IMPACT STUDIES IN LANGUAGE AND SOCIETY

Language Standardization and Language Change

The dynamics of Cape Dutch

Ana Deumert

Language Standardization and Language Change

Impact: Studies in language and society

IMPACT publishes monographs, collective volumes, and text books on topics in sociolinguistics. The scope of the series is broad, with special emphasis on areas such as language planning and language policies; language conflict and language death; language standards and language change; dialectology; diglossia; discourse studies; language and social identity (gender, ethnicity, class, ideology); and history and methods of sociolinguistics.

General Editor

Annick De Houwer
University of Antwerp

Associate Editor

Elizabeth Lanza
University of Oslo

Advisory Board

Ulrich Ammon
Gerhard Mercator University

Jan Blommaert
Ghent University

Paul Drew
University of York

Anna Escobar
University of Illinois at Urbana

Guus Extra
Tilburg University

Margarita Hidalgo
San Diego State University

Richard A. Hudson
University College London

William Labov
University of Pennsylvania

Joseph Lo Bianco
The Australian National University

Peter Nelde
Catholic University Brussels

Dennis Preston
Michigan State University

Jeanine Treffers-Daller
University of the West of England

Vic Webb
University of Pretoria

Volume 19

Language Standardization and Language Change: The dynamics of Cape Dutch
by Ana Deumert

Language Standardization and Language Change

The dynamics of Cape Dutch

Ana Deumert
Monash University

John Benjamins Publishing Company
Amsterdam/Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Ana Deumert
Language Standardization and Language Change : The dynamics of Cape Dutch / Ana Deumert.
p. cm. (Impact: Studies in language and society, ISSN 1385–7908 ; v. 19)
Includes bibliographical references and indexes.
1. Afrikaans language–Standardization. I. Title. II. Series

PF861.D386 2004
439.3'6--dc22 2003069657
ISBN 90 272 1857 9 (Eur.) / 1 58811 492 9 (US) (Hb; alk. paper)

© 2004 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

To Carel

Although it has to be considered in conjunction with many other issues, standardization certainly deserves closer study.

Elisabeth Eisenstein (1979)
The Printing Press as an Agent of Change

A standard in the sense of a set of defining characteristics arises from an awareness of a stereotype based upon a norm.

Robert B. Le Page (1988)
Some Premises Concerning the Standardization of Languages

Table of contents

List of figures xi

List of tables xv

Acknowledgements xix

Introduction: Standardization, language standards and standard languages 1
- The study of language standardization 2
- Linguistic focusing: From variation continua to language standards 2
- Codification and functional diversification: From language standards to standard languages 6
- Afrikaans historical sociolinguistics 8
- Outline of the book 10

I. History

Chapter 1

Afrikaans sociohistorical linguistics: Reconstructing language formation 15
- Historical corpora and their interpretation 15
- Language contact and language change at the Cape: Sociohistorical and linguistic evidence 21
- Mechanisms and outcomes of language change 38
- The insights of the acrolect 40
- Summary: Setting the stage 41

Chapter 2

Afrikaner nationalism and the discovery of the vernacular 45
- The rise of dialect writing 45
- Afrikaner nationalism and early vernacular standardization 54

Folk taxonomies and language attitudes 59
A diglossic community? 63
The linguistic marketplace and its entrepreneurs 71
Summary: The language question at the Cape 73

Chapter 3

The *Corpus of Cape Dutch Correspondence* and the social context of language use in the nineteenth century 77
The Cape Dutch speech community: Core and periphery 77
Literacy and writing practices 83
The *Corpus of Cape Dutch Correspondence* 87
Mapping the social universe: Age, gender, ethnicity and class 90
Summary: An acrolectal and mesolectal corpus 100

II. Variation analysis

Chapter 4

On the analysis of variability and uniformity: An introduction to multivariate clustering techniques 105
Statistics and variation studies: More than a numbers game 105
Numerical taxonomy 107
Hierarchical cluster analysis 110
Multidimensional scaling 120
Principal components analysis (PCA) 123
Focusing, diffusion and fixity: A statistical perspective 130

Chapter 5

The gradualness of morphosyntactic change 135
Variation analysis: Some caveats 135
The verbal system 137
Revisiting Conradie (1979) 146
Tense marking 156
Nominal gender agreement 165
The attributive adjective inflection 169
Summary: Morphosyntactic standardization as a process of rule extension 175

Chapter 6

Morpholexical and syntactic variation **179**
Personal pronouns **179**
The relativizer **189**
The demonstrative pronouns *hierdie* and *daardie* **193**
The negation **196**
The infinitive clause **204**
Objective *vir* **207**
The periphrastic possessive **211**
Summary: Comparing distribution patterns **215**

Chapter 7

The Cape Dutch variety spectrum: Clusters, continua and patterns of language alternation **221**
Identifying lects in the data **221**
Examining the social dimensions **236**
Linguistic patterns in the dialect writing tradition **240**
Standardization and diglossia revisited **243**
Afrikaans-Dutch code-mixing/switching **249**
Summary **257**

III. Establishing the norm

Chapter 8

***Engels, Engels, alles Engels*: Language contact, conflict and purism** **261**
British colonial rule: 1806–1910 **261**
English-Dutch/Afrikaans code-mixing/switching **263**
Anglicisms **270**
Language conflict and language purism: *Moenie jou languages mix nie* **272**

Chapter 9

Social networks and the diffusion of standard Afrikaans **279**
Networks, modernization and nationalism **279**
Reconstructing historical social networks **282**
Ties of coalition and cooperation: The Afrikaner nationalists **288**

The rise of Afrikaans: *Daar buite in die bloue lug* **292**
Summary **295**

Epilogue: Language standardization and language change 297

Appendix: The *Corpus of Cape Dutch Correspondence* 305

References 315

Index 355

List of figures

Figure 1.1 The chronology of language change at the Cape (adapted from Raidt 1983: 7, 28)
Figure 1.2 Free population and slaves at the Cape, 1692–1793
Figure 1.3 Historical demographics of a prototypical plantation society (adapted from Bickerton 1995: 317)
Figure 2.1 Publication of vernacular texts in newspapers and magazines in the Cape Colony 1826–1899 (number of texts per year; based on Nienaber 1966)
Figure 3.1 Oral reading practice in the Cape Colony: an itinerant teacher reads items from the *Zuid-Afrikaan* (drawing by C. Bells circa 1840; reproduced in Pells 1938)
Figure 3.2 Temporal distribution of the sample texts 1880–1922 in the *Corpus of Cape Dutch Correspondence* (number of texts per year)
Figure 3.3 The distribution of social classes in the *Corpus of Cape Dutch Correspondence*
Figure 4.1 Position of cases 1–4 in the two-dimensional space
Figure 4.2 Single linkage (adapted from Kaufman and Rousseeuw 1990: 47)
Figure 4.3 Complete linkage (adapted from Kaufman and Rousseeuw 1990: 47)
Figure 4.4 Tree diagram for single linkage (Euclidean distance)
Figure 4.5 Tree diagram for average linkage (Euclidean distance)
Figure 4.6 Tree diagram Ward's method (Euclidean distance)
Figure 4.7 Amalgamation graph (Euclidean distance, Ward's method)
Figure 4.8 Shepard diagram
Figure 4.9 Scatter plot (dimension 1 vs. dimension 2)
Figure 4.10 Correlation between [NP/del.] and [P/del.] with a least squares fit (regression line)
Figure 4.11 Scree plot
Figure 4.12 Scatter plot (based on component scores)

Figure 4.13 A schematic representation of diffusion, focus (language standards/incipient standard languages) and fixity (standard language; adapted from Smith 1996: 66)
Figure 5.1 Histogram summarizing the degree of inflectional loss (person and number distinction in the present tense) by individual writers ($n = 136$) (all verbs)
Figure 5.2 Apocope of [t]according to phonological and grammatical constraints: change in real time 1712–1922
Figure 5.3 Apocope of [t] according to phonological and grammatical constraints: Johanna Brümmer (born 1877)
Figure 5.4 Apocope of [t] according to phonological and grammatical constraints: F. S. Malan (born 1871)
Figure 5.5 Apocope of [t] according to phonological and grammatical constraints: P. van Huyssteen (born 1854)
Figure 5.6 Apocope of [t] according to phonological and grammatical constraints: B. J. Brümmer (born 1841)
Figure 5.7 Apocope of [t] according to phonological and grammatical constraints: Elisabeth Eksteen (born 1860)
Figure 5.8 Apocope of [t] according to phonological and grammatical constraints: Louis Botha (born 1862)
Figure 5.9 Apocope of *-e(n)* for plural verbs (present tense), infinitives and past participles change in real time, 1779–1922
Figure 5.10 Possible lines of development for the regularization of the past participle of strong verbs
Figure 5.11 Histogram summarizing the use of present perfect by individual writers ($n = 100$); main verbs only
Figure 5.12 Histogram for summarizing the degree of gender regularization by individual writers ($n = 60$)
Figure 5.13 Morphosyntactic regularization in the *Corpus of Cape Dutch Correspondence*
Figure 6.1 Histogram summarizing the use of nominative *ons* by individual writers ($n = 65$)
Figure 6.2 Histogram summarizing the use of *nie*-2 by individual writers ($n = 85$)
Figure 6.3 Generalization of the long infinitive with *om... te* by individual writers ($n = 74$)
Figure 6.4 Diffusion of gender regularization and *nie*-2 in the *Corpus of Cape Dutch Correspondence*

Figure 7.1 Tree diagram (Ward's method, Euclidean distances) for the variables GENDER, VERB, INF, PP_T, PP_EN, TENSE, PRONOUN, OMTE, NIE2 and VIR
Figure 7.2 Graph of amalgamation schedule for Ward's method
Figure 7.3 Multidimensional scaling (Guttman Lingoes) of the distance matrix for cluster I–III; based on Euclidean distances
Figure 7.4 Plot of means for the three cluster solutions
Figure 7.5 Scree plot
Figure 7.6 Multidimensional scaling (Guttman Lingoes) of the distance matrix for cluster I–III; based on Euclidean distances
Figure 7.7 Multidimensional scaling (Guttman Lingoes) of the distance matrix for cluster I and III; based on Euclidean distances
Figure 7.8 Distribution of Afrikaans forms across the three clusters
Figure 7.9 Distribution of social groups across clusters: age
Figure 7.10 Distribution of social groups across clusters: class
Figure 7.11 Distribution of social groups across clusters: sex
Figure 7.12 Multidimensional scaling (Guttman Lingoes); data basis: *Corpus of Cape Dutch Correspondence* and nine vernacular texts
Figure 7.13 Multidimensional scaling (see Figure 7.3). Position of the bi-dialectal speakers F. S Malan (A[1], A[2]) and Wynanda Hoogenhout (B[1], B[2])
Figure 8.1 Code choices in Johanna Brümmer's letters to F. S. Malan, 1893–1904
Figure 9.1 Sociogram of the functional relations among 35 speakers/writers (Krackplot 2.0; multidimensional scaling)
Figure 9.2 Cliques in the network
Figure 9.3 Code choices in Francois Malan's letters to Johanna Brümmer, 1891–1920
Figure 9.4 Code choices in Wynanda Hoogenhout's letters, 1905–1921
Figure 9.5 Publication of Afrikaans books, 1861–1925
Figure 9.6 Percentage of Afrikaans forms in the *Corpus of Cape Dutch Correspondence* across time

List of tables

Table 1.1 Origins of slaves imported to the Cape 1652–1808, in percentages (from Shell 1994: 41)
Table 1.2 Free population and slaves at the Cape, 1692–1793 (from Worden 1985: 11)
Table 2.1 *Wat* and *die* as relative pronouns in the Transvaal periodical press, 1866–1889 (from Van Rensburg & Combrink 1984: 115)
Table 3.1 A speech community typology (based on the degree of recognition of sociolinguistic norms; Santa Ana & Parodi 1998: 35)
Table 3.2 Urbanization in the Cape Colony (from Bickford-Smith 1995: 11)
Table 3.3 Census information (1891, 1904, 1911) for the category 'race'
Table 4.1 Two-dimensional data set
Table 4.2 Distance or dissimilarity matrix for cases 1–4 (Euclidean distances)
Table 4.3 Percentages of full, contracted, and deleted form of *is* with noun phrase (NP) subject or pronoun (P) subject (group style only; Labov 1969: 730)
Table 4.4 Internal structure of the pre-adolescent peer group cluster
Table 4.5 Euclidean distances between clusters
Table 4.6 Principal components (Eigenvalue > 0.5)
Table 4.7 Component loadings (unrotated)
Table 4.8 Component loadings: quatrimax rotation
Table 4.9 Component scores (based on unrotated components)
Table 5.1 Use of Afrikaans forms according to verb classes in the *Corpus of Cape Dutch Correspondence*
Table 5.2 Loss of verbal inflection for *schrijven* ('to write'), *leven* ('to live'), *regenen* ('to rain') and *beginnen* ('to begin') in the *Corpus of Cape Dutch Correspondence*
Table 5.3 Loss of verbal inflection for *maken* ('to make') and *spreken* ('to speak') in the *Corpus of Cape Dutch Correspondence*
Table 5.4 Regularization of the past participle, group I (<i>, <e> → <o> and <ē>, <ui>, <ie> → <ō>)

Table 5.5 Regularization of the past participle, group II (<ij>, <i> → <ē>)
Table 5.6 Regularization of the past participle for *schrijven* ('write') and *krijgen* ('get') in the *Corpus of Cape Dutch Correspondence*
Table 5.7 Present perfect/preterite ratio in the *Corpus of Cape Dutch Correspondence*
Table 5.8 Gender regularization in the *Corpus of Cape Dutch Correspondence*
Table 5.9 The emergence of the new system of adjectival declension in the *Corpus of Cape Dutch Correspondence*
Table 5.10 Frequency distribution for five monosyllabic adjectives in the *Corpus of Cape Dutch Correspondence*
Table 6.1 The pronoun system of Dutch
Table 6.2 The pronoun system of Afrikaans
Table 6.3 *Het, dit* and *dat* as referential subjects and objects in the *Corpus of Cape Dutch Correspondence*
Table 6.4 *Hulle* as subject and object pronoun in the *Corpus of Cape Dutch Correspondence*
Table 6.5 Attribute possessive pronouns in Dutch and Afrikaans
Table 6.6 *Mijn/mij* (in objective and attributive function) at the Cape 1710–1750 (from Ponelis 1993: 205)
Table 6.7 *Mij* and *zij/sij* as possessive pronouns in the *Corpus of Cape Dutch Correspondence*
Table 6.8 Summary: The subject/object pronoun variables in the *Corpus of Cape Dutch Correspondence*
Table 6.9 Relative marking in the *Corpus of Cape Dutch Correspondence*
Table 6.10 Use of objective *vir* with different lexical verbs preceding direct and indirect objects in the *Corpus of Cape Dutch Correspondence*
Table 6.11 Use of objective *vir* with different lexical verbs preceding direct objects in the *Corpus of Cape Dutch Correspondence*
Table 6.12 Distribution of individual writers in the *Corpus of Cape Dutch Correspondence*
Table 7.1 Multiple regression (beta correlation coefficient)
Table 7.2 Cluster means and corpus mean (all in %)
Table 7.3 Variance ($\Sigma(X–M)^2$ / N–1) for the three clusters and the corpus
Table 7.4 Correlation matrix for the ten variables
Table 7.5 Principal components analysis, factor loadings for factor 1 and factor 2
Table 7.6 Three-place implicational scale

Table 7.7 Implicational analysis including the variables VERB, INF, PP_T, PP_EN
Table 7.8 Implicational analysis including the variables GENDER, VERB, INF, PP_T and PP-EN
Table 7.9 Social characteristics of the three cluster solution
Table 7.10 Social characteristics of the three cluster solution for morphosyntactic variables
Table 9.1 Cliques in the network
Table 9.2 Adopter categories and the diffusion of Afrikaans

Abbreviations

AP	Afrikaanse Patriot
APO	African Political Organization
GRA	Genootskap van Regte Afrikaners
MCS	Medium for Community Solidarity
MIC	Medium for Interethnic Communication
PCA	Principal Components Analysis
SLI	Standard Language Ideology
VOC	Verenigde Oostindische Compagnie
ZAT	Zuid-Afrikaansche Tijdschrift

Acknowledgements

As any writer does, I have incurred many debts in the preparation of this book. I am thankful to friends and colleagues in and out of South Africa for their support, interest, encouragement and patience during the various stages of the research and writing process. Special thanks are due to the staff at the University of Cape Town Library, the South African Library, the Cape Archives and to I. H. T. Balie who made material available to me at the Moravian Mission Museum in Genadendal. I would also like to thank the University of Cape Town and the Graduate School 'Dynamics of Non-Standard Varieties' at the University of Heidelberg for research grants, and for providing an intellectually stimulating and also challenging atmosphere during the past few years. I am greatly indebted to Karel Schoeman who not only helped me to find my way to through the manuscript collection of the South African Library in Cape Town, but who also shared his knowledge and understanding of South African history with me. As a novelist and social historian he combines the best of two worlds. Thanks are also due to the native speakers of Dutch and Afrikaans who patiently answered my questions and shared their linguistic intuitions with me. I am especially grateful to Caren Dippenaar who explained the finer semantic nuances of Afrikaans idiomatic expressions to me, and to Mr. Allie and Mrs. Van Gend who told me their memories of growing up in Cape Town in the 1920s and 1930s, and tried to remember the languages they spoke and the languages they heard. Discussions with colleagues have been important. In particular, I would like to thank Carli Coetzee, the late Achmat Davids, Hans Den Besten, John Holm, Ernst Håkon Jahr, Nigel Love, the late Carla Luijks who I miss as a colleague and as a friend, Kay McCormick, Klaus J. Mattheier, Raj Mesthrie, Johan Oosthuizen, Paul Roberge, Joachim Scharloth, Peter Slomanson, Wim Vandenbussche, Hester Waher, and Evelyn Ziegler, who have contributed to the gestation of this study by sharing their own work and thoughts, by answering questions or by commenting on earlier drafts of the manuscript. I am also grateful to Olga Fischer and Susan Fitzmaurice for their comments on the original doctoral dissertation. Special thanks go Paul Roberge and Kees Vaes for their help in deciding on a subti-

tle for this volume. My thinking on the issues discussed here has been greatly influenced by Roger Lass, who supervised the doctoral thesis on which the research is based. In innumerable discussions he provided stimulation and encouragement as well as just and constructive criticism, not to mention laughter and intellectual enjoyment. Finally, thank you to the series editor Annick De Houwer and to Kees Vaes at Benjamins for their support and practical assistance, and to the two anonymous reviewers for their valuable comments on the manuscript.

This book was completed in a long Australian summer and a short Australian winter. Thanks are due to my new colleagues at the *School of Languages, Cultures and Linguistics* at Monash University, in particular Heinz Kreutz and Cree Olivier, for allowing me the time and space I needed to finish this book, and to Louisa Willoughby for her assistance with the final editing process.

A. D.

Introduction

Standardization, language standards and standard languages

> Afrikaans provides an ideal case for studying the latter two linguistic developments associated with modernization: vernacular elevation and standardization. For one thing, both processes are very well documented for Afrikaans. For another, vernacular elevation occurred so rapidly in Afrikaans that its very speed throws into sharp relief the essential characteristics of the process.
>
> D. Shaffer (1978). *Afrikaans as a Case Study in Vernacular Elevation and Standardization*

This book describes the formation of an early Afrikaans standard language in South Africa and the types of language change which shaped its development. Today, Afrikaans is spoken as a mother tongue by about six million speakers (Webb 2002:69), and is used as an inter-ethnic lingua franca in Namibia and parts of South Africa. It is the only language with a pidgin/creole ancestry which has been fully standardized, and which has succeeded in replacing its lexifier (Dutch) in all domains (law, science and technology, literature, education, government and administration). The standardization of Afrikaans was an exceptionally rapid process: in 1876 the first grammar was published by the *Genootskap van Regte Afrikaners* ('Society of True Afrikaners'), a language society with a strong nationalist agenda, in 1925 Afrikaans was granted full official status, alongside English and Dutch, and by the 1930s a formal and codified Afrikaans standard was firmly in place and was promoted through various institutions (education system, political administration, church, etc.).

This introductory chapter outlines the broad conceptual framework for the study of language standardization as an area of historical sociolinguistics, and presents a summary of the overall structure and organization of this book.[1]

The study of language standardization

The study of standard languages as a linguistic and socio-cultural phenomenon is today a well-established field of enquiry. A strongly sociolinguistic orientation to standardization research is implicit in the models which were proposed in the 1960s by Haugen (1966) and Kloss (1969). Standardization is concerned with linguistic forms (corpus planning, i.e. selection and codification) as well as the social and communicative functions of language (status planning, i.e. implementation and elaboration). In addition, standard languages are also discursive projects, and standardization processes are typically accompanied by the development of specific discourse practices. These discourses emphasize the desirability of uniformity and correctness in language use, the primacy of writing and the very idea of a national language as the only legitimate language of the speech community. In their 1985 study (second edition 1991) of the prescriptive tradition Jim and Lesley Milroy (1985a) described this set of discourses as a Standard Language Ideology (SLI), a notion which has since informed a number of studies (e.g. several of the papers in Bex & Watts 1999; also Crowley 1989; Lippi-Green 1994; Woolard 1998).

Linguistically-oriented approaches to language standardization have often concentrated on the identification of the regional and/or social dialects which form the phonological, morphological and syntactic basis of a standard language. However, standard languages which are based on a single dialectal source (monocentric selection) are rare, and most standard language histories have been shaped by dialect levelling and koinézation (cf. Haugen 1972: 266: 'by the time a norm has been codified and elaborated by its users, it has become virtually impossible to identify its base'). The majority of standard languages are thus composite varieties characterized by multiple selections; that is, the complex recombination of features from various dialects and varieties (polycentric selection; cf. Deumert 2003b). In other words, standard languages have 'multiple ancestors' and their history is shaped by various types of language contact (dialect convergence as well as spoken/written language contact; cf. Hope 2000; also Haugen 1972: 247; Van Marle 1997).

Linguistic focusing: From variation continua to language standards

From a variationist perspective standardization can be conceptualized as a movement towards linguistic uniformity through a competition-selection process: certain variants or linguistic habits are selected as part of the standard

norm and are generalized to new linguistic and communicative contexts. Linguistic variability and heterogeneity, on the other hand, become indexical of non-standard varieties.[2] Variability, however, does not imply the absence of norms. Non-standard varieties are characterized by a multiplicity of highly context-specific, particularistic norms which have emerged 'in response to the local needs of the loosely networked social groups which make up the speech community' (Lodge 1993: 95). The norms of standard languages, on the other hand, are universal and show little contextual, geographical or social variation. Both variation and standardization contribute to the formation of sociolinguistic groups: while the maintenance of variation marks social, ethnic and regional differences within the larger speech community, standardization promotes social and political unification and a common identity. Thus language, as Galli de' Paratesi (1977: 170, quoted in Joseph 1987: 42) has put it, fluctuates in a state of natural instability between 'una forza centripeda (la standardizzazione) e una forza centrifuga (la tendenza alla differenziazione)' ('a centripetal force (standardization) and a centrifugal force (the tendency towards differentiation)').

Language standardization, understood as a process of variant reduction, does not only include deliberate intervention by regulating authorities (such as language societies and academies, individual dictionary and grammar writers and also government institutions; i.e. the imposition of uniformity through authoritative acts), but also processes of cumulative micro-accommodation, levelling and dialect convergence, which are the outcome of the everyday linguistic activities of individuals. Le Page and Tabouret-Keller's (1985: 181–182, 187) Acts of Identity model of language change, and the concepts of 'focusing' and 'projection' have been useful in this context. Focusing is a type of language change which supports the gradual formation and stabilization of relatively uniform, well-defined varieties (spoken or written) through processes of sociolinguistic accommodation and inter-dialect levelling. Projection refers to the meaningful, identity-negotiating acts of interpretation which motivate the linguistic choices of speakers. According to Le Page (1988), projection and focusing are central aspects of language use: that is, speakers/writers adjust their speech either to reduce or to emphasize differences between themselves and their interlocutor(s). In a paper on language standardization in the Caribbean, Le Page (1988: 31) summarizes his view of language use and social identity as follows:

> Linguistic activity is a process of projecting onto others images of the universe as we perceive it, and by implication inviting others to share our symboliza-

> tions … If we feel that those we are speaking to are part of a group we wish to identify with, we may modify our behaviour so as to be more like our percept of theirs. In that case, the behaviour of a group will become more focused, and since our symbolization centers on the characteristics with which we have endowed groups which we imagine we perceive in our society, our own behaviour too will then become more focused, more regular. If, on the other hand, … we do not wish to identify with our interlocutors, the behavior of the group will remain diffuse.

However, although such identity-oriented linguistic choices constitute an important type of linguistic change, the development of focused linguistic systems is not entirely the result of intentional and meaningful speaker actions. Increased inter-group contact and exposure, for example, can bring about – or at least support – non-functional behavioural convergence, leading to stimulus-based imitation and repetition (cf. Häcki-Buhofer 2000 for observations of non-functional dialect convergence; Deumert 2003a: 28–33 for a discussion of the phenomenon of social contagion). Migration and urbanization, industrialization and the unification of the national (and international) economic market, the rise of a public school system and strongly centripetal social processes (such as nationalism or de-colonialization) have been described as 'focusing agents' in the context of standardization studies.

The norms of focused varieties are unlike the prescriptive rules of codified standard languages: they include those regularities or linguistic routines (language habits) which define language varieties as linguistic systems, and are the object of descriptive linguistic analysis. The identification of subsistent, uncodified norms (Gloy 1975) is usually based on a statistical understanding of norms in terms of distribution frequencies of variants. That is, normative are:

> diejenige(n) Merkmalsausprägung(en) einer Variable, für die empirisch mehr Ereignisse ermittelt werden können als für jede andere Merkmalsausprägung.
> (Gloy 1975: 26)

> ['those variants of a variable for which one can establish more tokens than for any other variant']

Coseriu's (1974) distinction between 'system', 'norm' and 'performance' is based on a similar line of reasoning: the linguistic system comprises the possibilities of expression, the norm the frequencies with which these possibilities are realized (that is, what is statistically 'normal' in a speech community at a given point in time), and the performance the actual speech acts executed by speakers within specific situational contexts.

A frequency-oriented approach to the description of norms is however not universally applicable. Henn-Memmesheimer (1986) has argued that there often exist within a speech community 'habitualized' speech patterns which, although infrequently realized, are nevertheless normal usage at a given point in time. Henn-Memmesheimer's approach to linguistic variation and language norms is strictly non-quantitative and provides an important counter-point to the frequency-based norm concepts which are commonly used in sociolinguistic research:

> Es geht darum generalisierte Regelformulierungen zu finden, die jedes belegte habitualisierte Muster erfassen, unabhängig von der Auftretenshäufigkeit. Im Gegensatz zu Leska 1966, Rosenkranz 1970, Stirnemann 1980 geht diese Arbeit nicht davon aus, daß sich kodifizierter Standard und Nonstandard im wesentlichen durch die Häufigkeit einzelner Muster unterscheiden, sondern will zeigen, daß es habitualisierte Muster gibt, die mit Standardgrammatiken nicht beschreibbar sind. In dieser Zielsetzung spielen weder pragmatische Verwendungsbedingungen eine Rolle noch ein sozialer Status, der der Verwendung eines Musters assoziiert ist, sondern relevant ist ausschließlich die Tatsache, daß ein Muster als Sprachgebrauch, als usuell belegbar gilt.
>
> (Henn-Memmesheimer 1986: 11)

> ['The aim is to find generalized rule formulations which include every attested habitualized pattern irrespectively of frequency of occurrence. In contrast to Leska 1966, Rosenkranz 1970, Stirnemann 1980, this study does not proceed from the assumption that codified standard and non-standard differ primarily in terms of the frequency of individual patterns, but will show that there are habitualized patterns which cannot be described with standard grammars. Neither the pragmatic conditions of use nor the social status, which is associated with the use of a pattern, play a role in this approach; the only relevant fact is that a pattern is attested as habitual language use.']

Both frequency norms and habitual norms are based on stable patterns of usage and can be described as 'customs' or 'conventions' (cf. Von Wright 1963) or, following Hechter and Opp (2001), as 'regularity norms'.

Focused and relatively uniform varieties which emerge via inter-dialect contact and inter-speaker accommodation frequently develop into what Joseph (1987) has called 'language standards'; that is, they come to constitute models of 'good' or 'appropriate' usage, and are imitated purposefully by individual speakers. However, unlike standard languages, language standards are not transmitted through institutionalized instruction (e.g. schools), and are not yet characterized by a prescriptive tradition. However, they are no longer mere habits or customs, but acquire a clear sense of 'oughtness', a moral imper-

ative, and thus come to resemble standard languages. The explanatory crux of this development is to understand how mere regularity norms or linguistic habits are transformed historically into language standards or 'oughtness norms' (Hechter & Opp 2001). Social norm theory assumes that oughtness norms are an emergent property of social interaction and behaviour; that is, if a specific type of behaviour is enacted by large numbers of people, people come to expect it and thus develop discursive rationalizations (ideologies in the sense of Woolard 1998) of what is perceived as 'normal', 'good' and 'appropriate' within their community (Horne 2001: 10; cf. also Schiffman's 1998 discussion of norm development through 'informal consensus'). However, to rephrase Jasso (2001: 348): humans are not only 'rule finders', they are also 'rule makers'. Standard languages are to a large extent the result of such self-conscious and deliberate rule-making.

Codification and functional diversification: From language standards to standard languages

While standardization in the sense of focusing, variant reduction and norm emergence appears to be a sociolinguistic universal, language codification and prescription (i.e. 'rule-making') is a non-universal phenomenon which shows clear signs of cultural invention (Joseph 1987: 19; also Schiffman 1998). The history of standard languages is closely linked to societal modernization and the functional diversification of social, economic and political life. Thus, standard languages are not merely uniform linguistic systems; they are also used in different modes (spoken and written), styles (public and private, formal and informal) and communicative domains (education/schooling, politics and administration, science and technology, religion, literature, law, mass media, conversation). Moreover, deliberate codification (i.e. the selection and formalization of prescriptive norms) has been identified as a necessary property of standard languages (Garvin 1991). Attention is thus drawn to standardization as an intentional, deliberate act; a perspective also reflected in Weinreich's (1954: 396) distinction between standardization and levelling:

> STANDARDIZATION could easily be used to denote a process of more or less conscious, planned and centralized regulation of language. Many European languages have had standardized varieties for centuries; a number of formerly 'colonial' tongues are undergoing the process only now. Not all levelling is equivalent to standardization. In the standardization process there is a division between regulators and followers, a constitution of more or less clear-

cut authorities (academies, ministries of education, *Sprachvereine*, etc.) and of channels of control (schools, special publications, etc.).

Codification types range from the 'academy-governed style' – which defines a single, monolithic standard code whose unambiguous norms are decided by decree (e.g. French) – to the 'free enterprise style' characterized by a diversity of (sometimes contradictory) authoritative texts (e.g. American and British English; cf. Garvin 1991).

The wide range of socio-cultural functions which standard languages fulfill in a society has linguistic repercussions, and 'flexible stability' and 'intellectualization' have been identified as central linguistic aspects of standard languages (cf. Havránek 1964; Garvin 1964 and 1991). The notion of flexible stability refers to the capacity of a language to be adapted to new functions, genres and styles without losing its linguistic identity and defining structure. Stability, as noted by Haugen (1972: 249), implies a 'slowing down' of language change. Intellectualization describes a specific characteristic of elaborated (written) standard languages, i.e. the ideal of the unambiguous expression of 'precise and rigorous, if necessary abstract statements' (Havránek 1964: 6). Terminological diversification, lexical monosemy, normalized syntax (e.g. no ellipsis, clearly identifiable syntactic constituents) and complex hypotaxis have been identified as typical linguistic results of the intellectualization process.

Unlike the norms of language standards, the norms of standard languages do not emerge spontaneously in social interaction; they are created by design and enforced formally through the education system, publishing houses, official spelling rules, grammar books and dictionaries. Moreover, the structural properties and social prestige of standard languages are legitimized by a specific type of language ideology; a metalinguistically articulated belief that there is *one and only one* correct way of speaking, leading to an ideological 'intolerance of optional variability in language' (Milroy & Milroy 1985a: 206), and to the political and social marginalization of varieties and dialects which are located outside of the standard norm (cf. Bourdieu 1991). Authors of pedagogical grammars and dictionaries as well as the public school system play a central and effective role in the diffusion of the ideologies and beliefs which legitimize the status of standard norms within a society (cf. Watts 1999; Von Polenz 1994: 337–342).

Afrikaans historical sociolinguistics

This book does not present a teleological history of Afrikaans from low prestige *kombuistaal* ('kitchen language') to codified, mature standard language, but traces the formation and existence of various (and sometimes competing) language standards in the decades prior to the official recognition of Afrikaans, and shows how these linguistic practices (and the discourses that accompanied them) prepared the ground for the rise of standard Afrikaans after 1925. In other words, the study takes Joseph's (1987: 15) hypothesis that 'standard languages come about through a historically stable, long-term sequence of developments' as a starting point, and describes the processes of linguistic focusing, evaluation, discursive rationalization and codification which shaped the early standardization history of Afrikaans, while paying due attention to continuing patterns of variation and persistent non-standard usages within the speech community.

As a case study the history of Afrikaans is an important example of a nineteenth century standard language and throws into relief many of the salient characteristics of standardization histories. These include: the process of variation reduction and variation selection, the interaction of focusing and early codification, the role played by linguistic ideologies and language-centred nationalism in shaping popular consciousness and civic ritual, the importance of élites and counter-élites, and the gradual spread of the new standard norms through social networks and institutional contexts. The history of Afrikaans shows clearly that, as a type of language change, standardization is the product of the emergent (conscious and unconscious) linguistic habits and normative activities of a collective (the speech community), as well as the creative construct of the intentional activities of specific groups or individuals within this collective (cf. Rubin 1977; Willemyns 1989). In addition, the history of Afrikaans provides a well-documented example of a standard language which emerged in the broad language-contact scenario of a colonial society. The history of the standard norm needs to be seen against the background of over 200 years of language contact and conflict, language imposition and language learning.

The linguistic and sociolinguistic analysis is based on a newly collected historical corpus: the *Corpus of Cape Dutch Correspondence* which includes 350 unpublished private letters and diary excerpts. The texts were written between 1880 and 1922; this was a time when the new standard of Afrikaans was being propagated in early grammars and model texts, but had not yet received formal recognition and codification. In addition, metalinguistic sources, such

as commentaries published in local periodicals, pamphlets and travelogues, allow for the historical reconstruction of language attitudes and sociolinguistic evaluations. Methodologically, the use of multivariate statistics (to investigate focusing) and network analysis (to investigate the gradual adoption of the new standard norm by individuals and social groups) opens new avenues for future studies of language standardization. The results of the study stand in sharp contrast to the standard view of Afrikaans historical linguistics, i.e. the received opinion that the standardization of Afrikaans occurred in a naturalistic fashion as the codification and elaboration of an already pre-existent and relatively uniform spoken vernacular which constituted the Low variety in a diglossic situation (cf. Scholtz 1963, 1980; Raidt 1983, 1991, 1994; Ponelis 1993; for a critique of the standard view see also Roberge 2003).[3] Instead, it will be argued that a dynamic social dialect continuum existed at the Cape until at least 1900. Language standardization involved the fundamental (strategic and intentional) re-organization of the Cape Dutch variety spectrum, the institution of a dominant standard norm (as a marker of social status and distinction) and the marginalization of the various non-standard language forms and varieties. The history of standard Afrikaans is closely linked to racist nationalism, the rise of Afrikaner hegemony in South Africa and the politics of apartheid. The early standardization history created the foundations of this development by establishing Afrikaans as a *witmans taal* ('white man's language'), an unambiguous marker of white Afrikaner nationalism and ethnicity (on the ideological dimensions of Afrikaans scholarship and language history cf. Roberge 1990 and 1992).

In sum: drawing on work on language standardization and language planning/policy, Le Page and Tabouret-Keller's Acts of Identity model, the tradition of sociolinguistic, philosophical and sociological norm theory (cf. Von Wright 1963; Gloy 1975; Joseph 1987; Hechter & Opp 2001), and the emerging field of historical sociolinguistics, this book provides a detailed sociolinguistic description of central aspects of the early standardization history of Afrikaans and the dynamics of the Cape Dutch variation continuum. Methodologically the study combines a variety of approaches (corpus analysis, multivariate statistics, network analysis, discourse and code-mixing/switching analysis), and situates language history firmly within the sociolinguistic paradigm.

Outline of the book

Chapter 1 presents an overview of the early development of Afrikaans; it includes a critical examination of the tradition of Afrikaans scholarship (in particular the work of J. Du P. Scholtz and Edith H. Raidt) and describes the language-contact situation in the colonial settlement.

Chapter 2 examines the role of the vernacular or dialect writing 'tradition' which emerged at the Cape from the 1850s, and summarizes the language-political discussions which took place during the nineteenth century in the context of growing Afrikaner nationalism. The chapter also discusses the conventional interpretation of the Cape speech community as diglossic and shows how attitudes and perceptions were shaped by an emerging Standard Language Ideology.

Chapter 3 describes the *Corpus of Cape Dutch Correspondence* (1880–1922) in terms of its social and quantitative coverage, letter quality and authorship. The chapter includes an overview of the social context of language use at the turn of the century, i.e. degree of bilingualism and language contact, sociolinguistic practices, literacy levels and literacy practices as well as the parameters of social, generational and ethnic stratification around 1900.

Chapter 4 introduces hierarchical and *k*-means cluster analysis, multidimensional scaling and principal components analysis which were used for the variationist analysis of the *Corpus of Cape Dutch Correspondence*. Multivariate clustering techniques analyze quantitative data in terms of similarity and co-variation. They are useful tools for the description of linguistic varieties as focused clusters of idiolects.

Chapters 5 and *6* examine the patterns of morphosyntactic, morpholexical and syntactic variation in the *Corpus of Cape Dutch Correspondence*. In order to understand the historical origin and development of the Afrikaans variants, the linguistic structures of the superstrate (metropolitan Dutch[4] and the Dutch dialects) and substrate languages (mainly Creole Portuguese, Malay, Khoe) are considered, as well as the dynamics of language contact and evolutive language change.

Chapter 7 presents a multivariate analysis of the variety spectrum discussed in Chapters 6 and 7. Sociolinguistic groups of speakers are identified in the *Corpus of Cape Dutch Correspondence* and their relative position in the standard/dialect continuum at the Cape is discussed. Chapter 7 also returns to the question of diglossia and shows that the patterns of standard/non-standard variation in the corpus correspond to relatively well-defined types of code-mixing/switching.

Chapter 8 discusses the role of English in the Cape Dutch variety spectrum and its influence on linguistic practices. Both code-mixing/switching into English and lexical and structural borrowing (anglicisms) are described with data from the *Corpus of Cape Dutch Correspondence*. The chapter further outlines the evaluation of such practices and the emergence of language purism from the 1930s.

In *Chapter 9* social network analysis is used to model the spread of the new standard of Afrikaans across speakers and groups of speakers. The chapter includes a discussion of the nature of network ties in standardization histories and the role played by Afrikaner nationalist networks in the successful diffusion of the emerging Afrikaans standard norm.

Terminologies for different cultural and ethnic groups have always been a problem for studies of South African history where ethnic labels functioned as ideological constructs in a polity obsessed with questions of race and racial categorization. Following Mesthrie (2002) and McCormick (2002) neither capitals nor scare-quotes will be used. The terms coloured, white and black will occur without further punctuation in reference to social and ethnic data from the official government census reports.

