

Sociolinguistics

Method and Interpretation

Lesley Milroy and Matthew Gordon



Blackwell
Publishing

Sociolinguistics

Language in Society

GENERAL EDITOR

Peter Trudgill, Chair of English
Linguistics, University of Fribourg

ADVISORY EDITORS

J. K. Chambers, Professor of Linguistics,
University of Toronto

Ralph Fasold, Professor of Linguistics,
Georgetown University

William Labov, Professor of Linguistics,
University of Pennsylvania

Lesley Milroy, Professor of Linguistics,
University of Michigan, Ann Arbor

- 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
(third edition), *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 Schröder
- 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*
- 13 Bilingualism (second edition),
Suzanne Romaine
- 14 Sociolinguistics and Second Language
Acquisition, *Dennis R. Preston*
- 15 Pronouns and People: The Linguistic
Construction of Social and Personal
Identity, *Peter Mühlhäusler and Rom Harré*
- 16 Politically Speaking, *John Wilson*
- 17 The Language of the News Media,
Allan Bell
- 18 Language, Society and the Elderly:
Discourse, Identity and Ageing,
*Nikolas Coupland, Justine Coupland and
Howard Giles*
- 19 Linguistic Variation and Change,
James Milroy
- 20 Principles of Linguistic Change, Volume I:
Internal Factors, *William Labov*
- 21 Intercultural Communication: A Discourse
Approach (second edition), *Ron Scollon
and Suzanne Wong Scollon*
- 22 Sociolinguistic Theory: Language
Variation and Its Social Significance
(second edition), *J. K. Chambers*
- 23 Text and Corpus Analysis: Computer-
assisted Studies of Language and Culture,
Michael Stubbs
- 24 Anthropological Linguistics, *William Foley*
- 25 American English: Dialects and Variation,
Walt Wolfram and Natalie Schilling-Estes
- 26 African American Vernacular English:
Features, Evolution, Educational
Implications, *John R. Rickford*
- 27 Linguistic Variation as Social Practice:
The Linguistic Construction of Identity in
Belten High, *Penelope Eckert*
- 28 The English History of African American
English, *edited by Shana Poplack*
- 29 Principles of Linguistic Change, Volume
II: Social Factors, *William Labov*
- 30 African American English in the Diaspora,
Shana Poplack and Sali Tagliamonte
- 31 The Development of African American
English, *Walt Wolfram and Erik R. Thomas*
- 32 Forensic Linguistics: An Introduction
to Language in the Justice System,
John Gibbons
- 33 An Introduction to Contact Linguistics,
Donald Winford
- 34 Sociolinguistics: Method and Interpretation,
Lesley Milroy and Matthew Gordon

Sociolinguistics

Method and Interpretation

Lesley Milroy and Matthew Gordon



Blackwell
Publishing

© 2003 by Lesley Milroy and Matthew Gordon

350 Main Street, Malden, MA 02148-5018, USA
108 Cowley Road, Oxford OX4 1JF, UK
550 Swanston Street, Carlton South, Melbourne, Victoria 3053, Australia
Kurfürstendamm 57, 10707 Berlin, Germany

The right of Lesley Milroy and Matthew Gordon to be identified as the Authors of this Work has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

All rights reserved. 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, except as permitted by the UK Copyright, Designs, and Patents Act 1988, without the prior permission of the publisher.

Some material in this book was originally published in Lesley Milroy,
*Observing and Analysing Natural Language:
A Critical Account of Sociolinguistic Method* (Blackwell, 1987).

First published 2003 by Blackwell Publishing Ltd

Library of Congress Cataloging-in-Publication Data

Milroy, Lesley.

Sociolinguistics : method and interpretation / Lesley Milroy and Matthew Gordon.

p. cm. — (Language in society ; 34)

Includes bibliographical references and index.

ISBN 0-631-22224-3 (alk. paper) — ISBN 0-631-22225-1 (pbk. : alk. paper)

I. Sociolinguistics. I. Gordon, Matthew. II. Title. III. Series: Language in society (Oxford, England) ; 34.

P40 .M544 2003

306.44—dc21

2002007988

A catalogue record for this title is available from the British Library.

Set in 10.5/12pt Ehrhardt
by Graphicraft Limited, Hong Kong
Printed and bound in the United Kingdom
by MPG Books Ltd, Bodmin, Cornwall

For further information on
Blackwell Publishing, visit our website:
<http://www.blackwellpublishing.com>

