Sali A. Tagliamonte

Variationist Sociolinguistics

Change, Observation, Interpretation

WILEY Blackwell

Variationist Sociolinguistics

Language in Society

Launched in 1980, *Language in Society* is now established as probably the premiere series in the broad field of sociolinguistics, dialectology, and variation studies. The series includes both textbooks and monographs by Ralph Fasold, Suzanne Romaine, Peter Trudgill, Lesley Milroy, Michael Stubbs, and other leading researchers.

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 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*

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 (second edition), *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*

35 Text, Context, Pretext: Critical Issues in Discourse Analysis, *H. G. Widdowson*

36 Clinical Sociolinguistics, *Martin J. Ball*

37 Conversation Analysis: An Introduction, *Jack Sidnell*

38 Talk in Action: Interactions, Identities, and Institutions, *John Heritage and Steven Clayman*

39 Principles of Linguistic Change, Volume III: Cognitive and Cultural Factors, *William Labov*

40 Variationist Sociolinguistics: Change, Observation, Interpretation, *Sali A. Tagliamonte*

Variationist Sociolinguistics

CHANGE, OBSERVATION, INTERPRETATION

Sali A. Tagliamonte

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2012

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered Office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Tagliamonte, Sali.
Variationist sociolinguistics : change, observation, interpretation / Sali A. Tagliamonte.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4051-3590-0 (alk. paper) – ISBN 978-1-4051-3591-7 (pbk. : alk. paper)
1. Language and languages–Variation. 2. Sociolinguistics–Research. I. Title.
P120.V37T348 2012
306.44–dc22
2011010578

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

This book is published in the following electronic formats: ePDFs 9781444344448; Wiley Online Library 9781444344479; ePub 9781444344455; mobi: 9781444344462.

Set in 10/12pt Ehrhardt by SPi Publisher Services, Pondicherry, India

1 2012

For
Anna Blanche Lawson
1930–2001

Love you forever Mum,
Sali

Contents

Acknowledgments		x
Foreword		xii
Series Editor's Preface		xiii
Preface		xiv
List of Figures		xvii
List of Tables		xx
1	**Sociolinguistics as Language Variation and Change**	**1**
	Sociolinguistics	1
	The Linguistic Variable	3
	Linguistic Change	8
	The Principle of Accountability	9
	Circumscribing the Variable Context	10
	Evolution of the Linguistic Variable	15
	The Importance of Accountability	19
	Language Variation and Change and Linguistic Theory	21
	Exercises	22
2	**Social Patterns**	**25**
	Social Class	25
	Sex (or Gender)	32
	Style and Register	34
	Mobility in Space and Mobility in Class	35
	Social Network, Communities of Practice	36
	Ethnicity and Culture	38
	The Mass Media	41
	Age	43
	Types of Change	56
	Principles of Linguistic Change	62
	Summary	65
	Exercises	66

3 Linguistic Patterns **71**
Sound Change 74
Morphological Change 76
Syntactic Change 80
Semantic Change 84
Grammaticalization 87
Lexical Effects 91
Exemplar Theory 94
Exercises 97

4 Data and Method **100**
The Speech Community 100
Corpus Building 101
Creating Sociolinguistic Corpora 102
The Individual and the Group 108
Constructing an LVC Study 110
Research Ethics 115
The Gold – Your Data 116
The Real World 117

5 Quantitative Analysis **120**
The Quantitative Paradigm 120
Distributional Analysis 121
Statistical Modeling 121
The Three Lines of Evidence 122
The Case Study – Variable (that) 124
Goldvarb Logistic Regression 126
Challenging the Variable Rule Program 129
Drawbacks to the Variable Rule Program 137
New Toolkits for Variationist Sociolinguistics 138
Summary 156
Exercises 157

6 Comparative Sociolinguistics **162**
Comparison 162
The Comparative Method 163
Comparison in Origins 165
Comparison in Language Contact 166
Standards for Comparison 167
Variable (did) 168
Exercises 173

7 Phonological Variables **177**
Variable (t,d) 179
Variable (ing) 187
Tips for Studying Phonological Variables 195
Exercises 203

8 Morpho-Syntactic Variables **206**
Verbal (s) 207
Adverb (*ly*) 217
Modal (have to) 228
Studying Morpho-Syntactic Variables 235
Exercises 241

9 Discourse/Pragmatic Features **247**
Quotative (*be like*) 247
General Extenders 258
Studying the Discourse/Pragmatic Variable 269
Exercises 277

10 Tense/Aspect Variables **279**
Grammaticalization and Tense/Aspect Variables 280
Future (going to) 281
Perfect (have) 296
Studying Tense/Aspect Variables 308
Exercises 311

11 Other Variables **314**
Variable (come) 315
Variable (Intensifiers) 320
Language and the Internet 336
Studying Unusual Variables 342
Exercises 345

12 Sociolinguistic Explanations **349**
What Are the Constraints on Change? 349
How Does Language Change? 350
How Is a Change Embedded in Social and Linguistic Systems? 350
Evaluation of a Change 351
Statistical Modeling 353
Traditional Explanations 354
The Principle of Interaction 355

Appendix A: Corpora Cited 358
Appendix B: Time Periods in the History of English 359
References 360
Subject Index 392
Index of Linguistic Variables 400

Acknowledgments

A person only ever stands somewhere along the ladder of life. I am indebted to many great minds and generous spirits who have helped me in my work. My students, my mentors, my colleagues, my friends, they are often the same people with no clear distinction among them. This is one of the truly gratifying aspects of doing sociolinguistics – you become part of a social network, a practice, a community.