Notes

1. This book is part of a larger research program in historical sociolinguistics and language history, developed most explicitly in 'Bringing speakers back in – epistemological reflections on speaker-oriented explanations of language change' (Deumert 2003a; cf. also Deumert & Vandenbussche 2003).

2. Cf. also Devitt's (1989:3) description of standardization as a 'complement to variation' and Haugen's (1972:107) classic definition of a standard language as being characterized by 'minimal variation in form' and 'maximal variation in function'.

3. The 'standard' view also underlies Abraham and Conradie's (2001:26) assessment that Afrikaans did not 'suffer' normative interventions prior to the twentieth century ('Weder das Afrikaans noch das Jiddische und ihre sprachhistorischen Ausbildungen haben an solchen normativen Eingriffen bzw. unter vereinheitlichenden Schriftnormen zu 'leiden' gehabt'; 'Neither Afrikaans nor Yiddish and their language-historical formations suffered from such normative interventions or unifying written norms'). On the standardization history of Yiddish see Peltz in Deumert and Vandenbussche (2003).

4. Following Roberge (e.g. 2003:16), the term 'metropolitan Dutch' will be used to refer to the 'emerging standard Dutch of the metropole (i.e. the 'core' areas of the Low Countries as opposed to the colonial 'peripheries' of New Netherlands and the Cape Colony)'.

Part I

History

Chapter 1

Afrikaans sociohistorical linguistics

Reconstructing language formation

> Every stage of language is a transition stage, every one just as ordinary as any other … and thus [one can see] plainly the fluid transitions of its spatial and temporal differences.
>
> Hugo Schuchardt (1885). *On Sound Laws: Against the Neogrammarians*

Historical corpora and their interpretation

Historical periodizations are central to the narrative structure of language histories. Frequently such periodizations follow a three-step model of old, middle and modern, thus clearly identifying a temporally and typologically transitional stage ('middle') which mediates between old and new linguistic systems (cf. Lass 2000 and Smith 2002 for a critical discussion). Although such periodizations can serve as a useful heuristic and help to structure the fields of historical enquiry, they often reflect an implicitly teleological view of language change: transitional states are frequently interpreted as formative periods for the modern language and the suggested punctuations deflect attention from the fundamentally transitional nature of all language states. By asking *Waneer is dit Afrikaans?* ('When is it Afrikaans?'), Roberge (1994a) drew attention to the epistemological and methodological problems involved in the often impressionistic identification of discrete language states in the history of Afrikaans. That the question of language demarcation and naming is of particular relevance in the context of standardization studies was pointed out recently by Mühlhäusler (2003; cf. also Le Page 1988) who identified the metalinguistic processes of naming and locating languages (i.e. situating a language within a certain geopolitical territory) as fundamental aspects of the standardization process.

The conventional chronology for the emergence of Afrikaans has been established by J. Du P. Scholtz and his student Edith H. Raidt who pioneered an

empirical approach to the history of Afrikaans. Their work focused on the careful diachronic description of individual aspects of grammar and phonology, based on the surviving philological record. Den Besten (1987b) has therefore dubbed this research tradition the 'South African Philological School'. External (i.e. sociohistorical) evidence, which had informed the work of many earlier scholars of Afrikaans and which had led to the formulation of (often highly speculative) theories about the language's origin, was only of secondary interest to those working within the philological paradigm (cf. Scholtz 1980: 34).[1]

Most of the linguistic data in the historical record reflects the language use of the Dutch superstrate. For the early years of the colony we have a number of journals kept during official expeditions and the official correspondence of the *Verenigde Oostindische Compagnie* (VOC; 'United East India Company', commonly referred to in English as the 'Dutch East India Company'). Although these texts represent by and large the norms and variation patterns of metropolitan Dutch, occasional examples of colonial usage (especially lexical items) found their way into these documents (cf. Kloeke 1950: 308–330; Raidt 1991: 119–124; 1983: 30–31).

For the eighteenth century there is the archival material collected and published by L.C. van Oordt (*Die Kaapse Taalargief* , 'The Cape Language Archive', 1712–1831; published 1947–48, 1949). Van Oordt's corpus contains over 300 official letters and reports sent by field cornets (*veldwagmeesters*) to the local magistrates (*landdrosten* and *heemraaden*). Most of the texts originate from the late eighteenth century and only twelve letters are available for the time before 1770. The *Taalargief* corpus is stylistically homogeneous and biographical information is available for the writers (most of whom were respected and relatively well-to-do farmers; for a detailed description of the corpus, cf. Raidt 1991: 131–135). During the 1960s Van Oordt collected further material (including bills, affidavits, letters, court proceedings) covering the first half of the eighteenth century. Unlike the data collected in the *Kaapse Taalargief*, the second, unpublished corpus is stylistically heterogeneous and little or nothing is known about the writers' social and economic background (for a critical assessment of the corpus cf. Raidt 1983: 36, 1991: 135–136). Additional archival material for the early eighteenth century exists in the documents written by French immigrants (collected in Franken 1953 and Pheiffer 1980). Important sources for the late eighteenth and early nineteenth centuries are the diaries by Johanna Duminy (1797; cf. Franken 1938) and Louis Trigardt (1836–1838; cf. Smuts 1968), and from the mid-nineteenth century popular dialect texts were published at the Cape. The Cape Dutch dialect literature constitutes an

important yet problematic linguistic corpus and will be discussed in detail in Chapter 2.

The indigenous inhabitants of southern Africa, the Khoe and the San, are commonly grouped together under the generic name Khoesan (in the historical sources they are referred to as *hottentots* and *boesjesmannen,* 'bushmen'). Their cultures, however, are different: the San are nomadic hunter-gatherers, while the Khoe live in a pastoral economy. Elphick (1972) has shown that the relationships between the two groups were complex: poverty-stricken Khoe were often integrated into San communities and San could 'become' Khoe if they acquired livestock (often through theft; for further details see Elphick & Malherbe 1989; Penn 1995: 45–47). The pre-settlement period saw extensive linguistic and cultural diffusion across Khoe and San groups. Linguistically, the Khoesan languages are thus best described as a *Sprachbund* phenomenon. Khoe groups were largely found on the coastal plains, whereas San groups lived in the northern interior. In early colonial times contact between Europeans and indigenous South Africans was largely limited to Khoe groups. Slaves were brought to the Cape from 1657.

Although early documentation is sparse and the written record is biased towards the linguistic usage of the European settlers, we nevertheless have examples of substrate varieties of Dutch in the historical record. Khoe Dutch utterances are reported by Willem Ten Rhyne (*Schediasma de Promontorio Bonae Spei*, 1686) and Peter Kolbe (*Caput Bonae Spei Hodiernum*, 1727). Examples of both Khoe and slave Dutch are also available in transcriptions of evidence given in court cases. For the nineteenth century we have two attestations of slave utterances in Teenstra's travelogue (1830 [1943]) and humorous imitations of non-European varieties in the popular press. Indirect evidence also comes from a small corpus of reported speech in Louis Trigardt's diary (1836–1838), where he describes his interactions with Isaak Albach, a Frenchman who lived among the Khoe at the colonial frontier. The Cape Muslim religious writings (in Arabic script, cf. Davids 1991), the diary kept by the Nama leader Hendrik Witbooi (Van der Zwan 1986), and the letters of Jan Jonker Afrikaner (Luijks 2000) provide further access to nineteenth century substrate varieties of Dutch. Following the methodological practice of traditional historical linguistics, older substrate forms are also believed to have been preserved in some rural, non-standard dialects spoken by communities of mixed ancestry (e.g. Griqua Afrikaans, Rehoboth Afrikaans or Malay Afrikaans; cf. Roberge 1994b: 66–67, 1995; also Van Rensburg 1989; Kotzé 1989).

Afrikaans language historians have traditionally focused on the analysis of the well-documented language use of the European settlers. Although both

Scholtz (1963) and Raidt (1983, 1991) acknowledge that different non-native Netherlandic varieties existed in the colonial settlement, these varieties are believed to have been marginal to the historical formation of Afrikaans. Substrate varieties are believed to have contributed only some lexical material and a limited number of grammatical features (e.g. objective *vir* as a borrowing from Creole Portuguese or reduplication from Malay). Scholtz and Raidt view the history of Afrikaans as an essentially gradual and linear development of seventeenth century dialectal Dutch: linguistic contact within the group of the European settlers led to an acceleration of internal tendencies towards deflection and regularization, but never interrupted the gradual process of normal linguistic change. Roberge (2002c: 84–85) has summarized this research tradition under the heading of the 'superstratist hypothesis'.

Within this model increasing grammatical simplification is dated to the early eighteenth century, leading to a transition period between 1740 and 1775. A relatively stable (spoken) vernacular, showing most of the relevant features of (modern) Afrikaans is believed to have been in existence after 1775. Full diffusion of all defining linguistic features (linguistically, socially and geographically) is assumed for 1850.[2] The three-step model of the history of Afrikaans is summarized in Figure 1.1.

Following Roberge (1992) the term 'Cape Dutch (Vernacular)' will be used to refer to the historical transition variety at the Cape. It is a terminological construct and describes a complex sociolinguistic variation continuum which comprises a wide range of varieties and variants.[3] Cape Dutch (Vernacular) thus represents what Le Page (1977: 222) has called a 'mediating system', that is, a typological, temporal and sociolinguistic intermediary between two clearly defined historical states (Early Modern Dutch and modern Afrikaans respectively). The temporal extension of this continuum will be discussed in more detail in Chapters 5 to 7.

seventeenth century Dutch dialects (regional and social)	→	Cape Dutch Vernacular		→	Afrikaans
		internal language change (accelerated by language contact), koinézation (levelling and simplification)			internal language change
1652		1700	1740		1775

Figure 1.1 The chronology of language change at the Cape (adapted from Raidt 1983: 7, 28)

Scholtz and Raidt interpret the continuing linguistic variation in the historical record after 1800 as a scribal artefact, a result of influence from the Dutch written norm which is thought to have existed in a diglossic relationship with the spoken language.[4] Writers are believed to have consciously avoided the use of variants which belonged to the Low variety and to have randomly borrowed standard features into their language in formal situations, leading to variation patterns in the written language which did not form part of the spoken vernacular. Thus, in her discussion of the pronoun system in the late eighteenth century Raidt (1995: 134) states:

> [I]n their written language women consistently used the personal pronoun *ik* ['I'] (following the Dutch spelling) and *wij* ['we'] for the plural, instead of *ek* and *ons* as spoken by them.

The paradigmatic assumption of a significant difference between written and spoken language was questioned by Roberge (1994a). The fact that the Dutch standard language, which slowly emerged in the course of the sixteenth and seventeenth centuries, was not socially diffused before the late nineteenth century (Van der Wal 1992) must be taken into account when interpreting the historical evidence, and one cannot assume general knowledge of, and overt prestige for, the emerging Dutch standard forms. The conventional assumption that writers were able (or even willing) to replace their local spoken language forms with the correct metropolitan form in the written medium is difficult to sustain in view of a sociohistorical reality characterized by marginal exposure to written language norms, minimal schooling and an emerging colonial identity.

Schooling was rudimentary before 1800 and historical research has shown that literacy levels at the Cape (estimated on the basis of signature marks in notarial records) were lower than in the Netherlands and the colonial settlements in North America (Biewenga 1996; De Wet 1981).[5] Even among those who mastered the skills of reading and writing, regular access to the norms of metropolitan Dutch was limited since books were an expensive commodity in the seventeenth and eighteenth centuries. Biewenga (1996: 116–119) estimated that between 1680 and 1730 only 25% of settler households in Stellenbosch owned books (including the Bible and devotional literature), a proportion considerably lower than in the Netherlands. There was no printing press or newspaper available in the colony until 1795 (Rossouw 1987: 131). Access to the norms of metropolitan Dutch was provided to some extent by private, mostly European-born tutors. However, their influence on colonial writing practices must have been limited: only 130 so-called teacher *knechts* are recorded up un-

til the late eighteenth century, and their teaching skills were reportedly rather poor (cf. Shell 1994: 19–20). The assumption that the majority of those who were to able to write had a reliable knowledge of the orthographical and grammatical conventions of metropolitan Dutch (or were in cases of doubt able to consult a dictionary or grammar) is thus difficult to substantiate from the sociohistorical evidence. Moreover, it is likely that new local, non-metropolitan prestige forms emerged at the Cape from the early eighteenth century when a distinct colonial identity developed among second generation settlers (cf. Katzen 1969: 197). Lass (1992, 1997: 61ff.; cf. also Muysken 1995) has suggested that under sociohistorical conditions of limited literacy and in the context of a largely oral society, written representations of language can be interpreted as 'utterance level phenomena'; that is, variation in the written medium can be taken as an indicator of the patterns of variation in the spoken language. With reference to Afrikaans historical linguistics, Roberge (1994a) has argued that the historical record was not an 'orthographic fiction' – as argued by Scholtz and Raidt – but should be interpreted as representing at least '*in some sense* a written reflex of intermediate forms of spoken language' (1994a: 156; emphasis in the original). In a detailed comparison of Louis Trigardt's diary (1836–1838) with an official letter written in 1823 by Trigardt to the magistrate in Grahamstown, Roberge (1994a) has shown that the written language was not an arbitrary mix of High and Low forms. Rather, individuals commanded different styles or varieties of (written and spoken) Cape Dutch Vernacular and used these depending on sociolinguistic dimensions such as purpose of text, formality and interlocutor.[6] Trigardt, for example, used a language much closer to metropolitan Dutch in the official letter.

That synchronic variation can indicate linguistic change in progress is a well-established principle of historical linguistics. The variationist model of language change can be summarized as follows: a change from variant A to B is typically characterized by an intermediate period in which both forms coexist, that is, A → {A,B} → B. Historical linguists have shown that the intermediate stage can stretch over long periods of time, or that variation can be stable without actually leading to language change (Romaine 1982; Lass 1997: 184ff.). From the perspective of variation analysis the procedure followed by Scholtz and Raidt, in which a change is dated as completed when the new variants have become conspicuous in their frequency, is disputable as under this criterion a period of slow change or even stable variation might be classified as a completed change (such as, for example, the loss of person and number distinctions in the Afrikaans present tense paradigm, which has been dated by Scholtz to 1740 although variation continued throughout the nineteenth century). Fol-

lowing Roberge (1994b: 29), the chronology established by the South African Philological School is best understood as a 'useful diachronic abstraction' (or a 'macro-story' in Lass' 1997: 288 terminology), describing the development of Afrikaans as a prototype – an idealized 'composite of all defining features' (for a critique of the early datings cf. also Den Besten 1987b: 68–69).

A second major critique of Scholtz's and Raidt's model comes from creolists who have argued that too little attention has been paid to the colonial contact situation of the seventeenth and eighteenth centuries. Although Scholtz and Raidt accept dialect contact within the European settler community as a factor which contributed to language change at the Cape, they described debates about a possible pidgin/creole ancestor of Afrikaans deprecatingly as 'vorwissenschaftliche Spekulation' ('pre-scientific speculation', Raidt 1983: 45). Thus, with regard to the early Khoe-Dutch data, Raidt (1991: 127) states categorically:

> [d]ie staaltjies van Khoi-Khoin se pidgin-Nederlands … leer ons nie veel omtrent die wording van Afrikaans nie, want ons het hier byna uitsluitend met gebroke taal te make wat nóg met Nederlands nóg met Afrikaans verband hou
>
> ['the sentences of Khoekhoe Pidgin-Dutch … don't teach us much about the formation of Afrikaans, because we are dealing here almost exclusively with broken language which is connected neither to Dutch nor to Afrikaans']

However, ever since Hesseling's (1899, 1923) study *Het Afrikaansch* where he described Afrikaans as a language characterized by incomplete creolization, linguists have repeatedly turned their attention to non-European varieties of Cape Dutch Vernacular in order to assess the role their speakers have played in the formation of Afrikaans.

Language contact and language change at the Cape: Sociohistorical and linguistic evidence

A salient typological feature of Afrikaans is the absence of productive verbal morphology, nominal gender marking as well as the reduction of case distinctions in the pronominal system (see Chapters 5 and 6 for details). Despite these far-reaching morphological changes, Afrikaans shows few prototypical creole characteristics and the general syntax of continental West-Germanic (excluding Yiddish) has been maintained: underlying SOV with V2 in main clauses and *Satzklammer* (sentence brace) in subordinate clauses. Afrikaans has also

maintained separable prefix verbs and unlike most creole languages Afrikaans has not developed a tense-mood-aspect (TMA) system, but retained the form of the Dutch present perfect and the Dutch future tense. However, Afrikaans does make use of adverbial tense markers to specify the temporal and aspectual dimensions of an utterance (e.g. *al/alreeds*, 'already', can be used as perfective markers; see Chapter 5 for details). The lexicon of Afrikaans is predominantly Netherlandic, with some important borrowings from Portuguese, Malay and Khoe. The patterns of derivational morphology are largely in line with Dutch (although there are some significant differences; cf. Van Marle 1994; Waher 1994b), the main innovation being the possibility of reduplication (as in *hulle het vroeg-vroeg aangekom*, 'they arrived very early'; cf. Donaldson 1993: 447ff.). The existence of a cluster of innovative syntactic features, i.e. features without direct Germanic antecedent which can be explained through positive substrate transfer (e.g. the objective marker *vir,* reduplication, the associative plural N-*hulle*; see Chapter 6), coupled with almost total morphological regularization, the comparatively rapid rate of change and a colonial setting involving European settlers practicing a slave-based economy, point to the pivotal role of language contact in the history of Afrikaans (Thomason & Kaufman 1988: 255). Taking the sociohistorical and linguistic evidence together, most creolists describe Afrikaans as an (almost prototypical) example of a semi-creole (e.g. Holm 1989: 339f., 2001, 2002; Bryn & Veenstra 1993; McWorther 1998) – thus drawing on a broad cover term which refers to borderline cases situated 'on a continuum between the clear non-creoles and the clear creoles' (Thomason 1997: 85; cf. also Mühlhäusler 1997: 11–12).

There is historical evidence that a rudimentary mixed trade jargon (involving English, Dutch and Portuguese as lexifier languages) developed from about 1590 when European merchant ships began to anchor regularly at the Cape to restock their supplies on their way to East India, and members of their crews engaged in barter for cattle and sheep with the indigenous Khoe. The European visitors made no serious attempts to learn the Khoe languages: communication proceeded either through the Dutch/English jargon or through Khoe interpreters who were trained by the Europeans from the early seventeenth century onwards. An extended jargon may have developed when the *Haerlem* stranded at the coast in 1647 and its crew stayed for more than a year at the Cape awaiting a ship to return home. The crew leader Janssen reported that the crew had daily contact with Khoe groups on whom they depended for cattle and sheep. According to Janssen's testimonial, the Khoe acquired a minimal Dutch vocabulary which formed the basis for inter-group communication (for details

regarding the historical and linguistic evidence of these early encounters see Raven-Hart 1967; Den Besten 1986, 1987a; Nienaber 1994a).

The first European settlers arrived in February 1657, five years after the Dutch joint-stock company *Verenigde Oostindische Compagnie* (VOC) had established the Cape as a permanent refreshment station for ships travelling between Amsterdam and Batavia. The transformation of the Cape into a settlement colony fundamentally changed the nature of inter-group contact: European farmers appropriated the land for agricultural use, thus ousting, displacing or subordinating the indigenous groups of nomadic herders and hunter-gatherers. Continuous immigration led to a steady increase of the settler population in the seventeenth and early eighteenth centuries (for demographic details, cf. Ponelis 1993).

The group of European arrivals comprised not only Dutch colonists (from the Netherlands as well as from the East Indies, where they had been in VOC employment), but also German settlers who arrived at the Cape in large numbers after the Thirty Year War (1618–1648). About 35% of settlers were of German origin and came from Low and Middle German dialect areas which form part of the German-Dutch dialect continuum. In 1688/89, the almost entirely West Germanic character of the settler community was complemented with the arrival of about 200 French Huguenots who had fled France after the revocation of the Edict of Nantes (1685). An isolated settlement of this linguistically homogenous group was prevented by Governor Simon van der Stel (1679–1707) who pursued (in accordance with the VOC) a policy of linguistic and cultural assimilation in order to maintain the Dutch character of the colony (Raidt 1983: 12). Pheiffer (1980) has shown that the Dutch documents written by the French Huguenots anticipate features which later became characteristic of Cape Dutch Vernacular (e.g. uninflected infinitives, plural verbs and past participles). However, the documents also show second language characteristics which are not part of Cape Dutch Vernacular (e.g. loss of plural marking in the noun and French word order). Morphological reduction is also noticeable in the texts written by German colonists. The German-Dutch documents show, for example, loss of neuter gender, loss of verbal concord and [t]-apocope (Ponelis 1993: 20–21).

Regarding the Dutch dialectal base of Afrikaans, Kloeke (1950) had argued that Afrikaans shows strong phonological affinities to the southern dialects of Dutch, in particular to the regions around Rotterdam, Delft and Gouda (cf. also Van Ginneken 1928: 207–208). This regional interpretation was challenged by Scholtz (1963: 232–256; cf. also Buccini 1996) who emphasized the importance of an incipient urban Holland koiné. This koiné had developed in port

cities such as Amsterdam and Utrecht where speakers of diverse Dutch dialects and refugees from France and Germany came into contact with one another. Demographic figures for the Dutch element of the settler population indicate the general importance of Holland varieties as well as the urban Amsterdam dialect at the Cape: about 28% of Dutch progenitors (1657–1820) came from Amsterdam, another 50% from North and South Holland (Ponelis 1993: 122).[7] Regional and social dialects, a partially stabilized port koiné (characterized by dialect levelling and simplification) as well as variants of nautical Dutch (which settlers encountered not only in the European ports but also during their voyage to the Cape) all played a role in the early history of Afrikaans (cf. also Mufwene 2001: 36 on the importance of port city koinés and nautical varieties in the history of contact languages).

From about the 1730s the settler population was mainly native born and new social status distinctions emerged in the context of colonial life. Shell (1994: 23), for example, describes the low social status of *knechts* (wage labourers of European descent) who constituted between 10% and 30% of the free population. Structures of social distinction and stratification defined the colonial society from early on:

> There is … no evidence whatsoever for the popular notion of an early colonial '*herrenvolk* democracy' among the 'master race' of European settlers. This golden age mythology found in the *plaasroman* (farm novel) of the 1930s could only have been latter-day compensation for the glaring inequality among early European settlers. Cape society perfectly illustrated Vilfredo Pareto's iron law: no matter how evenly goods are distributed initially, after a single generation ownership will be concentrated. (Shell 1994: 153)

Political and economic conflicts, skirmishes and wars characterized the relationship between the indigenous Khoe and the VOC during the late seventeenth century. Having lost their grazing rights (as a result of colonial land appropriation) and much of their livestock (due to re-occurring episodes of cattle disease), the Khoe were forced to take up wage labour on farms and in the colonists' households (for details cf. Elphick & Malherbe 1989: 10–18; Shell 1994: 26–34; Penn 1995: 75–76). The political subordination of the Western Cape Khoe and their integration into the socioeconomic structures of the colonial settlement was completed by 1700. The process was accompanied by acculturation and a gradual language shift towards the Dutch superstrate. The last traces of the Khoe's socioeconomic and political independence vanished when the smallpox epidemic of 1713 decimated the Khoe population at the Cape (Elphick & Malherbe 1989: 21; Shell 1994: 26ff.). By the mid-eighteenth

century most Western Cape Khoe languages had disappeared and the Khoe were fully absorbed as serfs into the cultural, social and political structures of the colonial society. According to Elphick (1977: 23), between 4000 and 8000 Khoe lived on the Cape Peninsula during the early years of the settlement. By the late eighteenth century about 20,000 Khoe and about 3000 San lived within the borders of the colony. They thus constituted roughly between half and one-third of the total population (Worden 1985: 11). The percentage of Khoe was particularly high in the eastern and northern districts, where the Khoe to settler ratio was around 3:1 (Ponelis 1993: 40; cf. also Shell 1994: 147).

In order to supplement the settlers' demand for labour, the Cape commandant, Jan van Riebeeck, advised the VOC directors in March 1657 to allow the import of slaves to the Cape. Cape slaves were drawn from a wide variety of geographical, linguistic and cultural origins and comprised 'a broad palette unparalleled in any other recorded slave population anywhere in the world' (Shell 1994: 6; see Table 1.1). Changes in shipping patterns and competition between slave trading companies shaped the import of slaves to the Cape: the first slaves came from Angola and Dahomey, between 1680 and 1730 most new slaves came from Madagascar, followed by slaves from India, the Indonesian archipelago and the Malayan peninsula. Import of slaves from Madagascar increased again in the 1750s and import of African slaves from about 1780 (Shell 1994: 41–46).

The linguistic diversity (including, in addition to the indigenous Khoe languages, Austronesian and Bantu, as well as Indo-European languages), which was characteristic of the colonial society at the Cape, was a prominent *leitmotif* in traveller reports such as in Michael Peter Kolbe's *Caput Bonae Spei Hordiernum* (1719, vol. 1: 70–71; also Sparrman 1785: 58; Swaving 1830: 304):

> Auf diese Condition also wisse Er, mein Herr, dass hier allerley Mund-Arten und Nationen anzutreffen syn, wenn auch gleich der Hottentotten nicht einmal gedacht wird. Denn man findet Holländer oder generaliter Niederländer, Hochteutsche, und aus denselben Preußen, Pommern, Brandenburger, Sachsen, Franken, Schwaben, Westphälinger und Schweitzer; ferner Schweden, Dänen, Pohlen, Portugiesen, Frantzosen, Savoyer, Italiäner, Engelländer,

Table 1.1 Origins of slaves imported to the Cape 1652–1808, in percentages (from Shell 1994: 41)

	Africa	**India**	**Indonesia**	**Madagascar**
Origin of slaves	26%	26%	23%	25%

> Hungarn und andere; welche, wenn man sie zugleich und einen jeden absonderlich in seiner Sprache reden hörte, auch etwan noch die Sclaven, die *vielerley Sprachen reden*, als die Malayen, Malabaren, Ceylonesen, Javanen, Banjanen, Amboinesen, Bandanesen, Boekier, Chinesen, Madagascaren, Angoler, Guinäer, Capo-verder etc. darzu zehlte, dürffte es gewiss nicht anders lauten, als ob man die Verwirrung der Sprachen by dem Babylonischen Thurme aufs neue vorstellig machen wollte.
>
> ['With respect to this he, my Sir, should know that all kinds of dialects and nations can be found here, even if one excludes the Hottentots. Because one finds people from Holland or more generally from the Netherlands, people from the High German areas, including Prussians, Pommerians, Brandenburgians, Saxons, Frankonians, Swabians, Westphalians and Swiss; moreover, Swedes, Danes, Poles, Portuguese, Frenchmen, Savoyards, Italians, Englishmen, Hungarians and others; who, if one could hear them together and each one speaking separately in their own language, and if one would also include the slaves *who speak many languages*, such as the Malays, the Malabaris, the Ceylonese, the Javanese, the Banians, the Amboinese, the Bandanese, the Bouginese, the Chinese, the Malagasys, the Angolans, the Guineans, the Cape-Verdians, etc., then it would not sound differently as if one would try to resurrect the confusion of tongues at the tower of Babel'; emphasis in the original.]

The chances for the emergence of a separate slave culture (including the development of a common linguistic code) were slight before the nineteenth century when the growing popularity of Islamic teachings united the now freed slave population in religiously-based, urban networks (cf. Shell 1994: 356–362 on the growth of Islam after emancipation in 1838). During the early and mid-eighteenth century, however, 'slaves acted very much as individuals' (Ross 1999a: 38; also Shell 1994: 49; Worden 1985: 86). Except for the VOC's slave lodge in Cape Town's *Herrengracht* which held about 600 (male and female) slaves at a time in overcrowded and unsanitary conditions, no large slave holders existed at the Cape. The number of slaves per household was low and the isolated location of many farms made interaction between slave groups difficult (Worden 1985: 32). The slave population was furthermore not self-reproducing and possible social and linguistic stabilization was regularly interrupted by the arrival of new slaves.

In addition to regular and often intense interethnic contacts under conditions of bondage (settlers and slaves) and wage labour (settlers and Khoe), sexual encounters were not uncommon, especially during the early decades of the settlement when there were only few women among the colonists and the VOC personnel (Shell 1994: 291). Ethnically mixed marriages are attested for

the seventeenth and early eighteenth centuries, and Heese (1984) estimated that between 1660 and 1799 over 1000 marriages took place between Khoe or ex-slave women and settlers of European descent. Furthermore, informal unions between Europeans and slave women were encouraged by slave owners: children born to a female slave were the property of the slave holder, who thus stood to gain financially as Mulatto slaves fetched high prizes on the local slave markets (Elphick & Shell 1989: 195; Shell 1994: 56). Children from unions between European settlers and Khoe women were known as *bastaards*, a term which came to indicate mixed parentage as well as a specific social status in the colonial society. So-called *bastaards* (or *basters*) were generally able to avoid menial labour and often worked as artisans or small farmers (Legassick 1989: 370). From this socio-ethnic nucleus a series of communities (the Griqua, the Oorlam and the Rehoboth Basters) emerged in the eighteenth and nineteenth centuries along the northern frontier of the colony. These communities were bilingual (Cape Dutch Vernacular and Khoe), and followed Khoe as well as colonial customs (cf. Rademeyer 1938 and Van Rensburg 1984 for studies of twentieth century Griqua and Rehoboth Afrikaans).

Another phenomenon of the frontier were the *trekboere* ('nomadic farmers' as opposed to *akkerboere*, 'crop farmers'; cf. Davenport & Saunders 2000: 29–33). These were livestock farmers of European origin who, when agricultural land became scarce, moved eastwards (across the Hottentots Holland mountains) and northwards (beyond the area of today's Tulbagh) to lay claim to what they saw as empty land. The *trekboere* lived a life of relative independence from the VOC and often interacted closely with local Khoe communities. Poor roads and long distances prevented close interaction with the colonial centre in Cape Town and surroundings. Nevertheless, farmers at the frontier were not entirely isolated from developments at the Cape (Guelke 1989; Penn 1995: 35). Their cultural and linguistic practices developed in a complex process, balancing relative geographic isolation with continuing influence of metropolitan forms. The language variety spoken by the *trekboere* is commonly referred to as Eastern Cape Afrikaans (*Oosgrens-Afrikaans*, 'Eastern Frontier Afrikaans') and is distinguished from the variety spoken at the Cape, i.e. Western Cape Afrikaans (Van Rensburg 1983; cf. also van Ginneken 1928).[8]

Malay and in particular Creole Portuguese (mostly in its Indo-Portuguese guise) were important lingua francas within the multilingual slave community and were spoken at the Cape up until the nineteenth century (cf. Franken 1953: 41–79; Shell 1994: 59–63). According to Den Besten's (2000: 958) overview of language use in court cases (up to 1772), about one-fifth of slaves spoke Dutch and over two-thirds spoke either Creole Portuguese or Malay

(other languages spoken by slaves in the courts included Malagasy, Buginese, Malabar and Tsonga).[9] However, knowledge of Creole Portuguese and Malay appears to have been limited among settlers and Khoe. Their use as lingua francas for interethnic communication within the Cape society was therefore unlikely (Ponelis 1993: 17). The numerical strength of the European element (Table 1.2) supported the adoption of a variety of Dutch as the main medium for communication between settlers, the multilingual slave population and the Khoe.

The colonial society at the Cape was a homestead society, i.e. the number of slaves per household was small and interaction between colonists and slaves was frequent and intense (cf. Mufwene 1996; Chaudenson 2001: 95–113). The pre-modern, colonial understanding of *familie* included slaves as well as Khoe serfs as an integral part of the immediate household (Shell 1994: 217–221). This is reflected in Lady Anne Barnard's (1797; cited in Shell 1994: 223–224) description of a colonial household near Paarl:

> … the room filled with slaves – a dozen at least – here they were particularly clean and neat, *Myfrow* [the mistress] sat like Charity tormented by a Legion of devils, with a black babie in her arms, one on each knee and three or four larger ones around her, smiling benign on the little mortals who seemed very sweet creatures and devilish only in their hue …

In those territories where prototypical creoles emerged (e.g. the Caribbean and the Gulf of Guinea), the homestead stage was followed by the plantation stage. Demographically this second stage was characterized by rapidly increased slave imports which 'brought along an early slave majority' (Mufwene 2001: 46). In Jamaica, for example, the proportion of slaves to settlers was

Table 1.2 Free population and slaves at the Cape, 1692–1793 (from Worden 1985: 11)

	slaves	**settlers**	**settler % of population**
1692	337	799	70%
1701	891	1,265	59%
1711	1,771	1,756	50%
1723	2,922	2,245	43%
1733	4,709	3,074	40%
1743	5,361	3,972	43%
1753	6,045	5,419	47%
1763	7,211	6,750	48%
1773	8,902	8,465	49%
1783	11,950	11,040	48%
1793	14,747	13,830	48%

roughly 3:1 in the seventeenth century, changing to 10:1 by the mid-eighteenth century (Mufwene 1996:92; Bickerton 1995; cf. also the statistics given by Singler 1995 for Haiti and Martinique and by Mufwene 2001:46 for Suriname). Baker (1993:137–138) has suggested that the typical scenario under which pidgin/creoles arise is characterized by a distinct demographic profile:

Event 1 when the number of slaves first surpasses the slave-owning population
Event 2 when the number of locally born slaves surpasses the slave-owning population
Event 3 when the regular supply of slaves comes to an end

Linguistic stabilization, according to Baker, usually takes place after event 2; that is, when the newly arriving slaves direct their language acquisition towards the contact language which emerged in the isolated and now self-reproducing slave community. The Cape demographic data (summarized in Table 1.2) indicate that event 1 occurred at the beginning of the 1720s, roughly 60 years after initial colonization. However, even then the slave population was not significantly larger than the settler population. Event 2 never took place: when the percentage of locally born slaves rose to over 70% in the 1830s, their numbers did not surpass those of the settler population (Shell 1994:47; note also that the percentage of slave women and children was consistently low at the Cape, cf. Shell 1994:48, 446–448). The historical demographics of the Cape homestead society and those of a prototypical plantation society are illustrated in Figures 1.2 and 1.3.

Bickerton (1981:4) has suggested that a proportion of about 20% superstrate speakers is the upper limit for creolization to occur, while higher per-

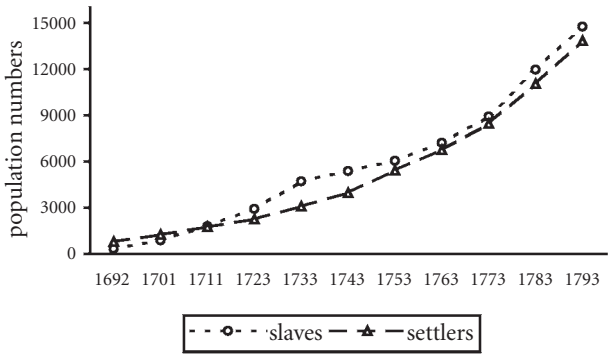


Figure 1.2 Free population and slaves at the Cape, 1692–1793

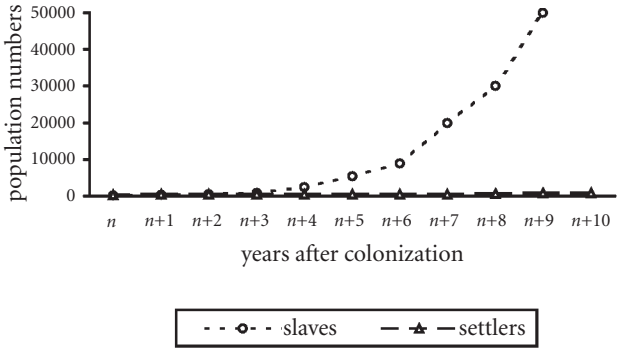


Figure 1.3 Historical demographics of a prototypical plantation society (adapted from Bickerton 1995: 317)

centages of superstrate speakers will support the gradual acquisition of the superstrate language by the substrate population (also Thomason & Kaufman 1988: 155–156).[10] Similarly Mufwene (1996: 100) emphasized that homestead societies usually do not lead to the formation of creoles, but to second language approximations (or interlanguages to use Selinker's 1972 term) of the dominant language:

> It is very unlikely that anything close to today's creoles was developing on a large scale, even if subsystems close to those of today's creoles may have been shaping up … in the speech of some individuals. Rather, approximations of European speech are likely to have been the trend among the non-Europeans living fairly closely with the European colonists. There is no reason why normal people (which the non-European labor generally were) living intimately with speakers of the lexifier would of necessity have developed creoles instead of close approximations of the lexifiers. (ibid.)

However, although the sociohistorical data on population demographics and social structure are suggestive of second language acquisition and the development of a dynamic, unstable and idiosyncratic interlanguage continuum (cf. Webb 1993; Ponelis 1988, 1993: 27–30, 1995; Van Rensburg 1989, 1994), the historical record suggests that pidginization – that is, the emergence of a structurally distinct, group-based and relatively stable contact variety within the substrate community – nevertheless played a role in the history of Afrikaans. The demographic factors explain, however, why basilectalization (i.e. the development of contact varieties maximally removed from the lexifier) was less pronounced at the Cape than in other Dutch settlements (such as the Virgin Islands or Berbice).[11]

Den Besten (1986: 198) has suggested that an early Dutch pidgin developed among the Khoe in the five to six years before the arrival of the slaves. On the basis of the existing trade jargon, Khoe speakers rapidly developed a relatively stable pidgin which was used in interactions with VOC officials and, from 1657, also in interactions (trade and labour) with settlers and slaves. The earliest examples of Khoe Dutch are recorded in an official journal written by Pieter Meerhoff in 1661 (cf. Den Besten 1986: 214–215, 1987a: 15–16). Meerhoff describes in his journal how (during a journey into the interior) his Khoe interpreter warned him:

(1) Namaqua boeba kros moscoqua
NAMAQUA OX SKIN CLEARLY ANGRY
'(The) Namaquas (with their) ox skin shields are clearly very angry'

The translation provided in (1) follows Den Besten (1987a: 15–16) and Roberge (2002a and forthcoming). *Boeba* is a pidgin word for 'ox' attested as early as 1615 and *kros* (< *karos*) is Khoe for 'skin rug' (*khō-b*). The linguistic interpretation of *moscoqua* is difficult. *Qua* can be interpreted as a derivation of Dutch *kwaad* ('angry'), and *-co-* might be cognate with Korana *//kō* ('be angry'), thus yielding the mixed double form *coqua*. *Mos* has been interpreted alternatively as an English loan (*much>mots>mos*; cf. Den Besten 1987), or as a derivation of Dutch *immers/ommers* ('indeed, in any event, after all'; cf. Kloeke 1950: 305; Scholtz 1965: 78). The substrate semantics of *mos* appear to be a result of its convergence with Nama *mûsa* ('visible'; cf. Roberge 2002b and forthcoming for details).[12]

Although the Khoe interpreter's utterance would not be easily comprehensible to speakers of metropolitan Dutch, Meerhoff does not offer a translation. This suggests that the readers of the journal (i.e. VOC officials at the Cape) must have been familiar with such speech styles (an alternative interpretation would be that Meerhoff was trying to impress the VOC officials with his local knowledge and the quotation of the sentence is symbolic rather than informative; the possibility of such an interpretation was suggested by Paul Roberge, personal communication).

Den Besten (2002: 14–15) has furthermore drawn attention to historical examples of a relexified Khoe variety which employs a Dutch lexicon and Khoe particles, postpositions and clitics (2). The historical record also contains Dutch-oriented pidgin sentences with SOV syntax (which is also the basic syntax of Khoe, example (3) to (5)). Unlike example (1), however, the latter are mutually intelligible with Dutch.

