Analysing sociolinguistic variation

The study of how language varies in social context, and how it can be analysed and accounted for, are the two key goals of sociolinguistics. Until now, however, the actual tools and methods have been largely passed on through ‘word of mouth’ rather than being formally documented. This is the first comprehensive, ‘how-to’ guide to the formal analysis of sociolinguistic variation. It shows step-by-step how the analysis is carried out, leading the reader through every stage of a research project from start to finish. Topics covered include fieldwork, data organisation and management, analysis and interpretation, presenting research results and writing up a paper. Practical and informal, the book contains all the information needed to conduct a fully fledged sociolinguistic investigation, and includes exercises, checklists, references and insider tips. It is set to become an essential resource for students, researchers and fieldworkers embarking on research projects in sociolinguistics.

SALI A. TAGLIAMONTE is Associate Professor of Linguistics at the University of Toronto. Her research focuses on variation and change in the evolution of English. She is author of African American English in the diaspora: tense and aspect (with Shana Poplack, 2001).
KEY TOPICS IN SOCIOLINGUISTICS

Series editor:
Rajend Mesthrie, University of Cape Town.

This new series focuses on the main topics of study in sociolinguistics today. It consists of accessible yet challenging accounts of the most important issues to consider when examining the relationship between language and society. Some topics have been the subject of sociolinguistic study for many years, and are here re-examined in the light of new developments in the field; others are issues of growing importance that have not so far been given a sustained treatment. Written by leading experts, the books in the series are designed to be used on courses and in seminars, and include useful suggestions for further reading and a helpful glossary.

Already published in the series:
Politeness by Richard J. Watts
Language Policy by Bernard Spolsky
Discourse by Jan Blommaert

Forthcoming titles:
World Englishes by Rakesh Bhatt and Raj Mesthrie
Language and Ethnicity by Carmen Fought
Bilingual Talk by Peter Auer
For Tara, Shaman, Freya and Dazzian
Love
Mom
## Contents

Preface ix  
Notes on codes and abbreviations xi  

1. Introduction  1  
2. Data collection  17  
3. The sociolinguistic interview  37  
4. Data, data and more data  50  
5. The linguistic variable  70  
6. Formulating hypotheses/operationalising claims  99  
7. The variable rule program: theory and practice  128  
8. The how-to’s of a variationist analysis  158  
9. Distributional analysis  191  
10. Multivariate analysis  217  
11. Interpreting your results  235  
12. Finding the story  254  

Glossary of terms  263  
References  267  
Index  277
The variationist approach to sociolinguistics began during the 1960s, when Labov, working with Uriel Weinreich, developed a theory of language change (Weinreich et al. 1968). Thereafter, Labov continued to advance the method and analysis of language variation and change, which today is often referred to as variation theory (e.g. Labov 1963, 1966/1982).

In the 1970s, one of Labov’s graduate students at the University of Pennsylvania was Shana Poplack. In 1981, Shana became a professor of Sociolinguistics at the University of Ottawa’s Department of Linguistics, the same year I entered the MA programme. I was fortunate to be Shana’s student until I completed my Ph.D. dissertation in 1991. Everything you will read in this book has come directly from what has been passed on from this lineage – training, techniques, insights, knowledge, and sheer passion for the field. The entire period from 1981 to 1995 was an invaluable apprenticeship through my studies with Shana and our many collaborations (e.g. Tagliamonte and Poplack 1988, Poplack and Tagliamonte 1989, 1991). I also benefited tremendously from the influence of David Sankoff, whose input to my questions of method and analysis was innumerable.

Most knowledge and learning in variation theory has been acquired like this, passed on through word of mouth, from one researcher to the next (see also Guy 1988: 124). In fact, it has often been noted that the practical details of how to actually do variation analysis are arcane, largely unwritten and, for the most part, undocumented (but see Paolillo 2002). This is precisely why this book was conceived and has now been written. The method needed to be recorded, systematically, thoroughly and straightforwardly.

I had originally intended this book to be completed by the mid-1990s, but academic life is unforgiving for time and relentless for energy. The advantage is that I have had that many more years of experience. Between 1995 and 2005, I have trained some of the next generation of variationist sociolinguists and, as has always been
the case, the students have helped the teacher learn a lot more than what she thought she already knew. Yet there is always room for improvement.

In completing this book, I requested ‘no-holds-barred’ comments from the best methodologists of my colleagues and students. I am indebted to Alex D’Arcy, Ann Taylor, Jennifer Smith and James Walker, who came through with the best feedback one could hope for – an intense amount of red ink. Through the Herculean efforts of my assistant Sonja Molfenter, who took on the unenviable role of ‘book bulldog’ during the revision process, I have taken all their comments into account, and then some. My own method has evolved in just this way, incrementally changing from one research project to the next, one student to the next, in my perpetual efforts to do things more usefully, more efficiently and more transparently. At the same time, the basics endure. The original ideas enshrined in Weinreich et al. (1968), built upon and elaborated by William Labov (Labov 1966/1982) herein are fundamental and pervasive. In sum, this book is simply one user’s tried-and-true manual of best practice.
Notes on codes and abbreviations

Codes in parentheses refer to the community from which the data come (see abbreviations), followed by a single character speaker code which identifies the individual speaker in each community. In some cases additional information may appear, e.g. the audio-tape number. Abbreviations for communities are: BCK = Buckie; CLB = Cullybackey, Northern Ireland; CMK = Cumnock, Scotland; DVN = Devon, Southeast England; ESR = Ex-Slave Recordings; GYE = Guysborough Enclave; GYV = Guysborough Village, Nova Scotia; KID = a corpus of child language acquisition, England; MPT = Maryport, England; NPR = North Preston, Nova Scotia; OTT = Ottawa; ROP2–4 = data collected in Toronto, Canada, in the years 2002–2004 through the Research Opportunities Programs at the University of Toronto; PVG = Portavogie; ROO = Roots Corpus, data collected in remote communities in Northern Ireland, Lowland Scotland and Northwest England; SAM = Samaná, Dominican Republic; TIV = Tiverton, England; TOR = Toronto, Canada; WHL = Wheatley Hill, England; YRK = York, England. Abbreviations for additional corpora include ST1, ST2 and ST3 – all narrative data sets of Canadian English.