

LO 11
286

Observing & Analysing Natural Language

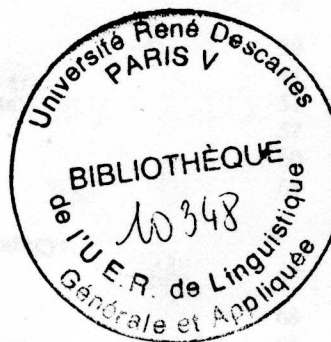


Handwritten signature

Observing and Analysing Natural Language

*A Critical Account of
Sociolinguistic Method*

LESLEY MILROY



BASIL BLACKWELL

Copyright © Lesley Milroy 1987

First published 1987

Reprinted 1989

Basil Blackwell Ltd

108 Cowley Road, Oxford OX4 1JF, UK

Basil Blackwell Inc.

432 Park Avenue South, Suite 1503

New York, NY 10016, USA

All rights reserved. Except for the quotation of short passages for the purposes of criticism and review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

Except in the United States of America, this book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

British Library Cataloguing in Publication Data

Milroy, Lesley

Observing and analysing natural language.
a critical account of sociolinguistic method.

I. Sociolinguistics—Methodology

I. Title III. Series

401'.9 P40.3

ISBN 0-631-13233-3

ISBN 0-631-13623-1 Pbk

Library of Congress Cataloging in Publication Data

Milroy, Lesley

Observing and analysing natural language.

(Language in society 12)

Bibliography: p.

Includes index.

I. Sociolinguistics—Methodology. I. Title

II. Series: Language in society (Oxford, England); 12.

P40.3M54 1987 401'.9'072 87-5202

ISBN 0-631-13233-3

ISBN 0-631-13623-1 (pbk.)

Typeset in 10/11½pt Times

by Alan Sutton Publishing Limited

Printed in Great Britain by

T.J. Press (Padstow) Ltd, Padstow, Cornwall

Language in Society

GENERAL EDITOR

Peter Trudgill, Professor of Linguistic Science,
University of Reading

ADVISORY EDITORS

Ralph Fasold, Professor of Linguistics,
Georgetown University

William Labov, Professor of Linguistics,
University of Pennsylvania

- 1 Language and Social Psychology
Edited by Howard Giles and Robert N. St Clair
- 2 Language and Social Networks (Second Edition)
Lesley Milroy
- 3 The Ethnography of Communication
Muriel Saville-Troike
- 4 Discourse Analysis
Michael Stubbs
- 5 The Sociolinguistics of Society
Introduction to Sociolinguistics, Volume I
Ralph Fasold
- 6 The Sociolinguistics of Language
Introduction to Sociolinguistics, Volume II
Ralph Fasold
- 7 The Language of Children and Adolescents
The Acquisition of Communicative Competence
Suzanne Romaine
- 8 Language, the Sexes and Society
Philip M. Smith
- 9 The Language of Advertising
Torben Vestergaard and Kim Schrøder
- 10 Dialects in Contact
Peter Trudgill
- 11 Pidgin and Creole Linguistics
Peter Mühlhäusler
- 12 Observing and Analysing Natural Language
A Critical Account of Sociolinguistic Method
Lesley Milroy

For Jim, David, Andrew and Richard

Editor's Preface

As is well known, William Labov, whose work was initially responsible for the development of research in those areas of sociolinguistics which are today sometimes referred to as *secular linguistics*, was initially not at all happy that the label *sociolinguistics* should be attached to work that in his view, and I believe in the view of all of us who are practitioners of this science, should really have been referred to simply as *linguistics*. This terminological battle has subsequently been lost, but the popularity of the term *sociolinguistics* does have the advantage of stressing the primary location of this form of research in the community rather than in the office or laboratory, and of indicating the somewhat independent traditions and objectives the subject has subsequently acquired within the wider field of linguistics.