(2) 1673
Was makom?
WHAT-YOU[13] MAKE?
'What are you doing?' (Den Besten 2002: 14)

(3) 1672
't Za lustigh, duitsman een woordt calm, ons V kelum
TZA QUIET, DUTCHMAN ONE WORD SAY, WE YOU CUT-THROAT
'Tza, be quiet Dutchman. If you say one word, we will kill you.'
(Franken 1953: 113)

(4) 1673
Duytsman altyt kallom: Icke Hottentots doot makom
DUTCHMAN ALWAYS SAY: I HOTTENTOTS DEAD MAKE
'The Dutch people always say: I will kill the Hottentots.'
(Ten Rhyne 1686: 140)

(5) 1727
Kenje die vieur in die buik zuypen? ... Die man
CAN-YOU THE FIRE IN THE STOMACH DRINK? ... THE MAN
toverman, die man ja tover makum zoo Gy ons immers dood
WIZARD, THE MAN YES MAGIC MAKE SO YOU US INDEED DEAD
maaken als ons die goeds zuipen
MAKE IF US THE GOODS DRINK
'Do you know the drink which sets your stomach on fire? The wizard, he makes magic. So you can indeed kill us if we drink that stuff.'
(Kolbe 1727, vol. I: 526)

The early Khoe Dutch pidgin data anticipate characteristic features of Afrikaans, such as the use of the object pronoun *ons* ('us') in subject position (example (3)), and in the Khoe sentences reported by Kolbe (see example (5)) the Dutch demonstrative *die* occurs as a definite article and demonstrative pronoun. We also find evidence for the use of tense adverbials to express completion (see Chapter 5 for a detailed discussion). Early Khoe Dutch makes use of a verbal marker *-um/-om* which has been interpreted as a generalization of the Khoe (Korana) verbal suffix *-m-* (Den Besten 1987a: 35, 2002: 16; note that the verbal suffix is particularly prominent in Kolbe's collection of Khoe sentences where it might have functioned as a stereotypical marker).

The historical corpus also includes utterances whose grammatical structures are fairly close to metropolitan Dutch. Example (6) was supposedly uttered by the Khoe Lubbert when five *swarte jongens* ('black boys') came to his

settlement trying to steal his livestock. Example (7) comes from the evidence given by a slave in a court case.

(6) 1706
Waarom neem je dat schaap, wagt wat ons sel jou kost
WHY TAKE YOU THIS SHEEP, WAIT A BIT WE WILL YOU FOOD
geeven
GIVE
(Lubbert, a Khoe man; Franken 1953: 93)

(7) 1721
dat francis aldaar tegen haar heeft gezegt, ons sal naar
THAT FRANCIS THERE TO HER HAVE SAID WE WILL TO
buiten loopen, jouwluij sal geen Cost mancqueeren
OUTSIDE RUN, YOU.PL SHALL NOT FOOD LACK
(evidence given by a slave, Scholtz 1963: 111)

These examples differ from the earlier pidgin data in particular with regard to the verb phrase, which shows variable morphological marking. In example (6) differentiation between subject (*je* 'you') and object pronouns (*jou*) is noticeable. The plural object pronoun *ons*, however, is still prominent in subject position. The infinitive is marked according to the norms of metropolitan Dutch by the inflectional ending *-en,* and even sentence brace occurs (*ons sel . . . geeven*). Example (7) includes an inflected present perfect (*heeft gezegt*) and two inflected infinitives (*loopen, mancqueeren*). It is unclear whether these acrolectal features reflect patterns of the spoken language (i.e. the gradual acquisition of morphosyntactic superstrate rules by substrate speakers), or whether these features should be interpreted as a result of scribal editorship, i.e. editorship by the Dutch speaking court scribes who transcribed the utterances according to their own linguistic and orthographic system and omitted unfamiliar non-European substrate or innovative features (cf. Stein 1995).

According to Den Besten, the early Khoe Dutch data suggest the existence of a relatively stable Khoe pidgin which formed the basis for the development of a Dutch pidgin among the slaves (Den Besten 1987a: 20–24, 1989: 222; also Roberge 1995, 2002b). Den Besten's proposal reflects aspects of the Founder Principle which was suggested by Mufwene (1996, 2001) to account for the influence of earlier populations (early settlers, first groups of slaves, indigenous people) in the history of contact languages. The historical connections between early Khoe Dutch and later slave Dutch are indicated by structural similarities between the two varieties: the use of nominative *ons*, the invariant determiner *die,* SOV syntax, use of tense adverbials, preverbal negation, and

some lexical continuities (e.g. *tecken*, 'take', an early Khoe loan from English; *maskie* from Portuguese *mas que*, 'even, nevertheless, never mind'; *kammene*, 'cannot', a form first attested in Khoe Dutch Pidgin; cf. Den Besten 1987a: 20–24; Roberge 2002b). Roberge (1995, 2002c) has furthermore drawn attention to a remark made by Otto Friedrich Mentzel. Mentzel (1921 [1785]: 49), who was in the service of the VOC between 1731 and 1733, described the close and regular interactions that took place between slaves and Khoe:

> Since the arrival of the Europeans the inhabitants of these kraals [the Khoe] that were near the new settlement greatly enriched their vocabulary by contact with the newcomers, they learned still more from the slaves, and borrowed some of the so-called Portuguese, or more accurately, of the lingua franca, common among all Eastern slaves.

Such interaction patterns would not only have supported the adoption of lexical items from the slaves' lingua francas (Creole Portuguese and Malay) by the Khoe, but also the acquisition, diffusion and stabilization of the Khoe Dutch Pidgin which, as a secondary hybrid (in the sense of Whinnom 1971), was a developing and unstable linguistic system. The 'motor' for tertiary hybridization (that is, the stabilization of the bilingual contact language through acquisition by a third group) was thus provided by the slaves. It is likely that language acquisition of the slaves was directed to differing degrees towards the acrolectal Dutch of their colonial masters (especially for house slaves) and the Dutch Pidgin spoken by the Khoe substrate population (for slaves working outside of the house alongside Khoe serfs; Roberge 2002c).

The outcome of these learning and acquisition processes was the formation of what Baker (1990, 1997, 2000) has called a 'medium for interethnic communication' (MIC), 'a new language, suited for their immediate interethnic needs … drawing at all times on the [linguistic] resources available' (Baker 1990: 111). Roberge (2002b: 69) dates the stabilization of an interethnic Afro-Asian Cape Dutch Pidgin to approximately 1710. This pidgin functioned as an MIC in the colonial society at the Cape. Baker (2000) has suggested that under certain conditions the MIC can develop into a 'medium for community solidarity' (MCS), i.e. a marker of group identity. The social and demographic conditions of plantation societies supported the transformation of the developing MIC(s) into cultural and social markers for the ethnically diverse but socially, economically and culturally increasingly homogenous slave population. At the Cape this transformation did not take place (Roberge 2002b). As noted above, a separate slave culture (comparable to the American South) did not develop at the Cape. It is, however, likely that strongly restructured forms

of Dutch developed into a MCS in cohesive substrate communities such as, for example, the Griqua and the Rehoboth Basters which had settled in peripheral and isolated areas at the colonial frontier. These groups speak strongly creolized varieties even today (cf. Van Rensburg 1984; Roberge 1994b). In other words, basilectalization (see Note 11), that is, the development of clearly identifiable group varieties maximally removed from the superstrate, paralleled the social, cultural and political disenfranchising of these groups which existed at the margins of the colonial society at the Cape.

According to Den Besten's (1989) convergence model the fully restructured contact varieties spoken by the Afro-Asian substrate (Proto-Afrikaans I, i.e. the interethnic Afro-Asian Dutch Pidgin) converged with the inter-dialectal Dutch koiné which had developed among the European settlers and which showed 'normal' inter-generational transmission (Proto-Afrikaans II): 'two types of Dutch, European and Pidgin Dutch, gradually coalesced to yield Afrikaans and its dialects' (Den Besten 1989: 234). Valkhoff (1966: 231; cf. also Franken 1953: 193, 202–203) has formulated a similar scenario which he calls the 'idea of a double origin' of Afrikaans:

> Hence, in the history of Afrikaans, too, it was not always either Dutch or Creole, but the two linguistic currents may well have met and the latter may have stimulated the development of the former.

Cape Dutch Pidgin, although crystallized to some extent, must have been a highly variable developing system, given the different substratal, adstratal and superstratal inputs and on-going processes of language learning. However, despite their extensive variability, contact languages show convention and predictability; a rudimentary grammar which facilitates interethnic communication and which becomes the target of acquisition for substrate and superstrate speakers. The complex reality between variation and stabilization as a defining characteristic of early language formation in contact scenarios was captured by Schuchardt (1883, cited in Bechert & Wildgen 1991: 12) in his description of Malay-Spanish Creole:

> Dieses *español de cocina* … besitzt für den Sprachforscher ein großes Interesse. Es ist kein fertiges Patois, wie das Portugiesische von Macao und Malacca, es existiert in den mannigfachsten Abstufungen, mit größerer oder geringerer Annäherung an die spanische Grammatik, mit größerer oder geringerer Verwendung malayischer Wörter, aber es ist kein immer wiederholtes individuelles Radebrechen mehr … es besteht eine Art Überlieferung, es macht sich ein breiter Durchschnitt bemerkbar, die Spanier müssen sich an das

> Küchenspanisch gewöhnen und sich ihrerseits zu einer vermittelnden Sprechweise bequemen.
>
> ['This *español de cocina* is of great interest to the linguist. It is not a finished patois, such as the Portuguese of Macao or Malacca, it exists in numerous shadings and degrees, with greater or lesser approximation to the Spanish grammar, with greater or lesser use of Malay words; but it is no longer an individual's broken language … there is a kind of tradition, one notices a broad average; the Spanish will have to get used to this kitchen Spanish and will need to employ an intermediate speech level for communication.']

As in other colonial societies (such as Brazil, the American South) there are reports of the acquisition of basilectal varieties by European children from their wet nurses and care takers as well as from their non-European playmates (Khoe or slave children; cf. Van Rheede 1685 for an early observation, discussed in Scholtz 1963: 257ff.; for the early nineteenth century cf. Worden, Van Heyningen, & Bickford-Smith 1998: 127; cf. also Shell 1994: 20–23 on regular interactions between Khoe, slaves and European-born *knechts* of low social status). Mentzel (1925 [1775]: 108) emphasized that the closeness of interaction between substrate and superstrate community was also function of social status:

> Kolbe's statement with regard to the bad upbringing of South African children is very misleading. It can apply only to the children of very poor or of dissolute parents. The average child is not allowed to run wild among the slave-children. On the contrary, he is tenderly cared for by his mother.

Historical research on racial, social and regional differentiation in the colonial society (cf. Guelke 1989; Armstrong & Worden 1989; Worden 1985; Shell 1994) suggests that in the eighteenth century the convergence of substrate (pidgin) and superstrate varieties must have been a complex and variable process which affected different groups to differing degrees. Roberge's (1995: 81–82) sociohistorical continuum model identifies geographical factors as well as sociolinguistic Acts of Identity (Le Page & Tabouret-Keller 1985) as central aspects of the variant selection process at the Cape.[14] Language change, according to Roberge, was a complex and multi-dimensional process:

> Several factors would determine the degree of influence of one code on the other. Members of the prosperous burgher class in Cape Town and the wealthiest wine and grain farmers in the Boland would not have known more than a smattering of the pidgin (if any). Settlers of lesser means in the Boland and along the frontier had at least a passive – often active – knowledge of the Cape Dutch pidgin. Acrolectal Dutch would have become more and more di-

> luted with increasing social and geographic distance from centres of power. The extent of dilution would naturally be greatest in the rural areas along the frontier (where speakers were simply not as familiar with prestige norms), within the slave community generally by virtue of multiple inputs, and among the Khoe and Bastards. In other words, the Dutch language at the Cape of Good Hope formed a continuum from the most creole-like variety within the Afro-Asian substrate to an uncreolised extra-territorial variety of the European superstrate. *The speech of individuals took on or avoided pidgin features depending on the interlocutor, the nature of their communicative networks, and the sociolinguistic circumstances* (*code-switching*). (ibid.; my emphasis)

The description of a continuum of language contact at the Cape invokes the notion of the (post)creole continuum – originally introduced by DeCamp (1971a) for his analysis of variation in Jamaican Creole – and Ponelis' (1988:126) description of Afrikaans as a 'behoudende akrolektiese kreool' ('conservative acrolectal Creole'), suggests such a line of thinking (cf. also Makhudu 1984:96; Den Besten 1989:225). According to DeCamp's original discussion, the (post)creole continuum emerges as the result of a historical process (decreolization) in which the original bi-systemic opposition of basilect (pidgin/creole; the variety most removed from the lexifier language) and acrolect (varieties close to the superstrate language, usually spoken by the settlers and a few private slaves) develops into a linguistic continuum via the emergence of intermediate varieties (mesolects which are thus historically 'younger' than basi- and acrolect). The formation of a (post)creole continuum is the result of acrolectal pressures on the linguistic performance of basilect-speaking individuals: decreolization is thus targeted language change in the direction of the acrolect and leads to the eventual disappearance of basilectal forms from the speech repertoire of individuals and the speech community at large.

Although it is possible that decreolization (loss of basilectal features in the speech of individuals) affected the speech of some individuals and social groups within the colony (such as, for example, the upwardly mobile group of freed slaves, the *vrije zwarten*, 'free blacks', in Cape Town, or migrants to missionary settlements such as Katrivier or Genadendal),[15] decreolization in the sense of DeCamp was not a central mechanism of language formation at the Cape.[16] Rather, from the early eighteenth century a complex Cape Dutch variation continuum, which comprised a wide range of varieties and variants, constituted a 'feature pool' of competing and co-existing forms (cf. Mufwene 2001:4–6). The variability of the continuum was gradually reduced through on-going processes of inter-group accommodation (focusing) in the context of inten-

sive and long-term language contact, 'leading to increasing isomorphism or equivalence of structure' across the continuum (Roberge 2002b: 58). Instead of acrolect-directed change of idiolects (i.e. decreolization), on-going interactions and accommodations between systems and codes eventually led to the selection and stabilization of phonological, grammatical and lexical features 'according to frequency, perceptual salience, semantic substantiveness, and plausibility as linguistic signs based on factors that are language- or situation-specific – this in addition to the intersecting dynamics of identity and network membership, which may not be retrievable from the philological record' (Roberge 2002a: 400).

Mechanisms and outcomes of language change

In their monograph *Language Contact, Creolization and Genetic Linguistics* (1988) Thomason and Kaufman developed a general, predictive theory of language contact which is based on a typological distinction between borrowing (i.e. 'the incorporation of foreign features into a group's native language by speakers of that language', p. 37), and shift (i.e. contact-induced change in which imperfect learning plays a role). While borrowing affects first and foremost the lexicon, shift-induced interference leads to the restructuring of the target language's (TL) phonology, morphology and syntax (via processes of reanalysis and rule extension). Eventually, if the contact between learners (the shifting population) and native speakers is intense, the learners' errors may affect the linguistic systems of native speakers of the target language. Den Besten's (1989) convergence model for the formation of Afrikaans is in line with Thomason's model of language genesis through shift:

> The process through which interference features are introduced by a group of learners into a second language – a target language – has two or three components, depending on whether or not the learners are integrated linguistically into the target-language speech community or not. First, learners carry over some features of their native language into their version of the TL, which will be called TL_2. Second, they may fail (or refuse) to learn some TL features, especially marked features, and these learners' errors also form part of TL_2. If the shifting group is not integrated into the original TL speech community, so that (as in the case of Indian English) its members remain as a separate ethnic or even national group, then the TL_2 becomes fixed as the group's final version of the TL. *But if the shifting group is integrated into the original TL-speech community so that TL_1 speakers form one speech community with TL_2 speakers, the*

> *linguistic result will be an amalgam of the two, a TL_3, because TL_1 speakers will borrow only some of the features of the shifting group's TL_2. In other words, TL_2 and TL_1 speakers will 'negotiate' a shared version of the TL and that will become the entire community's language.* (Thomason 2001:75; my emphasis)

Thus, in intensive contact situations (which homestead societies usually were) complex, reciprocal influences between TL systems result not only in direct substrate influence (transfer or borrowing), but support the stabilization of hybrid TL_3 structures which cannot be traced back unambiguously to either the substrate or the superstrate input. The type of linguistic restructuring that is characteristic of intensive contact ecologies supports multiple layers of system reorganization including loss, simplification and generalization of features (through imperfect learning), addition of new features (borrowed from the substrate languages) as well as 'complexification' (Thomason 2001:13), that is, modifications of rule conditions resulting from the interaction of developed and developing systems.

That the concept of linguistic hybridization or convergence (understood as an outcome not as a process, cf. Roberge 2002b) might help us to untangle some of the linguistic puzzles in the history of Afrikaans has been the working hypothesis underlying Roberge's academic project. Regarding, for example, the Afrikaans verbal hendiadys (a complex double verb construction expressing progressive aspect, as in *hy loop en sing*, 'he is walking along singing'; cf. Donaldson 1993:220), Roberge (1994c) suggested that serialization of Dutch lexical verbs, which is attested for Griqua Afrikaans and some varieties of Western Cape Afrikaans, was a common feature of basilectal varieties of Cape Dutch Vernacular. The serial verb structure was then reanalyzed by acrolectal speakers in terms of their knowledge of metropolitan structures (the early modern hendiadys construction with *en*). The reanalysis facilitated the incorporation of some aspects of the basilectal construction into the acrolect, i.e. optional affixation of the past participle marker *ge-* to the propositional verb, the extension of the construction to environments following finite modal and future auxiliary verbs as well as the use of *loop* as a hendiadys trigger. In other words, basilectal features fused in such a way with inherited structures that the resulting new rule cannot be traced back directly to either one of the different inputs. It is precisely this 'multilevel syncretism' leading to 'etymological opacity' which, according to Roberge, is characteristic of semi-creoles (and which at the Cape complemented the cultural syncretism described by social historians):

> This … seems to lie at the heart of what I understand a semi-creole to be. Measured against 'true' creoles, semi-creoles lack the massive restructuring of

> the grammatical systems. What we find instead is a cycle of incremental innovation, accommodation, and integration … The end result is often hybrids that are etymologically opaque to the extent that they are not directly relatable to any of their putative antecedents. (Roberge 1994c: 73)

Simplistic dichotomies between internally and externally motivated change which neatly group linguistic innovations as a result of either superstrate and substrate influence (including language universals) are of limited use for our understanding of language change in complex contact ecologies, which involve shifting populations with many different L1s and varying degrees of access to the superstrate. In such scenarios competition and selection of features are regulated not only by structural factors (e.g. congruence of forms, semantic or perceptual salience and frequency), but also by extra-linguistic factors such as group identity, access to norms and group interaction patterns (cf. Mufwene 2001).

The insights of the acrolect

Although there has been a tendency in the study of language contact to focus on the analysis of basilectal varieties, Rickford (1974) suggested in a paper titled *The Insights of the Mesolect* that 'rummaging in the mesolectal topsoil' could provide us with new data for a better understanding of the processes and outcomes of language contact. Similarly, Escure (1997) has argued that acrolects have been ignored in the study of contact-induced change and are often mistakenly equated with the metropolitan standard language. Acrolects, however, are better described as linguistic innovations characterized by the incorporation of linguistic features which have their origin in the contact situation itself. Unlike standard languages, acrolects usually have no overt set of linguistic norms and are pragmatically motivated (i.e. depend on the formality of the situation). In other words, the concept of the acrolect is both absolute (on the level of the speech community) and relative (on the level of the individual): '[a]n acrolect would, thus, be the *highest* variety an individual is able or willing to produce in the relevant formal contexts … Indeed, one person's acrolect may be another's mesolect' (ibid.: 66, 74; emphasis in the original). The boundaries between acrolect and mesolectal varieties are thus 'extremely difficult to establish' (ibid.: 67).

The European varieties collected by Van Oordt in the *Kaapse Taalargief* are probably best interpreted as examples of mesolectal/acrolectal usage. The same

applies to documents such as the diaries written by Johanna Duminy and Louis Trigardt (Roberge 1994b, 1995, 2002c). The linguistic usage in the Duminy diary is close to metropolitan Dutch (e.g. preterite and pluperfect are still intact, the distinction between finite and non-finite and strong and weak verbs is preserved), but also contains many innovative features (e.g. gender has almost disappeared and the verbal plural marker *-e(n)* is in the process of being replaced by singular forms). However, typical characteristics of Afrikaans – such as object pronouns in subject position, the brace negation and the demonstratives *hierdie* and *daardie* – are absent. In general, the 'formalizing effect' (Kytö & Rissanen 1983: 474) of writing probably skewed much of the historical sources in the direction of acrolectal and (upper) mesolectal varieties.

Just as there is no reason to exclude acro- and mesolectal varieties from the study of language contact, there is also no reason to exclude *a priori* any time period as irrelevant. Although historical linguists have so far identified the seventeenth and the eighteenth centuries as central for the historical development of Afrikaans, the continuing variation in the historical record suggests that the emergence of Afrikaans might have been slower than originally assumed and was shaped by different influences and various historical, cultural and socioeconomic conditions from the seventeenth to the late nineteenth century. Certain important characteristics of Afrikaans (such as the brace negation, the invariant possessive particle *se* and the demonstratives *hierdie* and *daardie*) appear in the sources only from the mid-nineteenth century. Moreover, it was during late nineteenth and early twentieth centuries that Cape Dutch Vernacular came into direct and intense contact with two fully-fledged standard languages (English and Standard Dutch) via the education system, leading to bilingualism and possibly bi-dialectalism in the speech community. Sociohistorically the period under investigation was characterized by industrialization and urbanization, and thus, by the increasing development of weak ties which (according to Milroy & Milroy 1985b) are particularly important for the spread of new linguistic variants (see also Chapters 3 and 9).

Summary: Setting the stage

This chapter has provided an overview of the history of Afrikaans in the seventeenth and eighteenth centuries, with particular attention to the various linguistic, demographic and social parameters of language contact in the colonial society. It has been argued that the research program of the so-called South African Philological School raises many questions regarding their interpreta-

tion of the continuing variation in the historical record. Details of their work will be addressed in Chapters 5 and 6.

Today, there is general agreement that all three population groups of the colonial society contributed to the formation of Afrikaans (i.e. the European settlers, the indigenous Khoe and the African and Asian slaves). Intensive and often intimate contact between people and their languages (including contact between and acquisition of various Netherlandic varieties, ranging from Cape Dutch Pidgin to approximations of metropolitan Dutch) was a central aspect of the homestead society at the Cape. The gradual convergence of languages was paralleled by the convergence of cultures, leading to the formation of a relatively homogenous Cape colonial culture by the late eighteenth century (Elphick & Giliomee 1989: 225). The standardization of Afrikaans as a national language, and its identification as a marker of white Afrikaner ethnicity, was a cultural innovation of the nineteenth and early twentieth century which deliberately disassociated the local Cape Dutch vernacular from its interethnic origins.

Notes

1. Those usually named as working within the paradigm of the South African Philological School include: J. Smuts, J. E. Loubser, J. Conradie, F. A. Ponelis and R. H. Pheiffer (cf. Roberge 1990: 148, fn. 15).

2. Raidt's datings differ slightly from those advanced by Scholtz, who generally saw the process of language change as completed by 1775.

3. Other terms used in Afrikaans scholarship to refer to this transition variety include *vroeë Afrikaans* ('early Afrikaans'), *Afrikaans Hollands*, *Kaaps Hollands* and *nie-standaard Netherlands* ('non-standard Dutch'; cf. Uys 1983; Griessel 1991; Ponelis 1993).

4. Cf. Scholtz (1963: 9–10), Raidt (1983: 29, 1994 [1985]: 311–313), Conradie (1986: 101f.), Ponelis (1993: 50, 1996: 134).

5. Literacy has been estimated at about 50% for the settler community (for the eighteenth century). Dutch-language literacy was rare among slaves and Khoe; some slaves were, however, literate in languages other than Dutch (cf. Davids 1991: 73ff.; Shell 1994: 59–61).

6. Those typically seen as belonging to the South African Philological School are, however, not entirely uniform in their interpretation of the variation found in the written record. Smuts (1968: 59–60) suggested that the variation in the written language did indeed reflect central aspects of the variation patterns in the spoken language, and Conradie (1979: 79) criticized Scholtz for giving 'te min ruimte vir die moontlikheid dat die variasie van woorde … ook variasie in die uitspraak kan verteenwoordig' ('too little room for the possibility that variation of words [i.e. spelling] can also reflect variation in pronunciation'; cf. also Franken

1953: 188). Scholtz and Raidt also offer detailed descriptions of the variation in the historical record.

7. Although we know little about the social origins of the settlers one can hypothesize that the colonists at the Cape had a social background similar to those recruited by the West India Company. Buccini (1996: 40) described these as individuals 'on the economic or social fringe' (cf. also Mufwene 2001: 35).

8. Van Rensburg (1983, also Van Rensburg, Du Plessis, & Klopper 1989) has suggested that Eastern Cape Afrikaans formed the dialectal base of modern standard Afrikaans. The hypothesis, however, is based on very limited historical evidence, and Grebe's (2002) recent analysis suggests that it cannot be upheld.

9. Cf. also Ponelis' (1993: 16) report of communication patterns in a 1726 court case: 'Nineteen slaves were involved; three were Portuguese speaking, four Malay speaking and twelve spoke Dutch, but among themselves they spoke Portuguese'.

10. Bickerton (1984: 176) distanced himself later from this rather rigid threshold, describing it as a somewhat 'simplistic approach'. Yet, the basic principle remains: a large percentage of superstrate speakers (possibly between one-third and half of the population) supports processes of second language acquisition in the substrate community rather than the emergence of a pidgin/creole language.

11. Following Mufwene (2001: 209, fn. 6) basilectalization can be understood as 'the process by which a language variety diverges structurally towards a pole at an extreme from its lexifier'.

12. In the early years of the colony around eleven (closely related) Cape Khoe varieties were spoken. Most of these languages were extinct by the middle of the eighteenth century, following the disintegration of Khoe society. In the nineteenth century Nama, Korana (!Ora) and Gri (Xri) were still spoken by Khoe groups in the Richtersveld and along the Orange River. Today the Khoe languages are represented only by speakers of Nama (Richtersveld, northern Cape, Namibia; for details see Traill 1995, 2002). In assessing possible substratum influence linguistic information available for Korana and Nama is usually seen as reliable (Den Besten 1986: 200–201; Holm 1988/89 II: 348). According to Den Besten (2002), the fact that Khoe is an SOV language supported the successful acquisition of Dutch syntax by the substrate population. The languages of the slaves from India (Indo-Aryan and Dravidian languages) were also SOV (ibid.).

13. Den Besten (2002: 14) reconstructs *was* as Dutch *wat* ('what') plus the Khoe clitic *-ts* ('you', second singular masculinum).

14. On variationist or selectionist models of language change in general see also Croft (2000) and Mufwene (2001) as well as the foundational work by Labov (1972) and Weinreich, Labov & Herzog (1968).

15. The group of *vrije zwarten* also included political exiles from Batavia and Chinese convicts, which were not enslaved (see Davids 1991: 32ff. and Shell 1994: 232).

16. Rickford (1983) proposed a quantitative model of decreolization which takes the speech community rather than the individual as the starting point. According to Rickford's model, the primary impact of decreolization lies in the declining proportion of people who speak the basilect, and not in a decline in the purity or creoleness of the variety itself. In other

words, during decreolization individuals acquire (additional) lects closer to the acrolect without necessarily loosing their knowledge of basilectal forms (cf. also Escure 1997: 64–65). Rickford's analysis of decreolization is more than a simple re-interpretation of DeCamp's original model – as Mufwene (1988: 114) has noted: 'Structural/formal change and change in the proportion of speakers using linguistic features ain't the same thing! If decreolization means the first kind of change, it cannot mean the second.' In the discussion presented here the term decreolization is used in accordance with DeCamp's traditional model (i.e. as describing the structural effects on the speech of individuals), Rickford's population-based model, on the other hand, might well be applicable to the situation at the Cape.

Chapter 2

Afrikaner nationalism and the discovery of the vernacular

> Every time that the language question appears, in one mode or another, it signifies that a series of other problems are beginning to impose themselves: the formation and growth of the ruling class, the need to stabilise the most intimate and secure links between that ruling group and the popular national masses, that is, to reorganise cultural hegemony.
>
> Antonio Gramsci (1935). *Quaderni dal Carcere*[1]

The rise of dialect writing

The nineteenth century saw the rise of a popular Cape Dutch writing tradition. Dialect literature constitutes an important, albeit problematic source for sociolinguistic analysis. In order to assess to what extent these texts reflect the patterns and norms of (spoken) language use, the conditions of their production and reception need to be evaluated carefully (for linguistic analyses of early Afrikaans dialect texts see, for example, Lätti 1978; Uys 1983; Van Rensburg 1983; Van Rensburg & Combrink 1984; Ueckerman 1987; Raidt 1994 [1992]).

The first deliberate imitation of Cape Dutch Vernacular speech exists in an anonymous poem ridiculing the citizens of Swellendam (*Lied ter eere van de Swellendamsche en diverse andere helden bij de bloedige actie aan Muizenberg in dato 7 Aug. 1795*, 'Song in Honour of the Swellendam and many other heroes at the bloody action at Muizenberg on 7 Aug. 1795'; the text is reprinted Raidt 1991: 151-152). Over 30 years later M. D. Teenstra included in his travelogue a short dialogue which portrayed the colonists' spoken language (1830; reprinted in Nienaber 1971:7–9). The text shows prominent features of Afrikaans: e.g. subjective *ons* ('us') instead of *wij* ('we'), loss of verbal inflection, and use of the brace negation (see Chapters 5 and 6). In the accompanying commentary Teenstra described the dialogue as an example deliberately devised to introduce the reader to the most salient features of Cape Dutch Vernacular:

> Ziedaar in een kort bestek de grootste en misschien de meeste taalfouten en voor den vreemdeling meest ongewone gezegden en eigene spreekwijzen deser inwoners.'
>
> ['See here in a short space the biggest and perhaps the most frequent linguistic mistakes and for the stranger the most unfamiliar sayings and peculiar ways of speaking of these inhabitants.']

The reliability of travellers' anecdotes and reports is difficult to assess. Nineteenth century travel accounts in particular are notorious for their insufficient attention to factual accuracy, focusing instead on the amusement of prospective audiences and financial success (Batten 1978: 30; Brenner 1989: 14). The commentary provided by Teenstra himself indicates the stylized nature of the dialogue, and Roberge (1994a: 164) therefore described the text as a literary artefact: a 'composite description of potential features' which occurred in the Cape Dutch Vernacular speech of individuals with varying frequencies.

In the 1820s short journalistic pieces written in what was meant to represent the general colloquial language started appearing in the Dutch press. One of the first Dutch periodicals at the Cape was *De Verzamelaar* ('The Collector'; first issue 7/1/1826), a satirical-political paper which has been described as 'a kind of Dutch *Punch*' (Meurant 1885: 75). *De Verzamelaar* was owned and edited by Joseph Suasso De Lima (1791–1858), who had arrived at the Cape in 1818 from the Netherlands. Using different humorous pseudonyms (e.g. *Jan Hennepikker*, 'John Henpecker'), De Lima published fictitious letters in which he commented satirically on social and political events of local importance. The letters were written in a highly variable, non-standard form of Dutch. While De Lima's pseudonymous characters represented members of the European colonist class, Charles Etienne Boniface (1787–1853) published vernacular letters and dialogues, which were meant to represent the speech of Khoe, in the 1830s. Well known are Boniface's fictitious dialogues with *Hendrik Kok*, a stereotypical Khoe character. The dialogues were published in the *Zuid-Afrikaan* which was at the time under the editorship of Boniface (for a linguistic analysis of these dialogues see Franken 1953: 188–207; Scholtz 1965: 76–103).[2]

Boniface's texts differ linguistically from De Lima's earlier texts. His language use is characterized by almost complete loss of verbal inflection, frequent use of the brace negation and a non-standard pronoun system. Cape Dutch Vernacular speaking Khoe characters also feature in Boniface's popular comedy *De Nieuwe Ridderorde of De Temperantisten* ('The New Knight's Order, or The Temperatists'; 1979 [1832]), while the European characters in the play speak metropolitan Dutch. The stereotypical portrayal of the Khoe as 'jolly

Hotnots' – that is, quick-witted, but irresponsible characters whose merriment was typically alcohol-induced (Van der Merwe 1994: 22–23) – is also evident in Andrew Geddes Bain's variety show piece *Kaatje Kekkelbek, or Life among the Hottentots* ('Cathy Cackle Mouth'; 1838) which enjoyed much success in the colony (reprinted in Nienaber 1971; cf. also Mesthrie 1993). However, unlike the texts by Boniface, Bain's version of vernacular speech is characterized by extensive code-mixing/switching between Cape Dutch Vernacular and English (see Chapter 8 for a discussion of Afrikaans/English language contact).

Louis Henri Meurant (1812–1893), who was acquainted with both De Lima and Boniface (Meurant 1885: 75), took up their earlier attempts at vernacular writing in the 1840s with his letters and dialogues published in *Het Kaapsche Grensblad* ('The Cape Frontier News', Grahamstown). Again the pieces were published either anonymously or under humorous pseudonyms, commenting on all kinds of local events and socio-political issues (cf. Scholtz 1965: 265). In 1860 Meurant published two vernacular dialogues in *The Cradock News* (he had left *Het Kaapsche Grensblad* in 1851), arguing for the political separation of the eastern and western Cape. Since the frontier wars (1834 to 1835), eastern Cape separatists had campaigned for political independence, the move of the capital to the eastern districts and for a more federal structure of the colony's government (cf. Davenport & Saunders 2000: 103). Meurant's dialogues were reprinted as a booklet in 1861. The dialogues were widely read and the two fictitious characters, *Klaas Waarzegger* ('Nicolas Truthsayer') and *Jan Twyfelaar* ('John Doubter'), were quickly integrated into the popular culture of the time (see below). Figure 2.1 shows that from the 1860s onwards dialect texts developed into a popular written genre and the stream of contributions (although fluctuating) increased steadily. The popularity of dialect writing was not limited to the Cape colony and similar texts were published in the periodical press in the northern republics (Scholtz 1964: 177f.; Nienaber 1967).

Judging from Figure 2.1, Van Rensburg and Jordaan (1995: 109) would seem to be mistaken in their assessment that the end of the separatist movement in the early 1870s also meant the end of the use of Afrikaans for socio-political commentary. It appears that it was not only the political content of Meurant's writings that mattered (i.e. the question of eastern Cape separatism), but the general popularity of his texts established the genre of dialect writing as a recognized and recognizable linguistic practice in the emerging mass media of the colonial society.

There is little doubt that the aim of these writers was to write what was considered the local Cape Dutch dialect and not metropolitan or standard Dutch

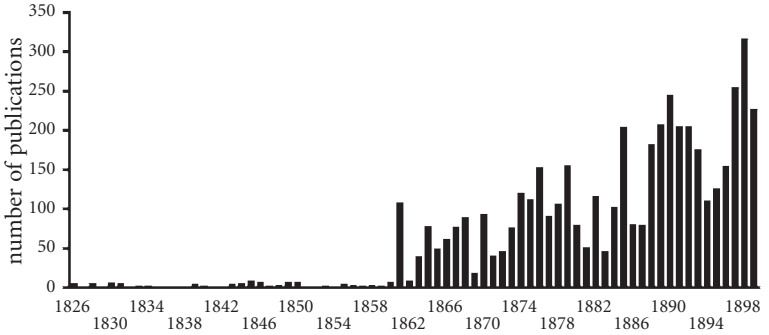


Figure 2.1 Publication of vernacular texts in newspapers and magazines in the Cape Colony 1826–1899 (number of texts per year; based on Nienaber 1966)

(Scholtz 1965: 277). The language name *Afrikaans* which locates the vernacular firmly in the colonial society and emphasizes its independence from metropolitan Dutch, occurred only sporadically in the first half of the nineteenth century. Up until the 1870s this variety was referred to as *Kaaps(ch) Hollands(ch)* ('Cape Dutch'; cf. Viljoen 1896: 17).

Neither De Lima nor Boniface, Bain or Meurant provided glosses or linguistic commentaries for their readers (as Teenstra had done for his European audiences). This suggests that their audiences had no particular difficulty in understanding the texts and were able to grasp their entertainment value without requiring additional information (possibly knowing the dialect themselves as a resource to be called upon in certain situations). Following Winer's (1984) discussion of Trinidad Creole writing, this might be seen as an indicator for the genuine character of the vernacular representations. However, comprehensibility alone does not guarantee linguistic authenticity and stereotypical exaggerations are a common feature of humorous and/or politically-oriented vernacular texts (cf. Ellis 1994; Baugh 1992; Janda & Auger 1992). Dialect texts frequently reflect a sociolinguistic stereotype of what constitutes 'the vernacular' in a general undifferentiated sense: a stylized Low colloquial, an amalgamation of different non-standard features which do not necessarily coexist (in general or in a given frequency) in the speech of any individual, but which endow the texts with the stereotypical characteristics of rustic, non-standard speech. Sociolinguistic terms such as 'stylization', 'variety imitation', 'double-voicing' and 'crossing' (cf. Preston 1992; Rampton 1995, 1999; Bell 1999; Coupland 2001; cf. also Bakthin 1981) are used to describe the strategic and self-conscious use of stylized linguistic forms in texts and social interac-

tions. Such purposeful and playful linguistic practices often project stereotyped identities and genres, and can become part of folk culture/popular culture (cf. Rampton's, 1995, study of adolescent language behaviour, and Preston, 1992, on imitations of African American English by white Americans; see also Lass', forthcoming, comments on 'dialect games' between standard-competent academic friends). In spoken discourse, the incorporation of such elements is usually flagged 'by pauses, hesitation phenomena, repetition and metalinguistic commentary' (Rampton 1995: 282). In the written texts discussed here linguistic stylization is flagged by the use of pseudonyms (representing social and ethnic stereotypes) and the overtly humorous and ironically self-referential content of the stories (see below). Although creativity is a central aspect of such practices, playful variety imitations can 'become institutionalized … in specific genres and situations' (Bell 1999: 524–525); for example, in jokes, anecdotes, gossip and interaction rituals such as greetings. In the case of Afrikaans the genre of humerous socio-political commentary in the growing print media became from the 1860s closely associated with these playful language practices.

The vernacular texts were generally characterized by irony and *spot* ('ridicule'), and a deliberately naïve and rustic style (referring, for example, to Queen Victoria as *tante*, 'auntie'; cf. Nienaber 1968: 11–16). They were interspersed with personal attacks on individuals of political, religious or cultural importance. Using the vernacular in the increasingly popular genre of *praatijes* ('conversations, dialogues') and *boerenbrieven* ('farmer's letters') became a way of speaking, an alternative (if not subversive) cultural practice when taking the point of view of those located outside of the dominant (economic and political) power structures of the colonial society, whose ideas differed from the orthodoxy of the dominant discourse. This is clearly visible in the vernacular letters published under the pseudonym *Jantje Eenvoudig* ('Johnny Simple'; alias T. F. Burgers 1834–1881). The letters were written in the context of the conflict between the progressive and orthodox forces in the *Nederlands Gereformeerde Kerk* (for the historical background cf. Raidt 1994 [1987]). While the orthodox party published its position exclusively in Dutch, the liberal party (represented by *Jantje Eenvoudig*) utilized the vernacular (Raidt 1994 [1987]: 275; Nienaber 1942: 8). A similar reasoning had also informed Meurant's language choice in his *Zamenspraak*. In a letter to Robert Godlonton (8/1/1861) Meurant explained his use of the vernacular as a strategic reaction to the conservative Dutch press in Cape Town which argued, using Dutch, against the proposals for political separation (quoted in Nienaber 1968: 6–7). However, not everyone was comfortable with the use of Cape Dutch Vernacular in the written domain. Criticisms of the vernacular writing practice abounded in the press

and many voiced discomfort in cases where the vernacular was used to discuss High culture topics such as education or church matters.