Contents

| | |
|--|-----------|
| List of Figures | ix |
| List of Tables | x |
| Series Editor's Preface | xi |
| Preface | xii |
| Acknowledgments | xiv |
| | |
| 1 Sociolinguistics: Models and Methods | 1 |
| 1.1 Data and Theory | 1 |
| 1.2 Earlier Approaches to Linguistic Description | 9 |
| 1.2.1 The American descriptivists | 9 |
| 1.2.2 Traditional dialectology | 11 |
| 1.3 Adaptations of the Traditional Model | 15 |
| 1.3.1 Between paradigms: Early urban studies | 16 |
| 1.3.2 Bridging paradigms: Adaptations of traditional dialectology | 19 |
| | |
| 2 Locating and Selecting Subjects | 23 |
| 2.1 Introductory | 23 |
| 2.2 Representativeness | 24 |
| 2.2.1 Some general principles | 24 |
| 2.2.2 Defining the sampling universe | 26 |
| 2.2.3 Stratification and sample size | 28 |
| 2.3 Quota and Judgment Sampling | 30 |
| 2.4 Research Objectives and Sampling: Some Examples | 33 |
| 2.5 Sampling and Age | 35 |
| 2.6 Sampling and Social Class | 40 |
| 2.7 Concluding Remarks | 47 |

| | | |
|----------|---|-----------|
| 3 | Data Collection | 49 |
| 3.1 | General Issues | 49 |
| 3.2 | Survey Approaches to Data Collection | 51 |
| 3.2.1 | Written questionnaires | 51 |
| 3.2.2 | Fieldworker-administered surveys | 54 |
| 3.2.3 | Rapid and anonymous surveys | 56 |
| 3.3 | Sociolinguistic Interviews | 57 |
| 3.3.1 | The structure and design of the interview | 57 |
| 3.3.2 | Interviews as speech events | 61 |
| 3.3.3 | Interview strategies for eliciting casual speech | 65 |
| 3.4 | Participant Observation | 68 |
| 3.5 | Balancing Fieldwork Strategies: The Belfast Project | 72 |
| 3.5.1 | The community studies | 73 |
| 3.5.2 | The Belfast doorstep survey | 76 |
| 3.5.3 | The rural hinterland study | 78 |
| 3.6 | Research Ethics | 79 |
| 3.6.1 | Informed consent | 79 |
| 3.6.2 | Preservation of anonymity and access to recordings | 80 |
| 3.6.3 | Surreptitious recording | 81 |
| 3.6.4 | The researcher's responsibilities to the community | 84 |
| 3.7 | Concluding Remarks | 87 |
| 4 | Language Variation and the Social World: Issues in Analysis and Interpretation | 88 |
| 4.1 | Introductory | 88 |
| 4.2 | Social Categories and Theories of Change | 92 |
| 4.3 | Social Class and Sociolinguistic Research | 95 |
| 4.3.1 | Models of social class | 95 |
| 4.3.2 | Linguistic markets | 97 |
| 4.3.3 | Interpreting correlations between language and social class | 98 |
| 4.4 | Sex and Gender | 100 |
| 4.4.1 | Introduction | 100 |
| 4.4.2 | Social class and gender | 101 |
| 4.4.3 | Gender and language variation: Some further issues | 103 |

| | | |
|----------|--|------------|
| 4.5 | Ethnicity and Race | 108 |
| 4.5.1 | Introduction | 108 |
| 4.5.2 | Language, race, and ethnicity in the United States | 110 |
| 4.5.3 | Language and ethnicity in Northern Ireland | 112 |
| 4.6 | Concluding Remarks | 115 |
| 5 | Social Relationships and Social Practices | 116 |
| 5.1 | Introductory | 116 |
| 5.2 | The Concept of Social Network | 117 |
| 5.2.1 | Social network and community of practice | 118 |
| 5.2.2 | Social networks and language variation | 119 |
| 5.2.3 | Network structure and language shift in bilingual communities | 124 |
| 5.2.4 | Weak network ties and theories of language change | 127 |
| 5.3 | Dialect Leveling | 130 |
| 5.4 | The Sociolinguistics of Mobility: Place as a Cultural Concept | 133 |
| 5.5 | Concluding Remarks | 134 |
| 6 | Investigating Phonological Variation | 136 |
| 6.1 | Introductory | 136 |
| 6.2 | Identifying and Selecting Variables | 138 |
| 6.2.1 | Types of phonological variables | 138 |
| 6.2.2 | Choosing and specifying variables | 139 |
| 6.2.3 | Pilot studies | 141 |
| 6.3 | Measuring Variation | 143 |
| 6.3.1 | Auditory techniques | 144 |
| 6.3.2 | Instrumental techniques | 145 |
| 6.3.3 | An evaluation of measurement techniques | 148 |
| 6.4 | The Linguistic Analysis of Phonological Variables | 152 |
| 6.4.1 | Defining the range of variation | 152 |
| 6.4.2 | Conditioning factors | 154 |
| 6.5 | The Relationships among Variants: Issues of Interpretation | 156 |
| 6.6 | Comments about Quantification | 161 |
| 6.6.1 | Counting matters | 161 |
| 6.6.2 | How many tokens are needed? | 163 |
| 6.6.3 | Use of statistical analysis | 164 |