Work of this type is, at its best, so insightful, productive and exciting that one might be surprised at how relatively little of it is actually performed, for all that it is by no means a totally minority activity. This should, however, really come as no surprise, since work in this paradigm, for all its importance, is most often difficult, complicated, time-consuming and expensive. It is no wonder that many of the weaker brethren confine themselves to the comfort of working with their own intuitions in the security of their own offices. With the publication of this book, however, work in secular linguistics will now be that little bit less daunting, because in *Observing and Analysing Natural Language* is distilled the collective wisdom of the first generations of sociolinguists, including not least Lesley Milroy herself, who is one of the best practitioners of this form of sociolinguistics there has ever been. Here, that is, is an author who really knows what she is talking about because in her own pioneering work she has confronted and overcome many of the thorniest practical and theoretical problems that this field has to offer.

The book itself deals not only with fieldwork methodology and data collection, and not only with the analysis and interpretation of this data, it also deals, crucially, with the interrelationships between these two aspects of secular linguistic work, and with the often neglected connections

between methodology, interpretation and linguistic theory. With this book we have, for the first time, a work which focuses on the history, objectives and achievements of secular linguistics to date, as well as on its methodology. It will be of enormous value to anyone studying, teaching or working in any sociolinguistic aspect of language in the community. Lesley Milroy's *Observing and Analysing Natural Language* is a very important milestone in this field. Indeed, it represents a coming of age in the development of sociolinguistics as a discipline.

Peter Trudgill

Preface

This book cannot in any sense be said to be a handbook or inventory of techniques, although it is certainly intended to be of practical value to those interested in studying the way people use language in naturally occurring social contexts. Sociolinguistic method is discussed in terms of its relationship to theory, in the belief that if this link is not acknowledged *interpretation* of research results may ultimately be difficult and unsatisfying. This is because apparently innocuous methods which are in fact associated with a specific theoretical paradigm can conceal important underlying assumptions. Methodological problems and principles will therefore be discussed not only in practical terms, but in terms of both assumptions underlying the chosen method, and the theoretical goal of the research. An account of method divorced from theory is not considered to be helpful, desirable, or even possible.

Chapters 1 to 4 focus chiefly on methods of data collection and chapters 5 to 8 on various aspects of data analysis and interpretation. But there are at all points areas of overlap, and underlying theoretical issues frequently emerge. Although a good deal of the discussion is placed within the general framework of the methods first developed by William Labov for use in urban settings (and indeed assumes some knowledge of his work), I very much hope that a number of principles will emerge which are of value to people who, while not necessarily researching within this paradigm, are concerned for a number of different reasons with the study of language in its social context. These include not only linguists with an academic interest in (for example) the language choice patterns of ethnic minorities in industrialized societies, in conservative rural dialects, in pidgin and creole languages and in 'exotic' languages. There are also professionals such as teachers, educational psychologists and speech therapists, who work extensively with language and who sometimes need to collect, analyse and interpret samples of naturally occurring speech. These wider applications are discussed in chapter 9, where a number of practical issues are explored which can be illuminated by sociolinguistic methods of data collection and analysis.

As far as seems possible without sacrificing style and clarity, I have avoided using the controversial generic pronoun forms *he*, *him*, *his*. But where alternative means of expressing generic reference seem awkward I have used the traditional forms, and hope that readers will accept this purely stylistic decision in the spirit in which it was taken.

Acknowledgements

Many people over the years have helped me, directly and indirectly, to write this book. Their assistance and influence on my approach to sociolinguistics is gratefully acknowledged here.

Most importantly, James Milroy has shared the responsibility for planning and directing the Belfast research projects where we both obtained 'hands on' experience. Without him, I would certainly not have embarked upon the work at all. Thanks are also due to the people who worked on those projects, contributing to them in different ways; they are Rose Maclaran, Domini O'Kane, John Harris, Linda Policansky, Zena Molyneux, Brendan Gunn, Máire Burke, Thomas Vogel and Ann Pitts. The financial support provided for both projects (HR 3771; HR 5777) by the Social Science Research Council is acknowledged with thanks.