My students have always been my best critics. Let them know that each one of them has helped immeasurably with this book. Derek Denis, Bridget Jankowski, Dylan Uscher, and Cathleen Waters: every question we considered over the past few years has made its way into these pages. My students in LIN1256, Advanced Language Variation and Change, January–April 2011 deserve special mention for their critical input to the prepublication version of the manuscript. Marisa Brook, Julian Brooke, Matthew Gardner, Heidi Haefale, Chris Harvey, Madeline Shellgren, and Jim Smith have shown me, yet again, how much teaching embeds learning.

My past has also woven its way through the chapters, as I have returned to my early research to integrate the present state of the field with its foundations. I am blessed by having been mentored by some of the greatest contributors to the field. Shana Poplack, David Sankoff, Jack Chambers, Peter Trudgill, and Jenny Cheshire: this book exists only because I have been able to stand on your shoulders. I am also lucky to have had a knowledgeable and attentive set of critics who scoured the draft manuscript and offered their insights, including four anonymous Wiley-Blackwell reviewers, a savvy team of Wiley editors, my new neighbour Victor Kuperman, my pal Paul Foulkes, and even the General Editor, Peter Trudgill, himself.

No field advances without change. Over the last ten years statistical methods have undergone a veritable renaissance. Chapter 5 evolved over several years of consultation on the state of the art in statistical methods in Variationist Sociolinguistics. I am thankful to Harald Baayen, Daniel Ezra Johnson, and John Paolillo for helping me in my ongoing efforts to model linguistic variation and change in ways that are not only insightful, but also statistically sound.

A sociolinguist is never alone in their research. I am lucky to have had a superb group of lab assistants and project coordinators. The latest, Michael Ritter, has the astonishing ability to manage, organize, interview, transcribe, extract, correct, code, copy, copy-edit, run *Goldvarb*, *R, and Ant-Conc*, and everything else I need doing.

I am immensely proud of my own academic progeny who have become my friends; Jen Smith and Alex D'Arcy, where would I be without your ongoing collaborations, savvy insights, and unabashed prodding? My wonderful colleagues Elizabeth Cowper, Elaine Gold, Alana Johns, Keren Rice, and Diane Massam: you have been exuberance, friendship, and community to me since I arrived at UofT in 2001. I have found in your model guidance and sanity. My confidant and best drinking buddy Anthony Warner listens and advises and tells me when I am being silly (this is a more important quality than you might think).

Since my last book, three of my children have become teenagers, the youngest one has started primary school, and I have gained a stepson in the early years of his professional life. This is a great learning ground for a sociolinguist. Dazzian, Freya, Shaman, Tara, and Adrian have taught me much of what I know about age grading, innovation, and incrementation. I am so very thankful to be part of the perpetual state of variation and change, love and commotion we live in. And to Duncan, who is the bedrock of my life, I am eternally grateful to have found in one man, husband, lover, gardener, and friend.

Finally, I would like to thank my mom. What I have been able to accomplish in my life was fostered in the love and support and many other intangible gifts she gave me.

Foreword

My grandparents lived in a small town in Southern Ontario. It was a farming hamlet in one of the oldest settled areas of Ontario, Canada, called Maple Station. They owned the general store, gas station, and post office. The store was always filled with locals. When I visited as a child, I would race to the store every time someone came in, trailing behind the adults to eavesdrop on the conversations. In the evenings, my great-aunts and -uncles would visit. Coming from farming stock, the families were huge. My grandfather had eight brothers and sisters and my grandmother had nine. There were people around all the time. They often talked long into the evening, playing Euchre or Crib. I can still hear the lilting cadence of those voices in my mind. This was a world of regularized past participles, double negation, all kinds of variation in vowels and diphthongs, and strange words and expressions. Little did I know of all that then! At the time, I only listened and marveled at how different they sounded.

My mother, who had grown up in that world, became a teacher, a specialist in early childhood education. Yet there were always aspects of her speech that were very different from the Canadian norms in my surroundings. When she talked to my grandparents or my aunts and uncles (her brother and sisters) on the phone, her voice would shift back toward the speech patterns I heard in Maple Station. Sometimes, when telling stories, I would even hear her use the occasional *I says* or *He come*. And when she quoted anyone in her family, her voice always changed.

While I sound just like any other Canadian, there are still parts of my speech that reflect my mother's vernacular, words like "wee" for "small", expressions like "it's a good job" for "it's a good thing". Even today, when my children make fun of some of the words I use and my pronunciations ("tiger" [tægr], "Saturday" [sɛrdeː], "southern" [sʌwðrn]) I blame my old-line Southern Ontario roots.

These are the realities of language variation and change. Our life histories are a study of continuities and changes, of ancestry and origins, of time and space, of uncommon similarities across time and remarkable differences across generation gaps. Our heritage follows us wherever we go and throughout our lives. For me, the world came alive when I discovered sociolinguistics because it made my experience make sense. The linguistic difference and variety around me had regularity and meaning, system and explanation. May this book make sociolinguistics – and the world of variation around you – more comprehensible to you.