The vernacular letters were published anonymously or using *noms de plume.* Many of these pseudonyms established prototypical colonial characters: *Arme Boer* ('Poor Farmer'), *Boerenseun* ('Son of the Land'), *Een Burger* ('A Citizen'); others appear to have been part of a comic popular repertoire: *Polly Lekkergoed* ('Polly Sweets'), *Grietje Beuzemstok* ('Grietje Broomstick'), *Man in die Maan* ('Man in the Moon'), *Koos Papegaai* ('Jacob Parrot'), *Jantje zonder naam en zonder geld* ('Johnny without name and without money'). Intertextual references were common and are reflected in pseudonyms such as *Maat van die Man in die Maan* ('Mate of the Man in the Moon') and numerous references to Meurant's popular character *Klaas Waarzegger*: *Gert* and *Piet Waarzegger, Klein Klaas Waarzegger* ('Little Klaas Waarzegger'), *Klaasie, Klaas Zoon* ('Klaas' Son') and *Klaas Waarzegger jun.* (cf. Ueckerman 1987:4–7; for a complete list of pseudonyms cf. Nienaber 1967:63–100). Joyce's (1991:163) comments on nineteenth century English dialect literature fit the situation at the Cape rather well:

> [A]uthorship is irrelevant: the uniformity of method, tone, and content … reflects intertextual influences, but far more than this it reflects a uniformity of expectation in their audience. To sell they had to correspond to the people's self-mythology, and in responding they also shaped that mythology. … Authors' *noms de plume* established the right to speak for particular localities.

Most pseudonyms used in the dialect literature indicate the category 'rural or small town European colonist', but (stereotypical) coloured characters similar to *Hendrik Kok* or *Kaatjie Kekkelbek* were also popular. However, while the vernacular in the mouths of the colonists was established as a form of entertaining, yet serious, socio-political commentary, its association with coloured speakers was generally of a more burlesque type (cf., for example, the sketches written by J. J. De Kock, published in *Het Volksblad* 1859–1863).

By the second half of the nineteenth century Meurant's *Klaas Waarzegger* dialogues had become an integral part of Cape popular culture and were drawn on as a point of reference and orientation. This is evident from the following letter (1873) to the *Zuid-Afrikaan:*

> Onder mijn schrijwe denk ek soo aan de briewe van 'Klaas Waarzegger' wat omtrent 12 jaar geleden in die koerante was. Die skrywer daarvan moet een verstandige kerel gewees het, want hij kan het nie beter aangeleg het, om te seh wat hij wou seh, dan in die taal van Afrika. Dinge wat kort gelede gebeur is het die mense al vergeet al; maar praat over 'Klaas Waarzegger' en dadelik

> begin hulle te lag; hulle kan hom nie vergeet nie. (*Zuid-Afrikaan,* 5/7/1873)
>
> ['When writing I sometimes think of the letters of 'Klaas Waarzegger' which appeared about twelve years ago in the papers. The writer of these must have been an intelligent fellow because he could not have done better to say what he wanted to say than in the language of Africa. People have already forgotten things which happened a short while ago; but just talk about 'Klaas Waarzegger' and immediately they will start laughing; they cannot forget him.']

The fact that most vernacular texts used pseudonyms makes it difficult to identify their authors. Those authors who have been identified belonged to the educated middle classes and the intelligentsia, i.e. they were individuals perfectly able to write 'correct' Dutch. For them the use of the vernacular was a conscious choice (Nienaber 1942:21).

Information about the targeted audience comes from the above-mentioned letter from Meurant to Godlonton in which Meurant identified 'Dutch Boers' as the targets of his propagandistic writings. There is little doubt that the category of 'Boer' is a wide one; however, in the context of the separatist movement those whose support was needed, and who Meurant addressed in his dialogues, were likely to be those farmers to whom cheap labour, a local deeds registry office and a better regional infrastructure mattered – in short independent farmers with at least moderate property holdings (Le Cordeur 1981:43, 284f.). Furthermore, the audiences must have been familiar with the norms of Dutch to grasp the linguistic humour of the texts, that is, to understand that these texts were vernacular or dialect imitations. Meurant's dialogues were instrumental in establishing a stylized linguistic and social prototype of the ordinary but honest and common-sense Afrikaner farmer as a symbolic counter-point to the English-dominated urban colonial society, and to the Dutch-dominated High culture of education, church and publishing (cf. also Van Zyl 2000).

The production of dialect texts also supported the formation of an increasingly focused typological conception of what constituted 'the vernacular' (or 'Afrikaans' as it came to be called). That is, people who decided to try their hand at this new genre had a model in earlier attempts and in turn their efforts constituted a (written) model for later attempts. One might term this (following Stein 1994) a type of 'initial standardization', i.e. certain selection decisions were taken, but the process was not yet coordinated or institutionalized (nor for that matter necessarily aimed at vernacular elevation). Table 2.1 shows the frequencies for the Dutch relative pronoun *die* (vs. Afrikaans *wat*) for two successive time periods. Dutch forms are clearly used much less in the later period, suggesting that forms marked as Afrikaans had gained

Table 2.1 *Wat* and *die* as relative pronouns in the Transvaal periodical press, 1866–1889 (from Van Rensburg & Combrink 1984: 115). The difference between the two periods is statistically significant at $p < 0.025$ (chi square).

	no. of relative clauses	***die* as relative pronoun**	**% of *die***
1866–1883	325	82	25.2
1884–1889	1170	220	12.4

popularity. Orthographic conventions also crystallized from the 1860s in the vernacular texts.

The popularity of vernacular writing was not limited to the European colonial community. Most notably the Muslim community at the Cape developed its own tradition of vernacular writing from the 1840s (Muller 1962; Davids 1991). Here the vernacular was used primarily to facilitate the teaching of religious practices (since proficiency in Arabic was low at the Cape). At least 72 Cape Dutch Vernacular texts written in Arabic script were published between 1869 and 1957 (Van Selms 1979: viii; Kähler 1971: 70–188; Davids 1991: 91–93), and with the publication of Abu Bakr Effendi's *Bayân-ud-în* ('Exposition of the Faith', 1869) a linguistic standard of how to write Arabic Afrikaans became available at the Cape (Davids 1991: 110ff.).[3] The *Bayân-ud-în* was initially circulated in manuscript form and was printed only in 1877. Earlier examples of Arabic Afrikaans (the *Koples*-books) show more variable language use with strong influence from Dutch. In addition, a folk musical tradition – the so-called *Ghommaliedjies*, picnic or street songs based on Dutch folk songs – made use of Dutch and varieties of Cape Dutch Vernacular (Du Plessis 1944: 57ff.; Bickford-Smith 1995: 188–189).

Support for vernacular writing was not uncontroversial in the Muslim community. In the *First Report* on the educational system in the Cape Colony, Abdol Burns, a cap-proprietor and General Secretary for the Muslim community, made the following submission to the question 'Do you want to see English and Dutch both taught?':

> Yes; but not what they call Cape Dutch. I prefer proper Dutch. I do not talk what is called the 'taal' in my house. My son is able to read English, Dutch, and Arabic. He was taught English and he picked up the Dutch.
>
> (*First Report and Proceedings*, 1892: 96)

A tradition of vernacular writing also developed on the Moravian mission stations.[4] In 1873 the conversion narrative *Benigna van Groenekloof of Mambre* ('Benigna of Groenekloof or Mambre') was published by the Moravian press in Genadendal. The main text which describes the story of Benigna's conver-

sion to Christianity is written in Dutch; the dialogues, however, are written in Cape Dutch Vernacular. The Moravian periodical *De Bode van Genadendal* ('The Messenger of Genadendal'; founded in 1859 and published under the title *Die Huisvriend* ,'The House Friend', from 1914), also printed vernacular texts which were typologically similar to those seen in other Dutch periodicals of the time, and offered social and political commentary in the guise of dialogues or fictitious letters (written under *noms de plume* such as *Oom Gezels*, 'Uncle Storyteller', *Vriend van Allen aan Oom Alspraat*, 'Friend of Everybody to Uncle Say-Everything'; cf. also Belcher 1987).

In her afore-mentioned assessment of Trinidad Creole writing Winer (1984: 194–196, also Winer 1993: 130–131) formulated two criteria for determining the linguistic reliability of literary vernacular representations:

i. there should be internal consistency of the linguistic features within a given text, and
ii. the representations should show external consistency with other evidence or the contemporary language.

When comparing the texts of the dialect literature with each other and with modern varieties of Afrikaans both criteria appear to be fulfilled, especially with regard to morphology, and to a lesser degree syntax and phonology (cf. Raidt 1994 [1992]). However, regarding the criterion of external consistency, the situation is less clear-cut if nineteenth century documents which are not deliberate and conscious vernacular representations (such as private letters and diaries) are included in the comparative analysis. Although the morphological and syntactic variants which occur in the dialect literature are also present in the private documents collected in the *Corpus of Cape Dutch Correspondence*, the differences in frequency are remarkable. For example, while the (post 1860s) vernacular writing tradition shows almost complete loss of verbal inflection and loss of gender, these aspects are still highly variable in nineteenth century family letters and diaries (see Chapters 5 and 6 for details).

Such observations and comparisons raise questions about the linguistic authenticity of the popular literature and its use by language historians. The Cape dialect writing tradition must be understood as a complex, socially mediated representation of 'the vernacular' as a cultural stereotype, rather than offering language historians direct access to the structures of the spoken language. As deliberate and conscious variety imitations these texts were characterized by various layers of stylizations and stereotypical overgeneralizations. Raidt's assessment of the letters of *Samuel Zwaartman* ('Samuel Blackman'; alias H. W. A. Cooper, a law agent and later advocate) as a historically reliable reflec-

tion of the *destydse omgangstaal* (the 'contemporary colloquial language'; Raidt 1968:117; also Lätti 1978) is thus highly debatable. That dialect literature in general is of limited use for identifying the quantitative aspects of linguistic variation was pointed out by Bailey and Ross (1992:519): 'while literary dialect is often useful in identifying the presence or absence of features, it frequently treats quantitative features qualitatively'.

However, the dialect literature is an important source for understanding the creative and strategic aspects of language change and language history. Coupland's (2001) discussion of dialect stylization and authenticity is relevant in this context. Based on his analysis of 'dialect performance' in radio talk, Coupland argued that dialect stylizations or variety imitations – despite their self-conscious artificiality and divergence from natural, everyday language use – 'can potentially deliver forms of personal and cultural authenticity that transcend local playfulness, so that the identificational effect is neither mere play nor outright parody' (2001:347). In the case of Afrikaans, the literary dialect stylizations became an important resource in the cultural construction of Afrikaner national identity, and provided a linguistic tradition on which early codification efforts could build.

Afrikaner nationalism and early vernacular standardization

Afrikaner nationalism with its romantic belief in the isomorphy of language and nation was of central importance for the sociolinguistic development at the Cape, where it led, as elsewhere in the world, to 'a golden age of vernacularising' (Anderson 1991:71).[5] Historians of nationalism have noted that although nationalist movements emphasize the historicity and continuity of the national community, few national symbols are older than the movements themselves (cf. Hobsbawm 1983; Gellner 1983; Hofmeyr 1987). Hobsbawm (1983:14) summarized the inventive and constructionist aspect of nationalism as follows:

> We should not be misled by a curious, but understandable paradox: modern nations and all their impediments generally claim to be the opposite of novel, namely rooted in the remotest antiquity, and the opposite of constructed, namely human communities, so natural as to require no definition other than self-assertion … because much of what subjectively makes up the modern 'nation' consists of such constructs and is associated with appropriate and, in general, fairly recent symbols or suitably tailored discourse (such as 'national history'), the national phenomenon cannot be adequately investigated without careful attention to the 'invention of tradition'.

Acknowledging that nationalism and national symbols (including language) are creative constructs emphasizes the importance of human agency and imagination, but this is not to say that national symbols are entirely or even primarily artificial. Nationalist movements typically work with what already exists (such as the vague and unspecified heritage of pre-modern memories, myths and traditions), leading to manipulation and re-construction rather than downright invention (Smith 1993). Their activities can be likened to the work of the *bricoleur* as described by Lévi-Strauss (1976: 19–22): the *bricoleur* is engaged in the creation of new structures and meanings through the novel combination and interpretation of existing elements.

Language, the hallmark of the romantic concept of a nation, did not escape the creative activities of nationalist movements. Deutsch (1968: 603–604) has shown that from the mid-nineteenth century such movements showed a clear tendency to manipulate available linguistic resources and national identities by increasing and even creating differences between closely related varieties (cf. also Hobsbawm 1990: 54, 101–111; Anderson 1991: 67–82). A language historian describing language variation in the context of nationalism needs to be careful not to make the mistake of accepting linguistic nationalism on its own terms by uncritically taking the statements of nationalist leaders as reflecting facts about actual communication patterns; yet at the same time one has to acknowledge that nationalist leaders did not invent linguistic traditions at whim. As Gal (1989: 349) put it in the context of a discussion on language and power: it is vital to maintain an approach to the historical sources which 'navigate[s] between a radical distrust of language as a conspiratorial distortion and a relativist confidence in its neutrality'.

Afrikaner political and cultural nationalism emerged in the 1870s in response to the challenges of British imperialism. The advent of Responsible Government in 1872 and the mineral revolution further intensified the struggle for political power and economic resources in the colony (Giliomee 1989; Davenport & Saunders 2000: 107–111). It was in this context that the first language society for Afrikaans, the *Genootskap van Regte Afrikaners* ('The Society of True Afrikaners'; henceforth GRA) was founded in 1875 in Paarl. The aim of the society was to promote Cape Dutch Vernacular, or Afrikaans as it was now called, as the national language of the Afrikaner nation.[6] Soon after its foundation the GRA brought out its own periodical, *Die Afrikaanse Patriot* ('The Afrikaans Patriot'; henceforth AP), the first grammar of Afrikaans (*Eerste Beginsels van die Afrikaanse Taal*, 'First Beginnings of the Afrikaans Language'; 1876), and a national history (*Geskiedenis van Ons Land in die Taal van Ons Volk*, 'History of our Country in the Language of our People'; 1877).

The publication of the grammar was closely linked to the GRA's nationalist project and aimed at establishing Afrikaans as an autonomous language – a language which had its own linguistic rules and regularities, and which was not parasitic on the grammatical system of Dutch. According to Du Toit's (1909 [1880]) history of the Afrikaans language movement, popular demand for the society's publications was strong: by 1895 over 70,000 copies of Afrikaans books and pamphlets, published by the society's own publishing house (D. F. du Toit & Co., founded in 1878) had been sold and the *Patriot* (now weekly) had a circulation of over 3000 copies (Du Toit 1909 [1880]: 78).[7] The grammar (offered for a mere half shilling) also sold well and by 1895 over 6000 copies had been sold (Du Toit 1909 [1880]: 173–176). Soon the type of language used by the GRA – a highly uniform variety characterized by complete regularization of the verbal system, loss of gender, a pronoun system different from Dutch and certain innovative syntactic features – became known as *Patriots Taal* ('Language of the Patriot') or *Du Toits Taal* (the latter term reflects the fact that S. J. Du Toit, one of the founding members of the GRA, dominated the movement's early codification efforts; cf. Dekker 1926: 77; Holliday 1993: 15–17). Although no GRA-meetings were held after 1878, individual members (especially S. J. Du Toit and C. P. Hoogenhout) continued with the publication of vernacular texts and language codification, including the second edition of the grammar in 1897 and the publication of a dictionary in 1902/1904.

The GRA rejected, at least rhetorically, prescriptive standardization and emphasized the primacy of linguistic usage (Du Toit 1891: 16). This has led to its codification work being viewed as primarily descriptive (cf. Raidt 1994 [1985]: 318). However, despite an overt *laissez-faire* rhetoric, poems sent to the *Patriot* for publication were corrected in accordance with the language use propagated by the GRA and occasionally entire letters were rewritten (cf. AP 17/4/1890). Most contributors, however, closely followed the rules formulated in the 1876 grammar and limited their linguistic experiments to spelling variations. Members and supporters of the GRA were well aware of the interconnections and continuities between its propagated language use and the earlier dialect texts (cf. Du Toit 1909 [1880]: 3ff.; De Waal 1932: 2). The conventions of the earlier dialect texts continued to function as a linguistic and stylistic model for writers.

The GRA published vernacular reading material not only for adults but also for children: Hoogenhout's *Geskiedenis van Josef* ('History of Josef') was published in 1873 (second edition in 1883), followed in 1878 by the primer *Spel- en Leesboek ver Afrikaanse Kinders* ('Spelling and Reading Book for Afrikaans Children'; second edition in 1896). In 1880 the *Eerste Afrikaanse*

printjis boek ver soet kinders ('First Afrikaaans picture book for well-behaved children'; second edition in 1885) was published and the *Patriot* also had a special reading section titled *Ver Kinders* ('For children').[8] The GRA also paid attention to what in today's terminology (cf. Cooper 1989) would be called acquisition planning by publishing English-Afrikaans teaching material (*Samesprake in Afrikaans and Engels*, 'Dialogues in Afrikaans and English', 1884–1890, second edition 1897; over 12,000 copies were sold before 1896; cf. *Di Tweede Afrikaanse Taalkongres* 1897:4). Prestige planning (cf. Haarmann 1990) was pursued not only through the publication of a grammar (thus bestowing on Afrikaans the status of a 'real' language), but also by appropriating arguments from popular linguistic science. This is illustrated in Du Toit's programmatic text *Afrikaans ons Volkstaal* ('Afrikaans our national language', 1891), which draws on arguments from historical-comparative linguistics to argue against the pejorative description of Afrikaans as a *Hotnotstaal* ('Hottentot's Language'), i.e. a mixed and 'bastardized' language. The label of *mengelmoes* ('mixed language') is instead applied to English, the dominant language of the Cape Colony under British rule:

> 1. Afrikaans is di suiwerste afstammeling van di Germaanse Tak … daarom geef ons dit as 'n direkte spruit daaruit, nader nog as Hollands of Duits. 2. Engels is 'n vermenging van die Germaanse en Romaanse Takke, daarom het ons dit moet voorstel as 2 lote van di 2 takke wat aan makaar gegroei het. Mar eintlik is Engels 'n mengelmoes, wat jy glad ni onder een of ander klas kan breng ni. (Du Toit 1891:4)
>
> ['Afrikaans is the purest offspring of the Germanic branch … we therefore give it as a direct shoot, even closer than Dutch or German … English is a mixture of the Germanic and Romance branches, therefore, we must imagine it as two lots of two branches which grew into one another. But actually English is a mixed language which you can't really group with one or the other class.']

In the contemporary press the GRA members were attacked as 'jong schoolmeesterjies en wynboertjies' ('young teachers and wine farmers'; Von Wielligh 1925b:15). Independent farmers, ministers and teachers were indeed prominent in the GRA. In addition, there were small entrepreneurs and independent artisans among the society's members (such as D. G. Roussouw and P. S. Toerien, both shop keepers, I. A. Peroldt, a wagonmaker, G. W. Malherbe, a cart-builder, and J. J. Uys, a butcher) as well as those belonging to the emerging professional class (such as J. M. Hoffman, a medical practictioner). On the whole the members of the GRA were prototypical supporters of nineteenth century nationalist movements, that is, they belonged to the rising, educated,

middle classes including the intelligentsia (Hobsbawm 1990: 116–122; see also the comment by Du Toit, 1909 [1880]: 77, on the impossibility of finding someone who knows about farming to write for the *Patriot* newspaper). All of the society's members came from the Boland area (primarily Paarl, Wellington and Malmesbury) and the majority of the founding members were born in the 1840s and early 1850s. This generational bias was maintained among those who joined the society later (a list of members can be found following the *Re'els and Bepalings,* 'Rules and Regulations'; reprinted in Nienaber 1974: 55–56).

Following the institutional demise of the GRA in the late 1870s and early 1880s, the nationalist agenda was taken over by the *Afrikaner Bond* ('Afrikaner Union'). The *Bond* was founded in 1879 by S. J. Du Toit to protect the economic and political interests of Afrikaner farmers and the petty bourgeoisie. Initially, the *Bond* advocated a strong neo-Calvinist, anti-English Afrikaner nationalism (against 'liquor, lucre and redcoats'; Davenport & Saunders 2000: 107). The movement was taken over in the 1880s by wealthy commercial farmers and under the leadership of Jan Hendrik Hofmeyr the *Bond* followed a more moderate policy: collaborating with, rather than challenging, British dominance (Giliomee 1989: 66; Bickford-Smith 1995: 41). The change in constituency also affected the linguistic orientation of the *Bond,* whose members now campaigned for the right to use Dutch in official functions; advocacy for Afrikaans was limited to a minority of its members.

Standardization efforts continued after the Anglo-Boer War (1899–1902), but were no longer in the hands of the GRA. The geographical centre of the promotion of Afrikaans shifted from the Cape to the northern interior where Johannesburg and Pretoria established their position as economic and increasingly also as cultural centres. New language societies took over: the *Afrikaanse Taalgenootskap* (the 'Afrikaans Language Society'; Transvaal) was founded in 1905, *Onse Taal* ('Our Language'; Orange Free State) and the *Afrikaanse Taalvereniging* ('Afrikaans Language Union'; Cape Province) both followed in 1906. While much work of the GRA had centred on codification and the publication of folkloric literature, the new language societies concentrated on the production of High culture fictional and non-fictional literature, leading to increasing popularity among members of the more traditional élite. The *Zuidafrikaanse Akademie voor Taal, Letteren en Kunst* ('South African Academy for Language, Literature and Art') was founded in 1909 with the explicit aim of promoting both Dutch and Afrikaans in South Africa. The tasks of the academy were specified as follows:

> De handhaving en bevordering van de hollandse Taal en Letteren en van de Zuid Afrikaanse Geschiedenis, Oudheidkunde en Kunst, en aan de ontwikkeling 'n wetenskappelike en artistieke leiding te geven. Onder het woord 'hollands' wordt verstaan de beide taalvormen gebruikelik in Zuidafrika.
>
> (quoted in De Villiers 1934: 278)
>
> ['The maintenance and support of the Dutch language and literature and of the South African history, archaeology and art, and to provide the development with scientific and artistic leadership. The word 'Dutch' is understood as including both language forms used in South Africa.']

Following the introduction of Afrikaans in primary schools in 1914, the Academy came to the fore as the central agency of language standardization and elaboration. The time around 1917 has been described as a transition period (*owergangstydperk*): new standard norms were slowly implemented but language use still exhibited considerable variation. Ultimately the elaboration of standard Afrikaans took place in the context of what has been termed *Vernederlandsing*, i.e. systematic adlexification from Dutch (see also Chapter 8).

Folk taxonomies and language attitudes

Even a brief look at the extensive secondary source material presented in Nienaber (1950) shows that the linguistic situation was generally perceived in terms of a rather static system of social and ethnic stratification. That speakers' perceptions and opinions are dependent on variables such as age, social group and/or sex, and might contradict linguistic analysis has been shown in empirical sociolinguistic research (cf. Mattheier 1985). Language users rarely perceive linguistic variation as continuous but typically report the existence of a clear system of two or three alternatives, often assigning everything which falls short of the standard or prestige norm to an amorphous low-prestige vernacular category. Linguists have, therefore, argued that what speakers think (language evaluation) should not be confused with the system of varieties actually in use (cf. Bloomfield 1944; Grootaers 1959; DeCamp 1971a; more recently also Patrick 1999: 275). Metalinguistically articulated (and socially mediated) beliefs about the structure of linguistic variation and folk linguistic taxonomies are, however, valuable sources for the reconstruction of the social meaning of variation, attitudes to language variation and varieties, linguistic stereotypes and ideologies; in short, beliefs about language which serve to rationalize and justify linguistic usages within a speech community.[9]

Social evaluations of language use and language varieties play an important role during processes of standardization, and establish clear and unambiguous boundaries between 'acceptable' (standard) and 'unacceptable' (non-standard) usages and speech forms. Binary classifications similar to Ross' (1954) popular distinction between U and non-U were common in the nineteenth-century language discourse at the Cape. Changuion (1844: 23) distinguished (excluding the metropolitan standard) 'educated' from 'uneducated' speech, based on the frequencies of Africanderisms (i.e. linguistic forms typical of Cape Dutch and unknown or rare in metropolitan Dutch; cf. also Swaving 1830: 302f.; Burchell 1953 [1822–1824]: 15; De Lima 1844: 9; Elffers 1903: 5). According to Changuion, 'uneducated' forms of the language were spoken primarily by the non-European population, but were also common among artisans and farmers of European descent. In the last quarter of the nineteenth century there existed a variety of terms for the description of basilectal Cape Dutch Vernacular, all of which show that it was associated in the popular consciousness of the time with poverty and colour: *Hotnotstaal* ('Hottentot's language'), *Griekwataal* ('Language of the Griquas'; see Chapter 1), *Kombuistaal* ('Kitchen language'), *plattaal* ('Low language'), *Brabbeltaal* ('Jabber language') and *Bastaardspraak* ('Bastard speech'; cf. Hofmeyr 1987: 97; Nienaber 1950: 24, 39–41).[10] That the educated/upper classes did not speak Afrikaans, but rather an extraterritorial variety of Dutch was emphasized by commentators throughout the nineteenth century. M. L. Wessels described the situation in 1880 as follows:

> [A]s far as I am able to tell, the majority of the better class … speak High Dutch in a slightly modified form … in a form that differs from the *purest* Cape Dutch as day does from night. Their grammar may now and then be at fault, and the distinctions of grammar may be disregarded, but in the main their language is defective High Dutch … That … the want of inflection is limited … to an indiscriminate use of the particle in one form to denote every case and gender, no one will deny; and it is only in the lower, I might say the lowest, the altogether uneducated classes, that the loss of inflection becomes generally perceptible in the verb. (*Cape Monthly Magazine*, 1880: 351, 45)

In an article written for the periodical *De Zuid-Afrikaan* (1/5/1890) Nicolaas Mansvelt distinguished *Kombuis-Hollands* ('Kitchen Dutch') from acrolectal Cape Dutch Vernacular speech which he called *Voorkamer-* or *Salon-Afrikaans* ('Drawing-room Afrikaans'). A more detailed description of the linguistic differences between educated or civilized (*beskaaft*) and uneducated (*plat*) vernacular speech was presented in a letter published in the *Zuid-*

Afrikaan (23/10/1890). The letter which claims to represent *beskaaft Afrikaans* shows a number of interesting linguistic features: maintenance of the preterite tense, absence of the brace negation, the adjective inflection still agrees with the Dutch system, the first person subject plural pronoun is *wy* (instead of Afrikaans *ons*), the third person subject plural pronoun is *sy* (instead of Afrikaans *hulle*), and plural verbs and infinitives are inflected. The use of the first person object plural pronoun *ons* in subject position, uninflected plural verb forms and infinitives are explicitly assigned to the system of *plat Afrikaans* ('low Afrikaans').

While most commentators reported only two forms of the vernacular (i.e. *plat* vs. *beskaaft*), the GRA leader S. J. Du Toit distinguished three varieties: *Heere-*, *Boere- en Hottentots-Afrikaans* ('Master-, Farmer- and Hottentots-Afrikaans'; *Zuid-Afrikaan*, 30/1/1875). This three-fold taxonomy was central to the language-political program of the GRA since it allowed the separation of *Boere-Afrikaans* (the language form promoted by the GRA and represented in the dialect literature) from upper-class speech, as well as from associations of colour and poverty (the three-fold scheme was, however, not an invention of the GRA and had also occurred in earlier commentaries; cf. Nienaber 1950: 10–11, 25; also Elffers 1903: 5; Van Rijn 1914: 13).[11] The *Patriot* was not only at pains to emphasize the differences between Afrikaans and the vernacular of people of colour, but also ridiculed *Heere*-Afrikaans as an unsuccessful effort to speak and write Dutch, and rejected such linguistic varieties as artificial and inauthentic (cf. AP 19/5/1882). While the variable nature of many acrolectal and upper mesolectal varieties (maintaining, for example, certain inflections such as *-en* for infinitive and plural marking, the neuter for certain nouns, etc.) was thus evaluated negatively, the highly regularized and uniform language described in the first GRA-grammar was defined as authentic and 'good' language. The metalinguistic discourse of the GRA thus shows the emergence of a clearly puristic response to linguistic variation (a Standard Language Ideology; cf. Milroy & Milroy 1985a), reflecting the need to establish clear, non-fuzzy boundaries between the linguistic norms of metropolitan Dutch and the new national language Afrikaans. By constructing Afrikaans as a uniform linguistic entity which could unambiguously be contrasted with metropolitan Dutch, an important pre-requisite for the development of language loyalty was achieved (cf. also the comments by Langenhoven 1935: 203–204, on the need to draw a *skerp lyn*, 'a sharp line', between Dutch and Afrikaans as a necessary condition for successful status planning).[12] The discursive construction of *Hottentot Afrikaans* served as similar purpose, clearly locating GRA Afrikaans in the European colonial society.

Perceptions of clear-cut boundaries and linguistic uniformity were however not shared by all of those who campaigned for vernacular elevation. This is evident from a letter (4/11/1894) by Francois S. Malan (who was at the time studying in Edinburgh) to W. J. Viljoen:

> Ek dink dat jou begrip van di Afrikaanse saak veul nouer en beperker is dan di van my. Di Afrikaanse taal wat jy voor werk is reeds gevormd in hoofsaak en het al meer of min syn bepaalde eigenaardighede ontvang … Mar dan is daar 'n menigte ander Afrikaner wat meen dat di Afrikaans van vandag nog maar di eier in di nes is wat nog moet uitgebroei worde, 'n pas gebore kindje van wi jy nog ni kan sé wat eigenlik di karaktertrekke van di man sal wees ni … Jy het al di woord 'purist' met betrekking tot di jong taal, terwyl di anderkant meen dat alle woorde en gebruike ewe veul reg het om in di stryd van bestaan op te tre. An di een kant word di gebruik van 'n vreemde woord afgekeur terwyl di ander sé: ge al di woorde kans en kyk watter een door di algemeen gebruik syn bestaan sal regvaardig.
>
> ['I think that your understanding of the Afrikaans question is much narrower and more limited than mine. The Afrikaans language for which you work is already fundamentally formed and has more or less received its defining characteristics … but there are many other Afrikaners who are of the opinion that the Afrikaans of today is just the egg in the nest which still must be hatched, a new-born child of which you cannot say what actually will be the characteristics of the man … You have already the attitude of a 'purist' with regard to the young language, while the other side is of the opinion that all words and uses have equal right to take part in the struggle for existence. On the one side, the use of a foreign word is rejected, while others say: give all the words a chance and see which one will justify its presence through the general usage.']

The perception of GRA-Afrikaans as being too extreme or exaggerated, and thus not representative of the variable nature of the actual spoken language, was not uncommon in the late nineteenth century (e.g. the letter by *Pietje Voorzichtig*, 'Pete Careful', to the *Zuid-Afrikaan* 13/11/1875). Malan also refers to the *onstandvastje karakter* ('fickle character') of Afrikaans in other places (cf. diary entry 23/10/1894), and a warning against a dogmatic and puristic approach to language codification was issued by Malherbe (1917: 17–18).

An important aspect of linguistic folk taxonomies is that they constitute evaluative hierarchies and allow for the historical reconstruction of language attitudes. It not surprising to find that in the context of the growing racialization of society a sociolinguistic category of *Hottentot Afrikaans* became a focal part of the popular discourse on language use (especially since such a category was already well-entrenched in the popular dialect texts). At the same

time more general class demarcational needs are evident in the separation of educated vs. uneducated Afrikaans. The growth of the educational system from the mid-nineteenth century (see Chapter 3 for details) supported the devaluation of popular modes of speech, and the recognition of the exogenous Dutch standard as prestigious. In general, the attitudes expressed in contemporary discourse are similar to those surrounding many non-standard codes, varying between rejection as an example of degeneracy and acceptance as a symbol of solidarity.

A diglossic community?

The linguistic situation at the Cape during the eighteenth and nineteenth centuries has usually been described as diglossic: Cape Dutch Vernacular (or Afrikaans) was the natively acquired Low variety and Dutch was the High variety, learned at school and used as a formal spoken and written language.[13] In the sociolinguistic literature a distinction is sometimes made between 'narrow' (or classical) diglossia and 'broad' diglossia. In speech communities characterized by narrow diglossia two varieties of one language exist in strict functional complementarity. The Low variety is learned as the first language by all members of the speech community; the High variety is never used in informal settings (cf. Ferguson 1959). In broad diglossia, on the other hand, two or more styles, varieties or languages exist as stable elements in the speech community's communicative repertoire, each variety is allocated to different functions, and yet there is no strict complementarity and the High variety can occur in informal conversations; the acquisition of the Low variety as a first language is not at issue (Fishman 1967/1980; Myers-Scotton 1986). Broad diglossia also includes cases where High and Low varieties are genetically unrelated languages and exist within a disjunctive society, that is, not all members of the society show diglossic behaviour, however, the overall societal distribution appears to be diglossic (cf. Fishman 1980: 7–8). While the extension of the notion of diglossia to scenarios where the languages in question are unrelated is generally accepted, extending the term to disjunctive speech communities has been questioned (Timm 1981).[14] In order to keep the different scenarios conceptually separate, Britto (1986: 35–40, 287) suggested the following classification of diglossic situations:

i. ***Use-oriented diglossia***: classical diglossia as described by Ferguson. The High variety is superposed for the entire speech community and everyone

learns it as a second language. Use of the High variety depends on domain specialization and is thus use-oriented; no section of the community uses the High variety for ordinary conversation. As noted by Hudson (2002) such scenarios are rare rather than common in language history.

ii. ***Dialectal or user-oriented diglossia:*** the High variety is not a superposed variety for all members of the speech community (i.e. certain groups acquire the High variety as their L1 and use it in informal conversation). Use of High and Low varieties depends not only on domain but also on social characteristics such as ethnicity, religion and class. The High variety is used as normal conversational language only by the élite. As a prestige model the language use of élite can 'provide the social impulse for shift away from L as the vernacular' (Hudson 2002:7).

iii. **Pseudo-diglossia:** the two varieties are used by separate speech communities within a given geographical or political organization; there is, however, no group-internal diglossia.

It was already noted in Chapter 1 that to describe the society at the Cape as diglossic has important consequences for the interpretation of the historical record. In cases of use-oriented diglossia the written language typically belongs to the High culture domain and thus offers the language historian little evidence for the reconstruction of patterns of variation in the spoken language. The situation is different in the case of dialectal diglossia where High and Low codes exist at the level of the speech community, but not necessarily at the level of the individual.

Language naming and writing are metalinguistic activities with the potential to shape perceptions about linguistic variation. The case of Romance provides an instructive example. Traditionally language historians have described the relationship between Latin and Proto-Romance during the Early Middle Ages as a case of diglossia. This interpretation was questioned by Wright (1982, 1991) who argued that the conceptual distinction between Latin and the Romance vernaculars was in fact 'created' by the Carolingian scholars when they developed a new system of distinctively non-Latin spellings to represent spoken varieties of Latin. Subsequently, these written varieties were interpreted as structurally distinct languages and the Latin-Romance variation continuum, which existed during the eighth and ninth centuries, was gradually transformed into an opposition of languages. The perception of language varieties as distinct, functionally, socially and linguistically well-defined entities is intricately related to the kind of metalinguistic reflections which emerge in the context of writing and formal language instruction. Wright concludes:

> [S]ystems of diglossia need not only be intentionally set up, but also to be continually reinforced subsequently – mainly by teaching the High variety in the education system – in order to arise at all, and do not arise naturally otherwise. What exists otherwise, and do indeed evolve unplanned, in a single wide speech community, are complex patterns of sociolinguistic variation.
>
> (ibid.: 107)

Within a speech community different groups of speakers are affected to differing degrees by such diglossic engineering, i.e. it affects first and foremost those with regular and intensive access to education. Although travellers had commented on the existence of a distinct Cape Dutch dialect from the mid-eighteenth century (e.g. Mentzel, Sparrman, Burchell, Swaving, Lichtenstein, Teenstra), the conceptual distinction of a clear High/Low dichotomy surfaces in the secondary sources only from about 1850, and thus developed in parallel to the dialect writing tradition, the emergence of the language name 'Afrikaans', and the expansion of the public education system under British rule (see Chapter 3).

The source material from the late nineteenth and early twentieth centuries includes various observations about differences between the casual spoken and the more formal spoken (and especially written) languages at the Cape (cf. Brill 1909 [1875]; Colquhoun 1906: 139–142; De Vooys 1913: 179; Langenhoven 1922 [1914]: 38–34, 52–53).[15] References to extensive differences between spoken and written/formal language were particularly common in the writings of those supporting the linguistic agenda of the GRA. Other commentators argued that the problems people encountered when reading Dutch were not a question of diverging grammars but of style, that is, what was needed was the use of a more simple style – a kind of plain Dutch – rather than vernacular elevation (e.g. letter by *conservatief*, 'conservative', to the *Zuid-Afrikaan* 24/2/1875; cf. also De Villiers 1934: 177–176).[16] At the same time, detailed remarks on the functional distribution of codes suggest that the notion of diglossia might nevertheless be applicable to the linguistic situation at the Cape. Viljoen's (1896: 24) description of the use of Cape Dutch Vernacular appears to be typical of a diglossic Low variety:

> Man hört sie unter den Dienstboten, auf dem Felde, in der Familie und im gewöhnlichen Verkehr. Sie ist auch die Sprache, welche der Bur seinen Untergebenen und Dienstboten gegenüber anwendet.
>
> ['One hears it [the Afrikaans language] among the servants, on the fields, in the family and in general conversation. It is also the language which is used by the farmer in conversations with subordinates and servants.']

Dutch, on the other hand, was *Kirchen- und Kanzelsprache* ('church and pulpit language'), used in prayer, writing and formal situations (Viljoen 1896: 21–22; Te Winkel 1897: 10; Hoogenhout 1904: 9, 13; also Nienaber 1950: 94–95). Viljoen, however, also notes dimensions of social stratification and the Low variety appears to have been exchanged asymmetrically, which is uncommon in diglossia (cf. Hudson 2002: 4).

Moreover, although functional specialization is indicated in some of the commentaries, not all members of the speech community perceived the informal, spoken language as a uniform vernacular category; many suggested that the spoken vernacular comprised a wide range of different varieties and variants, ranging from acrolectal Cape Dutch to more mesolectal and basilectal varieties.[17] Given such references to various intermediate forms of the spoken language it is debatable whether the situation at the Cape can or should be described as diglossic, or whether these observations indicate a sociolinguistic dialect continuum characterized by a 'gradasie van afwyking tussen Afrikaans en Nederlands' ('gradation of divergence between Afrikaans and Dutch'; Nienaber 1950: 34; also Besselaar 1914: 205, 211).

There is little doubt that metropolitan Dutch served as a Dummy High variety at the Cape (on the sociolinguistic concept of a Dummy High cf. Platt 1977: 373–374): it was seen as prestigious and learnt through secondary exposure (i.e. reading and formal schooling), but was rarely used by Cape-born speakers except in official documents. The growth of Afrikaner nationalism from the 1870s was at odds with the continued use of an exogenous High variety, and there is some evidence that from the second half of the nineteenth century the situation at the Cape was changing rather rapidly: knowledge and use of the norms of metropolitan Dutch was fading, a development which was possibly accelerated by the existence of an English-medium public school system from the 1830s (Mansvelt 1884: 180; cf. also Wessels in *Cape Monthly Magazine* 1880: 9, Viljoen 1891: 4; on the role of English in the Cape Colony see Chapter 8).