I have also profited greatly from discussing with other linguists various aspects of their work, as will often be apparent from references in the text. Warm thanks are due particularly to the linguists who treated me so hospitably during my visit to Australia during the spring of 1983. Our stimulating and often lengthy discussions of methods, goals and problems of field linguistics influenced considerably my approach to writing this book. But like many other sociolinguists, I have learnt more from William Labov than from anyone else, and am happy to acknowledge his support and continuing influence. Peter Trudgill's comments on an earlier version of this book were particularly appreciated, as was his tactful moral support throughout the entire task of producing it. Jack Chambers, Janice Kay, James Milroy and Kay Mogford also read and commented helpfully on parts of the manuscript.

An early draft of the first two chapters was written during my tenure of a Simon Senior Research Fellowship at the University of Manchester, during 1982-3. This financial support is gratefully acknowledged. At a much later stage of production, Anne Stals transformed an unprepossessing manuscript into an impeccable typescript for which she has my grateful thanks. Thanks are due also to Helen Moore, for help with proof reading.

I am grateful to publishers and authors for permission to adapt their

original figures and maps, the sources of which are specified in the text. Further details are as follows:

Figure 1.1 is adapted from A.E. Kibrik's *The methodology of field investigations in linguistics* (Mouton, 1977, p.7). Map 1.1 is adapted from David De Camp's 'Social and geographical factors in Jamaican dialects', which appeared in R.B. Le Page (ed.) *Creole language studies II* (Macmillan, 1961, p. 78). Figure 4.1 is adapted from a figure in W. Labov's 'Field methods used by the project on linguistic change and variation', *Sociolinguistic working paper* 81 (1981). Map 4.2 is adapted from Ann Pitts' 'Urban influence on phonological variation in a Northern Irish speech community'; the original figure appeared in *English world wide* 6,1 (Benjamins, 1985, p.66). Figure 6.1 is adapted from Barbara Horvath's *Variation in Australian English* (Cambridge University Press, 1985, p.71). Figure 8.2 is adapted from Allen Bells' 'Language style as audience design'; the original figure appeared in *Language in society* 13,2 (Cambridge University Press, 1984, p.171). Figure 8.4 is adapted from a figure in S.N. Parasher's 'Mother-tongue-English diglossia: a case study of educated Indian bilinguals' language use', *Anthropological linguistics* 22 (1980), 4. Figure 8.5 is adapted from Joan Rubin's *National bilingualism in Paraguay* (Mouton, 1968, p.109).

Finally, the spouses and children of authors have to put up with a great deal of disruption to the smooth running of their daily lives, and mine are no exception. I thank Jim, David, Andrew and Richard for their continuing love and support.

Lesley Milroy
University of Newcastle upon Tyne

Contents

Editor's Preface	ix
Preface	xi
Acknowledgements	xiii
1 Field Linguistics: Some Models and Methods	1
1.1 Data and Theory	1
1.2 Earlier Approaches to Linguistic Description	5
1.3 The Traditional Model: Some Adaptations and Criticisms	12
2 Sampling	18
2.1 Introductory	18
2.2 Random Sampling a Population	19
2.3 Judgement Sampling	26
2.4 Research Objectives and Informant Sampling	28
2.5 The Variable of Social Class	29
2.6 Informant Sampling and Social Network	35
2.7 Sampling Language in a Range of Situations	36
2.8 Concluding Remarks	38
3 Speakers: Some Issues in Data Collection	39
3.1 Introductory	39
3.2 The Interview as a Speech Event	41
3.3 Some Structural Limitations of Interview Data	51
3.4 The Idea of the Vernacular	57
3.5 Gaining Access to the Vernacular	60
3.6 Social Norms and Patterns of Language Use	64
4 Methodological Principles and Fieldwork Strategy: Two Case Studies	68
4.1 Introductory	68
4.2 The Project on Linguistic Change and Variation	68