Sali A. Tagliamonte
Toronto, Ontario

Series Editor's Preface

It is not often that one looks at a book and says "this is *it*." That, however, is precisely what I found myself thinking when I first received the text of Sali Tagliamonte's *Sociolinguistics: Variation, Change and Interpretation*. This *really is it* – this is the book that linguistic variation theory has been waiting for. It has not, however, been waiting too long. Now is exactly the right time for this book to appear; and, I like to think, the Blackwell's *Language in Society* series is exactly the right place for it to appear. The study of "Language Variation and Change" (LVC) has been with us now, as Professor Tagliamonte says, for about 40 years. My own first encounter with the field, at that time still without a name, was at what I believe to be the first ever academic meeting devoted to the topic, the Colloquium on New Ways of Analysing Variation in English held at Georgetown University in the USA in October 1972. This turned out to be the first of a series of annual NWAV conferences which continue to be held to this day – though without the word "English" in the title now – and indeed at the time of this writing, the next meeting is going to be the fortieth. I don't know what Sali was doing in October 1972, but she was certainly not nearly old enough to be at the meeting. In spite of her comparative youth, however, we are very grateful that it has fallen to her to produce in this book a distillation of all the advances that have been made and all the wisdom that has accrued in our now mature field over the last four decades. She is perhaps uniquely qualified – in terms of her erudition, her field-work experience, her analytical innovations, and the large amounts of data and the wide range of language varieties she has worked on – to write the first book which is truly an introduction to LVC, a summary of its main goals and achievements, and a springboard for future progress. She has done this, moreover, in a masterly fashion: not only will the reader of this volume learn how work in LVC is done, they will also learn why we do it, and what the benefits are. All languages are variable – variability is an essential component of human language. But it is only in the last 40 years that we have fully understood the degree to which this is so, have investigated the patterning in which variation is involved, and have developed the concepts and techniques for dealing with it – developments which Sali herself has played a very major role in advancing. As this book shows, any linguistic work which attempts to shed light on the nature of the human language faculty and on the nature of linguistic change, without taking account of language's inherent variability, will inevitably fall short.

Peter Trudgill

Preface

What this Book is About

This is a book about the fascinating, intricate and remarkable relationship between language and society, a field that is typically called sociolinguistics. However, this is not a book about everything in sociolinguistics because sociolinguistics is a very diverse and wide-ranging discipline. Taken broadly, sociolinguistics involves studying the interaction of language, culture, and society. This book cannot do all that. Instead, I focus on the type of sociolinguistics that has come to be known as Variationist Sociolinguistics, or "Language Variation and Change" (LVC). This is the type of sociolinguistics I have been practicing in my own research since 1981. This branch of the sociolinguistics tree is known for its focus on language change as well as its quantitative methods and its concern for accountable methodology. It is the study of linguistic variation and change through observation and interpretation.

Variationist Sociolinguistics deals with systematic and inherent variation in language, both in the present (synchrony) and in the past (diachrony). The goal of LVC studies is to understand the mechanisms which link extra linguistic phenomena (the social and cultural) with patterned linguistic heterogeneity (the internal, variable, system of language) (Sankoff 1988a: 157).

Here is the definition from the leading journal, aptly entitled *Language Variation and Change*:

> *Language Variation and Change* is the only journal dedicated exclusively to the study of linguistic variation and the capacity to deal with systematic and inherent variation in synchronic and diachronic linguistics. Sociolinguistics involves analysing the interaction of language, culture and society; the more specific study of variation is concerned with the impact of this interaction on the structures and processes of traditional linguistics. *Language Variation and Change* concentrates on the details of linguistic structure in actual speech production and processing (or writing), including contemporary or historical sources.

This book is written in this spirit, taking the details of variable linguistic structures of language in use and demonstrating how quantitative analysis can tell us something interesting about what we find, i.e. how variation patterns, why it exists, what explains it.

However, this textbook cannot even cover everything within the quantitative sociolinguistic enterprise. A number of subdisciplines have developed which involve specialized

methodological and data-specific practices. Some of these require very specific knowledge that extends beyond what can be covered in a single book. Therefore, I will confine myself to the area of sociolinguistics upon whose foundations these approaches to variation rest and with which I am most familiar. In so doing, I will leave to other experts certain subdisciplines in the field, including sociophonetics with its detailed methods of acoustic measurement and experimentation, discourse analysis with its elaborate qualitative component, aspects of historical linguistics which include variationist techniques, corpus linguistics, and the broad field of sociocultural linguistics. Nevertheless, I hope to demonstrate that variationist sociolinguistic principles and practices, the identification and study of patterns, and all the aspects of the methodology laid out here can be applied in virtually any study of language.

You will find me discussing the same old variable (ing) again. One of my students asked me once in exasperation: *Why do we have to keep talking about variable (ing)?* Why? It provides a familiar model and a good example of how to approach variation, interpret it, and understand it. Besides, there may be some things about variable (ing) we have not discovered yet. I will be sure to find some new variables to talk about too. To support creative thinking I have sprinkled "notes," "tips," and "mini quizzes" throughout the text. Tips provide advice for what the student might encounter in her own research and how to get around it. Notes are elaborations, often my own inner thoughts about research mentioned within the text. Mini quizzes embed learning by questioning the reader on some key point under discussion. I believe that teaching can best be accomplished by "doing." My approach will be to use the findings and observations arising from a series of case studies of "the linguistic variable," the key construct of the discipline, to demonstrate how Variationist Sociolinguistic theory is put into practice. I will embed these studies in the general research trends in the field over the past 40 years. The underlying goal is to show you the links between language and society as they arise from observation and interpretation of variable phenomena.

