different coun-·tries, to just a single news clip or radio programme:

Brunel 1970
  audio recordings of five French-language radio stations in Quebec, one on each day of the week
Glasgow University Media Group 1976, 1980
  video recordings of all 260 hours of television news broadcast in Britain throughout a five-month period
Bell 1977, 1982a
  35 hours of radio news, being a five-day, constructed week sample of all news broadcast in Auckland, together with additional recording of specific newsreaders

Burchfield, Donoghue and Timothy 1979
  extensive listening to BBC radio networks for a four-week period
Goffman 1981
  Erving Goffman's use of Kermit Schafer's lovingly gathered corpus of American radio 'bloopers', augmented by interviews, participant observation, 20 hours of recordings, and his own note-taking as a listener
Davis and Walton 1983a, 1983c
  video recordings of television news of a single event as reported in three different countries
Lerman 1983
  video recordings of commentaries on a Nixon Watergate speech, broadcast on the three United States television networks
Letiner 1983a
  audio recordings of three BBC radio phone-in programmes
Bell 1984a
  copy from local and international news agencies for a five-day con-structed week
Hartley and Montgomery 1985
  four feature articles on the conditions of the poor in England, published between 1860 and 1931
Coupland 1985, Selting 1985, Montgomery 1988
  recordings of individual radio programme presenters in Wales, Germany and England
Jucker 1986
  seven hours of interviews extracted from BBC Radio 4 programmes
Corbett and Ahmad 1986
  a 93-day sample of editorials in one Melbourne newspaper
Bell 1988
  a constructed week in each decade from 1920 to 1980 of the Lon-don *Daily Mirror*
van Dijk 1988a
  clips of one story reported in 250 newspapers in a hundred countries
van Dijk 1988b
  several days' sample of six national and two regional Dutch news-papers, national news agency copy, and source materials
Bell 1989
  six months of news on climate change published or broadcast in New Zealand.

In sample design, language researchers can take advantage of the long history of sampling media production in one tradition of com-

munication research – content analysis. The field of quantitative content analysis developed a range of sample designs in its heyday between the 1930s and 1960s. Leaving aside limitations of content analysis itself as an approach, the sampling methodologies developed are precisely of the kind needed for linguistic purposes, because they aim to control for the principal sources of fluctuation in media output. There are several standard surveys of content analysis and its techniques, including Holsti (1968, 1969), Danielson (1963) and Krippendorff (1980). The field has progressed little in the past 20 years, and Budd, Thorp and Donohew (1967) remains probably the clearest brief introduction, including a good chapter on sampling (and a 300-item annotated bibliography).

A sample must cope adequately with non-random fluctuations within the population, but not burden the researcher with a corpus of millions of words. In practice one often wants to satisfy a number of not quite compatible aims. In 1974 I designed and recorded a random sample of radio news in Auckland (Bell 1977, 1982a). It needed to be large enough to catch the stations' regular newsreaders, and to record one continuous week risked missing some them. But that played off against another requirement: collecting a sufficient amount of language from individual newsreaders. If I recorded non-continuous sample days, there was a risk of not gathering enough from the individuals. This problem is typical of those that come up in media sampling. I decided to gather a random sample of all news and record additional data from specific newsreaders.

Decisions on gathering a corpus of media language are required in three main areas: genres, outlets and outputs.

1   The genres are the particular kind of media content in which you are interested – news, classified advertising, game shows, weather forecasts, and so forth.
2   The media outlets are the publications, television channels or radio stations which carry the content.
3   'Outputs' are what the media outlets produce – specific newscasts, advertisements or programmes – and the time period to be covered.

2 WHAT'S NEWS: DEFINING GENRES

Study of media language usually begins with an interest in the language of a specific genre – in how a particular kind of news is

reported, how headlines differ from other language, how advertising seeks to persuade. Media content includes many genres, most of which have a language component: news, letters to the editor, display advertising, documentaries, soap operas, music, sports commentary. Of course, not all genres can occur in all media – the press cannot print phone-in programmes nor can the radio broadcast cartoons.

Two genres are common to all the primary media of mass communication – news and advertising. These have also been the focus of most research on media and on media language. Nearly all daily media carry news – hence their collective label of 'news media'. Most carry advertising. The absence of advertising in some public service broadcasting, for example BBC radio and television and the Public Broadcasting Service in the United States, is an exception. However, even PBS stations devote a significant proportion of their airtime to crediting sponsors and to regular, lengthy fund-raising drives.

Press advertising divides into two categories: the familiar small-print columns of classifieds, and display advertising scattered throughout the news pages and designed to attract the reader. Broadcast advertisements come in chunks of sound ranging from 5 to 120 seconds long. They may be gathered into 'commercial breaks', as in television, or dropped singly into the flow of radio programming.