Some forms of Cape Dutch Vernacular (*beskaaft*) were prestigious enough to be used in (spoken) High culture domains such as sermons and political speeches (see *Verslag van het Eerste Congres* 1897: 6; also Te Winkel 1897: 10; Elffers 1903: 5; Colquhoun 1906: 141). Basilectal Cape Dutch Vernacular (*plat*), on the other hand, is a typical example of a stigmatized Low variety, and terms such as *kombuistaal* ('kitchen language') unambiguously describe its social location. Judging from the secondary source material, varieties close to metropolitan Dutch were used by urban, educated members of the speech community in informal and casual interaction, while their use of basilectal

Cape Dutch Vernacular appears to have been limited to a rather restrictive set of domains (generally socially non-symmetrical interactions as noted by Viljoen). Elffers (1903:5) described the informal acrolectal varieties of Cape Dutch impressionistically as follows:

> a language lacking the grammatical niceties of the Dutch of Holland and the shades of meaning which necessarily adorn a tongue of which the learned make use, besides discarding much of the idiom of the North – but none the less expressive

Elffers distinguished acrolectal varieties from the 'real patois': 'low and undeveloped, dependent on circumstances and locality, easily influenced, and becoming more and more Anglicised'. Such descriptions support an interpretation of the situation at the Cape as an example of user-oriented or dialectal diglossia, i.e. varieties close to Dutch were not superposed for the entire speech community, and use of the Low variety depended on social variables such as class and ethnicity. In addition, the regular comments on intermediate language forms suggest that the nature of Cape Dutch Vernacular speech was more fluid and continuous than implied in the popular two- and threefold taxonomies which dominated much of the metalinguistic and language-political discourse. A clear continuum situation was described, for example, by Colquhoun: 'the most puzzling feature of this patois is found in the shades or types of vernacular which are encountered' (1906:139; cf. also Elffers', 1903:6–7, discussion of the 'Shades of language and where to expect them'). The protean and fluctuating nature of variation in user-oriented or dialectal diglossia was summarized eloquently by Paul in his *Prinzipien der Sprachgeschichte* ('Principles of Language History'; 1920:411–412). Paul describes a complex social dialect continuum, characterized by various processes of style-shifting and mixing:

> In jedem Gebiete, für welches eine gemeinsprachliche Norm besteht, zeigen sich die Sprachen der einzelnen Individuen als sehr mannigfache Abstufungen. Zwischen denen, welche der Norm so nahe als möglich kommen, und denen, welche die verschiedenen Mundarten am wenigsten von der Norm inficiert darstellen, gibt es viele Vermittlungen. Dabei verwenden die meisten Individuen zwei, mitunter sogar noch mehr Sprachen, von denen die eine der Norm, die andere der Mundart näher steht … Es kommen natürlich auch Individuen vor, die sich nur einer Sprache bedienen, einerseits solche, die in ihrer natürlichen Sprache der Norm schon so nahe kommen oder zu kommen glauben, dass sie es nicht mehr für nötig halten sich derselben durch künstliche Bemühungen noch weiter zu nähern, andererseits solche, die von den

> Bedürfnissen noch unberührt sind, die zur Schöpfung und Anwendung der Gemeinsprache geführt haben.
>
> ['In every area for which there exists a common linguistic norm, the languages of individuals show manifold gradations. Between those which approximate the norm as closely as possible, and those which use diverse dialects minimally infected by the norm, there are many intermediates. In this context most individuals use two, often even more languages, of which one is closer to norm, the other closer to the dialect … And of course, there are those individuals who use only one language; on the one hand, those who in their natural language approximate or believe to approximate the norm so closely that they don't consider it necessary to approach it further by artificial efforts; on the other hand, those who are still unaffected by the needs which led to the creation and use of a common language.']

The presence and use of acrolectal Cape Dutch varieties appears to have been common enough to make the idea of promoting a simplified, local form of Dutch as the future language of the colony a viable project.[18] The newspaper *Ons Land* ('Our country'; founded by Jan Hendrik Hofmeyr in 1892, united with the *Zuid- Afrikaan* in 1893), the church and several South African writers and intellectuals (e.g. Mansvelt, Brill and Viljoen) advocated the use of a simplified form of Dutch (cf. De Villiers 1934: 161). Frequently cited in this context is the position of the writer Melt Brink (1842–1915), an autodidact with little formal schooling, who wrote some of his work in what he described as *'n tussentaalvorm* ('an in-between language form'); that is, a variety located somewhere between Dutch and Afrikaans and, according to Brink, a more appropriate reflection of the spoken language than the highly uniform and stylized GRA norm (cf. also Te Winkel 1896: 432; Meyer-Benfey 1904: 232; D'Arbez (J. Van Oordt) in *Die Brandwag* 1/10/1914).

> Ik heb mij, wat de taal betreft, zooveel mij zulks mogelijk was, aan de middenweg tussen goed Hollands en Patriots gehoude. De eerste omdat ik denk dat deze voor velen gemakkelijker zal lezen, en ten tweede omdat ik het Patriots te overdreven acht, wij spreken zo niet onder ons. Ik heb het Kaaps-Hollands, zoals wij het gewoonlik spreken, gevolgd. (quoted in Kannemeyer 1984: 71)
>
> ['I have kept, with respect to language, as far as it was possible for me, to the middle way between good Dutch and the Patriot language. Firstly, because I think that this is much easier for many to read, and secondly, because I consider the Patriot language as exaggerated, we don't speak like this with one another. I have followed Cape Dutch as we ordinarily speak it.']

Intermediate linguistic forms also appeared in the periodical press. Interesting in this respect is not only *Ons Land*, but also *The Friend of the Free State* (cf. Raidt 1994 [1992]) and the newspaper of the *African Political Organization* (cf. Adhikari 1996), which published articles in what Ponelis (1996: 134) has termed 'approximate Afrikaans'. In the Muslim Cape Dutch Vernacular literature the use of intermediate forms (called *gemixte taal*, 'mixed language', cf. Davids 1991) was common in publications written in Roman script, while Arabic Afrikaans texts (i.e. texts written in the Arabic script) conformed from the 1860s rather closely to the model of Afrikaans as known from the dialect writing tradition (Ponelis 1981; Davids 1991: 96ff.).

The foundation of a society for the promotion of Dutch was first suggested by Jan Hendrik Hofmeyr in 1877, but nothing came of the proposed organization for Dutch until the 1880s when the aggressive nature of British imperialism ensured the political and cultural organization of Afrikaner nationalists in the *Afrikaner Bond*.

Dutch had been excluded from official functions and the state-sponsored school system since 1822 when a government proclamation granted official status to English only (see Chapter 8). In 1881 a petition signed by over 6000 individuals was handed to the government to demand equal rights for Dutch. In 1882 Dutch was re-instated in Parliament and the school system, in 1884 in the Courts of Justice, and in 1888 Dutch was made an obligatory subject for civil servants (De Villiers 1934: 109–120). The language variety promoted from the mid-1890s by the supporters of Dutch (who founded the *Zuid-Afrikaansche Taalbond*, 'South African Language Union', in 1890) was not metropolitan Dutch, but a simplified Cape Dutch (*eenvoudig Kaapsch Hollandsch*), which differed from Afrikaans as promoted by the GRA. *Eenvoudig Kaapsch Hollandsch* was characterized by local vocabulary items, loss of case inflection in the article and increasing loss of the neuter gender, the conjugation of verbs followed largely the paradigm of weak verbs and a partially different pronominal system was used (De Vos 1891: 6; ZAT 1892: 162–165; Engelenburg 1897: 360; cf. also Mansvelt in *The Cape Times* 4/11/1890).

The first congress for the simplification of Dutch took place in 1897 in Cape Town. Participants came not only from the Cape itself but also from the Orange Free State and the South African Republic/Transvaal. The discussion of simplification was influenced by the language reform movement in the Netherlands, where Kollewijn and his supporters had put the question of orthographical and grammatical simplification on the agenda (cf. Kollewijn 1903).

A central problem of the *Taalbond* and its efforts to establish a simplified South African Dutch as the language of education and administration, was that

codification was slow since its supporters did not want to act without backing from the *Vereenvoudigers* ('Simplifiers') in the Netherlands (*Verslag van Het Eerste Congres* 1897: 11, 17; Engelenburg 1897: 362f.; De Villiers 1934: 171; cf. also the commentary in the AP 26/12/1895, 6/2/1896). When the simplification proposals were finally accepted they were relatively modest (focusing largely on issues of spelling), and normative publications in the newly established standard did not appear before 1907. As a result of these delays, the variety propagated by the *Taalbond* remained diffuse and there was considerable insecurity regarding accepted usage; an untenable situation at a time when the challenge was quickly to gain control of the High culture domains (including the education system). The lack of a clear linguistic norm for Cape Dutch was exploited by those promoting Afrikaans as well as by those who favoured English, and Cape Dutch was dubbed a language with 'gen kop of stert' ('neither head nor tail', AP 24/12/1891; cf. also AP 9/10/1890 and *Cape Times* 17/11/1890). The general insecurity about the linguistic identity and status of Cape Dutch is aptly expressed in a poem published in the *Zuid-Afrikaansche Tijdschrift* in February 1890: Cape Dutch as a linguistic system can be identified only by stating what it is *not* (i.e. neither Dutch nor Afrikaans), no positive definition was available:

Ek zou nou graag wille weet,	I would quite like to know,
Wat ver 'n taal ons hier praat,	What language we speak here,
Want Afrikaans det is 't niet	Because this is not Afrikaans
Wat ek daarvan verstaat.	What I understand of it.
Ook is dit nie die paterjot	It is also not the language of the Patriot
Ni Boesman, ni Koraan	Neither Bushman nor Korana
Ni Portegeis, ni Makatees,	Neither Portuguese nor Makatees
Ni Kaffir, ni Javaan.	Neither Kaffir nor Javanese.
Di mense di det hollans noem,	The people who call it Dutch,
Is glad en al verkeerd,	Are absolutely wrong,
Geen hollander die praat en skrijf	No Dutchman talks and writes
As ons det word geleerd.	The way we are taught.

(*Zuid Afrikaanse Tijdschrift*, February 1890: 17)

The linguistic marketplace and its entrepreneurs

The metaphor of the linguistic market(place) is used in sociolinguistics to describe differences in the evaluation and usage of varieties (Sankoff & Laberge 1978; Woolard 1985; Gal 1989; Coulmas 1992; Eckert 2000). According to Bourdieu (1991), linguistic utterances (and, in particular, knowledge of the legitimate language or standard variety) are forms of symbolic capital which allow access to privileges and opportunities in society: languages (and language varieties) are always located on particular linguistic markets, and these markets accord them a certain 'value'.

Linguistic markets resemble economic markets in that their structure is under no one's deliberate control, but is the cumulative outcome of the activities of its participants. Deliberate intervention is possible, yet the results of such interventions are predictable only within limits (on the structure of markets in general cf. Schelling 1972: 17). The structures of social and political power and authority in a society shape the interactions on the linguistic market; that is, the value of language varieties and variants depends primarily on the power and authority of the groups they index (Bourdieu 1991: 67). The market value of an utterance can be measured most easily by its effects (whether the speaker is heard, attended to, imitated or, alternatively, ignored and pushed aside; cf. Bourdieu 1991: 55). The unification of the linguistic market is, however, never complete and in alternative markets different hierarchies exist (Woolard 1985: 743–746; Eckert 2000: 13–14, 25). Linguistic variants or varieties which are described as commanding covert prestige can be conceptualized as being exchanged on alternative markets, that is, on markets where the law of price formation is linked to solidarity rather than to educational privilege, status and socioeconomic class membership.

Using the linguistic market metaphor one might call those who try to manipulate the structural organization of the market 'language entrepreneurs', a term introduced by Rubin (1977) in her discussion of language change and language planning (Weinstein, 1979, 1982, used the more politically inspired term 'language strategists'). While at any time the linguistic repertoire of a speech community comprises a variety of ways of speaking (which are not necessarily sorted into well-defined systems of codes and language-related identities), language entrepreneurs shape the heteroglossia they encounter into 'voices' that symbolize community identity (Bakhtin 1981: 272, 293–294, 356–357).

The members of both the GRA and the *Taalbond* can be described as language entrepreneurs. Both language societies tried to change the structures of the linguistic market by challenging the historical legitimacy of metropolitan

Dutch, and by trying to establish a new national standard language (simplified *Kaapsch Hollandsch* and Afrikaans respectively) as a marker of (white) Afrikaner ethnicity. Linguistic nationalism, despite its rhetorical attention to *the people*, has always targeted first and foremost the formal linguistic market, aiming to win official recognition (and thus legitimacy and overt prestige) for the new national language in High culture domains. Gellner's (1983: 35–38, 57) argument that nineteenth century nationalism announced 'the age of a universal high culture' is of relevance to the study of language standardization and language promotion. Gellner draws attention not only to the High culture orientation of nationalist movements, but also to the stylizations and transformations of folk culture which were part and parcel of the process:

> [N]ationalism is, essentially, the general imposition of a high culture on society, where previously low cultures had taken up the lives of the majority, and in some cases of the totality of the population. It means the generalized diffusion of a school-mediated, academy-supervised idiom, codified for the requirements of reasonably precise bureaucratic and technological communication … If the nationalism prospers it eliminates the alien high culture, but it does not replace it by the old local low culture; it revives, or invents, a local high (literate, specialist-transmitted) culture of its own, though admittedly one which will have some links with the earlier local folk styles and dialects. But it was the great ladies at the Budapest Opera who really went to town in peasant dresses, or dresses claimed to be such. (ibid.)

From this perspective the conflict between the GRA and the *Taalbond* can be interpreted as a struggle for cultural hegemony and symbolic power, a struggle for the right to determine what constitutes the legitimate and authoritative language of the Afrikaner nation. Holliday (1993: 26–27) has interpreted the competition between the GRA and the *Taalbond* in terms of a conflict between traditional élite (*Taalbond,* supported primarily by the grand bourgeoisie, wealthy farmers, the church and *Gelehrten*, 'scholars/academics'; Viljoen 1896: 23), and counter-élite (GRA, supported primarily by the emerging professional class and the petty bourgeoisie). Counter-élites challenge the established power basis of the traditional élite (in this case the symbolic capital of standard Dutch) and represent oppositional (cultural and linguistic) practices and discourses (cf. Joseph 1987: 51–56). However, despite differences in constituency and goals, personal continuities existed between the GRA and the *Taalbond*. Supporters of Dutch were present at the language congresses for Afrikaans, and members of the GRA actively supported efforts to re-establish Dutch in official functions in the colony, for example, C. P. Hoogenhout, one of the founder members

of the GRA, was vice-chair of the *Taalbond* in Wellington, other GRA members involved in the work of the *Taalbond* were D. F. Du Toit and S. J. du Toit (De Villiers 1934: 134, 151).[19] Furthermore, the GRA never positioned itself openly in opposition to the efforts to reinstate Dutch (or varieties close to Dutch), but supported (in parallel to the political unification in the *Bond*) a broad opposition to English ('Net maar teeno'er Engels moet ons eensgesind bly en ni onder mekaar stry', 'Only against English must we remain unified and not fight among each other', AP 11/5/1877; also AP 5/3/1891, 1/7/1881, 6/5/1887). Halliday's interpretation of a straightforward élite conflict thus simplifies a complex reality of ambiguous and conflicting loyalties and interests which characterized language-political activities at the Cape up until the early twentieth century.

Summary: The language question at the Cape

The second half of the nineteenth century saw what historians call the *questione della lingua* (i.e. a complex combination of issues about language and power) rise to prominence in the popular and political discourse at the Cape. Questions of code identification and differentiation (i.e. what counts as a language and what are its boundaries) featured prominently in the metalinguistic commentaries. Contemporary folk taxonomies suggest that the linguistic situation was often perceived as relatively focused, although there is some indication that the actual patterns of variation were probably more fluid. For most language activists (as well as writers of dialect texts) it was a question of either Dutch or Afrikaans, while the ambiguity of possible in-betweens was rejected, ridiculed and branded as non-authentic. The mental make-up of the time resembles what Zerubavel (1991) has termed the 'rigid mind', a mind which shuns the greys, the twilight, and which perceives the world (in Aristotelian fashion) as a composite of clear-cut entities:

> The most distinctive characteristic of the rigid mind is its unyielding commitment to the mutual exclusivity of mental entities. The foremost logical prerequisite of rigid classification is that a mental item belongs to no more than one category. (Zerubavel 1991: 34)

Nationalist movements *per se* are exemplifications of the rigid mind, as the setting of boundaries, the unambiguous definition of Us and Them is vital to their cultural and political agenda. The symbolic association between language and group is constructed through discourses which firmly establish the belief that

groups of people and forms of speech exist as distinct and natural entities; yet the identification of such entities and the marking of the boundaries is generally obscured in the nationalist discourse which propagates the naturalness of the distinctions made. Sociolinguists investigating linguistic variation and change under such historical conditions need to be aware that conscious manipulation of linguistic resources is pervasive during such times, and that folk taxonomies, rather than describing the realities of language use, must be understood as part of a discourse reflecting the struggle for ideological domination (Foucault 1984; Blommaert 1999).

Notes

1. Translation by Steinberg (1987:206).

2. The anecdotes narrated in the Hendrik Kok dialogues are based on an incident mentioned by John Philip in his book *Researches in South Africa* (1828). Philip reported that the magistrate W. M. Mackey from Somerset (now Somerset East) punished a Khoe, who had emptied a small cask of brandy belonging to the magistrate, not only with imprisonment, but also kept him and his family for further three years under conditions of forced labour (for the historical background cf. Nienaber 1971:43–44). The *Zuid-Afrikaan* was known for its opposition to what at the time was perceived as the 'negrophile' position of the British humanitarians (Bosman 1930:3). Boniface's satirical reworking of the incident clearly reflects the editorial bias of the paper.

3. Effendi, a Turkish religious scholar, tried to introduce Hanafite teachings to the Cape, which so far had followed the Shafi'ite school.

4. The Moravians on St. Thomas (Virgin Islands) had openly encouraged the use of Dutch Creole (*Negerhollands*) for religious purposes (cf. Stein 1989, 1995), and were also actively engaged in the compilation of dictionaries and grammatical descriptions. No indication of similar codification activities was evident in the work of the Moravian church at the Cape.

5. By the second half of the nineteenth century the model of the independent nation state based on cultural and linguistic unity was common place in European intellectual circles – a political/cultural model ready for 'pirating' (Anderson 1991:81). It can be assumed that young Afrikaners who went to study in Europe in the nineteenth century were introduced to these discussions.

6. The original motivation for the foundation of the GRA is usually attributed to Arnoldus Pannevis' suggestion to translate the Bible into what he termed *Afrikaans Hollands*. However, under the leadership of S. J. Du Toit a clearly nationalist agenda and discourse soon came to dominate, and without much discussion the GRA postponed the issue of the Bible translation after initial attempts at convincing the *Britse en Buitenlandse Bybelgenootschap* ('British and Foreign Bible Society') to publish a translation had failed.

7. Letters to the editor of *Die Patriot* indicate that the paper was not only read in the Cape Colony, but also in the northern republics. The paper had agents in the Free State and the Transvaal from 1876 (cf. also Du Toit 1909 [1880]: 77).

8. E.g. AP 2/2/1877, 16/2/1877, 3/8/1877, 7/12/1877. This section was continued irregularly during the 1880s. Cf. Bester (1989) on early Afrikaans children's literature.

9. Cf. Woolard and Schieffelin (1994: 57–64), Woolard (1998), Gal (1995), Davis (1996), and Preston's (1989) work on perceptual dialectology; a research programme for the study of folk linguistics was already outlined by Hoenigswald (1966; for a recent discussion see Niedzielski & Preston 2000), and is central to the ethnography of speaking (Hymes 1974).

10. Commentaries on geographical differences are rare during the nineteenth and early twentieth centuries. Regional differences are discussed in some detail by Von Wielligh (1925a).

11. The need to dissociate Afrikaans from the language variety used by people of colour is still evident in early twentieth century texts (e.g. Van Wielligh 1925a: 96). See also Francken's historical play *Susanna Reyniers* (1908) in which the language of the slave character *Platneus* is described as a *armoedig dialect* 'poor dialect', while that of the commander is characterized as *beskaafd Afrikaans* (for a discussion of the play cf. Coetzee 1996).

12. On language loyalty as being dependent on the availability of a clearly defined code and linguistic norms, see Weinreich (1968: 99–100), also Garvin (1964) and Hill (1993). On the interconnectedness of linguistic purism, social movements and social change see Jernudd (1989), Annamalai (1989). Wexler (1971: 343) predicts linguistic purism in cases where the emerging standard is closely related to the old standard norm: 'In such a situation, we can expect widespread puristic efforts directed against any influences from the rival written norm'.

13. E.g. Ponelis (1993: 50), Raidt (1991: 231–232, 1994 [1984]: 187–188), Ueckerman (1987: 51), Uys (1983: 136–139), Shaffer (1978: 57), Van Wyk (1978: 47), also Puddu (1996, I: 347).

14. Much of the discussion surrounding the question of what qualifies as diglossia is related to the fact that Ferguson's original paper was deliberately tentative and exploratory, as explicitly acknowledged in the closing remarks: 'Perhaps the collection of data and more profound study will drastically modify the impressionistic remarks of this paper, but if this is so the paper will have had the virtue of stimulating investigation and thought' (1959: 340). Nevertheless, the notion of diglossia is central to sociolinguistic theory and Hudson's (2002) recent review affirms its status as an analytical and explanatory concept.

15. Although such metalinguistic commentaries are valuable sources for language history, they always also reflect assumptions and intellectual preoccupations of the popular and scientific *Zeitgeist*. From the mid-nineteenth century linguists had begun to emphasize the primacy of the spoken language and the philologists' traditional preoccupation with written sources (now categorized as 'artificial') came under attack (cf. Noordegraaf 1985: 339–343, 412; Steinberg 1987: 201–202; Joseph 1989: 252–253; for Dutch cf. Roorda 1856: 97ff.; Verdam 1890: 57; Te Winkel 1904: 279). In the Netherlands as elsewhere in Europe, the perception of an opposition between spoken and written language was soon identified as a social and educational problem which limited children's success at school and kept adult literacy levels down (Moledijk 1992: 170, 173). However, although linguists and education-

alists argued emphatically that the Dutch spelling rules obstructed the acquisition of literacy, illiteracy was at an all time low (Moledijk 1992:170). It is a truism of social research that what is perceived as a social problem (i.e. a problem which is believed to be harmful to society as a whole) does not necessarily correspond to objectively harmful realities. Relatively benign conditions are frequently catapulted into the floodlight of social and political attention, while extremely harmful conditions are often ignored (Goode & Nachman 1994).

16. Locally produced primers (such as De Villier's *Fondamentssteenen* 1884–1894) commented on pronunciation features (such as /sk/ for /sx/ and apocope of /t/) and noted lexical peculiarities of the spoken language at the Cape. However, they rarely remarked on divergent grammatical and syntactic features.

17. See, for example, the evidence given by Prof. Van der Turk to the education commission (*First Report* 1892:102), De Wet (1876 [quoted in Du Toit 1909[1880]:102–103), Besselaar (1914:193–194, 219–220), Van Rijn (1914:13), the testimony of J. J. Willemsen (quoted in Nienaber 1950:34–35), and the letter by W. to *De Goede Hoop* (November, 1906). The Cape-born student C. H. Persoon who went to study in the Netherlands, reported (in 1775) in a letter to his parents that frequent comments were made about his *Caaps krom spreeken* (lit. 'Cape corrupted speaking', that is, the 'corrupted' Cape Dutch dialect; Raidt 1983:40). This can be interpreted as indicating the existence of a marked extraterritorial variety, while the linguistic differences (although sufficient for ridicule) seem to have been small enough not to lead to communication problems.

18. It has been suggested that the relatively strong position of varieties close to Dutch in the nineteenth century was stimulated through increased immigration after 1850 (cf. Ponelis 1993:47). However, although Dutch immigrants played an important role in nineteenth century language-political debates and activities (e.g. Pannevis, Hoogenhout and Mansvelt were all born in the Netherlands), their numbers were on the whole marginal. According to the Census of 1891 (p. 80) only 0.23% of persons born outside of the colony came from the Netherlands. The total percentage of European-born persons was 13.2% – the vast majority of these came from the British Isles. Raidt (1991:241) attributes the strong position of Dutch not so much to immigration patterns, but to the fact that with increasing attention to educational matters from the 1830s, the question of language standards (in this case the standard of metropolitan Dutch) became an important issue.

19. Such seemingly contradictory linguistic loyalties were described in detail by Fishman (1987) in his biography of Nathan Birnbaum (who supported first Hebrew, then Yiddish and finally no particular language as a marker of Jewish identity), and based on his analysis Fishman urged researchers 'to question the usual functional analysis of élitist language ideology in terms of self-serving or class-related bias' (1987:129).

Chapter 3

The *Corpus of Cape Dutch Correspondence* and the social context of language use in the nineteenth century

> The population of the Western Province is partly English and partly Boer or Dutch-Huguenot, the descendants of the Dutch East Indian Company's servants and settlers … The labouring classes are, as elsewhere in South Africa, coloured, and here largely half-casts, the descendants of the first Dutch residents and their slaves, or much more rarely of blended Dutch and Hottentot blood. In Cape Town itself are found also Malays, Chinamen, Hindus, and the representatives of all European nations.
>
> Olive Schreiner (1896). *Thoughts on South Africa*

The Cape Dutch speech community: Core and periphery

The speech community is an analytical category which, despite its conceptual fuzziness, is central to sociolinguistic research. While some have approached its definition from a strictly linguistic point of view (e.g. Lyons 1970: 326: 'all people who use a given language (or dialect)'), others have defined the speech community in terms of regional and cultural identity (Frings 1957; Hymes 1974), interaction patterns and shared (socio-)linguistic knowledge (Bloomfield 1933; Gumperz 1962/1968; Gumperz & Levinson 1996), linguistic and attitudinal commonalities (Labov 1991 [1972]), or have taken a radical subjectivist view and dismissed the concept altogether (Hudson 1996).

In multilingual and multi-dialectal societies, where notions such as joint sociolinguistic knowledge and shared patterns of speech performance cannot be taken for granted, questions of demarcation and definition are particularly salient. Varying degrees of proficiency and knowledge, as well as the existence of diffuse and overlapping norms make it difficult to identify the borders of the speech community. Three case studies (from Canada, East Sutherland and Mexico) will serve to illustrate some of the problems involved in defining the criteria for speech community membership in these contexts.

(i) ***Canada*** Mougeon and Beniak's (1995) study of the French speech community in Ontario shows that bilingual speech communities can include marginal members whose language use, although grammatically adequate, is functionally restricted. In this case the marginal members of the speech community use French predominantly, if not exclusively, within the formal setting of the French medium school and only infrequently in informal conversations. Variation between the standard preposition *de* and the non-standard preposition *à* (*l'auto de mon père* vs. *le char à mon père*, 'the car of my father'), for example, which is socially stratified for core and semi-core speakers (i.e. speakers who use French also in informal situations), is absent from the speech of marginal speakers who only make use of the standard variant *de*. Mougeon and Beniak have termed the non-acquisition of socially significant variation by marginal members 'sociolectal reduction' (ibid.: 92). Apart from 'sociolectal reduction' the presence of marginal speakers in the community can also lead to the emergence of new cases of variation (typically grammatically simpler variants; ibid.: 83–85; cf. also Thomason & Kaufman 1988: 51).

(ii) ***East Sutherland*** Labov (1991 [1972]) circumvents the problem of multiple performance norms by explicitly allowing for variable patterns of production ('heterogeneity of production'), which are mediated by collectively uniform evaluative norms or language attitudes ('homogeneity of interpretation'). However, Dorian's (1982) work on the Gaelic-speaking community in East Sutherland has shown that some speech communities show heterogeneity of production as well as of interpretation. In East Sutherland marginal individuals such as 'low proficiency "semi-speakers" and near-passive bilinguals in Gaelic and English' (ibid.: 26) lack the linguistic knowledge of important linguistic variants and do not participate in the patterns of social evaluation shared by core speakers. They are nevertheless seen by others and themselves as integral members of the Gaelic speech community.

(iii) ***Mexico*** Further evidence for heterogeneity of interpretation or evaluation within a speech community comes from Santa Ana and Parodi's (1998) study of linguistic variation in non-metropolitan Mexico. When conducting their interviews Santa Ana and Parodi (1998: 26) noticed that 'some individuals, in contrast to the majority of interviewees, apparently demonstrated no awareness of any of the social evaluation patterns of language variation'. Santa Ana and Parodi (1998: 38–39) summarize the linguistic behaviour and social position of these marginal members as follows:

> [They] move in relatively restricted and close-knit social networks … limited to a handful of extended families and to very local interactions. This

> social network and its isolation from the linguistic hierarchy do not seem to be a single-generation accident; this has been the prevailing social structure for generations … These speakers exhibit little conscious or unconscious recognition of differential use of language by other speakers, and they minimally modify their speech to accommodate their interlocutor … They may acknowledge some variation in speech; however, they are indifferent to the social judgements that are linked to such variation by other speakers … Contacts with the socializing and evaluative social institutions that strongly affect people's sense of self and their speech, such as schooling and work outside the home, have been tenuous and brief. Other contacts with the larger social world are superficial, such as the passive reception of radio and television programs. … these individuals speak the regional dialect. They use non-standard lexical items and phonological forms, as well as stigmatized items of non-taboo content; but they do not demonstrate mindfulness of the reactions of other speakers to their language use.

The possibility that these differences in evaluation could be explained by assuming the existence of two distinct speech communities was rejected for Mexico: no separate linguistic feature was found which identified these close-knit, localized networks as a separate speech community (1998: 26–27). Santa Ana and Parodi thus propose 'a model of nested speech community configurations' (1998: 34): those individuals who did not share the evaluative norms of the majority of the sample were only partially integrated into the large-scale, national speech community, but were nevertheless part of it.[1]

Extending social networks to include wider regional (as opposed to local) ties leads to a second group in the Mexican speech community. Members of this second group recognize stigmatized variants, 'are sensitive about the way they speak with outsiders' (1998: 39–40), and are aware of the normative linguistic hierarchy. The third extension of the speech community includes a large set of mutually anonymous speakers exhibiting significant class and educational differences. Members of this group recognize linguistic stereotypes and markers and are able to evaluate speakers on the basis of their language use. Finally, the largest group is the 'national speech community' (Santa Ana and Parodi termed the sub-groups 'speech locale', 'speech vicinity' and 'speech district'). Individuals who participate in this group are aware of the social prestige attached to the standard variety and usually aim at producing standard forms in their own speech.

The four groups can be ordered on an implicational scale. As the individual speakers come to recognize and evaluate types of variables (stigmatized forms,

Table 3.1 A speech community typology (reflecting the degree of recognition of sociolinguistic norms; Santa Ana & Parodi 1998: 35)

	stigmatized forms (stereotypes)	**regional forms (markers)**	**standard forms**
I locale	–	–	–
II vicinity	+	–	–
III district	+	+	–
IV national speech community	+	+	+

regional forms, standard forms), they will obtain 'membership in a widening sequence of speech community configurations' (1998: 41).

Gradient membership and the existence of conflicting sociolinguistic norms also characterized the Cape Dutch speech community. The complexity of the speech community was noted, for example, by Van Ginneken (1928: 213ff.) who distinguished for the early twentieth century several different language groups and varieties:

i. Western Cape Afrikaans
ii. North-Eastern Afrikaans
iii. Khoe Afrikaans
iv. Griqua Afrikaans
v. Malay Afrikaans
vi. *Masbieker-Afrikaans* (spoken by the descendants of slaves from the African east coast, i.e. Mozambique and Madagascar; in the nineteenth century these were regarded as a distinctive group, cf. McCormick 2002: 28)
vii. *Kaffer-Afrikaans*

Processes of language learning and language shift (see Chapter 1) continued to form a central aspect of the sociolinguistics of the Cape Dutch speech community in the nineteenth century. The term *Kaffer-Afrikaans* describes second language varieties used by speakers of South African Bantu languages. Some of these speakers were, according to Van Ginneken (ibid.: 233), in a process of language shift. For this group, proficiency was largely a function of exposure:

> Velen echter hebben hun eigen taal verleerd en spreken gebroken Afrikaansch. Het Afrikaansch der Kaffers … verschilt … naar het aantal jaren dat zij in dagelijksch verkeer met Afrikaansch-sprekende kleurlingen of blanken hebben doorgebracht.'
>
> ['Many have indeed lost their own language and speak broken Afrikaans. The Afrikaans of the Bantu-speakers … differs … depending on the number of

> years that they have been in daily contact with Afrikaans-speaking coloured or white people.']

Social fragmentation, lack of cohesion and the existence of schisms between cultural and ethnic groups were also emphasized by Olive Schreiner in her essay collection *Thoughts on South Africa* (1896/1923):

> There are in the Colony, roughly speaking, a million-and-a-half of men. One million of these are natives, Hottentots, and half-castes, but mainly Bantus of the Chuana [Tswana] or Kaffir races; the remaining half million are divided between men of English and other European descent speaking English and the men of Boer descent, often speaking the 'Taal'. Now not only are these peoples who form our population not united to each other by race, language, creed or custom but, and this is a far more important fact, each division forming our population is far more closely connected by all these ties to masses of humans beyond our border than to their fellow Cape Colonists within. … Race, language, creed and tradition, which in the true national state form centripetal forces, binding its parts to one centre, in such a state become centrifugal forces, driving them from it; and political boundaries are so crossed and recrossed by these lines of union that they are rendered void.
>
> (Schreiner 1992 [1896]: 52–53)

Schreiner's comments bring to mind Bakhtin's (1981) thinking on language history (cf. Crowley 2001 for a discussion of Bakhtin's work). Bakhtin's concepts of heteroglossia (i.e. the dialogic and heteroglot reality of speech communities) and monoglossia (i.e. a linguistic scenario which is dominated by a concern for uniformity and purity, basically a standard language scenario) are useful conceptual additions to sociolinguistic theory and language history. Situations of monoglossia and heteroglossia are brought about by the centripetal (unifying) and centrifugal (diversifying) forces noted by Schreiner for the society at the Cape. Language histories, according to Bakhtin, are shaped by the on-going interaction and dialectic of these forces (also Ferguson, 1988: 121, on the notion of a 'standardization cycle', and Jespersen, 1925: 51, on 'unifying forces' in language history; cf. also Chambers 2002).

Although the colonial society was highly heterogeneous with regard to ethnic origin, language use and cultural traditions, from the mid-nineteenth century social and economic change (including rapid urbanization, immigration, industrialization and economic diversification, improved transportation and the rise of a public school system) facilitated inter-group contact and linguistic and societal convergence. A complex trading network now connected the political centre of Cape Town to the growing agricultural areas of the eastern Cape where a number of smaller and bigger towns served as political,

Table 3.2 Urbanization in the Cape Colony (adapted from Bickford-Smith 1995: 11)

	Cape Town	Port Elizabeth	East London	Kimberley	Colony total
1865	27,000	11,300	–	–	470,000
1875	45,000	13,000	2,000	13,600	720,000
1891	79,000	23,000	7,000	29,000	1,500,000
1904	170,000	33,000	24,000	34,000	2,400,000

economic and cultural centres for the surrounding regions. The discovery of diamonds in the 1860s and of gold in the 1880s shaped colonial life in the second half of the century and brought with it new types of prosperity and poverty. Increasing numbers of European immigrants, the import of indentured labourers from Asia and internal migration of labour to the mining areas and to the growing towns (cf. Table 3.2) contributed to the formation of complex and linguistically/culturally diverse urban communities and created 'new conditions of language contact and language learning' in these environments (McCormick 2002: 29).

The Cape Dutch speech community at the Cape can be defined loosely as including all those individuals who made use of a Netherlandic variety in at least some of their interactions. Behind this broad definition lies a complex reality of diverse acquisition trajectories, ranging from monolingual first language or near-native second language to rudimentary second language jargon, and finally to balanced or shifting bilingual language use in some groups. English-Dutch bilingualism was common in the urban areas (see Chapter 8) and Khoe-Dutch bilingualism was a characteristic feature of ethnically mixed communities such as the Griqua or Rehoboth Basters (see Chapter 1). The Khoe languages found alongside varieties of Dutch in these communities were Gri (until the mid-nineteenth century), Korana (until about 1900) and Nama (still spoken in many communities; see Traill 1995: 7–8). Among ex-slaves and their descendants Malay was spoken until the late nineteenth century (especially within the Muslim community; Franken 1953: 116ff.; Valkhoff 1966: 261ff.; cf. also the above mentioned comments on shifting speakers of Bantu language background).

The Cape Dutch speech community resembled the bilingual communities described by Dorian and Mougeon and Beniak in important aspects: speakers had varying degrees of membership in the speech community and differing degrees of exposure to varieties of Cape Dutch Vernacular. This supported the non-acquisition of certain variants by marginal and semi-core speakers, and the development of group-specific variants among more peripheral groups.

The speech community typology developed by Santa Ana and Parodi (1998) is also helpful here as it acknowledges the fact that the norms of the supra-regional (and exogenous) standard might not be actively (or passively) known by a large number of individuals, who are nevertheless full members of the speech community. It also emphasizes the importance of localized, group-specific norms and knowledge, that is, the existence of multiple overt and covert prestige norms and linguistic markers of diverse social identities.

Literacy and writing practices

Not all speech styles used by the members of a speech community are of equal interest for the study of language variation and change. Labovian variation studies have given special attention to the study of 'the vernacular': a specific linguistic style used in informal speech 'in which minimum attention is given to the monitoring of speech' (Labov 1991 [1972]: 208). This relaxed or casual style is believed to offer the most systematic data for sociolinguistic analysis and the study of language change. A major challenge for studies in historical sociolinguistics has therefore been to get as close as possible to the casual, spoken language of the time. Although variation in the written language is worthy of sociolinguistic inquiry in its own right (cf. Romaine 1982: 16; Vachek 1989), the question of whether and to what extent the sociolinguistic patterns observed in the written record can be interpreted as paralleling variation in the casual, spoken language remains a central question for historical sociolinguistics (Lippi-Green 1994: 1–3; see also Kytö & Rissanen 1983: 473–475).

There can be little doubt that people write quite differently from the way they speak and that the process of writing supports the 'formalization' of language use (Kytö & Rissanen 1983). From a broadly psycholinguistic perspective Maas (1988: 16) has argued that the 'strict linearity of speech forces all structuring activity on the short-term memory … [the] short-term memory is relieved during writing'. In other words, when writing speakers have access to linguistic forms which are stored in long-term memory, and writing might therefore facilitate the retrieval of lesser used material or structures (including the strategic use of prestige or standard variants which are not normally used in the spoken language). It has been argued (e.g. by Havelock, Goody, Ong) that writing as such supports the objectification of language and abstract thought, and that written language (literacy) is thus qualitatively different from spoken language (orality). Gee (1994: 172) summarized the central idea underlying this position as follows:

> What had been written could be seen as an object (a 'text') and not just heard or felt. You cannot stop and review what you are listening to, especially if you are 'caught up' in its rhythm, but writing allows one to take a second look and thereby to notice contradictions and inconsistencies.

However, the cognitive, social and linguistic effects of literacy are not uniform across societies and research since the early 1980s has shown that differences between spoken and written language may have been overstated in the past. Besnier's (1988) and Biber's (1988) comparative studies of spoken and written language genres failed to establish a clear boundary between the two channels and found patterns of linguistic variation to be largely independent of medium. Drawing on a wide range of studies of literacy practices, Roberts and Streets (1997: 168) have argued that the assumption of a 'great divide' between writing and speech 'is increasingly less convincing once contextual and actual social uses of the channels are taken more fully into account'. In other words, literacy is not an autonomous, uniform technology which stands in a dichotomous relationship to orality, but a social practice which is shaped by the political, social and historical context in which it occurs. To adequately account for the differences between spoken and written language therefore requires an ethnographic and sociohistorical perspective on literacy practices (see Basso's, 1989 [1974]: 432, outline of a research programme for an ethnography of writing; cf. also Gee 1994).