The book takes as a foundation the major findings of sociolinguistics as put forward in broad-based introductory level textbooks (Wardaugh 2002), with a focus on "Variationist Sociolinguistics" in particular, as synthesized in Chambers (2003). I define "classic" research in sociolinguistics as that conducted by William Labov, Peter Trudgill, Walt Wolfram, Ralph Fasold, and Lesley Milroy. This early research exposed persistent, regular sociolinguistic patterns that have given rise to "sociolinguistic principles." This will be my departure point.

The discoveries of this early body of knowledge is already consolidated in the leading introductory sociolinguistic textbooks in the field. Each chapter ends with a reading list of the major sources I have drawn from. My goal for this book is to put the cumulative findings of the last 30–40 years into context with this foundational work. The findings I will report are meant to broaden and enrich classic sociolinguistic research by bringing the latest evidence to bear on fundamental sociolinguistic observations. Therefore I will focus more on developments to sociolinguistics as put forward in Labov's most recent research as synthesized in his important *Principles of Linguistic Change* volumes (Labov 1994, 2001a, 2010). This work will be brought to the forefront, in the context of, and with reference to, other major research advances in the field, particularly those arising from the journal *Language Variation and Change*. Then, to make practical exactly how this research is done, I will turn to a series of choice linguistic variables. This research encompasses analyses of multiple levels of grammar – phonology, morpho-syntax-semantics, and discourse-pragmatics. Each case study presents findings and observations about how different types of variants are used and how they pattern at the community level and within the systems of grammar of which they

are a part. Each case study interprets the findings within the context of sociolinguistic inquiry as I have defined it above.

The textbook is organized as follows. Chapter 1 introduces "sociolinguistic variation theory" (Sankoff 1988a: 140) as Language Variation and Change (LVC). Chapters 2 and 3 synthesize the observations and findings of LVC research that have led to sociolinguistic principles and sociolinguistic theory. These chapters present a synthesis of the pervasive "patterns" both sociolinguistic and linguistic, as, for example, elucidated in Chambers (2003) or Trudgill (2000), from which LVC has developed. Chapter 4 considers issues relating to data collection, field work, and the key methodological issues of how to deal with the effect of individuals and lexical items. Chapter 5 summarizes the state of the art in quantitative methods and statistical practice. Chapter 6 outlines the comparative sociolinguistic approach. Chapters 7–11 present case studies of linguistic variables from phonology to discourse. Each chapter introduces the variable(s), issues arising from studying them, solutions, and findings. Observations are evaluated both from the perspective of sociolinguistic principles as well as in the context of the prevailing knowledge of the variables in the field. Chapter 12 synthesizes the observations so as to provide explanations for both internal and external patterns of language variation and puts them into the perspective of their social and historical context.

Mini Quiz

Q1 Variation Sociolinguistics is the study of systematic and inherent variation in language, past and present.
(a) True
(b) False
Answer = TRUE

Figures

1.1	Rates of *avoir* and *être* usage with "tomber" per thousand lines of transcription	17
2.1	Idealized pattern for sharp stratification by social class	26
2.2	Idealized pattern for gradient stratification by social class	27
2.3	Idealized pattern for stratification by social class and style – indicator	28
2.4	Idealized pattern for stratification by social class and style – marker	29
2.5	Curvilinear pattern for social class when change originates from the middle class	31
2.6	Idealized pattern of stratification by sex and social class	33
2.7	Idealized pattern of female-led linguistic change	33
2.8	Frequency of phonological and grammatical variables	42
2.9	S-curve of linguistic change	44
2.10	An idealized pattern of linguistic change in progress (generational change)	45
2.11	Overall distribution of quotatives by age in Toronto English, c. 2002–2004	46
2.12	An idealized pattern of age-graded change	47
2.13	An idealized pattern of the adolescent peak	49
2.14	Ch-lenition in Panama c. 1969 and 1982–84	51
2.15	Pattern of a feature increasing in use over 60 years in real time for Jane Doe and apparent time for the speech community	54
2.16	An idealized pattern of a stable linguistic variable	56
2.17	An idealized pattern of linguistic change from across the branches of the family tree, i.e. from outside the community	57
2.18	Distribution of variable (h), York English, c. 1997	67
2.19	Distribution of variable (that), Toronto, c. 2003–2004	67
2.20	Distribution of [f], [t], and Ø variants as opposed to [θ]	68
2.21	Distribution of non-RP variants for three linguistic variables in Norwich English, c. 1972	69
3.1	Idealized pattern for sound change via weakening	75
3.2	Idealized pattern for morphological change via analogical extension – leveling	78