4.3	The Belfast Projects	75
4.4	A Note on Ethics	87
4.5	Conclusion	93
5	Analysing Variable Data: Speaker Variables	94
5.1	Introductory	94
5.2	Speaker Variables and Theories of Change	96
5.3	Social Class and Sociolinguistic Research	97
5.4	Sex of Speaker	101
5.5	Ethnicity	103
5.6	Variables Associated with Speaker Identity	104
5.7	Social Network	105
5.8	Concluding Remarks	111
6	Analysing Phonological Variation	113
6.1	Introductory	113
6.2	Identifying Linguistic Variables	114
6.3	The Measurement and Quantification of Variables	117
6.4	Problems in Quantifying Variables: (a) in Belfast	121
6.5	Problems in Quantifying Variables: (ε) and (o) in Belfast	122
6.6	Types of Linguistic Variable	129
6.7	Determining the Lexical Input to Phonological Variables	131
6.8	More on Quantification	134
7	Analysing Syntactic Variation	143
7.1	Introductory	143
7.2	Data Collection Problems	143
7.3	The Use of Elicitation Techniques	146
7.4	Syntactic Analysis and the Sociolinguistic Variable	150
7.5	The Problem of Semantic Equivalence	158
7.6	Semantic Equivalence between Standard and Non-standard Forms	162
7.7	Syntactic Variation: The Semantic and Discourse Contexts	164
7.8	Conclusion	169
8	Style-shifting and Code-switching	171
8.1	Introductory	171
8.2	The Quantitative Analysis of Stylistic Variation	172
8.3	Code-switching and Code-mixing	184
8.4	Analysis and Interpretation of Variable Data	190
8.5	Conclusion	197
9	Sociolinguistics: Some Practical Applications	199
9.1	Introductory	199
9.2	Interactional Sociolinguistics	201
9.3	The Educational System	202

9.4 Formal Language Assessment	205
9.5 Concluding Remarks	212
References	213
Index	227

Field Linguistics: Some Models and Methods

1.1 Data and Theory

It is illuminating to begin a discussion of sociolinguistic method by considering two general issues which have implications for the field linguist at all stages of data collection, analysis and interpretation; these are the relationship of the investigator to data and the relationship between data and grammars.

There are a number of ways in which an investigator might proceed in carrying out a piece of synchronic linguistic description. The term 'description' is used rather widely here (following Kibrik (1977)) to cover either descriptions of languages unknown to the investigator or descriptions of the investigator's native language, both of which can be based either on introspection or on some sort of field investigation. By considering the task of linguistic description in this way, we can examine rather radically the relationship between investigator and object of study in terms of the type of data used, assuming that the output will always be (to a greater or lesser extent) an idealized model of a fragment of a language.

Models might be of many different kinds, examples being an account of some aspect of 'core' grammar (such as the English Noun Phrase); the grammatical categories of an entire language (such as Dixon's (1971) grammar of Dyirbal); or an account of the systematically variable use by members of a speech community of a portion of the linguistic system (such as Labov's (1966) study of the distribution of five sociolinguistic variables in the Lower East Side of New York City). Although the models produced by sociolinguists are often felt in some sense to be closer to the data base than those of other types of linguist, it is important to remember that a representation such as Labov's famous graph of the variable realization of /r/ in New York City is actually *an idealized model of sociolinguistic structure*; the figures upon which it is based are the product of a long process of sociological, mathematical and linguistic abstraction. Idealized models of any kind, whatever the differences in method,

theoretical goal and assumption which underlie them, bear an indirect relationship to data.

Gumperz has discussed Labov's methods and theoretical goals in such a way as to make the abstract character of his models very clear. He points out that although Labov rejects Saussurian and Chomskyan assumptions of uniformity in grammatical systems, he shares with other linguists an interest in understanding the general character of *grammars*, believing these to be affected by the social characteristics of human groups. Gumperz then goes on to argue that the relatively abstract approach associated with this theoretical goal entails a neglect of the *speaker as participant in interaction*, and that quite different methods are needed to investigate issues arising from the ability of speakers to interact, such as the co-occurrence (or otherwise) of their judgements in the interpretation of discourse: 'A speaker-oriented approach to conversation . . . focuses directly on the strategies that govern the actor's use of lexical, grammatical, sociolinguistic and other knowledge in the production and interpretation of messages in context' (Gumperz 1982: 35).