Protestant religious instruction provided basic literacy at the Cape from the early European settlement, and the ability to sign one's name, to read the Bible, and to recite the articles of faith was a prerequisite for church membership. Religious literacy was taught primarily by recitation and emphasized memorization; this resulted in the mastery of only a very limited set of devotional texts and oral repetition from memory frequently masqueraded as reading. Reading itself was rarely a silent, solitary experience, but a social activity with a strong oral component: the daily portion of the Bible, newspapers and other texts was usually read aloud in the colonial household (Figure 3.1).

Book ownership was rare in the colony and the conventional view that the Bible was available in all households and had a conservative influence on the language of the colonists (a position originally voiced by Hesseling 1923: 59–60, 128; see also Maeder 1919) is not supported by the historical evidence. For the eighteenth century, Biewenga's (1996) research has shown that possession of books (including Bibles and other devotional texts) was limited (see Chapter 1); a situation which hadn't changed much a century later – at least not in the rural *hinterland* where books were still a 'treasure' rather than a com-

Figure 3.1 Oral reading practice in the Cape Colony: an itinerant teacher reads items from the *Zuid-Afrikaan* (drawing by C. Bells circa 1840; reproduced in Pells 1938)

modity. Schoeman (1995: 16) quotes the following anecdote reported by John Campell of the London Missionary Society in 1819 after a visit to Swellendam:

> The two men who lived there were father & son – that his name was Maretts … We gave him some tracts which he received with great pleasure – observing this we asked if he had a bible, he feelingly answered no, and that when he was last at the Cape about two years ago he was about purchasing a bible, but he found he had not as many dollars with him as would pay for it. Mr. Moffat immediately brought his octavo bible from which he preached and presented it to him. He hugged the bible in his arms with the apparent affection as if it has been his own child. I mentioned that it would show him the way to heaven, which he said he believed it would do. Nothing which I witnessed on the journey delighted me more than the manner in which the father and son acted towards this copy of the scriptures. After our evening worship was over they carried the treasure in triumph to their home.

In the first half of the nineteenth century a public school system was established by the British colonial authorities, and with the inception of the position of the Superintendent General of Education in 1839 a centralized authority in educational matters was created (Malherbe 1925: 95–97; Pells 1938: 31–32). Although legislation was not yet in place, the school system was practically racially segregated: public non-denominational schools, which taught a wide range of subjects were frequented almost exclusively by white children, while the vast majority of black and coloured children (as well as a fair number of so-called 'poor white' children) received their education from mission schools, where instruction focused almost exclusively on the three Rs of reading, writ-

ing and arithmetic (Bickford-Smith 1995: 25–29). According to a report of the Education Commission, daily attendance of those enrolled at school was highly irregular (Education Commission, *Third and Final Report* 1892: 4–6), and the so-called Ross Report – which was compiled by Donald Ross who had been sent out from Scotland to report on the educational system of the Cape Colony – described the low quality of instruction across school types (1883: 4):

> In a large, by far the largest number of schools it would be an absolute waste of time to examine the pupils in detail, so very elementary were their attainments and so inferior was the quality of instruction imparted.

The educational system improved significantly after 1892 when Dr. Thomas Muir was appointed Superintendent General of Education. School inspections now took place regularly, educational surveys were conducted, curricula were extended, suitable textbooks were issued by the Department of Education, and the training of teachers became a focal point of the department's activities (Malherbe 1925: 172; Pells 1938: 94–95). Not only was school attendance sporadic and the quality of teaching questionable before the 1890s, but opportunities to practice the literacy skills acquired at school were not many in the colony. The situation was similar to that described by Graff (1979: 310) for nineteenth century Canada:

> There is little reason, in fact, to suspect that the daily culture of a nineteenth century city overemphasized the printed word, or that much literacy was required to learn its ways. Residence, commerce, and industry intermingled more than not; walking was sufficient to find one's way about. The city was a place of sights and sounds more than of print and text, with structures, both obvious and hidden, to be 'read' and explored with all of the senses.

Despite declining costs of production, books remained something of a luxury in the late nineteenth century, with prices varying between two and ten shillings. Cheap books were generally badly bound and printed, and for the same amount of money one could have purchased five pounds of bread, two pounds of meat or half a gallon of wine.[2] There is no evidence that anything comparable to the popular street literature that existed in England and North America was available in the colony (on street literature and its importance for the development of literacy, cf. Neuburg 1973: 206–207). At the Cape, cheap and easily digestible reading matter was found primarily in newspapers and periodicals, however, compared to the sensationalist broadsides (the 'penny dreadfuls' and 'shilling shockers'), the former constituted rather difficult reading matter for the many semi-literates in the colony. Moreover, living conditions were often anything but conducive to the development of reading habits,

and long hours of physical work left little time or inclination to exercise one's literacy skills.

Although literacy remained peripheral for many, reading and writing were nevertheless deeply embedded in Cape society. Laws were no longer orally transmitted but written down, political changes and decisions were disseminated and publicized in writing, and the ability to manipulate pen and paper became increasingly important for individuals to protect and increase their social and political rights. Oral culture and oral practices interacted as elsewhere in the world with the emerging culture of the written word.

The nineteenth and early twentieth centuries are of particular interest for historical sociolinguistics, since an expanding school system enabled more individuals (including women and non-elite members of society) than ever before to acquire at least the basic skills of reading and writing, while at the same time the schooling most children received was too rudimentary to guarantee long-term exposure to standard norms. The description and analysis of the writing practices and language use of what one might call (following Fairman 2000) the 'minimally' or 'intermediary schooled' classes is today an important research area (cf. Gracía-Bermejo Giner & Montgomery 1997; Mihm 1998; Schneider & Montgomery 2001; Elspaß 2002; Vandenbussche 2002). However, despite their strong oral component, the texts cannot be taken as a direct, unmediated representation of speech. They reflect various written language competencies, indicative of the intensity of contact with the school system and written texts in general. However, as holograph manuscripts produced under conditions of limited and variable exposure to standard norms they are the closest the language historian has to something which approaches 'utterance tokens' (Lass forthcoming).

The *Corpus of Cape Dutch Correspondence*

Following Kytö and Rissanen's 'hypothesis of the closeness of informal speech to informal writing' (Kytö & Rissanen 1983:474), the *Corpus of Cape Dutch Correspondence* was designed to include personal documents only (see Appendix). Unlike the Helsinki Corpus, it is not a multi-purpose corpus (cf. Rissanen 1991), and its diachronic coverage is limited to the years 1880 to 1922 (roughly two generations). The vast majority of texts are private letters to family members and friends; however, two private memoirs (M. Heroldt and A. Schabort) and excerpts from one diary (F. S. Malan) were also included. Some of the letters are only personal in the wider sense of the word, i.e. the recipient

occupied a position of institutionalized authority in relation to the writer (typically brethren – minister/missionary); yet from the style and tone of the letter it is clear that the writer knew the recipient personally and that their relationship was trusting and affectionate. The informality of the texts in the corpus can thus be described as a function of a close bond between writer and recipient, and a communication situation characterised by privacy (cf. also Rissanen 1986: 101). The corpus also includes several letters by individuals who achieved considerable influence within Cape colonial and Afrikaner national politics, such as Louis Botha, J. H. De Villiers, Jan De Waal, N. M. Hoogenhout, F. S. Malan and J. C. Smuts.

All original documents are handwritten (ink or pencil) and no typed material was included. The quality of the paper varies from heavy, good-quality paper to cheap, yellowish note-book paper. Many of the writers were clearly inexperienced in the written medium, exhibiting what historians have called the 'illiterate scrawl' (cf. Franken 1953: 179). However, texts were on the whole legible and morphological units (including inflectional endings) identifiable.[3] Spelling conventions were highly variable in the documents. Most common were inconsistencies between *ij* and *y*, variation between *e* and *i, z* and *s,* and final *sch* and *s*. Letter writing, as reflected by the sheer volume of some of the archival collections, was an important activity around 1900. The corpus includes intimate love-letters, affectionate notes between husband and wife, concerned enquiries and advice from parents to children, friendly gossip between siblings and friends, and letters requesting help, support or advice from a friend, parent or benefactor. The tone of most letters is informal and close to spoken language conventions. Most letter writers make use of a rather limited vocabulary, short sentences, a great proportion of first person references, exclamations and a fragmented style with idea units often strung together without connectives (cf. Tannen 1982; Chafe 1982).

The question of what constitutes an adequate sample size for variationist research remains a moot point (cf. Romaine 1980; Milroy & Gordon 2003: 28–30). Sociolinguists have generally worked with judgement samples: Labov's (1966) New York City study was based on eighty-eight speakers, Trudgill's (1974) study of Norwich included sixty speakers, Milroy's (1980) analysis of Belfast vernacular included forty-six speakers, Patrick's (1999) study of Jamaican Creole sixty speakers, and Santa Ana and Parodi's (1998) investigation of Mexican Spanish thirty-five speakers. Sankoff (1980: 51–52) has suggested that a sample of around 150 speakers is adequate for the description of even complex and relatively heterogeneous speech communities, while 'samples of

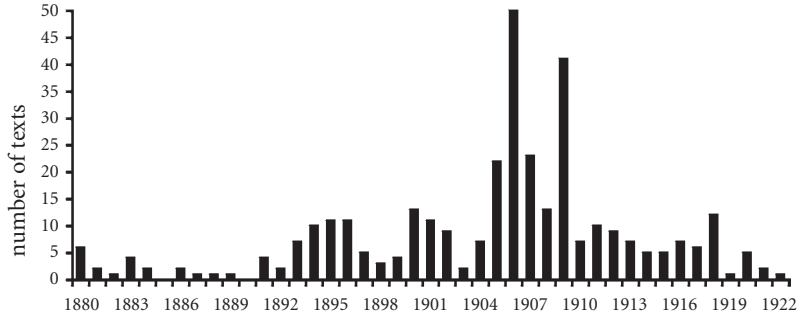


Figure 3.2 Temporal distribution of the sample texts 1880–1922 in the *Corpus of Cape Dutch Correspondence* (number of texts per year)

more than about 150 individuals tend to be redundant, bringing increasing data handling problems with diminishing analytical returns' (ibid.).

The *Corpus of Cape Dutch Correspondence* focuses on language use in the Cape Colony (western and eastern districts) and includes 136 individual writers whose 350 letters comprise approximately 130 000 words. In order to assess the relative size of the sample it can be compared to the Helsinki corpus for Early Modern English which (covering a period of roughly 200 years, 1500–1700) consists of 550 000 words (Nevalainen & Raumolin-Brunberg 1989: 167). Raidt's eighteenth century corpus of Cape Dutch Vernacular texts included 57 writers and text length varied between 30 and 1000 words (1994 [1984]: 186). The boundary dates for the *Corpus of Cape Dutch Correspondence* are 1880 and 1922, with most texts written between 1890 and 1910. As outlined in the previous chapter, these years were characterized by intense public discussions of language issues and increasing attention to Cape Dutch Vernacular/Afrikaans, culminating in the recognition of Afrikaans as medium of instruction in 1914 and as the third official language (alongside Dutch and English) in 1925. Figure 3.2 summarizes the temporal distribution of sample texts in the *Corpus of Cape Dutch Correspondence*.

The social background of each writer was reconstructed on the basis of the available historical sources. Often the writers remarked on their personal circumstances in their letters, explaining, for example, why a debt hadn't been paid or asking for financial support from a friend or family member. In a letter (3/9/1906) to C. P. Hoogenhout, [Mary] Basson described the socioeconomic circumstances of her family as follows:

> wy leven zoo zuinig als wezen kan, van klederen koopen weten wy byna niet meer in de laatste paar jaren, want wy hebben geen inkomen hier en gy weet

> zelf wat voor voedsel uitgaat, ik verkoop niet eens meer een hoender [,] het wordt opgeslacht want wy zyn dikwyls in gebreken met vleesch
>
> ['we live as cheaply as we can, we almost know nothing anymore of clothes buying in the last few years, because we don't have an income here and you know what one pays for food, I don't even sell a chicken any more, it will be slaughtered because often we lack meat'] [4]

Death notices, where available, were an important source of extra-linguistic information. These notices gave the place of birth and ethnic group of the deceased, the place and date of death, age upon death (which allowed the calculation of the year of birth), names of surviving and/or deceased spouse(s), names of children and last residential address. Also listed in the death notices was information on the legacy of the deceased (or this information was filed separately in the estate accounts), giving an indication of their social standing in terms of wealth and property. However, death notices were not filed regularly and many people died without their death being officially recorded.[5] Other sources used for the collection of extralinguistic information were the voter's list for 1903 (indicating residential address, race and occupation, for men only) and the *South African Directories* (1883/4, 1903 and 1911), which list members of the municipal councils, ministers and individuals with business activities for each town or district (for a summary of the sources available for genealogical research in South Africa, cf. Lombard 1984). An interesting characteristic of the corpus is that sub-groups in the data form relatively close-knit networks and it was possible to reconstruct aspects of the network relationships between writers (see Chapter 9).

Mapping the social universe: Age, gender, ethnicity and class

In this study four extralinguistic aspects of variation are considered: age, gender, ethnicity and class.

Age

While the measurement of age as a continuous variable (measured in years) is useful for biological (including psycholinguistic) analysis, for sociolinguistic purposes an anthropological understanding in terms of life stages, which reflects the social consequences of age better than a continuous scale, is usually considered more appropriate (cf. Eckert 1997). Conventionally, four stages

are distinguished: childhood, adolescence, adulthood and old age (cf. Eckert 1997, 2000: 8ff.). The stages of adolescence and adulthood are of particular interest for standardization research as it is during this time that the individual (at least in modern, urban societies) comes into regular and prolonged contact with the public and formal standard varieties of a language and is expected to fulfil a large number of social roles (son/daughter, friend, pupil and student, husband/wife, parent, employee/employer adviser, etc.). This supports the acquisition and development of new forms of behaviour, including an increased awareness of the standard language (cf. Eckert 1997; Eckert 2000: 14–15).

The *Corpus of Cape Dutch Correspondence* includes adolescent and adult writers, i.e. writers who had been exposed to varying degrees (depending primarily on educational background) to the norms of the standard linguistic market. Two generations (or age cohorts) can be distinguished:

i. individuals born before 1865 (age group one; 42% of the sample), and
ii. individuals born in or after 1865 (age group two; 58% of the sample).

The rationale for this division lies in the sociohistorical circumstances: individuals born after 1865 clearly profited from an improving public school system which accompanied the socioeconomic processes of industrialization and urbanization (leading to greater exposure to the norms of both Dutch and English), and were exposed to a fully-fledged language nationalistic agenda in public discourse (cf. Chapter 2). Individuals born before 1865 grew up in a largely agricultural society characterized by considerable geographical fragmentation and limited access to educational institutions. They weren't exposed to the language-political ideas and discussions of Afrikaner nationalism (which gained momentum in the 1880s) until during early adulthood.

Gender

The variable gender is primarily treated as a biological category in variation studies. However, gender is also a cultural category and conceptions of femininity and masculinity are transformed, mediated and defined through social discourse and cultural norms (cf. Eckert & McConnell-Ginet 1992, 2003; Wodak & Benke 1997).

Gender played an important role in language contact and change at the Cape. In Chapter 1 the importance of slave wet-nurses in lower-class colonial households was mentioned, and McCormick (2002: 18) has argued that the divisions of labour

> in the households of slave-owning families were such that it would have been women – of all social groups – who had the greatest intensive exposure to languages other than their own … Women worked in the home or garden and were in constant daily contact with one another as they performed a variety of household tasks necessitating verbal communication, some of it quite extended.

With regard to the history Afrikaans, Raidt (1994, 1995) has argued that women were both leaders of linguistic change and, at the same time, maintained archaic Dutch dialect forms in their written language, and Roberge (2002c: 84) has pointed out that the historical record frequently shows women as 'instrumental in the preservation of Dutch'. The role of women in language change at the Cape thus appears ambiguous: they participated in language change as well as language maintenance.

The statistical information provided by the census reports suggests that at the Cape class and ethnicity interacted in important ways with gender. The 1911 census classified 47% of white women as wives or daughters exclusively occupied with household duties; compared to only 12% of those described as 'other than white'. On the other hand, 43% of black women were classified as holding an occupation in the category 'agricultural', as opposed to only 2% of white women. Furthermore, during the 1880s and 1890s women joined the work force in increasing numbers, thus challenging the moral values of Victorian society. These working women found employment not only in the domestic sphere (as servants), but also entered into independent employment in the newly emerging manufacturing sector (Ward 1991: 19; cf. Ross 1999a: 47 on gender and respectability; cf. also MacMillan 1930: 38 on urban migration of women, and Bickford-Smith 1995: 19, 94–95 on the traditional division of labour between 'men's work' and 'women's work' in the nineteenth century). Based on the census of 1904, the population of the Cape Colony can be estimated at approximately 1,870,000 people (the number is rounded and excludes dependent children younger than 14). 49% of these were female, 51% male (Census 1904: 302–303). The distribution in the *Corpus of Cape Dutch Correspondence* is comparable: 64 of the writers were female and 72 were male.

Ethnicity

Around 1900 about half of the population of the Cape Colony were classified as African, living mostly in the rural eastern territories of the colony. Those classified as white comprised about a quarter of the total population and those classified as coloured about one fifth (see Table 3.3). As noted above,

Table 3.3 Census information (1891, 1904, 1911) for the category 'race'

	African	white	coloured
Census 1891	55%	25%	20%
Census 1904	59%	24%	17%
Census 1911	59%	23%	18%

the Cape Dutch speech community at the Cape comprised coloured, black and white speakers.

The Census administrators' category of coloured is characterized by its diversity, including not only descendants of the heterogeneous slave population, but also the indigenous Khoe as well as anyone who was of mixed (European/African/Asian) descent (cf. Census 1904: xxi). Ross (1982) has argued that the specific nature of Cape slavery – characterized by the import of new slaves (rather than biological reproduction) and the distribution of slaves over isolated farms – inhibited the emergence of a cohesive social identity comparable to that of the American South (see Chapter 1). Instead, different ethnicities continued among the slaves into the nineteenth century. At the same time, however, the collective experience of slavery and subjugation provided the foundations for the formation of a distinct and cohesive coloured identity (cf. Adhikari 1992; Bickford-Smith 1995: 187–189).

In terms of religion the category coloured included both Christians and Muslims (also referred to as Malay). Apart from their distinct and colourful dresses and Islamic religious practices, the development of a separate Malay/Muslim identity was strengthened by the existence of strong occupational (washerwomen, tailors and fishermen) and residential networks (Bickford-Smith 1995: 35). The Muslim community came to prominence in the Cape Town cemetery riots in 1885/86 (which erupted after the city authorities had closed the Muslim cemeteries on Signal Hill), and in the late 1880s members of the Malay/Muslim élite put forward their own parliamentary candidate (Bickford-Smith 1995: 195ff.).[6] However, since Malays/Muslims 'shared similar experiences of racial exclusion and marginality with other proto-coloured groups' (Adhikari 1989: 10), they were ultimately incorporated into a broader political and cultural coloured identity. At the same time the foundation of the *African Political* (later *People's*) *Organization* (APO) in 1902 signalled – at least in its name – the possibility of an even more inclusive identity which embraced not only the various coloured groups but also black South Africans. However, the APO remained strongly oriented towards coloured interests, and in 1910

the then president Abdullah Abdurahman categorically stated that it was 'an organization of the coloured people only' (cited in Adhikari 1996: 9).

In the early twentieth century attempts were also made by mission churches to include Christians classified as coloured as marginal members into the Afrikaner national community. The aim was to prevent the development of an independent and oppositional identity, and to promote the model of Christian respectability and sobriety within the larger society (cf. Ross 1999a: 47). Such a position was advocated, for example, in *De Ebenezer,* a religious monthly paper edited by F. N. Van Niekerk in the early twentieth century (1911 to 1923 available in the South African Library, Cape Town). An inclusive language- and religion-based definition of Afrikaner identity was also supported by the Moravian paper *De Bode van Genadendal* (see Chapter 2). However, in parallel to the folk-linguistic divisions and taxonomies of Cape Dutch speech forms along ethnic lines, Afrikaner identity was increasingly defined in terms of racial purity and European descent (cf. Van den Berghe 1968). From the early 1900s the term Afrikaner came to denote not only those who spoke Afrikaans, but also excluded all those who were not white (the historical semantics of the term 'Afrikaner' are outlined by McCormick 2002: 213).

White ethnicity was no less heterogeneous than other ethnicities. Although public discourse emphasized shared European ancestry and cultural heritage as a uniting factor, both English and Afrikaner ethnicities (defined primarily in terms of language and religion) were on the rise from the mid-nineteenth century and interacted with other dimensions of social structure, such as property ownership, occupation and the growing polarization of urban and rural interests and needs (cf. Bickford-Smith 1995). Although there was political cooperation between Afrikaners and British colonists (as evidenced most clearly in Jan Hofmeyr's support for Cecil Rhodes), Afrikaner support for the Dutch-speaking northern republics (the Orange Free State and the South African Republic/Transvaal) was always a potentially dividing element. The Jameson Raid, a failed coup against the South African Republic (1895), finally led to the break-up of the Hofmeyer-Rhodes coalition. After the raid Afrikaner and British public opinion became increasingly polarized and an ethnically based two-party system emerged at the Cape. The relationship between British imperialists and the Boer republics remained tense and culminated in the South African or Anglo-Boer War (1899 to 1902), which further polarized British imperialist and Afrikaner nationalist interests in the Cape Colony.

The neat racial classifications of the Census reports conceal the fact that by the late nineteenth century the conceptual divisions and borders separating the different ethnic groups were blurring, and that intermarriages were taking

place with increasing frequency – especially among the working classes. This is reflected in the following report from the *South African News* (4/4/1901):

> We have been astonished and from the standpoint of our social prospects disheartened, to find how surprisingly large is the percentage of such settlers who marry coloured women. Let the Imperial Government take a Census on this point in, say, District Six of Cape Town, and the result will astonish them.
> (quoted in Bickford-Smith 1991:27)

Politics of racial segregation were initiated from the 1890s in most urban areas in response to health concerns about over-crowed inner-city slums, and to satisfy the capitalist's need for an available and controllable African labour force.

Since membership in the Cape Dutch speech community was not limited to those classified as white, an attempt was made to collect letters by coloured and African speakers in, for example, the archives of the Moravian Church and the *Sendingskerk* archives of the Dutch Reformed Church. However, the decision to include only personal documents prevented much material from inclusion in the corpus: only 16 out of 136 writers are coloured or African. Further archival work should aim at closing this gap. This will, however, only be possible within limits as literacy itself was racially skewed: according to the 1891 Census 90% of whites older than 15 years claimed to be able to read and write, as opposed to about 15–20% of those classified as other than white (see also Nevalainen & Raumolin-Brunberg 1998 on illiteracy and social representativeness in language history).

Social class

Nineteenth-century government statistics show a general interest in the socio-economic classification of the population, and a variety of economic classes and sub-classes are distinguished in the Cape Census data (1875, 1891, 1904 and 1911). The occupational groups identified by the census are, however, of little use for a sociological analysis of class structure. Individuals are classified not according to market position and capital assets, but according to economic sector (i.e. domestic sector, commercial sector, agricultural sector, etc.) and profession. Individuals grouped together under, for example, 'mercantile persons' (a suborder of 'commercial class') embrace a variety of market positions (in terms of skills as well as property/capital relations):

> Merchants, Wholesale and Retail Dealers, all Speculators, Traders, Hawkers, Pedlars, Totgangers, all Capitalists, Financiers, Bankers, Bank Officers, Joint Stock, including Insurance, Company Managers and their Clerks, all Com-

> mercial Agents, Travellers, Commercial Assistants, Clerks, Salesmen, all Auctioneers and Valuers. (Census 1891: ixxv)

While some of the merchants might have owned property and accumulated considerable economic capital, the same is most certainly not true for hawkers and pedlars. Within the agricultural sector, socioeconomic differences were also more divisive than suggested by the all-embracing census category of 'people possessing, working or cultivating land/or animals'. This group included the wealthy Cape gentry which had consolidated its socioeconomic power in the eighteenth century ('almost inevitably white and in general considerable employers of labour'; Ross 1981: 1), independent farmers of different ethnic origins from well-off to poor, the so-called *bywoners* (farm tenants who were given the use of land in exchange for either a share of the crop or seasonal labour services) and, finally, wage labourers. Buchan's (1903: 62–63) humorous description of the South African farmer of independent, yet limited means – the prototypical South African *Boer* – captures several of the social and ethnic groups interacting within the agricultural domain:

> Let us take an average household. Jan Celliers (pronounced Seljee) lives on his farm of 3000 morgen with his second wife and a family of twelve. … Two cousins of his mother squat as bywoners on his land, and an orphan daughter of his sister lives in his household. The farmhouse is built of sun-dried bricks, whitewashed in front, and consists of a small kitchen, a large room which is parlour and dining room in one, and three small chambers where the family sleep. Twelve families of natives live in a little kraal, cultivate their own mealie-patches, and supply the labour on the farm, while two half-caste Cape boys, Andries and Abraham, who attend to the horses, have a rude shanty behind the stable. Jan has a dam from which he irrigates ten acres of mealies, pumpkins, and potatoes. … His land yields him with little labour enough to live on, and a biscuit-tin full of money, buried in the orchard below the fifth apricot-tree from the house, secures his mind against an evil day.

At top end of the Cape society was the grand bourgeoisie which was predominantly English-speaking and included the owners of great mercantile firms in the cities, many of whom owned shares in the gold and diamond mines in the north. Members of the landed gentry and leading civil servants (with incomes of over £1000 a year) also belonged to this group, which was close-knit and exercised considerable political and economic power (Bickford-Smith 1995: 20). An emerging privileged class consisted of white-collar professionals such as doctors, lawyers, teachers and ministers. These were members of an educational élite who possessed few liquid assets or property, but scored high

on what Bourdieu (1984 [1979], 1985) has called cultural capital (for a general discussion of the rise of the professional classes from the late nineteenth century cf. Perkin 1989; also Hofmeyr 1987: 93).

Although the first factories were built at the Cape in the 1880s, the manufacturing sector was still embryonic at the turn of the century (Ward 1991). This meant that (unlike in nineteenth century England or Germany) the industrial proletariat was small in numbers (Bickford-Smith 1995: 37). The casual labour force of the colony was heterogeneous, including urban low-skilled and low-paid wage earners as well as the landless rural poor, many of whom tried for better employment opportunities in the cities (Bundy 1986). Skilled workers (such as the railway workers in Cape Town and artisans) were in a slightly better position: some managed to enter permanent employment or were able to establish their own small workshops (Bickford-Smith 1995: 21). Although the pool of manual workers was heterogeneous in terms of ethnicity and skills, spontaneous strike actions (including desertion and boycotting), and a growing trade unionism from the 1890s onwards point to an incipient working class identity (Mabin 1983; Bickford-Smith 1991, 1995: 96–99, 164–185).[7] The increasing polarization of the interests of employers and workers was reflected, for example, in the *Masters and Servants Act* from 1873 which specified penalties (ranging from £1 to £5 and/or imprisonment for up to three months) for employees. 'Careless work', 'neglect', 'refusal to obey' and 'bad language' were specified in the act as sufficient for severe punishment – in short, 'virtually any behaviour … that threatened an employer's authority' (Bickford-Smith 1995: 94).

The fact that most manual work was casual employment brought with it a blurring of the boundaries between manual workers and the unemployed poor. People both in the cities and rural areas drifted in and out of employment, and the plight of the large pool of casual labourers and their families was poverty, bad sanitation, housing shortages and high rates of mortality. For members of this group upward social mobility was unlikely, and the economic depressions in the second half of the nineteenth century, the Anglo-Boer War and the mechanization of agriculture all contributed to increasing poverty at the turn of the century.

Economic and cultural assets had a direct impact on the distribution of political power in the Cape Colony as political rights were granted only to male and literate property owners. Franchise qualifications (municipal and provincial) were altered several times in the late nineteenth century, always in the direction of reducing the popular vote. In 1892 the *Franchise and Ballot Act* raised the property qualification from £25 to £75 (thus disenfranchising many

artisans; Bickford-Smith 1995: 144). According to the census of 1904, 69% of white and 41% of coloured and black men qualified for the vote.

The class scheme used in this study is eclectic and broadly based on Bickford-Smith's (1995) socio-historical analysis of nineteenth century Cape Town, as well as on Bourdieu's (1984 [1979], 1985) work on social structure and forms of capital (cf. also Chapter 2). Capital, according to Bourdieu, exists in both material (economic, monetary) and immaterial forms (social, cultural, symbolic). Bourdieu conceives of social structure as a multidimensional space in which individuals are positioned (i) according to the different forms of capital (and thus sources of power) they control, and (ii) according to their trajectories in social space, i.e. how their initial capital is likely to be transformed through their life histories (Bourdieu 1984 [1979]: 114, 128–129). The following classes were distinguished in terms of capital and property holdings:

1. *The Dominant Class* (*Grand Bourgeoisie*)
   Owners and managers of large amounts of economic capital and/or property, i.e. high ranking politicians, top civil servants; some professionals. Class position is determined by the possession of high amounts of economic capital, as well as by significant amounts of social/ symbolic/ cultural capital.
2. *The* (*New*) *Professional Class*
   The majority of civil servants, clerical workers, clerics, teachers and other professionals (e.g. doctors, lawyers, journalists). Often propertyless but also including owners of some property. Class position is determined primarily by the possession of cultural/ symbolic capital; geographical and social mobility is common.
3. *The* ('*Traditional*') *Middle Classes* (*The Petty Bourgeoisie*)
   a. Independent farmers owning moderate amounts of land but not commanding much disposable economic capital. Employers of labour. Low on cultural capital, but relatively high on symbolic and social capital, often occupying a position of authority in the area; deeply integrated into close-knit local networks.
   b. Small scale production and ownership (including shop owners and some economically successful independent artisans). Small to moderate property owners, some disposable economic and cultural capital. Often deeply integrated into close-knit local networks.

4. *The Working Classes*
    a. Most artisans, skilled manual workers and domestic servants, small-scale farmers and *bywoners.* Marketable skills, but generally low on all forms of capital.
    b. Unskilled manual workers and farm labourers. Subsistence farmers. Non-owners. No or extremely small amounts of any form of capital.

Based on the occupational information provided in the Cape Census the respective size of the different classes can be estimated (cf. Stedman Jones 1976 for a similar approach to historical class analysis). Class 1 comprised only a very small minority of the population. Judging from the Census reports class 2 was small but growing and constituted about one-tenth of the population. Class 3 can be estimated at about one-quarter to one-third of the population. The vast majority of the population, however, belonged to the 'labouring' or working classes.

Class position interacted closely with ethnicity. Class 1 was exclusively white; classes 2 and 3 were also predominately white but also included coloured and black individuals (mission school teachers, doctors, lawyers, journalists also farmers and shop owners; cf. also Bickford-Smith 1995: 23). Class 4a (skilled workers) was ethnically more diverse than classes 1 to 3. Many artisans were ex-slaves who had acquired their craftsmanship during their time of bondage. For example, of the over 8000 masons in the colony more than 5000 were classified as 'other than white' (Census 1904: cxlv), as well as about 2000 of the over 9000 carpenters and joiners (Census 1904: cxl). Class 4b was predominately black but also included many of the so-called 'poor whites'.[8]

In the context of language convergence and standardization the interaction between class membership and urbanization is also relevant. Regardless of economic standing, exposure to standard norms (and English) was greatest in the urban areas, where schools and print media were more easily accessible. Classes 1, 2 and 3b were predominantly urban (including smaller towns, especially in the case of class 3b); class 3a was rural; members of class 4 were found in rural as well as urban areas. Age (or rather generation) also interacted with class membership, in particular with respect to class 2 (the new class of white-collar professionals), which emerged as a distinct social group only from the mid-nineteenth century. As a result most of the representatives of class 2 in the *Corpus of Cape Dutch Correspondence* belonged to age group 2 (born after 1865).

Writers in the *Corpus of Cape Dutch Correspondence* were grouped into social classes on the basis of occupational information, financial assets and ed-

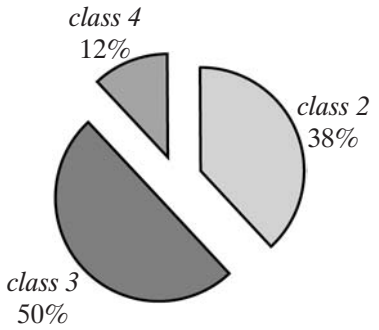


Figure 3.3 The distribution of social classes in the *Corpus of Cape Dutch Correspondence*

ucational background. In the case of women, class position was assigned on the basis of the father's or husband's position if the women had no independent employment; however, women who had employment were classified individually. Although the classification was fairly straightforward for the majority of individuals, there were some ambiguous cases. For example, D. G. Malan started out as a teacher in a small school in the Boland, but later bought a farm near Paarl in the Western Cape and became a farmer. Should he be assigned to class 2 (white-collar professional class) or class 3? A similar problem existed with respect to the classification of Louis Botha: should he be classified as a member of class 3 (farmer) or rather as a member of class 1 (politician)? In such cases the first occupational status of the respective individual was used, i.e. D. G. Malan was classified as a member of class 2, Botha as a member of class 3. Looking at the overall distribution of the sample across the different social categories (Figure 3.3), it is not surprising, given the social and ethnic distribution of literacy, that the corpus is not representative of the Cape society at large and that certain groups (i.e. class 2 and 3) are clearly over-represented.

Summary: An acrolectal and mesolectal corpus

The *Corpus of Cape Dutch Correspondence* is well-defined but biased. It includes only certain sectors of the complex, multilingual Cape Dutch speech community, which was characterized by degrees of membership, bilingualism, bi-dialectism and a diversity of highly localized sociolinguistic norms. At the same time, processes of urbanization led to increased language contact and supported linguistic and cultural convergence in urban-based networks.

The corpus includes mostly writers belonging to the white petty bourgeoisie and the emerging professional classes. The ethnically more diverse working classes are clearly under-represented. The sample bias is a consequence of the very nature of the historical record which reflects the ethnically and socially skewed distribution of literacy at the Cape. With regard to its social composition, the corpus thus resembles the material collected by Van Oordt in the *Kaapse Taalargief* for the second half of the eighteenth century (with the exception of the 'new' professional class which only emerged in the nineteenth century). The two corpora, however, differ stylistically: while the *Taalargief* comprises semi-official reports and letters, the corpus collected for this study only includes private material reflecting an informal writing style.

Absent from the corpus is the 'true' basilect (or rather the basilectal varieties). Focusing in the linguistic analysis on acro-and mesolectal varieties should not be interpreted as a 'failure to reach the vernacular' (Milroy 1979: 91 in his critique of Macaulay 1977) since basilects are not in any theoretical sense superior to other linguistic varieties; a point emphasised by Macaulay (1988: 110, 113) in his response to Milroy's critique (cf. also the comments in Chapter 1 on studies of mesolectal and acrolectal varieties):

> What troubled me was that the criticism implied that the samples of speech recorded in the interviews were somehow not legitimate examples of Glasgow speech … I believe that they were all speaking the vernacular in Illich's sense of the language used by 'people who mean what they say and say what they mean in the context of everyday life' … I agree that none of them was using the basilect in its most casual style; but for most of the respondents the basilect would not be the most casual form of speech … The search for a pure vernacular is potentially dangerous, if it leads to the undervaluing of other varieties and a consequent lack of attention to them. The aim of sociolinguistic investigation should be to describe the totality of speech use in the community.

Like all historical data the linguistic corpus used in this study is a relic: a collection of texts which have survived into the present by accident. Unlike present-day sociolinguists, language historians have to make do which what statisticians call 'convenience samples'. The limitations of socio-historical research were noted by Labov (1972: 100) who dismissively characterized historical sources as 'bad' (i.e. incomplete) data and advocated the study of the past through the present. However, work in historical sociolinguistics and the social history of language (cf. Romaine 1982; Lippi-Green 1994: Nevalainen & Raumolin-Brunberg 1996, 2003; Burke & Porter 1987, 1993, 1995) has since shown that historical data, although always fragmentary, nevertheless allow the

language historian to reconstruct important aspects of language use, attitudes and variation patterns in historical speech communities.

Notes

1. See also Jespersen's (1925: 389) definition of the speech community: 'Leaving the individual, we turn to the study of speech communities. These are of different magnitudes and may be arranged in an ascending series – the family, the clan, the tribe, the people or nation, and finally the super-nation, by which term we mean such community as the English-speaking community with its 150 millions distributed over five continents'.

2. See for example *Catalogues van Boeken enz.* (J. Dusseau & Co. 1897) and *Prijslist van Boeken bij die Paarl Drukpers Maatschappij, Beperkt* (October 1908), for general price and wage index see Burnton (1903: 234–326; cf. also Bickford-Smith 1995: 91–102, 168, 179).

3. The use of capital letters and punctuation was arbitrary in many documents and was standardized in the process of transcription, i.e. capitals were used for proper nouns and words at the beginning of sentences, punctuation symbols were added in square brackets to clarify the syntactic structure. Formulaic expressions at the beginning and end of letters (such as *Liewe* XYZ, 'Dear XYZ', or *Jou liefhebbende* XYZ, 'Your loving XYZ') were excluded from the quantitative analysis.

4. Such remarks must, of course, be interpreted carefully and cannot be taken as direct evidence. It is always possible that the writer exaggerated the situation to obtain financial support from the addressee. However, the inclusion of additional sources (e.g. death notices and deeds office register) allowed for the assessment of the reliability of the reports found in the letters and diaries.

5. Furthermore, women were filed not under their maiden name but under the name of their husband. This made it impossible to trace some of the young female writers such as Mimmie Laubscher or Catharina E. Beyers (who were still unmarried at the time of writing their letters). However, their family background could be assessed by locating the death notices of their parents.

6. However, Adhikari (1989) has argued that these episodes were not examples of an exclusively Malay/Muslim resistance as both Malay/Muslim and other coloured groups were involved in the riots. Similarly Effendi, the Malay/Muslim parliamentary candidate, did not emphasize Malay/Muslim interest in his campaign but 'was careful to project himself as a working man's candidate' (ibid.: 14).

7. Bickford-Smith (1995: 165) notes that some of the nineteenth century workers' unions were ethnically mixed. However, the division of labour at the workplace (typically white workers as supervisors and black or coloured workers as subordinates) prevented the development of a broad unionist movement.

8. Whites had always formed a significant part of the poor. However, from the late nineteenth century poverty was defined in racial terms, i.e. the popular Victorian distinction between the deserving and undeserving poor was now drawn along racial lines (Bickford-Smith 1995: 126–127).

Part II

Variation analysis

Chapter 4

On the analysis of variability and uniformity

An introduction to multivariate clustering techniques

> It is widely acknowledged that the science of taxonomy is one of the most neglected disciplines … The practice of taxonomy has remained intuitive and commonly inarticulate, an art rather than a science.
>
> Robert R. Sokal & Peter H. Sneath (1963). *Principles of Numerical Taxonomy*

Statistics and variation studies: More than a numbers game[1]

Linguistic variation, although sometimes categorical, is often a question of frequency. In other words, while there are situations in which one person (or group of people) uses only variant X and another person (or group of people) only variant Y, often speakers (or groups of speakers) are distinguished by the relative frequency with which they employ the two variants. It appears to be uncontroversial that such situations of variable frequency can be described quantitatively by counting the number of occurrences of each variant, and by comparing the relative frequencies (usually expressed in percentages) across individuals or groups. In this way numbers are assigned to originally qualitative observations, and as a result a new type of data is produced which is numerical in nature and can be subjected to statistical analysis.