3.3	Idealized change in progress that exhibits the Constant Rate Effect	83
3.4	Idealized functional effect	87
3.5	Idealized pattern of a grammaticalizing feature according to a relevant linguistic context	90
3.6	Three variants in apparent time in Texas, USA, c. 1980s	97
3.7	Frequency of *have got* for possession by nature of the complement in real time	99
3.8	Frequency of *have to* for deontic modality by verb type in apparent time in York English	99
4.1	Distribution of intensifiers by speaker sex, Toronto, c. 2003–2004	109
4.2	Distribution of intensifiers by individual males in Toronto, c. 2003–2004	109
4.3	Distribution of intensifiers by individual females in Toronto, c. 2003–2004	109
4.4	Distribution of zero copula in real time by year and month – Shaman	118
5.1	*R*, random forest, linguistic, and social factors, all verbs	153
5.2	*R*, conditional inference tree, social factors	154
7.1	Constraint ranking of morphological categories on variable (t,d)	182
7.2	Overall distribution of simplified clusters of variable (t,d) in York, UK (c. 1997) and Toronto, Canada (c. 2003–2004)	184
7.3	Distribution of simplified clusters for variable (t,d) by gender and age in Toronto, Canada	184
7.4	Distribution of simplified clusters for variable (t,d) by gender and age in York, UK	184
7.5	Distribution of simplified clusters for variable (t,d) by education in Toronto, Canada	185
7.6	Distribution of simplified clusters for variable (t,d) by education in York, UK	186
7.7	Comparison of frequency of simplified clusters by grammatical category across communities	186
7.8	Overall distribution of alveolar variants of variable (ing) in York, UK (c. 1997) and Toronto, Canada (c. 2003–2004)	191
7.9	Distribution of alveolar variants for variable (ing) by gender and age in Toronto, Canada	192
7.10	Distribution of alveolar variants for variable (ing) by gender and age in York, UK	192
7.11	Distribution of alveolar variants for variable (t,d) by education in Toronto, Canada	193
7.12	Distribution of alveolar variants for variable (ing) by education in York, UK	193
7.13	Comparison of frequency of alveolar variants of variable (ing) by grammatical category across communities	194
7.14	Pattern of alveolar variants of variable (ing) among nouns compared with indefinite pronouns in York and Toronto	194
7.15	Distribution of simplified clusters for *just* and all other contexts by following phonological segment	200
7.16	Distribution of simplified clusters in apparent time in Toronto English	200
8.1	Constraint ranking for verbal -s in pronouns vs. NP contexts	212
8.2	Distribution of variable (s) by grammatical person, Devon and Samaná	213

8.3	Inter-variety comparison of the type of subject constraint	214
8.4	Inter-variety comparison of the subject type constraint, including northern Englishes	214
8.5	Overall distribution of *-Ø* adverbs in the history of English	219
8.6	Distribution of zero adverbs in York English by age	224
8.7	Distribution of zero adverbs in York English by age, sex, and education	225
8.8	Distribution of zero variants by adverb semantics and age in York English	226
8.9	Overall distribution of deontic modal forms across dialects	230
8.10	Distribution forms for deontic modality in contexts of subjective obligation by generation in York	232
8.11	Distribution of deontic *have to* in Toronto, Canada and York, UK in apparent time	233
8.12	Factor weights for the probability of deontic *have to* by type of modality across varieties, York, Buckie, Wheatley Hill, and Toronto	234
9.1	Distribution of quotatives in Canadian English, c. 1995 and c. 2002–2003	254
9.2	Overall distribution of quotatives across the generations in Toronto English	255
9.3	Cross-study comparison of GE frequency per 10 000 words	263
9.4	Test of decategorization of *and things like that*	266
9.5	Distribution of GE types in apparent time	267
9.6	Comparison of token count and proportion count for *like*	274
9.7	Distribution of quotative *like* in real time (Clara) and in apparent time (all 9–19 year olds in Toronto)	275
10.1	Pathway for grammaticization of *going to*	282
10.2	Distribution of the major future variants in each of the communities	284
10.3	Distribution of future variants in York, UK, c. 1997 by generation	289
10.4	Hierarchy of constraints for semantic function across varieties	306
11.1	Use of past reference *come*	317
11.2	Distribution of past reference *come* by age and sex	318
11.3	Distribution of major intensifiers in York in apparent time	326
11.4	Distribution of intensifier *so* by sex in *Friends*	329
11.5	Distribution of major intensifiers in Toronto in apparent time	330
11.6	Distribution of *so* and *pretty* by sex of the speaker in apparent time, Toronto	331
11.7	Delexicalization process	332
11.8	Distribution of *so* by adjective type across generations in Toronto	333
11.9	Distribution of *so* by emotional value of the adjective	333
11.10	Distribution of laughter variants among adolescents in IM, c. 2004–2006	341

Tables

1.1	Count of all quotative types with *be like* as a quotative	20
1.2	Distribution of *be like* according to type of quotative, i.e. viewed as a proportion of the total of each type	20
2.1	Difficulty of acquisition of linguistic variables	60
2.2	The gender paradox	63
3.1	Leveled paradigm for past tense "to be"	77
3.2	Distribution of zero marking of tense on stative verbs across varieties	92
3.3	Frequency of unmarked verbs in Class I across varieties	93
3.4	Bybee's model for types of change underlying lexical diffusion	96
4.1	York English Corpus (c. 1997)	103
4.2	Sampling strategy for Toronto neighborhoods	104
4.3	Toronto English Archive of Spoken Materials (c. 2003–2010)	105
5.1	Explanations for *that*/Ø variability	125
5.2	Logistic regression of the linguistic factors conditioning *zero* complementizers in York English	126
5.3	Logistic regression of the social factors conditioning *zero* in York English	129
5.4	Non-orthogonal factor groups. Worst case scenario	133
5.5	Non-orthogonal factor groups. Likely case scenario	133
5.6	Idealized logistic regression showing nonindependence of factor groups	134
5.7	*Rbrul* modeling menu	140
5.8	*Rbrul*, mixed effects model, individual random, age continuous	142
5.9	*R*, comparison of marginals for the zero complementizer	146
5.10	*R*, mixed effects model, individual random, age as continuous	147
5.11	Test of interaction between education and occupation	151
5.12	Comparison of calculations for matrix verb as a 4-way categorical factor group	152
5.13	Five logistic regression analyses of the contribution of factors selected as significant to the probability that strong verbs will surface as stems (all factor groups selected as significant)	158
5.14	Wald statistics for prepositional dative of the verb "give" – New Zealand English	159
6.1	Standards for comparison	167