| | | |
|----------|--|------------|
| 7 | Beyond Phonology: Analyzing and Interpreting Higher Level Variation | 169 |
| 7.1 | Introductory | 169 |
| 7.2 | Data Collection Issues | 172 |
| 7.3 | Grammatical Variables and Historical Linguistics | 176 |
| 7.4 | Syntactic Analysis and the Sociolinguistic Variable | 179 |
| 7.4.1 | Some early work | 179 |
| 7.4.2 | The principle of accountability and the envelope of variation | 180 |
| 7.5 | Semantic Equivalence and the Discourse Context | 185 |
| 7.6 | Variationist Analysis and Syntactic Theory | 190 |
| 7.7 | Concluding Remarks | 196 |
| 8 | Style-Shifting and Code-Switching | 198 |
| 8.1 | Introductory | 198 |
| 8.2 | Style as a Response to Situation | 200 |
| 8.2.1 | Stylistic variation as a function of attention paid to speech | 200 |
| 8.2.2 | Style as audience design | 204 |
| 8.3 | Style as Initiative and Strategic | 206 |
| 8.4 | Code-switching | 209 |
| 8.4.1 | Introductory | 209 |
| 8.4.2 | “Allocational” and “interactional” paradigms | 212 |
| 8.4.3 | Language choice and code-switching: three case studies | 214 |
| 8.4.4 | Bilingual conversations | 216 |
| 8.5 | Concluding Remarks | 222 |
| 9 | Epilogue | 223 |
| | Notes | 227 |
| | References | 230 |
| | Index | 257 |

List of Figures

| | | |
|-----|--|-----|
| 3.1 | Characteristic network of modules for adolescent or young adult speaker (adapted from Labov 1984: 35) | 59 |
| 3.2 | Location of five speech communities in Belfast | 74 |
| 6.1 | A spectrogram showing “hotdog” as produced by a male speaker | 146 |
| 6.2 | F1 and F2 frequency values of vowels as produced by a male speaker from upstate New York (from Labov 1994: 180) | 147 |
| 6.3 | The vowel changes of the Northern Cities Shift (after Labov 1994: 191) | 159 |
| 6.4 | Another view of the Northern Cities Shift (after Gordon 2001b: 197) | 160 |
| 6.5 | The Sydney speech community: core and periphery (after Horvath 1985: 71) | 167 |
| 7.1 | The use of -(E)S (%) as opposed to -(E)TH in the third-person singular in a corpus from 1540–59, excluding HAVE and DO (male writers) (from Nevalainen 1999: 519) | 178 |
| 7.2 | Multiple negation according to social rank. Upper ranks: royalty, nobility, gentry; middle ranks: professionals, merchants, social aspirers; lower ranks: other ranks below the gentry (from Nevalainen 1999: 523) | 179 |
| 8.1 | Percentage of intervocalic /t/ voicing by four newscasters on two New Zealand radio stations, A and B (adapted from Bell 1984: 171) | 204 |

List of Tables

| | | |
|-----|--|-----|
| 4.1 | (h) in Bradford and Norwich: percentage of zero realization (after Chambers and Trudgill 1998: 59) | 89 |
| 4.2 | (o) as realized by Newcastle women, by age and social class | 90 |
| 4.3 | (o) as realized by Newcastle men, by age and social class | 90 |
| 4.4 | Percentage of dental stop realizations of (th) and (dh) according to age and sex split by type of network (after Dubois and Horvath 1998: 254) | 91 |
| 7.1 | Frequency of multiple negation in the Milton Keynes data set (Cheshire 1999: 61) | 172 |
| 8.1 | Frequency scores for two variables in three styles for 13 Belfast speakers | 202 |
| 8.2 | Correlation of raising of (aj) with combined sex and social category, separating two clusters of burnout girls (figures are Varbrul scores) (Eckert 2001: 125) | 208 |
| 8.3 | Networks and codes of <i>el bloque</i> | 215 |

Series Editor's Preface

In Lesley Milroy's *Observing and Analysing Natural Language*, published earlier in this series, there was distilled, as I said at the time, the collective wisdom of the first generation of sociolinguists about how to *do* sociolinguistics. This wisdom included, most of all, that of Lesley Milroy herself, who was one of the best practitioners of this form of linguistics there had ever been. Here was an author who really knew what she was talking about because, in her own pioneering work, she had confronted and overcome many of the thorniest practical and theoretical problems that the field had had to offer. Like most other people teaching sociolinguistics, I had often been approached by students who were excited about work they had read in sociolinguistics, who wanted to carry out work of this type themselves, but who had no idea how to set about doing it. Lesley Milroy's book was therefore a godsend; it was a book to refer such students to with gratitude and with confidence that their questions would be answered.

Many things have changed since that very welcome publication. Sociolinguistics has acquired new concepts, new problems, new concerns, new methodologies, new theories, and new analytical tools. One thing has not changed, however. Lesley Milroy remains one of the best practitioners of sociolinguistics there has ever been. In this new book, therefore, we have once more an expert, insightful, exciting and up-to-date guide to sociolinguistic methodology and interpretation that will be invaluable to all those concerned with the carrying out of sociolinguistic fieldwork, and the meaningful analysis of the data obtained through such fieldwork. Lesley Milroy and Matthew Gordon deal with the history of sociolinguistic methodology, modern methods in fieldwork and data collection, recent approaches to the handling of data, and the relevance of social categories, social relationships and social practices. Once again, we know which book to refer our students to with gratitude and confidence.

Peter Trudgill

Preface

Over the past decade and a half the field of sociolinguistics has experienced remarkable growth which is marked not simply by the continuing attraction of new scholars to the field but more importantly by the expanding range of approaches now practiced by sociolinguists. Evidence of the expanding interests of sociolinguistic researchers can be seen in the kinds of linguistic phenomena they investigate, the data they consider, the analytical tools they employ to uncover patterning, and the linguistic and social theories they draw upon to interpret their results. This book seeks to provide readers with a sense of the range of this research.