However, to acknowledge that speech is variable does not mean that one attaches theoretical importance to the observed frequency variation. In other words, quantitative fluctuations might simply be performance phenomena which do not impinge on the categorical nature of language *per se*. Bloomfield (1933: 37–38), for example, has suggested that the 'speech habits of a community' can adequately be described as the sum of all linguistic forms available to the members of the community, regardless of their frequency patterns. 'Genuinely statistical observation', on the other hand, is necessary only for the description of the process of language change. Chomsky (1965, 1980)

has repeatedly stated that frequency variability is part of performance (or E-language) and not competence (or I-language) – which alone constitutes the object of linguistic theory. This is also the position of Lightfoot (1991), who interprets the surface variability observable in historical sources as a reflection of changing 'triggering' experiences which ultimately (i.e. once they change in some critical fashion) lead to new parameter settings in the minds of individual speakers (and thus to changes in the I-language). Sankoff (1980: 77–78), however, has argued for a strong and psychologically real connection between observable performance and underlying competence. She accordingly interprets performance variability as a central aspect of linguistic competence:

> My position has been that statistically fluctuating performance data need not be interpreted as reflecting underlying competence that is categorical in nature, and that a paradigm representing competence as containing some probabilistic and nondeterministic components is a better approximation to linguistic reality than one that insists on categoriality and determinacy. Note that this does *not* imply the nonexistence of categorical rules, but simply the existence as well of probabilistic rules.
> Such an assumption does not counter the principle that (socio)linguistic competence is what exists in people's heads; rather, it takes the position that people can internalize rules that are not categorical.

Describing frequencies and establishing probabilities of occurrence is thus not simply a descriptive exercise; it is of theoretical importance for our understanding of the nature of language. Probabilities of occurrence describe an important aspect of the individual's linguistic competence(s) and ultimately contribute to the formation and change of language structures and norms (cf. Bybee & Hopper 2001).

Once the linguistic data are re-ordered in terms of frequencies and proportions, descriptive and inferential statistical procedures can be used to summarize and interpret the numerical information. That the discovery of 'patterns of usage, which pertain to the relative frequency of occurrence or co-occurrence of structures' is a central aim of sociolinguistic and linguistic research was noted by Poplack (1993: 258; emphasis in the original). To investigate such patterns the tools of descriptive statistics are useful. These include numerical summaries (such as mean, standard deviation, median, inter-quartile range etc.), correlational methods and different types of visual representations (stem-and-leaf plots, line plots, histograms, etc.) – in short all techniques which fall under the heading of Exploratory Data Analysis (cf. Tukey 1977; Diaconis 1995). Descriptive statistics are best described as 'tools of discovery' (Anderberg 1973: 4)

as they allow the researcher to explore unknown data and to generate hypotheses about the structure of the data.

The strength of the quantitative evidence derived from the data can be assessed further with the help of inferential statistics.[2] Briefly, inferential statistics are used to calculate the probability that the relations observed in the sample are likely to occur in the population from which the sample is drawn and are not due to a sampling error. The probability is represented by the *p*-level. Results that yield $p < 0.05$ are generally considered statistically significant, i.e. in this case the patterns observed in the sample can be generalized to the population with a specifiable degree of confidence (the probability of error is less than 5%).[3] Inferential statistics can be divided into parametric and non-parametric methods; the former demand a moderate to large sample size and are based on the assumption that the data scores are normally distributed. Nonparametric tests, on the other hand, do not rely on the estimation of population parameters, nor do they need large samples. In other words, they are specifically designed for the analysis of low-quality data, that is, small samples drawn from populations about which little is known. The most frequently used nonparametric technique in linguistics is probably the χ^2-test which measures the significance of the deviation of a given sample distribution from the expected frequency of events: $\chi^2 = \Sigma(f_{obs} - f_{exp})/(f_{exp})$.

The analysis presented in Chapters 5 and 6 uses, in addition to descriptive statistics and data summaries, inferential statistics (χ^2-test) to assess the reliability of the quantitative distribution patterns in the *Corpus of Cape Dutch Correspondence*. Chapter 7 uses multivariate clustering techniques (which are useful tools for data classification) to investigate (i) how individual variables (and their varying frequencies) are combined into linguistic sets or varieties, and (ii) whether writers can be grouped into clusters on the basis of similarities in their language behaviour. In its quantitative approach to language standardization this study follows earlier work by Devitt (1989) and Lippi-Green (1994).

Numerical taxonomy

Classification is a basic human conceptual activity and is fundamental to the practice of science. Observation is followed by the systematic cataloguing of objects in terms of their properties and relations, and the resulting taxonomies help to structure the domains of scientific inquiry. More than 30 years ago

the biologist G. G. Simpson noted the ubiquity and importance of scientific classifications in his *Animal Taxonomy* (1961):

> Scientists do tolerate uncertainty and frustration, because they must. The one thing they do not and must not tolerate is disorder. The whole aim of theoretical science is to carry to the highest possible and conscious degree the perceptual reduction of chaos … In specific instances it can well be questioned whether the order so achieved is an objective characteristic of the phenomena or is an artificial construct by the scientist … Nevertheless the most basic postulate of science is that nature itself is orderly … All theoretical science is ordering and if systematics is equated with ordering, then systematics is synonymous with theoretical science. (quoted in Lévi-Strauss 1972: 9–10)

Taxonomists have frequently based their classifications on their experience with the data and their resulting intuitions about its structure. In the 1960s, however, objections were raised against such subjective methods. A multivariate, systematic and strictly empirical approach to classification (relying on explicitly formulated statistical methods rather than intuition) was advocated by the biologists Sokal and Sneath. In their book *Principles of Numerical Taxonomy* (1963) Sokal and Sneath argued that in order to construct adequate biological classifications one should gather all information available on a set of organisms, code this information numerically and then use a clustering algorithm to group similar objects together into groups (cf. also Panchen 1992: 132–151; for a history of the field see Hull 1988: 117–130). Sokal's and Sneath's book was a major stimulus for the development of automatic clustering techniques. The increasing availability of high speed computers, which reduced computing time and made it possible to handle large data sets, further encouraged interest in and application of clustering techniques in many different fields. Today cluster analysis, multidimensional scaling and principal components analysis (PCA) are widely used for establishing classifications, not only in botany and zoology, but also in the medical sciences, sociology, engineering and economics.

In quantitative sociolinguistic research clustering techniques have also been shown to be powerful heuristic tools for exploratory data analysis, however, their use is by no means as widespread as in other fields. Cluster analysis was used, for example, in the Tyneside sociolinguistic survey to classify speakers into groups based on similarities in their linguistic behaviour (Pellowe 1976: 210–212). The sociolinguistic survey of multilingual communities carried out by Le Page et al. in Belize and St. Lucia also made use of cluster analysis (cf. McEntegart & Le Page 1982; also Le Page & Tabouret-Keller 1985). Bernstein (1993) used cluster analysis in her study of phonological variation

in Texas to reduce the original twelve linguistic variables into a smaller number of clusters. Multidimensional scaling and PCA were applied by Berdan (1978) to Macaulay's Glasgow data for the investigation of intra-group variation, and were used by Cichocki (1988) for the analysis of regional differences in Canadian speech (reported in Lee & Kretzschmar 1993). Multidimensional scaling was also used by Sankoff and Cedergren (1976) in their re-analysis of several data sets, including the data on complement *que*-deletion in Montreal French. PCA was used by Ma and Herasimchuk (1972) in their analysis of speech styles of Puerto Rican bilingual speakers to show the interdependency of phonological variables, and by Horvath (1985) in her study of variation in Australian English. Clustering techniques have also been used to group different speech and text genres in terms of their linguistic similarities (Biber 1988; Besnier 1988), and in dialect classification (Nerbonne & Heeringa 1998). However, despite the successful application of multivariate clustering techniques in sociolinguistic research, uni- and bivariate data summaries have retained their prominence in research designs and introductory statistics courses for linguists. While statistical concepts such as mean, standard deviation, *p*-level and significance certainly belong to the academic jargon of linguists, the same is not true for terms such as linkage distance, amalgamation graph, alienation coefficient, eigenvalue or component loading, which are integral to the clustering techniques used in this study.[4] It therefore remains necessary for linguists who decide to make use of these lesser known techniques to introduce them carefully. Pellowe's (1976: 203–204) comments on the need for descriptive transparency in sociolinguistic research are still relevant today:

> Unclearness in the specification of aims and assumptions, allusiveness in a statement of methods, suppression of exceptions, and many other features of what may be normal presentation methods in other research areas, are particularly disadvantageous for the proper growth of sociolinguistic research, since precisely this information must be the basis for the formation of a critical apparatus by which to evaluate competing approaches.

The following sections offer a conceptual and operational, rather than mathematical, description of the three multivariate techniques used in Chapter 7 for the variationist analysis of the corpus data. Furthermore, based on a small linguistic data set (Labov 1969) a step-by-step description of these methods, including all statistical outputs, will be given to illustrate the application of these techniques to the analysis of linguistic variation.

Hierarchical cluster analysis[5]

The conceptual and mathematical simplicity of cluster analysis is striking when compared to the complex deductive mathematics underlying other statistical methods. In cluster analysis individual cases (such as the speakers or groups of speakers studied in sociolinguistic research) are grouped together into relatively homogenous clusters based on a number of measured characteristics. The similarity between cases is expressed numerically in terms of their metric distance (or proximity). The metric concept of similarity can be illustrated by using a two-dimensional example (Table 4.1; following the discussion in Romesburg 1984).

The rows in this small data set represent cases, that is, objects whose similarity to each other we want to estimate. The columns stand for variables, that is, the properties of the cases. We can treat the two variables as coordinates that fix the positions of cases 1 to 4 in a two-dimensional space (Figure 4.1). The distances between the cases 1 to 4 are indicated by dotted lines and can be interpreted as indices of similarity or dissimilarity. Examining the distances visually, the cases 1, 4 and possibly 3 can be described as being relatively similar to one another and thus forming a cluster, while the distance of case 2 indicates a high degree of dissimilarity to the other three cases.

The Euclidean distance (that is, the geometric distance in the *n*-dimensional space) is a convenient measure for the distances between cases 1 to 4 (for details regarding the mathematical procedure cf. Romesburg 1984: 13–14; Kaufman & Rousseeuw 1990: 11–12). Briefly, following Pythagoras' theorem ($a = \sqrt{b^2 + c^2}$) the distance *a* between two points *i* and *j* with the coordinates (x_i, y_i) and (x_j, y_j), equals the length of the hypotenuse of the triangle and takes the following value:

$$a(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

This expression can be extended to an *n*-dimensional space (i.e. situations where the position of *i* and *j* is determined by more than two variables) as

Table 4.1 Two-dimensional data set

	variable 1	variable 2
case 1	10	10
case 2	20	30
case 3	5	20
case 4	15	10

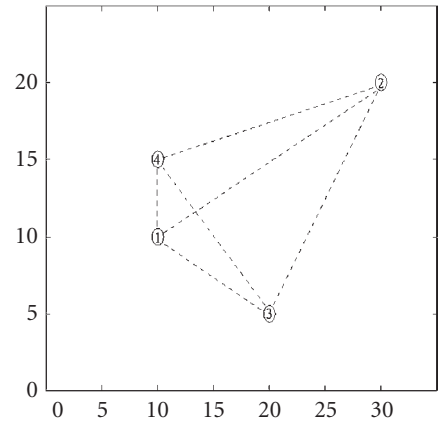


Figure 4.1 Position of cases 1–4 in the two-dimensional space

follows: $d(i, j) = \sqrt{(x_{i1}-x_{j1})^2 + (x_{i2}-x_{j2})^2 + \ldots + (x_{in}-x_{jn})^2}$, where $(x_{ix}-x_{jx})$ represent the different variables which define *i* and *j*. Thus, two variables create a two-dimensional space, three variables a three dimensional space, and so forth.[6] The Euclidean distances between the cases are then displayed in a distance or dissimilarity matrix (Table 4.2). The calculation of a distance matrix is the first step in the development of numerical classifications.

After the distances between the objects have been established a clustering algorithm mechanically searches the distance matrix for the lowest distance value between pairs of cases. These pairs of cases are then combined into larger clusters. In the example data set cases 1 and 4 would be linked first. Now the distances between the remaining cases and the cluster {1,4} will be searched by the algorithm, and again the item with the smallest distance value will be merged with the existing cluster; that is, case {3} and cluster {1,4} will be combined into a new, larger cluster {1,4,3} (for more details cf. Everitt 1986). Hierarchical clustering techniques are thus agglomerative: they successively combine cases into larger and larger clusters by grouping at any step the two

Table 4.2 Distance or dissimilarity matrix for cases 1–4 (Euclidean distances)

	1	2	3	4
1	0			
2	22.4	0		
3	11.2	18.0	0	
4	5.0	20.6	14.1	0

cases (or groups of cases) that are closest to each, reducing the number of cases at each step until only one cluster is left.

Two main types of cluster linkage (or merger) rules can be distinguished: linkage of nearest neighbour (single linkage) and linkage of furthest neighbour (complete linkage). Linkage of nearest neighbour (illustrated in Figure 4.2) means that an object (or a cluster of objects) is joined to an existing group based on a high level of similarity with *any* member of the existing group; that is, only a single link is required for a case (or groups of cases, i.e. a cluster) to merge with an existent cluster.

Complete linkage (i.e. linkage according to the furthest neighbour; illustrated in Figure 4.3) is the opposite of single linkage: a case will only be merged with an existing cluster if this case shows a minimum level of similarity to the furthest member of that cluster. As a result the data set will be patterned into relatively compact clusters containing highly similar cases (Aldenderfer & Blashfield 1984: 40). Complete linkage is usually recommended in data sets where the cases belong to highly distinct clusters; it should be used with care in data sets where no obvious patterning is to be expected.

As a compromise between single and complete linkage, average linkage was developed by Sokal and Sneath (1963: 182–185). This algorithm computes the average of the similarity between two clusters (by calculating the means for the

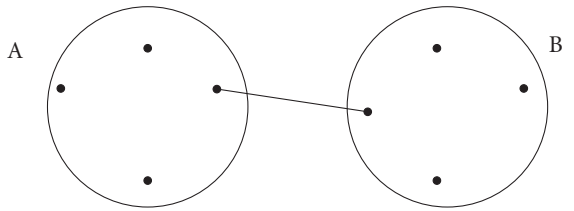


Figure 4.2 Single linkage (adapted from Kaufman & Rousseeuw 1990: 47)

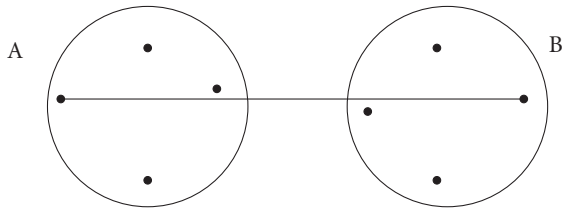


Figure 4.3 Complete linkage (adapted from Kaufman & Rousseeuw 1990: 47)

distances between all objects of cluster A and all objects of cluster B). Clusters are joined only if a minimum level of similarity can be established on the basis of this average value. Average linkage has been shown to perform well on many types of data and is widely used (STATISTICA 1995 III: 3171; Romesburg 1984: 126). Apart from average linkage, Ward's method is a frequently used linkage rule (Ward 1963; Romesburg 1984: 129; Jobson 1992: 524). Ward's method was designed to minimize within-cluster variability by using an analysis of variance approach to evaluate the distances between clusters. In several studies Ward's method has been shown to outperform other clustering methods, including average linkage (Aldenderfer & Blashfield 1984: 61).

A hierarchical tree diagram (also called dendrogram) is the graphical output of cluster analysis (see Figures 4.4, 4.5 and 4.6). The diagram represents the hierarchical relationships that exist between the different cases and clusters, and illustrates the series of steps which led to the final classification; that is, each step at which a pair was merged is represented as a node in the tree. Unlike the evolutionary tree diagrams used in biology and linguistics, the clustering tree only represents a classification of cases based on affinity, but does not necessarily reflect phylogeny (descent). The major steps of cluster analysis can be summarized as follows:

INPUT Data matrix
↓
STEP 1 Selection of a distance measure (most commonly Euclidean) and computation of the distance matrix
↓
STEP 2 Choice and computation of a linkage rule (e.g. single linkage, complete linkage, average linkage or Ward's method)
↓
OUTPUT Dendrogram representing the classification of cases into hierarchical clusters

The decision which of the distance measures or linkage rules to use lies entirely with the researcher, who therefore needs to be aware of the strengths and weaknesses associated with the different methods. The interpretation, and especially the validation of the clustering results, are the most difficult steps in cluster analysis:

> An inherent problem in the use of a clustering algorithm in practice is the difficulty of validating the resulting data partition. This is a particularly serious issue since virtually any clustering algorithm will produce partitions for any data set, even random noise data which contains no cluster structure. Thus,

> an applied researcher is often left in a quandary as to whether the obtained clustering of a real life data set actually represents a significant cluster structure or an arbitrary partition of random data.
>
> (Milligan 1981, quoted in Jobson 1992: 563/564)

In other words, in cluster analysis one cannot test the null hypothesis (i.e. the hypothesis that there is no structure in the data). It is thus difficult to determine whether the clusters identified by the algorithm reflect 'true' or 'natural' typologies, or whether they constitute 'administrative clusters' which simply present more or less convenient groupings of the data set.[7] The validity of a cluster structure can be assessed by considering the following criteria (for summaries of validation procedures see also Romesburg 1984: 256–259; Aldenderfer & Blashfield 1984: 62–74; Jobson 1992: 563–568).

i. External criteria can be used to validate a given cluster solution. Validation through the use of external criteria is particularly feasible in sociolinguistic research where clusters are established on the basis of the linguistic variables alone, and the results can then be checked against extra-linguistic information (i.e. whether the clusters are also well-formed in terms of social group membership). A high degree of 'predictivity' (understood as 'the degree to which a specific classification agrees with characters not used in the formulation of that classification'; Fitch 1979) is seen as a strong indicator of naturalness (Panchen 1992: 149).
ii. Internal criteria can also be used to validate a cluster solution and to assess the naturalness of the identified clusters. Natural clusters are assumed to satisfy the conditions of external isolation (i.e. the clusters are clearly separated from one another; the criterion of cluster separation) and internal cohesion (i.e. the cases in a cluster are very similar; the criterion of cluster density; cf. Jobson 1992: 519). The internal structure of the cluster solution can be investigated with the help of *k*-means clustering (also called iterative partitioning method), as well as by inspecting the visual representation of the distance matrix using multidimensional scaling (see below). The input for *k*-means clustering is a specified number of clusters *k*, and based on this input the program will compute exactly *k* clusters that are as distinct as possible. Based on the descriptive output of *k*-means clustering, the internal structure of each cluster (density, i.e. distances between cluster members) and the differences between the clusters (separateness, i.e. distances between clusters) can be described.
iii. The replicability of a cluster solution based on split samples of the original data set is also used to validate a cluster solution. Sokal (1986: 435) distin-

guishes between 'character stability' (stability of the classification despite the addition or subtraction of variables) and 'case stability' (the classification is robust even if cases are randomly added or subtracted). If the solution cannot be replicated, this is reason enough to reject the solution entirely. However, successful replication alone does not guarantee the validity of a solution (Aldenderfer & Blashfield 1984: 65).

iv. A comparison of cluster solutions obtained by different clustering algorithms, including *k*-means clustering and different multivariate techniques (such as principal components analysis and multidimensional scaling), can also be used to validate a cluster solution. In this case validity would be assumed if the results obtained from the different techniques agree with each other. The use of *k*-means clustering for the validation of a cluster solution obtained from the hierarchical analysis also helps to compensate for a general shortcoming of hierarchical methods, that is, the fact that these methods cannot repair what was done in a previous clustering step: once two objects have been joined into a cluster, they cannot be separated (Kaufman & Rousseeuw 1990: 44–45). *K*-means clustering, on the other hand, makes several passes through the data and can therefore modify a poor initial partition (Aldenderfer & Blashfield 1984: 46).[8]

Example (Labov 1969)

In order to illustrate in a little more detail how cluster analysis works and what kind of insights it can provide in the context of variation studies and sociolinguistics, Labov's (1969) data on copula deletion in African American English (AAE) will serve as an example. AAE is a variety of English which is learned by most African Americans as a first language, and which is used by the majority of Black adults in relaxed settings when conversing with friends and family (Labov & Harris 1986: 4). AAE differs from other varieties of American English with regard to grammar, pronunciation and lexicon. A well-known linguistic feature is the optional deletion of the copula in a number of syntactic environments. Labov has shown that variation in the copula is the result of a series of grammatical and phonological rules that are parallel to those of colloquial contraction in standard English (Labov 1969: 722ff.).[9]

An important variable constraint on deletion in AAE (as well as on contraction in Standard English) is the nature of the subject, i.e. whether the subject is a pronoun (P) or some other type of noun phrase (NP). The data set given by Labov lists the percentages for full, contracted and deleted forms of the copula for six groups of informants. The first four groups are differ-

Table 4.3 Percentages of full, contracted, and deleted form of the copula with noun phrase (NP) subject or pronoun (P) subject (group style only; Labov 1969:730). Due to rounding some percentages do not add up to 100.

	NP/full	NP/contracted	NP/deleted	P/full	P/contracted	P/deleted
T-Birds	44	15	42	7	33	60
Cobras	45	19	36	0	23	77
Jets	54	19	27	0	42	58
OscarBros	51	23	26	4	33	64
Adults	61	26	14	1	72	27
Inwood	41	59	0	1	99	0

ent pre-adolescent peer groups (with members aged between 10 and 18 years), Adults refers to a sample of working-class adults in the Cobra and Jet territory, Inwood to a contrasting sample of white working-class speakers from the Inwood neighbourhood of upper Manhattan (Table 4.3).

Cluster analysis can be performed on the raw or on the standardized data matrix.[10] Standardization is necessary only in cases where different units of measurement are used. In this case the raw data matrix was used as input into the program as the values of the variables are expressed in percentages and are thus comparable across the six groups of speakers.

Concerning the choice of a linkage rule, it was noted above that it is common practice to use different linkage rules for exploratory data analysis. Three different linkage rules were used for data exploration: single linkage (Figure 4.4), average linkage (Figure 4.5) and Ward's method (Figure 4.6). The distance measure employed was the Euclidean distance. The map given by the tree shows a hierarchical ordering of similarities that begins at the bottom of the tree where each object or case is separate, in its own cluster and similar only to itself. As one moves upwards individual cases are merged into clusters on the basis of their similarity.

A comparison of the three diagrams suggests that we either have two clusters: {Jets, T-Birds, OscarBros, Cobras}, {Adults, Inwood}; or three clusters: {Inwood}, {Adults}, {Jets, T-Birds, OscarBros, Cobras}. Furthermore, all three diagrams identify OscarBros and Jets as forming a subgroup within the cluster of pre-adolescent peer-groups. The fact that the pre-adolescent peer groups are clustered together was to be expected from our knowledge about the social context of language use in American cities. A second (expected) result is that Inwood is clearly separated from the adolescent peer group cluster. The position of black working class adults, however, is somewhat ambiguous: do

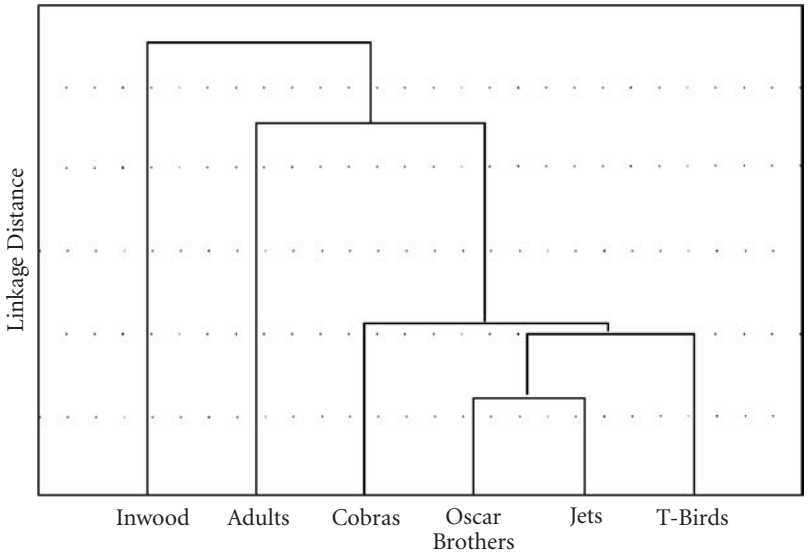


Figure 4.4 Tree diagram for single linkage (Euclidean distance)

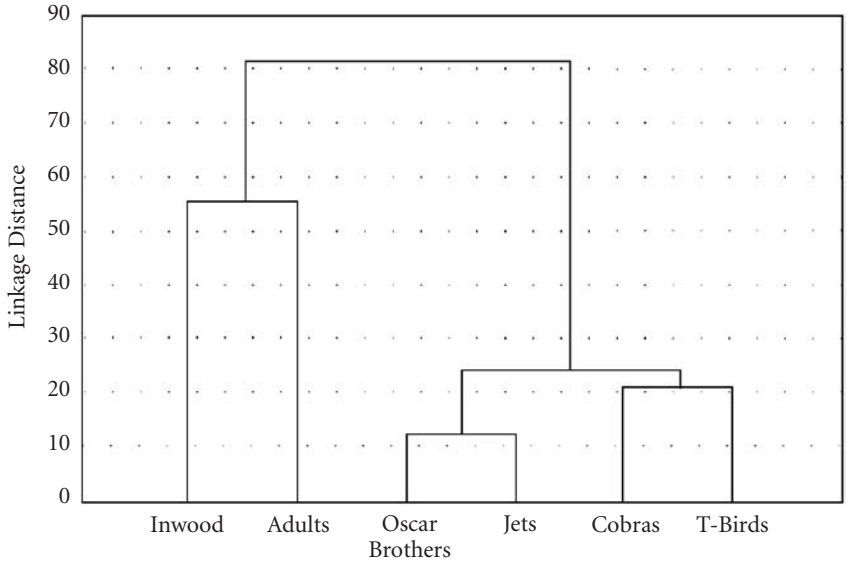


Figure 4.5 Tree diagram for average linkage (Euclidean distance)

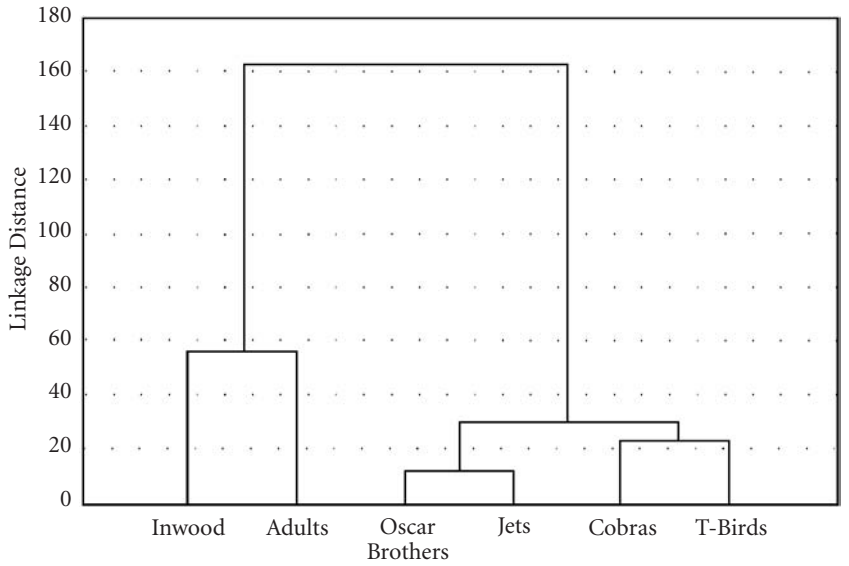


Figure 4.6 Tree diagram Ward's method (Euclidean distance)

they form a linguistic cluster of their own? Or should they be grouped together with Inwood?

The interpretation of the hierarchical tree diagram is not always as straightforward as this example might suggest. The more cases are involved in the analysis the more difficult it is to identify groups in the data. The nested structure of the diagram suggests that many different groups exist in the data and the analyst has to decide where to cut the tree in order to obtain the optimal number of clusters (Aldenderfer & Blashfield 1984: 53). Frequently the use of external criteria as well as expert intuition is used to make this decision (Anderberg 1973: 176). A more formal approach is found in the interpretation of the graph of amalgamation (linkage) schedule. The graph plots the successive clustering steps against the distances at which the clusters were formed (Figure 4.7).

The amalgamation graph is read from right to left. A steep jump in the graph indicates that two relatively dissimilar clusters have been merged, a flattening of the graph implies that after this point clusters were formed more or less at the same linkage distance. The number prior to the point where the graph levels out into a plateau is therefore the most probable solution.

In this example a three cluster solution is indicated by the amalgamation graph: {Inwood}, {Adults}, {Jets, OscarBros, Cobras, T-Birds}. That this solu-

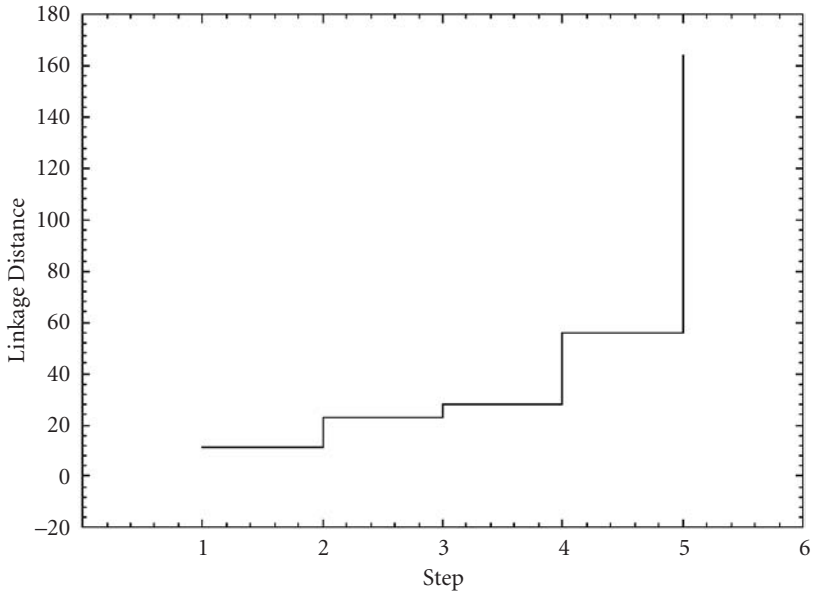


Figure 4.7 Amalgamation graph (Euclidean distance, Ward's method)

tion is plausible in terms of external criteria does not need detailed explanation. But what can be said about the internal structure of the clusters? How similar or dissimilar are their objects and how distinct are the clusters from each other? In other words, do they fulfil the requirements for natural clusters?

To investigate the internal structure of the three clusters more closely, *k*-means clustering with $k = 3$ is used. The clusters identified by the partitioning method are identical with those obtained from the hierarchical analysis: {Inwood}, {Adults}, {Jets, OscarBros, Cobras, T-Birds}. The first two clusters are by definition homogeneous as they only contain one case. The structure of the cluster comprising the pre-adolescent peer groups (Table 4.4) shows that the Cobras, which were also singled out in two of the dendrograms, are furthest removed from the cluster centre.

Table 4.4 Internal structure of the pre-adolescent peer group cluster

	T-Birds	**Cobras**	**Jets**	**OscarBros**
Distance from cluster centre	5.2	6.8	5.8	3.4

Table 4.5 Euclidean distances between clusters

	Inwood	Adults	Peer groups
Inwood	0		
Adults	22.9	0	
peer groups	43.4	24.2	0

The distances between the three clusters, which are given in Table 4.5, indicate that the peer group cluster is separated clearly from Inwood. The group of Adults is located somewhere half-way between Inwood and Peer groups.

Multidimensional scaling

Unlike cluster analysis, multidimensional scaling is not designed to link elements together into clusters. Multidimensional scaling simply visualizes the similarity relationships between cases in an *n*-dimensional space by arranging these cases in such a way that the rank order of distances between cases (on all measured variables) are maintained.[11]

For example, multidimensional scaling could be used to create a two-dimensional map based only on information (summarized in a distance or dissimilarity matrix) about the road distances between cities or other geographical locations. To achieve this, the program moves the cases in the space defined by the requested number of dimensions (i.e. number of coordinate axes in the *n*-dimensional space), and checks how well the recorded distances between the cases are reproduced by the configuration. On the first iteration, all objects are ordered along a single line and the reproduced distances are compared with the original distances. Typically a large measure of error is found. On the following iterations the cases are moved about until the differences between original and reproduced distances are close to zero (for a detailed description of the iteration procedure see Schiffman, Reynolds, & Young 1981: 7–10). To measure what is called the 'goodness of fit' of a configuration (i.e. the degree of correspondence between the reproduced distances in the visual configuration and the distances in the original matrix input), two different indices are used: the stress coefficient and the alienation coefficient (for details of the computation see STATISTICA III 1995: 3237, 3255). The smaller the value for both coefficients, the better the fit of the reproduced configuration. The goodness of fit of a particular solution can also be assessed with the help of the so-called Shepard diagram (named after Roger Shepard's seminal paper of 1962). A Shepard

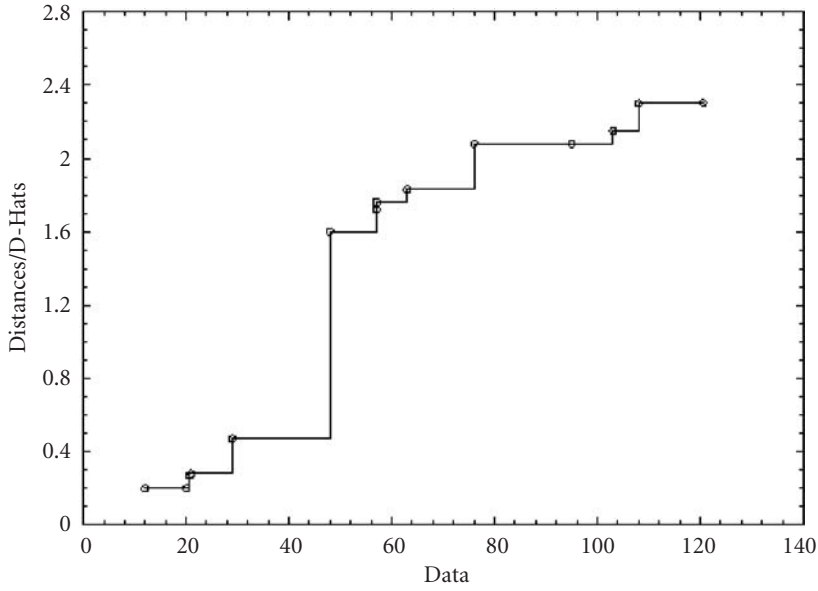


Figure 4.8 Shepard diagram

diagram is a scatter plot in which the reproduced distances are plotted against the original matrix distances (Figure 4.8).

Multidimensional scaling requires as input a distance matrix, which can be obtained from the cluster analysis module. Next, the analyst has to decide on the number of dimensions to be computed. At this point the decision is purely intuitive. Only after the dimensions have been established can their adequacy be investigated. Generally one would start with as few dimensions as possible. Only if a two-dimensional solution cannot be shown to be adequate should more dimensions be included. If no more than three dimensions are specified, a two- or three-dimensional scatter plot can be used to represent the configuration of cases visually. With help from the results of cluster analysis, or simply through visual inspection of the graph, groups of cases can then be identified in the scatter plot. Multidimensional scaling thus comprises the following steps:

INPUT Distance matrix (usually Euclidean)

↓

STEP 1 Specification of n dimensions and computation of the n-dimensional configuration

↓

STEP 2 Review the goodness of fit of the configuration (Shepard diagram, stress and alienation coefficient)
↓
OUTPUT Scatter plot of the configuration

The interpretation of the results includes the visual inspection of the scatter plot, as well as the interpretation of the actual dimensions (i.e. the coordinate axes) established by the program. To interpret the dimensions one looks at the properties of the cases at each end of the axes in order to determine whether there is some characteristic that changes clearly along the dimension. Replication of the analysis based on split samples and correlation with external criteria can be used for the validation of the results.

Example (Labov 1969)

The dissimilarity matrix (Euclidean distances) obtained from the cluster analysis module was used as input for multidimensional scaling. Two dimensions were specified and eight iterations were computed. In the final configuration the stress and alienation coefficients were zero, which suggests a perfect representation of the original data. The Shepard diagram, given in Figure 4.8, also indicates a flawless re-scaling of the distance matrix. The steep function depicted in the Shepard diagram represents the monotone transformation of the data: each data point represents a combination of the original data dissimilarity (plotted on the X-axis) and the reproduced distances (plotted on the Y-axis). A close fit of the steep function to the data points indicates a good fit between original and reproduced configuration (for details regarding the computation, see STATISTICA 1995 III: 3255).

The clusters established with the help of the cluster analysis algorithm can then be located in the scatter plot which depicts the final two-dimensional configuration (Figure 4.9). Each point in the plot corresponds to one of the six groups of speakers (i.e. each point represents a case). Dimension I distinguishes the groups with regard to the variables involving contraction and deletion of the copula, i.e. those on the left {Peer groups} show high scores for deletion and low scores for contraction; the opposite is true for Inwood and Adults which are located on the right. Dimension II separates Adults from the other cases.

Based on the pictorial representation in the scatter plot, the previously identified clusters can be described in terms of density, variance, shape and separation (Aldenderfer & Blashfield 1984: 34). A thick swarm of data points in a space constitutes a cluster of high density, while data points are widely sep-

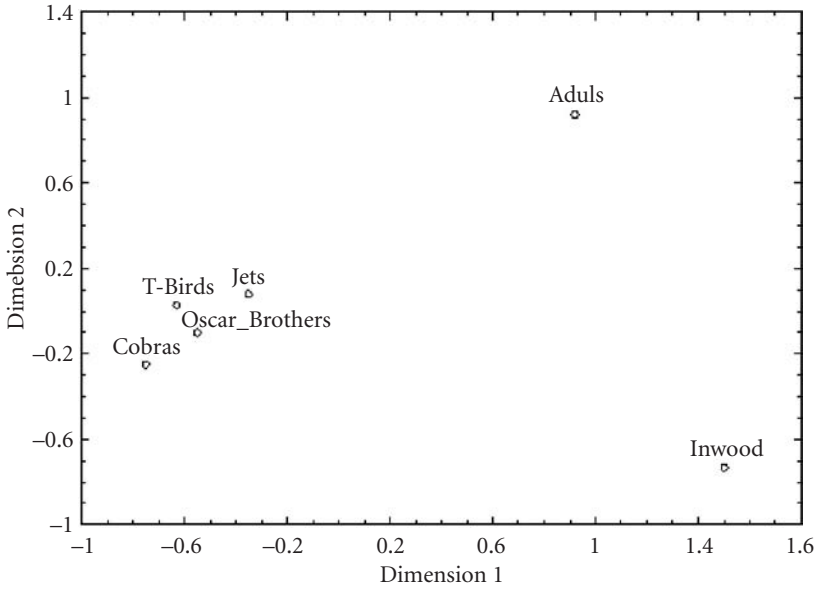


Figure 4.9 Scatter plot (dimension 1 vs. dimension 2)

arated in a cluster of low density. Variance refers to the dispersion of the data points from the centre of the cluster. In tight clusters all data points are near the centroid, while the data points are dispersed from the centre in loose clusters. Finally, shape is simply the arrangement of points in the *n*-dimensional space, and separation is the degree to which clusters are separated from one another.