6.2	Logistic regression of internal factors contributing to the probability of *did* in past habitual contexts in Somerset and Samaná	171
6.3	Comparison of similarities and differences in internal linguistic features in Samaná and Somerset	171
6.4	Logistic regression analyses of nonstandard marking on Ukrainian and English origin nouns in monolingual Ukrainian conversations	174
6.5	Comparison of constraints on absence of free N/pronoun objects in Tamambo and Bislama	176
7.1	Distribution of simplified clusters for *just* in comparison with grammatical categories in York, UK and Toronto, Canada	198
7.2	Distribution of simplified clusters in *just* and other preceding [s] words	199
7.3	Distribution of simplified clusters in *just* compared to all other contexts for following phonological context	199
7.4	Logistic regression analyses of the contribution of social factors selected as significant to the probability of [ʔ]	205
8.1	Variable *-s* in Tiverton, southwest England	216
8.2	Outline of the development of deontic modality variants	229
8.3	Logistic regression analysis for nonstandard "be" (*is/was*) in nonexistentials in Early New Zealand English	244
9.1	Predictions for increasing grammaticalization of *be like*	250
9.2	Overall distribution of quotative verbs, university students	251
9.3	Logistic regression of the contribution of internal and external factors to the probability of *be like* vs. all other quotative verbs. Factor weights not selected as significant in square brackets	252
9.4	Overall distribution of quotative verbs, Canadian youth, 9–19 years of age, c. 2002–2003	253
9.5	Logistic regression of the contribution of internal and external factors to the probability of *be like*, Canadian English, Toronto, Canada, 2003–2004, 9–39 years of age	255
9.6	Contribution of external and internal factors on the use of *be like* in Toronto English, 17–39 years of age, 2003–2004 reordered	257
9.7	Timeline of earliest attestation of general extenders (OED)	260
9.8	Overall distribution of GEs and fixed expressions in Toronto, c. 2003–2004	265
9.9	Summary of findings for distributional tests of grammaticalization	266
9.10	Four independent logistic regression analyses of the main GEs in Toronto English	268
9.11	Variable rule analysis of the contribution of speaker sex to the probability of different quotatives	278
10.1	Logistic regression analysis of *going to* in five North American varieties. Factor groups not significant in square brackets	287
10.2	Logistic regression analysis of *going to* in York English. Factor groups not significant in square brackets	290
10.3	Overall distribution of surface forms in past temporal reference in Samaná English	303
10.4	Overall distribution of surface forms in present perfect contexts in Samaná English	304

10.5 Overall distribution of surface forms found in present perfect contexts across corpora 305
10.6 Three logistic regression analyses showing the results for discourse position in habitual past contexts where all three forms are possible 312
11.1 Logistic regression analyses of the contribution of factors selected as significant to the probability of past reference *come* 318
11.2 Frequency of intensifiers in York, UK, c. 1997 (N ≥ 10) 325
11.3 Three logistic regression analyses of the contribution of factors to the probability of *really* in York English 327
11.4 Frequency of intensifiers in *Friends*, 20 year olds, USA, c. 1990s 328
11.5 Frequency of intensifiers in Toronto, c. 2003–2004 (N ≥ 10) 330
11.6 Rough overview of intensifiers in the history of English 334
11.7 Distribution of characteristic IM forms among Canadian teenagers, c. 2004–2006 339

1

Sociolinguistics as Language Variation and Change

> *Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity.* (Weinreich, Labov, and Herzog 1968: 188)

In this chapter I introduce fundamental concepts and key constructs of the study of Variationist Sociolinguistics that will be detailed in later chapters. Why approach the study of language from this perspective? What can be learned from this method that cannot be learned from other sociolinguistic methods? A major component of this approach to language is that it is linguistic, but also social and statistical. Why is a combined socioquantitative method useful and desirable?

Sociolinguistics

Sociolinguistics has its roots in dialectology, historical linguistics, and language contact with considerable influence from sociology and psychology (Koerner 1991: 65). This is why it has evolved into an exceptionally broad field. An all-encompassing definition would be that the domain of inquiry of sociolinguistics is the interaction between language, culture, and society. Depending on the focus, virtually any study of language implicates a social connection because without this human component language itself would not exist. However, the scope of sociolinguistics in this expansive interpretation is vast. Sociolinguistics has as many different facets as its roots. Some areas of the discipline put more emphasis on one area (culture); some disciplines put more emphasis on another (education). There is no one sociolinguistics other than the overarching unity of language in use. Depending on which aspect of language in use comes to the fore, sociolinguistics diverges into innumerable subdisciplines.

Every day we speak and write and use a complex, structured system to communicate but at the same time that system is evolving. The fundamental LVC (Language Variation and

Variationist Sociolinguistics: Change, Observation, Interpretation, First Edition. Sali A. Tagliamonte.

Change) question is, *How does this happen?* Weinreich, Labov, and Herzog (1968: 100–101) answered the question by saying, "the key to a rational conception of language change is the possibility of describing orderly differentiation." This order, yet differentiation, as the normal state of affairs (Labov 1982: 17), the idea that variation is an inherent part of language (Labov 1969: 728), is the foundational maxim of the LVC approach. Differentiation, anomalies, and nonstandard features are easy to spot. In fact, just about everyone likes to talk about the wacky, weird, and/or reprehensible bits of language.