In a newspaper, everything other than advertising is called 'editorial'. [1] Most editorial content is written 'copy'. Some is visual, but may have a subsidiary language component (cartoons, graphs). We can divide editorial copy into three broad categories: service information, opinion and news. Service information consists of lists rather than continuous copy: sports results, television programmes, share prices, weather forecasts. It is often associated with specialist sections such as sports or business pages.

Opinion copy includes what are often called 'editorials' or 'leaders' – a statement of the newspaper's own views on an issue, usually appearing on an inside page under a reduced banner of the paper's 'masthead'. Most of the remaining opinion copy is regular contributed columns, letters to the editor and reviews. By journalistic tradition, opinion and news reporting are supposed to be kept separate. Opinion copy is usually flagged by devices such as a standard heading or 'mugshot' above a columnist's copy, and 'bylining' with the writer's name. Although numerous media researchers have shown that fact and opinion are by no means easy to separate, this has made little difference to how newsworkers perceive – or newspapers present – these categories.

Types of press news

For language analysis, I divide the genre of press news into four categories, which are generally the categories newsworkers themselves use under various labels:

- hard news
- feature articles
- special-topic news, e.g. sports, racing, business/financial, arts, agriculture, computers
- headlines, crossheads or subheadings, bylines, photo captions.

Newsworkers' basic distinction is between hard news and features. Hard news is their staple product: reports of accidents, conflicts, crimes, announcements, discoveries and other events which have occurred or come to light since the previous issue of their paper or programme. The one-off, unscheduled events such as fires and disasters are sometimes called 'spot news'. The opposite to hard news is 'soft' news, which is not time-bound to immediacy. Features are the most obvious case of soft news. These are longer 'articles' rather than 'stories' covering immediate events. They provide background, sometimes 'editorialize' (carry the writer's personal opinions), and are usually bylined with the writer's name. A newspaper's feature articles may be gathered together on feature pages, often together with opinion material. Features are often produced by a different group of journalists from those who write the day-by-day hard news. A lot of newspapers acquire their features from outside services provided by news agencies or major newspapers like *The Times* of London, and the *Washington Post/Los Angeles Times* service.

For both newsworkers and researchers, the boundaries between hard and soft news are unclear (see the insightful discussion in Tuchman 1978: 47). Indeed, journalists spend much of their energy trying to find an angle which will present what is essentially soft news in hard news terms. Journalists and media researchers both recognize hard news as the core news product, the typical against which other copy will be measured. Hard news is also the place where a distinctive news style will be found if anywhere. In features, journalists are allowed more liberty of style, and many features are written by non-journalists. Research problems can arise when language users' own basic categories are overlooked. The million-word Brown corpus of American English and Lancaster-Oslo-Bergen (LOB) corpus of British English appear to make no distinction between hard news and

features, treating all news as 'reportage'. Burger notes (1984) that the German 'boulevard press' deliberately blurs the boundary of soft and hard news in its reporting.

My third category, special-topic news, normally appears in sections of the paper explicitly flagged for their subject matter. Such pages are generally produced by separate groups of specialist journalists under the control of their own editor, such as the business or sports editor. In research I always distinguish special-topic news such as sports or financial from general news. Many newspapers allocate news to pages according to its geographical origin – for example, local, national, international. I treat these as simple divisions of general news. And because politics is so much the dominant topic of general news, I count pages flagged as 'political' in with general news.

The fourth category is a miscellaneous or residual one which cuts across the first three. The 'body copy' – the main text of a story or feature – is classified in the first three categories. The adjuncts to it are headlines, crossheads (subheadings within a story), writers' names or bylines, and captions to photographs. All these are usually visually distinct, set in different typefaces or sizes to the body copy.[2]

Broadcast news

News and advertising make up the great majority of press content, but they occupy a small proportion of broadcast airtime. In television and radio, staple news consists of a number of short items of hard news gathered into a bulletin usually three to five minutes long. In addition, once or more each day, many stations broadcast news programmes, between 15 minutes and two hours long. These combine short news items with longer, 'softer' backgrounders, interviews and the like, plus weather forecasts and other material. Broadcasting's equivalent of the feature is current or public affairs (commentary and documentary are more akin to opinion material). While the core of broadcast hard news tends to be gathered into short bulletins, the boundary between current affairs and news is a grey one, particularly for the audience.

Different sections of what listeners hear as a single, seamless news programme may in practice be compiled by two editorially independent sections of the broadcasting organization. A programme such as BBC Radio 4's *World at One* lunchtime programme carries bulletins originated by News personnel, while the surrounding, magazine-like material originates with Current Affairs. There is little liaison between the two groups and they are usually unaware of the specifics of what