Labov himself has contrasted in a similar way two alternative approaches to linguistic variation: one can start by examining linguistic forms (variables) and their distribution, or by examining speakers and the kind of behaviour appropriate to different situations. Labov prefers the first type of framework because it gives a better idea of the *system* as a whole, although it is not capable of yielding optimal information about *speakers* (Labov 1966: 209).

Labov's focus on system rather than speaker leads to consideration of a widely accepted principle of scientific linguistics: the language itself is an abstract object not amenable to direct observation. However, as Kibrik notes, 'the concrete utterances which represent the realisation of the linguistic competence of speakers who know the language can be observed' (1977: 2). Such an emphasis seems uncontroversial, and is useful in that it helps us to remember the relative abstractness of (for example) the variable language patterns which are of interest to urban dialectologists. Nor does Kibrik's remark imply acceptance of Chomsky's controversial competence/performance distinction, which has not proved to be a useful one for sociolinguistics (see further L. Milroy 1985).

Labov has observed that the general programme of all linguists – not only those who are writing competence grammars – begins with a search for invariance (Labov 1975: 7). The context of this remark was an attempt to focus on precisely how one kind of linguistic enterprise differs from another, and Labov argued that good methods and theories could best be developed by considering the important assumptions which linguists shared before examining those which divided them. Commenting on the theory/data relationship in a manner which also tends to emphasize the similarities between different types of linguistic enterprise, Kibrik lists what he

considers to be three crucial concepts in any conceivable descriptive linguistic activity (but compare the comments of Gumperz, quoted above):

- 1 The *subject* of investigation (the language or part of the language).
- 2 The *object* of the investigation (written texts or tape-recorded data).
- 3 The *product* of the investigation. This is the *model* of the subject of the investigation which is usually called the *grammar*. Thus, Labov's graphic representations, which model patterns of language variation in New York City, may reasonably be described as grammars.

By using the term 'grammar' in this extension of its usual sense, we can begin to compare and contrast coherently the aims, methods and procedures of, for example, a descriptive linguist and an urban sociolinguist both working on a portion of the verb phrase (see Cheshire 1982a and J. Harris 1984 as examples of sociolinguists with those interests). It is not profitable to see these differences (as Chomsky apparently has done; Chomsky 1980: 248) in terms of differences in the amount of idealization of the data base; models are always abstract and indirectly related to data. But using Kibrik's framework, and assuming with Labov (1975) that co-operation is essential if linguists are to benefit from each other's insights, differences between descriptive linguists and sociolinguists may be analysed as differences in the relationship between the investigator, the subject of study and the object of study in the process of arriving at the final product (or model). Differences in the character of these relationships also give rise to methods which differ in their potential to achieve particular goals.

Following this general line of thinking, we can present in graphic form (figure 1.1) three different models of the process by which an investigator arrives at a product:

Figure 1.1(a) models an investigator/data/grammar relationship whereby the investigator directly accesses the target language by means of his or her own linguistic competence. Since the description is based on introspective self-observation (sometimes checked against the introspections of others), a body of data (the object of the investigation) is absent. One point which might be made is that this method cannot be used to study any language or language variety not known to the investigator, and since academic linguists are seldom competent speakers of non-standard dialects or uncoded languages, can in practice be used for describing only fully codified standard languages. This is not of course to deny that those who have grown up as native speakers of a dialect (for example, Peter Trudgill in Norwich) may have intuitions about its structure; so also might non-native speakers who have developed an intimate knowledge of the structure of a dialect (see J. Milroy 1981 for an example). But descriptions of non-standard dialects generally use intuition as an aid to focusing the