While questions of method constitute a major focus of our discussion, the book is not intended to be a handbook or an inventory of techniques, although it certainly is designed to be of practical value to anyone interested in studying the ways people use language in various 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. Some apparently innocuous methods – which are in fact associated with a specific theoretical paradigm – can often conceal important underlying assumptions. Methodological problems and principles will therefore be discussed not only in practical terms, but in terms of the 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.

The origins of this book lie in Lesley Milroy's *Observing and Analysing Natural Language* (OANL) which first appeared in 1987. While the general orientation of that work has been maintained, and some of its material has been reproduced here, the tremendous expansion of the field has necessitated that the original work be substantially revised and updated for the current project. A good deal of new material has also been included to treat issues that have since emerged as significant (see, for example, the discussions of instrumental techniques for analyzing phonological variation (section 6.3.2)

and the treatment of style-shifting as a strategic maneuver (section 8.3)). The additional perspective provided by the co-author, Matthew Gordon, serves to distinguish further the current work from *OANL*.

The basic structure of the book partly follows that of *OANL*. Chapter 1 offers a theoretical introduction to the general framework of variationist sociolinguistics, and is followed in chapters 2 and 3 by a discussion of study design and methods of data collection. Chapters 4 and 5 explore issues related to the social dimensions of language variation, and chapters 6 and 7 focus on linguistic issues, discussing various aspects of data analysis and interpretation related to phonological variation, and grammatical variation. Finally, style-switching and code-switching are examined in chapter 8.

Acknowledgments

A great many people have contributed in various ways to making this book a reality, and we gratefully acknowledge their help and influence. First of all we are indebted to our colleagues and students at the University of Michigan and the University of Missouri for providing moral support and an intellectually stimulating environment. Various conversations with Charles Boberg, Dave Britain, Jack Chambers, Gerry Docherty, Penny Eckert, Naomi Nagy, Robin Queen, Natalie Schilling-Estes, Peter Trudgill, and Dominic Watt have been particularly helpful and, of course, our debt to William Labov is apparent throughout the book. We also wish to thank Vicki Carstens, Aidan Coveney, Paul Foulkes, Janet Fuller, Jim Milroy, Acrisio Pires, and Robin Queen for their thoughtful comments on various chapters. We are particularly indebted to Keith Walters who read and commented on the manuscript in its entirety, and to Bridget Anderson and Jennifer Nguyen for their assistance in preparing the text for publication. Jennifer also undertook the task of compiling the index.

The authors and publisher gratefully acknowledge the following for permission to reproduce copyright material:

Figure 6.5 from B. M. Horvath, *Variation in Australian English: The Sociol dialects of Sydney*. © 1985 by Cambridge University Press;

Figures 7.1 and 7.2 from T. Nevalainen, "Making the best use of 'bad' data: Evidence for sociolinguistic variation in Early Modern English," in *Neuphilologische Mitteilungen*. © 1999;

Figure 8.1 from A. Bell, "Language Style as Audience Design," *Language in Society* 31 (2). © 1984 by Cambridge University Press;

Table 4.1 from J. K. Chambers and P. Trudgill (eds.), *Dialectology* (2nd edn.). © 1998 by Cambridge University Press;

Table 4.4 from S. Dubois and B. M. Horvath, "Let's Tink about Dat: Interdental Fricatives in Cajun English," *Language Variation and Change* 10 (3). © 1998 by Cambridge University Press;

Table 8.3 from P. Eckert, “Style and Social Meaning,” in P. Eckert and J. Rickford (eds.), *Style and Sociolinguistic Variation*. © 2001 by Cambridge University Press;

The Panjabi excerpt and translation on p. 211, from V. Chana and S. Romaine, “Evaluative reactions to Panjabi-English code-switching,” in *Journal of Multilingual and Multicultural Development* 5. © 1984 by Multilingual Matters.

Lesley Milroy would also like to thank the US National Endowment for the Humanities and the College of Languages, Science and the Arts at the University of Michigan for financial support during the process of writing this book.

1

Sociolinguistics: Models and Methods

1.1 Data and Theory

This book focuses primarily on the methods and theories underlying the quantitative paradigm of sociolinguistic research pioneered by William Labov, with the goal of providing a resource for investigators who are setting up a research project, large or small. This tradition of research is sometimes called *variationist*, to distinguish it from other sociolinguistic subfields. We consider in subsequent chapters data collection, analytic procedures, and interpretation of results, with continuing attention to the theories and assumptions underlying research practice. Like all fields of enquiry, variationist theory has developed a distinctive orientation to its object of investigation (i.e., human language), and a distinctive set of research questions which, while not always explicitly articulated, provide the characteristic focus of those investigations. Variationists do not of course operate independently of other branches of linguistic science, nor indeed of other kinds of sociolinguistics. Furthermore, their orientation, and sometimes the assumptions underlying their theories, are often best understood with reference to historical antecedents.