> The normal condition of the speech community is a heterogeneous one … Moreover this heterogeneity is an integral part of the linguistic economy of the community, necessary to satisfy the linguistic demands of every-day life. (Weinreich, Labov, and Herzog 1968: 17)

> Variability [is] not.. a nuisance but is a universal and functional design feature of language. (Foulkes, 2006)

Variation in language is most readily observed in the vernacular of everyday life. For example, a teenager says: "that were like sick"; an elderly man recounting a story to his granddaughter says: "you was always workin' in them days." Are these utterances mistakes? Are they slang? Are they instances of dialect? An LVC-oriented sociolinguist views such instances of language in use as an indication of the variable but rule-governed behavior typical of all natural speech varieties. The vernacular was first defined as "the style in which the minimum attention is given to the monitoring of speech" (Labov 1972c: 208). Later discussions affirmed that the ideal target of sociolinguistic investigation is "everyday speech" (Sankoff 1974, 1980b: 54), "real language in use" (Milroy 1992: 66). Variation in language can be observed just about everywhere from a conversation you overhear on the street to a story you read in the newspaper. Sociolinguists notice such variations too. In undertaking sustained analysis, what they discover is that people will use one form and then another for more or less the same meaning all the time the language varies. The harder part is to find the order, or the system, in the variation chaos. The way LVC undertakes this is by means of the "linguistic variable." A linguistic variable is the alternation of forms, or "layering" of forms, in language. A basic definition is "two or more ways of saying the same thing." A more nuanced, early, definition also mentions that linguistic variables should be structural and "integrated into a larger system of functioning units" (Labov 1972: 8).

Linguistic variables in a given speech community, whether morphosyntactic, phonological, lexical, or discursive, do not vary haphazardly, but systematically. Because it is systematic, this behavior can be quantitatively modeled (Labov 1963, 1969). Analyses of heterogeneous structures within the speech community rest on the assumption that whenever a choice exists among two (or more) alternatives in the course of linguistic performance, and where that choice may have been influenced by any number of factors, then it is appropriate to invoke statistical techniques (Sankoff 1988a: 2). The statistical tools used in the study of variation will be discussed in Chapter 5.

The combination of methods employed in Variationist Sociolinguistics forms part of the "descriptive-interpretative" strand of modern linguistic research (Sankoff 1988a: 142–143). Large-scale studies of variation in speech communities from New York to Norwich have produced extensive bodies of data. The descriptive component requires detailed, critical observation of variation and change. The patterns that have emerged from these undertakings have demonstrated that linguistic change is not only the result

of universal principles but is also shaped by the social context in which it occurs (Labov 1963: 74). This is where the interpretive component of LVC has proven critical. Descriptions of variation can only be understood in context. While sociolinguistic principles prevail wherever you go, each situation provides a unique interpretation. In the case studies in Chapters 7–11 I will demonstrate how the study of different types of linguistic variables must take into account historical, contemporary, and social facts to explain language use.

The Linguistic Variable

LVC research begins with the observation that language is inherently variable. Speakers make choices when they speak and they alternate among these choices. Take, for example, the use of forms which strikes the ear as nonstandard, unusual, dialectal, or new, as in the examples in Example 1.1.[1]

Example 1.1

(a) And then next *mornin'* [In] they were all brought back again. (YRK/002)
(b) Our car was *like* seven miles from where the entrance was. (TOR/021)
(c) There was a supply boat *Ø* came down to our cottage everyday. (TOR/036)
(d) He was like *so* funny and *so* nice. (TOR/054)

These features can only be fully understood if they are examined alongside the relatively unremarkable alternates with which they vary, as in Example 1.2.

Example 1.2

(a) And I started work on an *evening* [ŋ]. (YRK/012)
(b) We were oh probably *about* six miles from it. (TOR/054)
(c) The people *that* did it were brainwashed. (TOR/069)
(d) She's *really* funny, and I think she's *really* pretty too. (TOR/021)

Some variables may even have three or more alternates, as in Example 1.3.

Example 1.3

(a) I can't remember what that *building* [in] is called. (TOR/008)
(b) I was on vacation for *approximately* six weeks. (TOR/038)
(c) I'm only exposed to the people *who* speak the same way that I do. (TOR/016)
(d) He's *very* funny; he's *very* generous. (TOR/023)

In other words, speakers may vary among various pronunciations of "ing" at the end of words. They may signal approximation with *like* or *about* or *approximately*. They may choose among relative pronouns *that* or *who* or leave it out entirely. They may select *so* or *really* or *very* to intensify an adjective. These choices are potential "linguistic variables."

NOTE Linguistic variables are typically referred to by inserting the phoneme or morpheme or word that is variable inside parenthesis, i.e. variable (ing), (ly), (that), (so), etc. Phonetic realizations are represented inside square brackets, e.g. [n]. Phonemes are represented inside forward slashes, e.g. /n/.