Mindful of these observations, we have approached the task of writing this book with the conviction that effective researchers need to develop an awareness of the assumptions underlying practice in their fields, so that they may, if necessary or appropriate, coherently query those assumptions. They also need to develop an ongoing awareness of the relationships between their own fields and others – and of the historical antecedents that have shaped their field – sometimes by providing a framework against which practitioners react. A clear example of such a reaction is the critical stance of Labov's early work to what Chambers (1995) describes as *the axiom of categoricity* – the traditional assumption in linguistics that language structure should be treated as invariant. In accordance with this principle, variability

has often been dismissed as unstructured and thus of little theoretical value, but Labov's classic sociolinguistic studies in Martha's Vineyard (1963) and New York City (1966) demonstrated that the trajectories of specific linguistic changes could be inferred from the observation of patterns of variation in contemporary speech communities (see further section 2.5 below). He thus reacted quite radically not only against the axiom of categoricity but against the Saussurian dichotomy, fundamental to structural linguists, which held that the synchronic study of language states was an endeavor entirely separate from the diachronic study of language change. Labov also reacted in these influential early works against the methods and assumptions of contemporary dialectological research – a field quite distinct from the mainstream linguistics of the 1960s. These intersecting reactions still inform many of the assumptions and practices that characterize variationist theory as a subfield.

In this chapter, we expand on these observations, attempting in the remainder of this section to locate the methods and assumptions of variationist theory in relation to those of adjacent fields. In section 1.2 we examine two important historical antecedents to quantitative sociolinguistics – the work of the American descriptivists and that of traditional dialectologists in both the United States and Europe – and in section 1.3 we explore further the relationship between the dialectological and variationist traditions by considering examples of projects that bridge the two approaches.

Johnstone (2000b: 1) points out that contemporary sociolinguistics comprises a great many different traditions of research which address correspondingly different sets of research questions. However, all sociolinguists share a common orientation to language data, believing that analyses of linguistic behavior must be based on empirical data. By this we mean data collected through observation, broadly defined, as opposed to data constructed on the basis of introspection. The most commonly studied data among sociolinguists are those representing speakers' performance – the way they actually use language. Still, researchers may observe elements other than language use. Sociolinguists are often interested in subjective responses to particular linguistic behaviors (e.g., a specific feature or a variety/dialect) and may observe them by eliciting evaluations of speech as is done in perceptual dialectology (see section 1.3.2). Researchers may also make use of speakers' self-reports of their usage (see, e.g., our discussion of written questionnaires in section 3.2.1). Such data can be useful in examining the effects of language ideology (see Milroy and Milroy 1999), but it is important to recognize, however, that such reports are not generally accepted by sociolinguists uncritically as "true" reflections of actual usage.

This general orientation to language data is shared with adjacent fields such as linguistic anthropology and conversation analysis (see Psathas 1995; Pomeranz and Fehr 1997). Qualitative traditions of research such as those of Gumperz (1982a) and Hymes (1972), which emerge from linguistic

anthropology, have been influential from the earliest days of sociolinguistics, continuing to influence contemporary subfields such as discourse analysis (Schiffrin 1994) and interactional sociolinguistics (Gumperz 1982a; Brown and Levinson 1987). Although qualitative research characteristically does not focus on patterns of linguistic variation or employ quantitative methods (but see Schiffrin 1987 for an example of an extensive project which combines quantitative and qualitative procedures), it continues to exert an important influence on variationist theory. Particularly important is the emphasis of linguistic anthropologists on ethnographic methods of observation which attempt to uncover patterns of informal social organization in speech communities, with a primary focus on speakers as social actors, rather than on abstract language patterns. This focus enables a richer and more accountable interpretation of social information than is found in much of variationist research (see further section 3.4).

Like sociolinguists in general, variationists are trained as linguists and routinely use descriptive and analytic tools that are common throughout the field of linguistics. In fact, one regularly finds in the variationist literature descriptive accounts of particular linguistic phenomena which employ standard concepts from syntactic or phonological theory. For example, Martin and Wolfram (1998) examine negative concord and other syntactic features in African American English using a Government Binding framework. Guy and Boberg (1997) offer an account of coronal stop deletion in English that draws on the obligatory contour principle – a construct borrowed from formal phonology. Nevertheless, despite sharing some analytical common ground, such accounts, and those of variationists generally, differ from the accounts of contemporary theoretical linguistics in at least two fundamental ways: (1) they involve differing orientations to data, and (2) they derive from distinct approaches to linguistic variation.

As noted above, variationists, like other sociolinguists, tend to base their analyses on observed data. Traditionally these data have often been gathered in the context of conversational interviews in which the subject (or informant) remains unaware that his or her linguistic usage is the focus of investigation (for alternatives to such methods of data collection see chapter 3 and section 7.2). This source of data clearly contrasts with those sources often used by mainstream theoretical linguists. In this tradition, investigators may rely on data that they themselves construct, drawing on their own intuitions. Alternatively, when dealing with languages which they do not command natively, the investigators may elicit forms from, or verify their own constructions with, native-speaker informants. Thus, the data arise from an explicitly metalinguistic context, one in which the investigator and any informants are thinking about language. Here the questions are of the form “Can you say X?” By contrast, the sociolinguist’s questions are closer to “Do you say X?” though, since the data are usually gathered through observation rather than

elicitation, such questions are not made explicit. The sociolinguist's orientation toward data is further distinguished from that of the theoretical linguist's by the former's adherence to the principle of accountability, which is discussed in the following section.

A more fundamental distinction between theoretical linguistics and variationist sociolinguistics relates to their respective approaches to variability. The principles set out by Weinreich, Labov, and Herzog (1968) are still taken as axiomatic by variationists; namely, a language system that did not display variability would not only be imaginary but dysfunctional, since structured variability is the essential property of language that fulfils important social functions and permits orderly linguistic change. Chambers (1995: 12–32; see also Chambers and Trudgill 1998: 127) has discussed the role of the *linguistic variable* as a structural unit, parallel to such units as the phoneme and noun phrase in linguistic theory. Chambers points out that, from the earliest days of structural linguistics, analysts produced descriptions and generalizations based on an underlying assumption that linguistic structure was fundamentally categorical. Following the Axiom of Categoricity, language is seen as operating with a kind of mathematical consistency. Still, it has always been known that speakers varied in their realizations of particular abstract linguistic structures (intraspeaker variation) and, furthermore, that usage varies across speakers (interspeaker variation). Thus, for example, /æ/ (the vowel of *cat* and *bad*) in many varieties of American English is realized with a range of vowel qualities from a low front monophthong [æ] to a high front diphthong [ɪə] (see Labov 1994; Gordon 2001b). It is also common for plural subjects sometimes to take singular and sometimes plural forms of the verb BE (as in *we was sleeping*/*we were sleeping*) (Tagliamonte and Smith 2000; see also Tagliamonte 1998). In the past many linguists (notably, Edward Sapir) have displayed sensitivity to the pervasiveness of variation. Nevertheless, the mainstream linguistic approach to the plain fact of variability has often been to exclude it from consideration in the interests of providing a coherent and elegant descriptive and theoretical account. This orientation was captured in Chomsky's oft-quoted statement that the primary concern of linguistic theory is "an ideal speaker-listener, in a completely homogeneous speech-community" (1965: 3). In this tradition variability is treated as a methodological complication: it introduces a kind of noise which obscures the important underlying invariance. The perceived peripherality of variation to theoretical matters is reinforced by a dichotomy that segments off language structure (competence) from language use (performance) and assumes many kinds of variation to be the purview of the latter (see further below).

Such an orientation to variability describes alternative realizations like [æ] versus [ɪə], or *was* versus *were* following a plural subject as either the outcome of *dialect mixing*, held to be a temporary situation of instability, or

instances of *free variation*. However, as the psychologist Fischer noted “Free variation is a label, not an explanation. It does not tell us where the variants came from nor why the speakers use them in differing proportions, but is rather a way of excluding such questions from the scope of immediate enquiry” (Fischer 1958: 47–8; cited by Chambers and Trudgill 1998: 128). As their name implies, variationists view variation as central to their immediate enquiry; rather than treating, and often dismissing, variation as free, accidental, unconstrained, or temporary, variationists begin with the assumption that variation is structured and seek to uncover patterning. In truth, many theoretical linguists approach variation with similar assumptions, and treatments of variation have made their mark within theoretical linguistics from time to time. For example, Pollock (1989) used evidence of the variable position of infinitival verbs in French and in English to propose a radical revision of sentence structure which resulted in major changes to mainstream syntactic theory. Similarly, Diesing (1992) presented a highly influential proposal for how interpretation relates to syntactic position, based on variable word orders in German and their semantic correlates. Moreover, alternations like the English dative shift (*give the book to her* vs. *give her the book*) and the variable position of objects in verb particle constructions (*turn off the light* vs. *turn the light off*) are also time-honored topics of study within the generative paradigm.

Nevertheless, one obvious difference between treatments of variability within the generative tradition and within sociolinguistics is that the latter make reference to social (i.e., extralinguistic) as well as linguistic information in specifying the constraints on variability. Thus, frequencies of particular variants are constrained not only by different linguistic contexts (type of following consonant in the case of (æ); type of grammatical subject in the case of (BE PAST)) but also by social characteristics of the speaker such as gender, age or status and the kind of social context (interview talk vs. peer interaction, for example) in which language samples are embedded.

A somewhat more subtle distinction between generative and variationist approaches to variability stems from the emergence of the linguistic variable as a structural unit. Chambers (1995) elaborates in some detail the significance of this development for the sociolinguistic enterprise and for linguistic theory more generally. Examples of underlying linguistic variables would thus be (æ) or (BE PAST), as mentioned above. Variants which realized those abstract variables would include, respectively, [æ] and [ɪə] (and countless intermediate forms) and *was* and *were*. A major goal of the variationist enterprise is to specify and order the constraints which lead to one choice rather than another. The linguistic variable works in terms not of categorical use, but of greater or less frequency of one variant than another, so marking the abandonment of the axiom of categoricity. However, like other structural

linguistic units, it is understood as an abstraction underlying actual realizations (see further section 6.5 below).