A linguistic variable in its most basic definition is two or more ways of saying the same thing. An important question is, What does it mean to say two things mean the same thing? One time a student asked this question: what is the difference between a synonym and a linguistic variable? Let us explore this distinction. Synonyms are different lexemes with the same referential meaning as in Example 1.4:

Example 1.4

(a) *car, automobile, vehicle, wheels*
(b) girl, lass, chick, sheila, babe, doll, skirt

A more restrictive definition of synonymy would require that two synonyms are completely interchangeable in every possible context. In reality, most are not. For example, *lass* is primarily used in Scotland and northern England, *chick* is used in North America, *sheila* in Australia, whereas *girl* is not confined to a particular variety of English. For many practical purposes, such as with the production of dictionaries, it is customary to adopt a looser kind of definition for synonym. Near synonyms are lexemes that share an essential part of their sense, as in Example 1.5:

Example 1.5

(a) *interesting, intriguing, fascinating, absorbing, spellbinding, engrossing*
(b) *striking, arresting, unusual, out of the ordinary, remarkable, salient*

But this is not the whole story. Linguistic variables must also be alternatives (i.e. options) within the same grammatical system which have the same referential value (meaning) in running discourse (Sankoff 1988a: 142–143). Although some variants may differ subtly in meaning and distribution, if they are part of a linguistic variable they will be members of a structured set in the grammar. Moreover, the choice of one variant or the other must vary in a systematic way – this is what is meant by structured heterogeneity. There is difference, but there is structure to it. Different ways of saying more or less the same thing may occur at every level of grammar in a language, in every variety of a language, in every style, dialect, and register of a language, in every speaker, often even in the same discourse in the same sentence. In fact, variation is everywhere, all the time. This is why it is referred to as "inherent" variation (Labov 1969: 728). Now, consider a more in-depth definition of the linguistic variable:

- two different ways of saying the same thing;
- an abstraction;

- made up of variants;
- comprising a linguistically defined set of some type:
  - a phoneme
  - a lexical item
  - a structural category
  - a natural class of units
  - a syntactic relationship
  - the permutation or placement of items
- although its delineation can be at any level of the grammar, the variants of the variable must have a structurally defined relationship in the grammar;
- they must also co-vary, correlating with patterns of social and/or linguistic phenomena.

Synonyms could be a linguistic variable. However a linguistic variable is more than simply a synonym. Deciding which forms co-vary meaningfully in language is actually a lot trickier than you would think.

Mini Quiz 1.1

Q1 How would a variationist sociolinguist explain the following example?
"There *was* two of us. Yeah, that's right there *were* two of us."
(a) Alternation in styles.
(b) Free variation.
(c) Linguistic variation.
(d) Random differentiation.
(e) Bad grammar.

Q2 Which of the following provides an example of two variants of a linguistic variable?
(a) And we said, "if you join the club, you must go to church."
(b) He'd light a furnace for to wash the clothes.
(c) He was awful homesick, you know, my Uncle Jim.
(d) To prove I could do it, I had to prove that I could do it.
(e) There's two girls on my street who have pink hats.

To this point this discussion has focused on the technical description of the linguistic variable. However, there is an entirely different side to linguistic variation that does not come from the mechanics of the linguistic system but involves issues of stigma and salience that come from the external evaluation of language by its users – us humans. There is no reason for a velar sound to be superior to an alveolar sound. There is no reason for a synthetic construction to be better than an analytic one. There is no inherently terrible thing about a double negative. However, there is an absolutely insidious view that certain ways of saying things are better than others. This comes down to the social interpretation of language use.

Most people are convinced that linguistic features are good or bad. For example, here is Sara Kempt, aged 49, in Toronto, Canada (c. 2003), in Example 1.6.

Example 1.6

> … and I think the natural inclination of anybody is to get lazy and sloppy and not think. So I th– there's more and more slang, and people dropping their Gs and things like that, just that … frankly grates on me. I hate it! *Then again, I find myself doing it sometimes*. (TOR/027)

Another fascinating thing about linguistic variables is that people are often completely unaware that they use them, particularly when certain of the variants are not part of the standard language. For example, this is Gabrielle Prusskin, aged 55, in Toronto, Canada (c. 2003). The interviewer has just asked her what she thinks about the word *like*, as in Example 1.7.

Example 1.7

> It's usually young females um when every other word is "like" and it drives me insane. I just *like* I hate it. (TOR/054)

TIP One way to find a linguistic variable is to look for the words that occur most frequently in data. Are there other ways of saying the same thing? If language is always in flux, then it is just a question of finding out what is on the move in a particular place and time.

Linguistic variables inevitably involve variants that have social meaning. These are typically called "sociolinguistic variables." Sociolinguistic variables are those which can be correlated with "some nonlinguistic variable of the social context: of the speaker, the addressee, the audience, the setting, etc." (Labov 1972c: 237). One variant might have overt social stigma, e.g. "I *ain't* got it", another might entail authority, e.g. "You *must* listen", or prestige, e.g. "I *shall* tell you a story." Yet another variant may be neutral, e.g. "I *have* it." These social evaluations may differ markedly from one community to the next, from one country to the next, from one variety to the next, from one social situation to the next. It may even be the case that one person's admired pronunciation will be another person's loathed one. The patterns of a linguistic variable in the speech community tell the story of how the speech community evaluates the variants of the variable and in so doing this reveals how society is organized and structured. Which groups talk to each other? Which groups do not? How a linguistic feature is socially evaluated often has to do with its history as well. Which groups have been in the community a long time? Which groups are new? Language use is a reflection of the society in which it is embedded and the time period in which it occurs.

NOTE One time I went to a conference in the United States with my then current group of British graduate students. One of them had a strong accent from a variety of somewhat modest prestige in the United Kingdom. She was shocked to be told, repeatedly, how lovely her accent was. Similarly, I was chagrined to discover that my own middle-class Canadian accent – unremarkable in Canada – was heard as an entirely unbecoming American accent in the United Kingdom.