Chambers suggests that the abandonment of the axiom of categoricity renders variationist theories irreconcilable with those of contemporary generative linguistics, since the latter paradigm abstracts away from variability while variationists treat it as central. It is worth considering for a moment some reasons for Chambers's rather pessimistic verdict. First, the constructs of mainstream generative linguistics have become more rather than less abstract over the years. Sidnell (2000) notes the continuing relevance of Givón's (1979) critique, namely that the abstraction that was originally devised as a point of methodological convenience has become progressively more prominent. Thus the focus is no longer on the detailed linguistic rules of early generative theory, but on the specification of broad principles and parameters constraining the form of universal grammars (Cook and Newson 1996). For this reason, the hope for fruitful collaboration between different traditions of linguistic research expressed by Labov (1975) has become less likely to be realized in the intervening quarter century. Labov's (1996) examination of problems in the use of intuitions (introspective judgments of grammatical well-formedness), either of the analyst or of the individual whose dialect is being studied, is both an update of and sequel to his 1975 monograph. In the course of this more recent article, he contrasts the approaches of variationists and generativists, and examines the roles of intuition and observation in deriving valid linguistic generalizations. While he finds many points of overlap between the two traditions, the optimistic hopes of collaboration expressed in his 1975 paper are noticeably absent.

Second, the distinction between *competence* and *performance*, first expounded by Chomsky in 1965, remains problematic to all sociolinguists. A speaker's competence is the underlying ability to produce and interpret well-formed sentences in a given language and to distinguish well-formed from ill-formed strings. The specifics of such competence are generally established by eliciting intuitions (or using the analyst's own intuitions) of grammaticality. Performance, on the other hand, covers not only the manifestation of competence on actual occasions of language use, but the effects of memory, perception, and attention on language behavior. In 1986, Chomsky revised the competence/performance dichotomy, preferring a distinction between I(nternal) and (E)xternal language. As Sidnell (2000) points out, this change in terminology involved no significant alteration in the underlying abstraction except a slight change of focus on what constitutes E-language. While generativists are interested exclusively in competence/I-language and have not elaborated any coherent theory of performance/E-language, the distinction is problematic to sociolinguists, most obviously because it treats language as intrinsically asocial (see, again, Labov 1996 for

a wide-ranging discussion of the issues). Much systematically variable language behavior is treated globally as performance/E-language, along with the linguistic effects of memory and attention. For variationists, not only is variation essential and intrinsic to human language, but the detail of systematic, socially embedded variable behavior is the key to an understanding of the dynamics of language change. In an account of variability of pronominal reference in Spanish, Cameron (1996) aligns himself with Prince (1988), arguing specifically for an enlarged conception of competence to include memory and attention phenomena.

In the early days of sociolinguistics Hymes (1972) pointed out that Chomsky's competence was only one kind of linguistic competence. Not only did competent speakers produce and interpret well-formed sentences, but they also used varieties of language from a systematically structured community repertoire to perform social actions in contextually appropriate ways that were meaningful to other members. They also recognized particular utterances as ironic, teasing, serious, etc. (Hymes 1972, 1974). Any socially informed linguistics concurs with Hymes in conceiving of knowledge "with a view to its fundamental role in communication between socially located actors in continuously changing human societies" (Sidnell 2000: 41).

While these rather fundamental incompatibilities need to be acknowledged, it is important not to exaggerate the impenetrability of the boundaries between sociolinguistics and theoretical (usually generative) linguistics or to further polarize the two research traditions. On the generative side, Henry (1995) has produced an account of dialect acquisition in Belfast, Northern Ireland within a principles and parameters framework, which takes account not only of variability but is based largely on observed data (see also Wilson and Henry 1998). Prince (1988) has argued for a much enlarged concept of competence that takes account of observed and naturalistic language data. Schütze (1996) provides a critique of the role of intuition in syntax research, arguing for a radical rethinking of its empirical base. Conversely, variationists regularly work with frameworks developed by theoretical linguists. For example, Cornips (1998) examines syntactic variation within a principles and parameters framework; Nagy and Reynolds (1997), Guy (1997), and Zubritskaya (1997) work with Optimality Theory from theoretical phonology; and Docherty et al. (1997) examine the descriptive adequacy of theoretical accounts of glottalization phenomena from a variationist perspective.

We conclude this review of the interrelationships between linguistics, sociolinguistics, and cognate disciplines with some comments on the orientation of variationist theory toward the social dimension of language behavior. Gumperz (1982a) has pointed 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. Our discussion so far has in fact assumed this orientation by locating variationist theory as a subdiscipline of linguistics. However, Gumperz then argues that the relatively abstract approach associated with Labov's 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 judgments 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 1982a: 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 behavior 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).

The distinction drawn here between the properties of a variable system and the behaviors of the speaker is an important one that still gives rise to tension in the field. It embodies the chief distinction between qualitative and quantitative, interactional and variationist traditions of sociolinguistic research, the former type being influentially exemplified by Brown and Levinson (1987). Johnstone (2000b) has recently provided a clear account of the methods, goals and assumptions of qualitative sociolinguistics. Over the years, Labov's work has become increasingly oriented to the linguistic system rather than to the speaker, attempting primarily to specify universal patterns of change in vowel systems (Labov 1991, 1994) and to map out the large-scale spatial distributions of these systems (see further section 1.3.2; Labov, Ash, and Boberg, forthcoming). These analyses make little if any reference to social information or to the behaviors of speakers. A more generally unsophisticated treatment, or even neglect, of social factors by variationists has given rise to criticism not only by linguistic anthropologists but by variationists themselves (see, for example, Rickford 1999; Eckert 2000). As we will see in chapter 5, Eckert and others have recently argued strongly for a more socially sophisticated approach to language variation that systematically takes into account the behaviors and motivations of speakers. The chief point we make here, however, is the continuing relevance of the distinction articulated by both Gumperz and Labov between an approach that primarily addresses the properties of variable linguistic systems and one that primarily addresses the behaviors of speakers in their speech communities.

1.2 Earlier Approaches to Linguistic Description

1.2.1 *The American descriptivists*

The American linguists, commonly known as “structuralists” or “descriptivists,” placed a high premium on the development and practice of a rigorous and accountable set of field and analytic methods. In this section we outline the characteristics of their approach and philosophy only insofar as they are particularly relevant to the central concerns of this book. For fuller accounts, the reader is referred to Lepschy (1982), Hymes and Fought (1980), Robins (1967), and Sampson (1980).

The concern with method in mainstream American linguistics from the early decades of the twentieth century until the emergence of Chomsky’s generative grammar is largely attributable to a desire to describe as rapidly and efficiently as possible a large number of dying native American languages. Gumperz (1982a: 12) contrasts the atmosphere of empiricism at that time in America, where scholars were concerned with working in the field, with that in Europe, where they worked in offices. He apparently sees contemporary sociolinguistics as a continuation of this empirical tradition. Following the line of reasoning elaborated by Bloomfield (1926; 1933), the accreditation of a *scientific* status to linguistics was associated with the development of rigorous methods of description. American linguists strove to obtain objectivity by developing accountable procedures for inductively deriving linguistic generalizations from observable data, and an important methodological principle springing from this concern was that the phonological, morphological, syntactic and semantic patterns of a language should be analyzed separately. They should, moreover, be analyzed in that order so that the analyst could remain in touch with the “observable” part of language – the sequence of sound segments with which the description began.

A similar concern with accountability to the data subsequently became the hallmark of variationist work; Labov’s *principle of accountability* extends the general philosophy of accountability to a specifiable procedure which is the cornerstone of quantitative method (see further section 6.1). In this respect his views resemble those of earlier American linguists but differ sharply from those of Chomsky and others working within the generative paradigm. Replacing induction with a hypothetico-deductive mode of reasoning, the generativists argued that no corpus of data, however large, can usefully serve as a basis for linguistic generalizations since any corpus is a partial and accidental collection of utterances (Chomsky 1965: 15). Chomsky’s general point about the inadequacy of corpora as the *only* source of information is surely correct, but in practice seems uncontroversial since intuition

and introspection (either of the investigator or more usually of the informant) have always been used by linguists, including sociolinguists, who work in the field. Voegelin and Harris (1951) discuss the relative roles of observation and intuition (of the informant) as data-gathering procedures, and the same theme is revisited by Rickford (1987). Johnstone (2000b: 71–9) discusses at some length the role of intuition and introspection in sociolinguistic research. Although Chomsky seems to be attacking a straw man in his critique of corpus-based research, the effect of his remarks has been a shift of focus from observation to introspection, and a corresponding removal from mainstream linguistics of the need to be accountable to an independently collected body of data.

Despite this major paradigm shift in linguistics, American descriptivist methods still provide the basis for procedures of data collection and analysis in the field. The extensive studies of Australian languages by Robert Dixon and his associates are obvious examples, as are the continuing efforts of linguists working on the indigenous languages of the Americas (see Mithun 1996 for a review). Rather less obviously, structuralist methods – the most influential of which are outlined below – have been developed in various ways for use in quantitative sociolinguistic work (see particularly Labov 1984).

To get a sense of descriptivists' methods, we can consider their approach to establishing which sounds were contrastive – a procedure that they considered to be the major task of a phonological analysis:

We take an utterance recorded as DEF. We now consult an utterance composed of the segments of DA'F where A' is a repetition of a segment A in an utterance which we had represented as ABC. If our informant accepts DA'F as a repetition of DEF . . . and if we are similarly able to obtain E'BC (E' being a repetition of E) as equivalent to ABC, then we say that A and E . . . are mutually substitutable (or equivalent) as free variants of each other . . . If we fail in these tests we say that A is different from E and not substitutable for it. *The test of segment substitutability is the action of the native speaker: his use of it or his acceptance of our use of it.* (Harris 1951: 31 – our italics)

In the absence of any alternative framework capable of application to a substantial body of data, linguists studying unknown languages still need to establish contrastivity in a similar way (see Healey 1974: 8 for a New Guinea example). And so do variationists; but one respect in which they have advanced the substitution method is by querying the assumption of objectivity in pair testing and showing that native speaker judgments of “same” and “different” do not necessarily correspond in a straightforward way with independently observed phonological patterns (see section 6.5; Milroy and Harris 1980). Harris's painstaking account in the above quotation gives an idea of the care with which the descriptivists formulated their “discovery procedures” as they were called – this basic method of substituting one element for another