Principal components analysis (PCA)[12]

Unlike in cluster analysis where the researcher can decide on the setting of parameters such as distance measures and linkage rules, most of the framing decisions in PCA have already been made by statisticians and the method can, therefore, not easily be tailored to meet the specific needs of a given data set or research question. The central idea of PCA is to reduce the dimensionality of the data set, which is believed to consist of a large number of interrelated variables. By transforming a set of related variables into a smaller set of uncorrelated components, the number of variables is reduced and the structure of the relationship between the variables can be described. To understand how variables are combined into components, it is best to start again with a two-

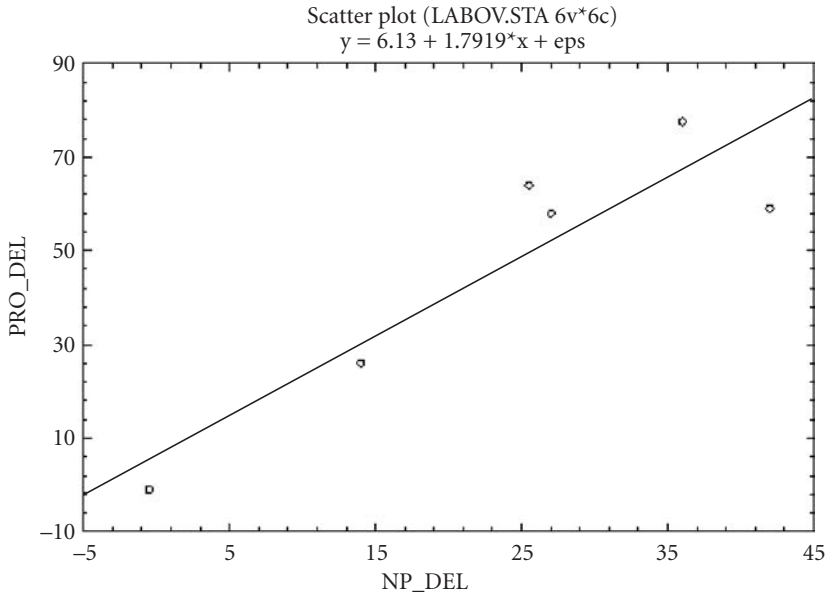


Figure 4.10 Correlation between [NP/del.] and [P/del.] with a least squares fit (regression line)

dimensional example. Figure 4.10 plots the scores for the deletion of the copula in AAE after pronoun subjects and after noun phrases in a scatter plot with [NP/del.] on the horizontal axis and [P/del.] on the vertical axis.

The graph indicates that the two variables are highly and positively correlated, i.e. high variable values on [NP/del.] will also yield high variable values on [P/del.] (correlation coefficient r = 0.9).[13] A regression line (eigenvector), which captures most of the variance displayed by the two variables, has been fitted to summarize the linear relationship between the two variables. In order to account for the remaining variability a second line can be drawn through the data, then a third one, and so forth. In this way consecutive eigenvectors (or components) are extracted. If the two-variable plane is extended to a multidimensional space (i.e. if more than two variables are included), the computations become more involved; the principle, however, remains the same.

For each eigenvector a separate eigenvalue is computed. The eigenvalue indicates the amount of variance extracted by the component.[14] The eigenvector with the highest eigenvalue becomes the first principal component and accounts for the greatest percentage of variance in the data. Each consecutive component is defined to capture the maximum amount of the variability

not captured by the preceding component. Generally a small number of these components will account for most of the variance in the data set.

However, if the objective of PCA is to replace *n* variables by a smaller number (*m*) of principal components, it is important to know how small *m* can be without serious information loss. Various guidelines have been proposed for making the decision of how many principal components to retain:

i. The most commonly used criterion is probably Kaiser's rule. Kaiser (1960) proposed retaining only components with eigenvalues greater than 1 since components with eigenvalues lower than 1 extract less variance than any of the original variables.[15] Several studies have shown that the criterion is fairly accurate, especially when the number of variables is low (10–15) or moderate (< 30; Stevens 1992: 378). Blind use of Kaiser's rule can, however, lead to the preservation of principal components which only account for a very small percentage of the overall data variability (Stevens 1992: 378).
ii. A graphical method, the scree test, has been proposed by Cattell (1966): the eigenvalues of the principal components are plotted against their ordinal numbers (i.e. whether it is the first principal component, the second, etc.). Typically the graph shows a sharp drop at one stage and then levels off (see Figure 4.11). The interpretation of the scree graph is analogous to the interpretation of the amalgamation graph in cluster analysis: the point beyond which the graph levels out is identified as the last component to be retained (see the example below). Components identified by this method generally account for most of the variance in the data.
iii. A rather straightforward criterion for the retention of principal components is to define a percentage of variation (75, 80 or 90%), and to retain as many components as are necessary to account for the specified amount of variance.

These guidelines are *ad hoc* and their justification lies primarily in the fact that they are plausible and work in practice (Jolliffe 1986: 93). They are often used in combination, i.e. the analyst compares the results obtained by different criteria and then decides on the number of components to retain.

The interpretation of the components is based on the component loadings. These loadings indicate the strength of the correlation between original variables and components, and thus allow one to trace the structure of the relationships between variables. Often, however, the component loadings are not easily interpretable; in this case the analyst can employ certain techniques to simplify the structure of the component loadings. The most common technique is to rotate the eigenvectors in different directions without changing

the relative locations of the data points. So-called quatrimax rotations, for example, are performed so that each original variable loads mainly on one component (for further information on component rotation cf. Stevens 1992; Rietveld & Van Hout 1992).

So far PCA has been described as a tool for the grouping of variables, i.e. it enables the analyst to understand the correlations or relationships between the individual variables. However, PCA can also facilitate the classification of cases into groups. Given a high correlation between two or more variables, it can be concluded that they are quite redundant. In other words, if it is possible to summarize thirty variables into three components, the classification of cases can be based simply on their component values. The component scores can be used as input data for cluster analysis, or they can be represented in a scatter plot (as done by Horvath 1985).[16] The most important steps of PCA can be summarized as follows:

INPUT	Raw data matrix (or correlation matrix)	
	↓	
STEP 1	Computation of PCA	
	↓	
STEP 2	Review eigenvalues, scree test and % of variance accounted for and decide on how many components to retain	
	↓	↓
STEP 3	Review component loadings	Review component scores
	↓	↓
STEP 4	Rotation of factors (optional)	Scatterplot (s)
	↓	↓
STEP 5	Interpretation of factors as combinations of variables	Identification of clusters of cases (e.g. speakers)

Generally, sample size poses no problem for cluster analysis and multidimensional scaling as these methods are purely descriptive; however, the situation with regard to PCA is more ambiguous. Several popular guidelines suggest that the ratio of cases to variables is the crucial criterion (Horvath 1985: 56). There should be more cases than variables, and the recommended case/variable ratio varies from two cases per variable to 20 cases per variable (Stevens 1992: 384). However, a Monte Carlo study by Guadagnoli and Velicer (1988, reported in Stevens 1992: 384) indicates that the magnitude of factor loadings is more important than case/variable ratio considerations. Based on Guadagnoli and Velicer's study, Stevens (1992: 384) recommends the following guidelines:

a. Components with four or more loadings above .60 are reliable, regardless of sample size.
b. Components with about 10 or more low (.40) loadings are reliable if sample size is greater than 150.
c. Components with only a few low loadings should not be interpreted unless sample size is at least 300.

Furthermore, any component with at least 3 loadings over 0.80 will be reliable (Stevens 1992: 384). Regarding the validation of the results, the criteria mentioned with respect to cluster analysis, especially validation through external criteria and replication, are also used in PCA.

Example (Labov 1969)

The data set given by Labov for copula deletion in AAE has a case/variable ratio of only 1:1. However, the data show very high and therefore (probably) reliable component loadings and can be used to illustrate PCA. Only components above an eigenvalue of 1 were extracted (Table 4.6).

The structure of the scree plot also suggests that to retain two main components would be appropriate (Figure 4.11). Unlike the graph of amalgamation schedule (Figure 4.7), the scree plot is read from left to right.

The examination and interpretation of the (unrotated) component loadings constitutes the next step in the analysis (Table 4.7).

Table 4.6 Principal components (Eigenvalue > 0.5)

	Eigenvalue	**% of variance accounted for**
Component I	3.90	65
Component II	1.32	22

Table 4.7 Component loadings (unrotated)

	component I	**component II**
NP/full	0.07	–0.90
P/full	0.42	0.64
NP/contracted	–0.94	0.28
P/contracted	–0.98	0.01
NP/deleted	0.98	0.14
P/deleted	0.96	–0.07

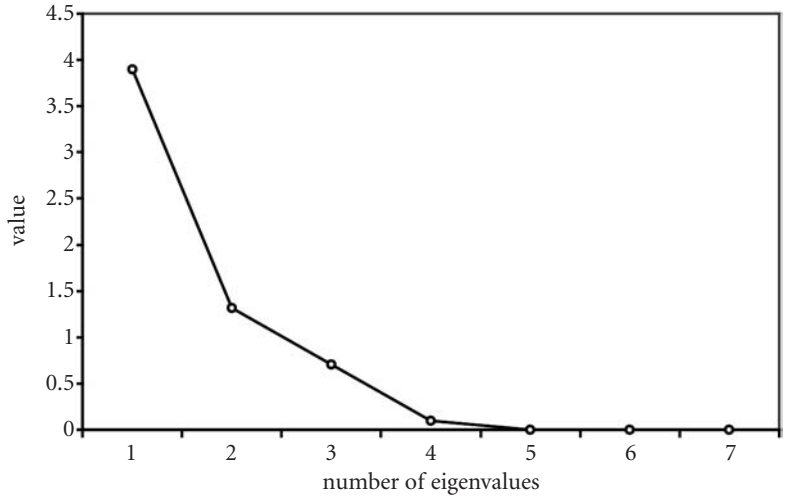


Figure 4.11 Scree plot

Component I shows a mixture of high positive and high negative loadings and can be described as a bipolar component, contrasting contraction and deletion of the copula. That is, a pattern of high scores on the deletion variables and low scores on contraction variables accounts for most of the variability in the data, thus separating the peer groups from both Adults and Inwood. The second component shows only one very high (negative) and one moderately high (positive) loading, both on the two variables which were not accounted for in component I. The loadings suggest that low scores on variable [NP/full] correlate with relatively high scores on variable [P/full]. The pattern can be clarified with the help of factor rotation (Table 4.8).

The bipolar structure of component I was not affected by the rotations. The structure of component II, however, is now more pronounced: component

Table 4.8 Component loadings: quatrimax rotation

	component I	**component II**
NP/full	0.06	–0.99
P/full	0.30	0.15
NP/contracted	–0.93	0.34
P/ contracted	–0.99	–0.05
NP/deleted	0.96	0.10
P/deleted	0.99	0.04

Table 4.9 Component scores (based on unrotated components)

	component I	**component II**
T-Bird	0.90	1.27
Cobras	0.71	–0.60
Jets	0.25	–0.91
OscarBros	0.45	0.09
Adults	–0.54	–1.31
Inwood	–1.77	0.91

II now loads highly (but negatively) only on variable [NP/full]. However, a component defined by only one loading is in fact not much of a component; in this case the actual variable could simply be substituted. Component I and II together account for roughly 87% of the variance. Finally, PCA can also be used to group cases on the basis of their component scores. The component scores for the data set are given in Table 4.9.

Component I shows positive scores for the preadolescent peer groups, i.e. they show high frequencies of copula deletion together with low frequencies of contraction, a pattern most visible in the linguistic performance of the T-Birds; Adults and particularly the Inwood control group show the opposite behaviour (negative loadings on component I). It has been mentioned above that high loadings on component II (unrotated) indicate low scores on variable [NP/full] together with comparatively high scores on variable [P/full]. Only T-Birds (and to a lesser degree Inwood) show this pattern. Jets, Cobras and Adults, on the other hand, demonstrate the opposite pattern, i.e. relatively high scores on variable (NP/full) and low scores on variable [P/full].

The component scores can now be used as the input for cluster analysis, or simply be plotted in a two-dimensional scatter plot. Figure 4.12 shows the scatter plot for the first two principal components (thus accounting for around 90% of variance in the data). Note that the internal structure of the pre-adolescent peer group cluster has changed when compared with the results from cluster analysis and multidimensional scaling, i.e. the cluster is less compact. The underlying mathematics of PCA influence the representation of the original data in scatter plots.

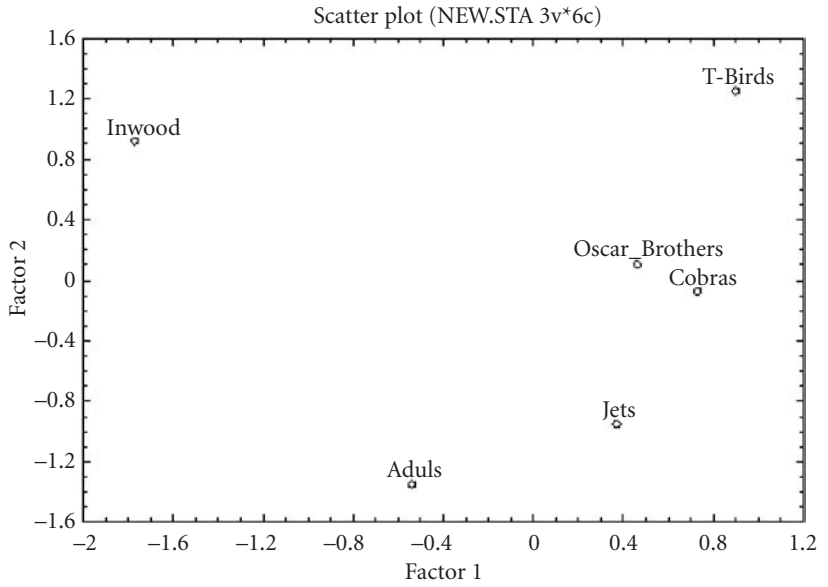


Figure 4.12 Scatter plot (based on component scores)

Focusing, diffusion and fixity: A statistical perspective

The multivariate techniques discussed in this chapter are best characterized as classificatory or taxonomic, grouping cases and/or variables on the basis of either similarity (cluster analysis and multidimensional scaling), or correlations (PCA). The importance of multiple similarity relations for the identification of linguistic sets or varieties was emphasized by Lieb (1993: 59–61), and Hudson (1996: 51, 68) has argued that language varieties are best defined in terms of co-occurrence rules (i.e. co-variation or positive correlation of variables; cf. also Labov 1971: 462; Horvath 1985: 154; Berutto 1987: 264). Statistical methods such as cluster analysis, multidimensional scaling and PCA translate such concepts into mathematical models and provide promising tools for exploring the structures of linguistic variation. A central advantage of multivariate clustering methods is that no *a priori* hypothesis regarding the existence of sociolinguistic groups is needed; groups are established purely on the basis of linguistic similarities between speakers. This approach differs from the classical Labovian tradition where the population is divided according to pre-established socio-economic categories, and it is taken for granted that these social groups share a common linguistic behaviour.

Similarity relations also form the basis for Le Page and Tabouret-Keller's (1985) sociolinguistic concepts of focusing and diffusion. As outlined in the introduction, focusing describes language change through convergence or levelling, that is, individuals gradually adjust their own speech patterns so that they resemble the speech of the larger group. Processes of linguistic focusing can be transitional (i.e. *ad hoc* linguistic accommodation in a given situational context; cf. Coupland 1984), but – if repeated across time and speakers – they can result in the stabilization of relatively uniform (spoken and/or written) group-based varieties (i.e. the data points representing speakers would move closer to one another in the multidimensional scatter plot). The term diffusion describes language or dialect contact scenarios where the pool of variants is maintained in its diversity, and intra- as well as inter-speaker language use is highly variable (the data points would be dispersed). Finally, modern codified and written standard languages are characterized by what Smith (1996:65–66) calls 'fixity', i.e. a belief that there exists 'a fixed collection of prescribed rules from which any deviation at all is forbidden' (p. 66; see Figure 4.13 for a schematic representation of these concepts). While focusing and diffusion are dynamic processes of language change, fixity is an abstract type, an ideal state (cf. Gloy 1997 on standard languages as 'Institutionen im Reich der Gedanken', 'institutions in the realm of thought'; and Milroy 2001:543 on standard languages as 'an idea in the mind'). In a scatter plot this would be represented by a near complete overlap of data points. Diffusion, focus(ing) and fixity are illustrated in Figure 4.13.

In the following chapters the patterns of linguistic variation in the *Corpus of Cape Dutch Correspondence* will be described with the help of uni-variate and multi-variate statistical techniques. The aim of the quantitative analysis is

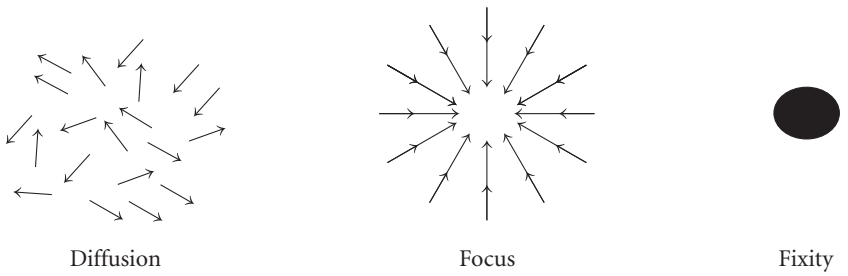


Figure 4.13 A schematic representation of diffusion, focus (language standards/incipient standard languages) and fixity (standard languages; adapted from Smith 1996:66)

twofold: (i) to provide detailed descriptions of individual variables and their variants (morphological, lexical and syntactic), and (ii) to illustrate processes of focusing and sociolinguistic norm development at the Cape before the onset of formal standardization efforts in the early twentieth century.

Notes

1. Smith (1989: 178–186) dismissed quantitative variation studies as a mere 'numbers game', irrelevant to linguistic theory whose sole concern is the understanding of language as a mental/psychological entity.

2. Early sociolinguistic work in the 1960s has relied almost exclusively on descriptive statistics. Inferential methods, however, have been used increasingly from the early 1970s (Davis 1982). The application of probability theory in variation studies is not limited to the use of general statistical techniques, but also led to the development of Varbrul. Varbrul (cf. Cedergreen & Sankoff 1974) assigns probabilities to the linguistic and social constraints which affect a variable rule, and thus redefines the notion of linguistic optionality in terms of statistical probability theory.

3. In more technical terms, inferential statistics are used to prove or disprove what is called the null hypothesis (H_o). The null hypothesis states that the relations observed in the sample do not exist in the population from which the sample is drawn. The alternative hypothesis (H_a) states that the distribution observed in the sample is also true for the population. The aim of inferential statistics is to accumulate enough evidence to reject the null hypothesis; a p-level of < 0.05 indicates that the probability of erroneously rejecting the null hypothesis is 5%.

4. Earlier statistical textbooks for students of linguistics (such as Butler 1985; Woods, Fletcher, & Hughes 1986) generally focused on uni- and bi-variate techniques, and offer little guidance to the student who intents to use multivariate procedures (cf. also Van Hout 1988: 978). The textbook by Rietfeld and Van Hout (1992), however, includes a comprehensive chapter on PCA, and Oakes *Statistics for Corpus Linguistics* (1998) covers multivariate clustering techniques.

5. Cluster analysis is known under a variety of names: numerical taxonomy, automatic classification, botryology, phenetics and typological analysis (Kaufman & Rousseeuw 1990: 3). An accessible description of cluster analysis for the mathematically untrained is given by Aldenderfer and Blashfield (1984), a more technical discussion with extensive references can be found in Arabie and Hubert (1995). Romesburg's *Cluster analysis for researchers* (1984) offers detailed information and gives a summary of the mathematics involved.

6. Other types of distance measures include: squared Euclidean distance, Manhattan (city block) distance, Chebychev distance, power distance.

7. The notion of 'natural' clusters does not imply a realist ontology, but simply reflects criteria such as separateness and density (cf. Panchen 1992: 148).

8. However, it should be noted that partitioning methods have not been studied in as much detail as hierarchical methods, and there is more insecurity about how one should interpret the results (Aldenderfer & Blashfield 1984: 45). Most statistical packages allow researchers to control the way in which the initial cluster centres are computed in *k*-means clustering. The following options are available: 'maximize between-cluster distances', 'sort distances and take observations at constant intervals' and 'choose the first N (number of clusters) observations' (for a discussion of the different methods see STATISTICA III: 3187–3188). For the analyses in this study the option 'maximize between cluster distances' was used as it yields well-defined and maximally separated clusters.

9. Labov's interpretation has since been modified by Baugh (1980) who showed that unlike standard English, AAE favours zero copula in adjective environments.

10. In STATISTICA standardization is computed as follows: Std. Score = (Raw Score – Mean)/Std. Deviation (STATISTICA 1995 I: 1369).

11. There are two different types of multidimensional scaling: metric and non-metric scaling. While the aim of non-metric scaling is to reproduce the rank-order of the original distances, metric scaling reproduces the actual distances. In this study non-metric scaling (Guttman Lingoes) was used.

12. Jackson's *User's guide to principal components analysis* (1991) gives a detailed description of the procedure, including its mathematical foundations. A more conceptual description is given in Stevens (1992). PCA differs from the related method factor analysis in that no assumptions are made about the distribution of the data. The results are thus purely descriptive. Factor analysis, on the other hand, assumes normal distribution and has an inferential aspect (for a detailed comparison of factor analysis and PCA, cf. Rietveld & Van Hout 1992: 251ff.).

13. Correlations can be positive (i.e. as one variable increases so does the other) or negative (i.e. as one variable increases, the other decreases). Correlation coefficients range from 1 (complete positive correlation) through 0 (no correlation) to –1 (complete negative correlation).

14. For details concerning the computation of eigenvalues, see STATISTICA (1994 III: 3231) and Jackson (1991: 7–10).

15. Many statistical programs use this criterion automatically. STATISTICA, however, gives the option of retaining factors with an eigenvalue < 1.

16. For a general discussion of the graphical representation of data obtained from PCA see Jolliffe (1986: 64–71). Note that principal components analysis tends to blur the relationship between clusters: widely separated clusters are normally maintained, while the distances between groups that are not widely separated are reduced (Aldenderfer & Blashfield 1984: 21).

Chapter 5

The gradualness of morphosyntactic change

> Boswell: Sir Alexander Dick tells me he remembers having a thousand people in a year to dine at his house …
> Johnson: That, Sir, is about three a day.
> Boswell: How your statement lessens the idea.
> Johnson: That, Sir, is the good of counting. It brings everything to a certainty which floated in the mind indefinitely.
>
> J. Boswell (1791). *The Life of Samuel Johnson*

Variation analysis: Some caveats

The patterns of sociolinguistic variation and the development of an Afrikaans proto-standard will be described with regard to three main points of reference: Early Modern Dutch and its dialects, nineteenth and twentieth century (standard and non-standard) Dutch, and nineteenth and twentieth century (standard and non-standard) Afrikaans. A range of morphosyntactic as well as morpholexical and syntactic variables, which are of diagnostic value for describing Afrikaans as a linguistic system distinct from Dutch, are investigated. These include: verbal inflection, tense marking, nominal gender agreement, use of Afrikaans pronoun variants, generalization of the long infinitive clause, use of the brace negation, the object marker *vir* ('for') and the possessive *se*. The variation analysis describes distribution patterns of individual variables in the *Corpus of Cape Dutch Correspondence,* and provides an overview of the structure of the Cape Dutch variation spectrum as well as the linguistic behaviour of writers.

The concept of the linguistic variable is central to quantitative sociolinguistics where it has been utilized as an analytical tool and a procedural device. Briefly, a linguistic variable is a linguistic unit consisting of two or more identifiable variants which constitute social and/or stylistic alternatives. Their referential meaning is, however, identical (Labov 1966; Wolfram 1993). Most sociolinguistic research has focused on phonological variables where the cri-

terion of referential sameness does not pose serious theoretical and practical problems. Phonological variables also have the additional advantage of high frequency. Whether linguistic variability of other sorts (morphological, syntactic, lexical, pragmatic/discourse) is equally amenable to variation studies utilizing the concept of the linguistic variable has, however, been questioned. While morphological variation (e.g. the absence or presence of a grammatical inflection as in *he goes* vs. *he go*) is generally seen as fulfilling the criterion of referential sameness, assumptions of syntactic or pragmatic equivalence, as well as lexical synonymy, have been controversial (cf. Lavandera 1978; Harris 1984; Winford 1984). Following Dines' (1980:15), variants of syntactic and lexical variables (which are discussed in Chapters 6 and 7) will be interpreted as being weakly equivalent, but as having nevertheless common semantic grounds.

Turning to the more technical issues of quantification, there is the question of how many tokens of a variable are needed to be confident that the variation patterns shown by a speaker are not random, but constitute a general and reliable pattern which indicates the speaker's variable competence. Romaine (1980) who discusses the issue in some detail, emphasizes that although research (e.g. Guy 1980) suggests that figures based on a minimum of 10 tokens per variable appear to be reliable for phonological variables (the ideal seems to lie around 30 token), it is not always possible to generalize this as a guideline: 'In some cases 1000 occurrences may not be enough, while in others fewer than 10, and even as few as two, might show contrastive patterns of usage' (Romaine 1980:192).

In this context the notion of salience, referring to socially marked variants (markers and stereotypes according to Labov's terminology) is worth noting. Dines (1980:16) has argued that in certain sociolinguistic and sociohistorical contexts a single occurrence of a variant can be 'socially diagnostic'. It is likely that the sociolinguistic salience of variables increases during processes of language standardization when certain variants are assigned to the new norm of the emerging standard language, while the non-selected variants are branded as 'vulgar' and 'dialectal' (cf. Stein 1994:7–9). Since variants belonging to the emerging system of Afrikaans featured prominently in popular discourse from the 1870s (see Chapter 2), a considerable degree of sociolinguistic salience can be assumed for these forms.

Only writers who showed a minimum of six tokens per variable were included in the distribution histograms (Chapters 5 and 6) and the final multivariate analysis (Chapter 7). The threshold agrees with Scholfield's (1991:381) guideline of a minimum of five occurrences per variable, but is lower than

Guy's (1980) recommended minimum of ten. However, most writers had considerably higher frequencies (well over 100 on some high frequency variables).

Variables (with the exception of objective *vir*) were standardized as percentages. Since sociolinguists often cannot avoid working with small numbers (especially in the context of historical linguistics, or when investigating syntactic variables), the practice of calculating percentages on the basis of comparatively low frequencies is a generally accepted practice (cf., for example, Mesthrie 1992 or Patrick 1995, 1999).

The verbal system

In (standard) Dutch the present tense paradigm for main verbs (1) has three distinct forms: verb stem, stem plus inflectional *-t* and stem plus inflectional *-e(n)* (pronounced as schwa; cf. Geerts et al. 1984: 431–434, 441–446; Booij & Van Santen 1995: 70–75). Final *-n* can be heard in front of vowels and in what is called 'reading pronunciation' (*leesuitspraak*) as well as in some regional varieties (Geerts et al. 1984: 424; Debrabandere 1997). Although verbal inflection codes meaning (i.e. person and number), it is essentially redundant in Dutch which is not normally a pro-drop language.

(1) Infinitive: *werk-e(n)*, 'to work'

First singular	*werk*	First/third plural	*werk-e(n)*
Second/third singular	*werk-t*	Second plural	*werk-e(n)*, *werk-t*

In written Dutch, inflectional *-t* is not added to verbs with a stem ending in *-t* (*zitte(n)/hij zit*, 'to sit/he sits'). Verbs with stems ending in *-d* do add the *-t* in the written language; this, however, does not affect their pronunciation (e.g. *rijde(n)/hij rijdt*, 'to ride/he rides'). Inflectional *-t* is dropped in the second person singular in syntactic inversion (*je leest een boek/lees je een boek?*, 'you read a book'/'do you read a book?'). The second person plural variant ending in inflectional *-t* is today rare and the form *-e(n)* is commonly used (Donaldson 1981: 113).

In addition there exists a small set of five high-frequency monosyllabic main verbs (*gaan*, 'go', *staan*, 'stand', *slaan*, 'hit', *doen*, 'do' and *zien*, 'see') and their derivatives (such as *verstaan,* 'understand', or *aanzien*, 'look at'). Their infinitive and plural forms do not end in *-e(n)* but in *-n*. Otherwise their conjugation is regular (2).

(2) Infinitive: *gaa-n*, 'to go'

First singular	*ga*	Plural	*gaa-n*
Second/third singular	*gaa-t*		

The present tense conjugation of the auxiliary *zijn* ('to be') has suppletive forms rather than inflections in the singular, while the paradigm of the auxiliary *hebben* ('to have') is inflectional (albeit irregular in the third person singular; see (3)). Note that Dutch has a second infinitive form (*wezen*) which can replace *zijn* in double infinitive constructions (cf. Donaldson 1981: 140; Geerts et al. 1984: 444, 578). The forms *bent/hebt* are variants of the second person plural but are less common than *zijn/hebben* (Donaldson 1981: 137, 139).

(3) Infinitive: *zijn (wezen)/hebben*

First singular	*ben/heb*	First/third plural	*zijn/hebben*
Second singular	*bent/hebt*	Second plural	*zijn/hebben; bent/hebt*
Third singular	*is/heeft*		

In spoken, informal Dutch the forms *je heb/ben* and *hij heef/heb* are often heard, and with the second person formal plural pronoun *u* both *hebt* and *heeft* as well as *bent* and *is* are acceptable in metropolitan Dutch (cf. Talen et al. 1908: 112f.; De Vooys 1953: 132–134; Geerts et al. 1984: 442, 444; on the acceptability of these forms in nineteenth century South African Dutch cf. Elffers 1893: 111f.).

The Dutch modal auxiliaries (with the exception of *willen,* 'to want/wish') belong to a historical class called preterite present verbs, which are inflected like the preterite of strong verbs in the present tense (De Vooys 1953: 127–129; LeRoux & LeRoux 1973: 173; see (4)).

(4) Infinitive: *mogen*, 'may'

Singular	*mag*	First/third plural	*mogen*
		Second plural	*mogen, mag*

For the second person plural two forms exist: *mogen/mag*. As in the case of *zijn* and *hebben*, the monosyllabic variant is less common than the variant ending in *-e(n)*. For *zullen* ('shall') and *kunnen* ('can') a variant ending in *-t* exists for the second person singular (*je zult*, *kunt*) and varies with the invariant singular form (*je zal, kan*). *Willen*, although not a preterite present verb, follows a similar inflectional paradigm, i.e. the third person singular is formed without inflectional *-t* and two forms exist for the second person singular (*je wil/wilt*; Talen et al. 1908: 114f.; Geerts et al. 1984: 442–446).

In comparison to its lexifier Dutch the verbal system of Afrikaans (5) is characterized by a drastic reduction in the number of forms. The forms of the Afrikaans present tense paradigm are based on either the verbal stem (main

verbs), the inflected but invariant present tense singular form (*hebben/zijn* and the modal auxiliaries), or the plural/infinitive form (verbs of the set *gaan*). The Afrikaans form of the auxiliary 'to have' is based on a variant of the third person: *het/heet* (<*heeft*). Use of *het* was widespread in seventeenth century Dutch and is still characteristic of some contemporary dialects of Dutch (Kloeke 1956; cf. also Van Loey 1959; Scholtz 1963; also Smits 1996: 130–131 on the use of *het* in Iowa Dutch). Some relics of the old inflectional system exist in contemporary Afrikaans, but they are rare and regularized forms are used in the spoken language (Grayson 1962; Ponelis 1993: 391ff.).

(5)

Infinitive: *werk*	infinitive: *wees/hê, het*[1]
present tense (singular/plural): *werk*	present tense (singular/plural): *is/het*
infinitive: *gaan*	infinitive: *mag*
present tense (singular/plural): *gaan*	present tense (singular/plural): *mag*

West-Germanic languages show a general tendency towards loss of material in unstressed syllables (cf. Ferguson 1995 for a discussion of loss of agreement as an example of Germanic drift; also Lass 1987: 330–331). Phonetic attrition in posttonic syllables has commonly been attributed to the adoption of initial expiratory stress (Meillet 1970: 53–55; Kastovsky 1994: 145–146). Within the Germanic family, English has lost all verbal inflection (except for the third person singular). Swedish and Norwegian (*Bokmål*) do not distinguish person and number in the present tense; however, they maintain the finite-infinite distinction. Loss of inflectional material is also common in spoken German whose standard (written) system shows comprehensive inflectional paradigms. The full pronunciation of *-e(n)* is often lost in speech (e.g. the plural verb in *wir kommen*, 'we come', would typically be realized as /kom:/; cf. Lütke 1989: 135). However, adoption of initial stress and phonetic attrition are not universally related, since, for example, Icelandic, Finnish, Hungarian and Czech show initial stress without reduction of inflectional suffixes (Roger Lass, personal communication).

According to the superstratist hypothesis of language change at the Cape (see Chapter 1), inflectional loss is a straightforward result of inter-dialect levelling in the early settlement, where it interacted with evolutive language change as outlined above. Variation in the present tense paradigm has been reported for early Dutch (cf. Raidt 1983: 118) and is prominent in the Dutch dialects, which utilize a variety of inflectional systems. Stem forms are used in the singular, for example, in the dialects of South Holland and De Velduwe. In

the plural, stem plus inflectional *-t* has been reported for Drenthe, the eastern part of Overijsel and the east of Gelderland. Stem forms in the plural are today common in De Veluwe and have also been reported for Zuthpen and Zelhem (Combrink 1978: 75–76; Smits 1996: 70–71; cf. also Van den Berg 1949; Hol 1955; Sayre 1972; Goeman 1976 [1983], 1993).

The existence of an unmarked verb phrase in Afrikaans has also been attributed to contact between colonial Dutch (the superstrate variety) and the local pidgin/creole variety spoken by the substrate population (Valkhoff 1966: 25; Den Besten 1989: 235–236; De Kleine 1997). Proponents of this position have emphasised not only the relative speed of the process, but also the fact that the Afrikaans verbal system is not a straightforward example of inflectional levelling and loss. That is, the fact that some verbs continue Dutch stems, others Dutch infinitives, or Dutch third person singular forms, has been interpreted as reflecting reanalysis of morpheme boundaries by non-native speakers who acquired the verb forms as holistic chunks (on the acquisition of inflected verb forms as 'chunks' during early second language acquisition see, for example, Diehl et al. 2000; cf. also Wray 2002). The use of invariant verbs in basilectal varieties was further supported by substrate influence since neither Khoe nor Malay or Creole Portuguese inflect verbs for person and/or number.

Large-scale morphological loss is not only characteristic of imperfect language learning and pidgin/creole development, but is a general feature of extraterritorial varieties. Reinecke (1937: 70–72) reported inflectional loss for the German dialect spoken in Virginia and for New Jersey Dutch. Maher (1984) has argued that languages used in extraterritorial settlements show striking similarities regarding the degree and nature of their morphological reduction. Thus, Louisiana French, Canadian French and Virgin Island French all show simplified pronominal systems, loss of inflectional endings in the verbal system, loss of auxiliary verb alternation in the perfect tense (*avoir* vs. *être*), and a predominance of periphrastic constructions over synthetic ones.[2]

New Netherland Dutch provides a point of reference. The social and regional origin of the settler population was very similar in the two settlements. However, settlers in New Netherland had only sporadic contact with non-European groups and the number of slaves was much lower than at the Cape (Buccini 1996). Although New Netherland Dutch shows clear signs of morphological reduction, the process was never completed: Prince (1910) described the marking of person/number as variable and Van Ginneken (1928: 292–301) remarked that generalized loss of inflection was prominent only in the varieties spoken by ex-slaves (cf. also Van Marle & Smits 1993: 321, 323–324).

Buccini (1992) summarized the linguistic situation in New Netherland Dutch as follows:

> While verbal inflection was reduced … the reduction was largely phonologically motivated: *the principle remained until the end.*
>
> (my emphasis; also Buccini 1996: 44)

The *Corpus of Cape Dutch Correspondence* comprises over ten thousand finite present tense forms (n = 11231). Forms which are identical in Dutch and Afrikaans and which cannot be classified unambiguously as representing either system, were excluded from the analysis. This affected the following forms:

i. first person singular of main verbs (Dutch *ik werk* – Afrikaans *ek werk*), as well as second and third person singular of main verbs with stems ending in *-t/-d* (i.e. which either do not inflect in the singular or inflection is not reflected in the spoken language);
ii. singular forms of the modal auxiliaries (Dutch *ik zal, je zal/zult, hij zal* – Afrikaans *ek, jy, hy zal*);
iii. third person singular for *hebben/zijn* (Dutch *hij heeft (*dialectal *het)*; *is* – Afrikaans *hy het; is*);
iv. plurals and infinitives of the verbal set *gaan* (Dutch *gaan* – Afrikaans *gaan*).

This left 5118 finite verb forms (1296 *hebben*, 871 *zijn*, 494 modal auxiliaries, 1905 main verbs and 552 verbs of the set *gaan*), and over half of these (54%) showed morphological reduction in accordance with the emerging system of Afrikaans and would have been ungrammatical in Dutch. The degree to which the individual writers in the sample used a verbal system conforming to the emerging norms of Afrikaans can be summarized by using a block histogram representing a grouped frequency distribution (Figure 5.1). The height of the blocks in the diagram is proportional to the number of individuals in each class interval; e.g. the block furthest to the right indicates that 34 writers conform to the emerging system of Afrikaans with over 90% of their present tense verbs forms.

The linguistic behaviour of writers is polarized as well as continuous. Two large clusters are situated on either side, that is, many of the writers conform closely either to the norms of metropolitan Dutch (i.e. maintain the person/number distinction with few exceptions; about one-third of writers fall into this category), or to the emerging system of Afrikaans (i.e. show almost complete loss of the person/number distinction; about one-quarter of writers fall into this category). Nevertheless, variable usage is pronounced and about half of the writers are located in the middle field of the distribution; that is,

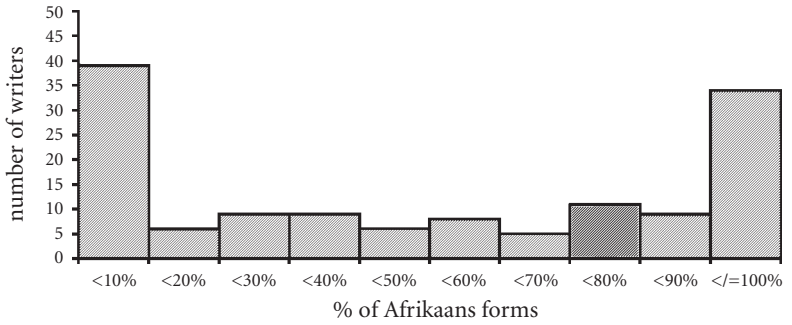


Figure 5.1 Histogram summarizing the degree of inflectional loss (person and number distinction in the present tense) by individual writers ($n = 136$) (all verbs)

their language use represents neither the traditional norms of Dutch nor the new norms of Afrikaans.

Further analysis of verb forms in the *Corpus of Cape Dutch Correspondence* shows that with regard to inflectional reduction (statistically significant) differences exist between verb classes. The numerical differences between modals, the auxiliary *hebben* and main verbs are, however, marginal when compared to the considerable gap between this group and *zijn*/*gaan*. The quantitative data given in Table 5.1 translate into a smooth, yet rather flat S-curve, illustrating the spread of the change (loss of the person/number distinction) across environments (verb classes).

Unlike in earlier Cape Dutch texts, there is no longer a clear quantitative difference between singular and plural verbs. In the case of main verbs 57% of singular verbs and 53% of plural verbs show inflectional reduction. In the Duminy diary (1787) stem forms were still rare in the plural, but common in the singular (cf. Franken 1953: 171). In her work on Iowa Dutch Smits (1996: 121) observed a similar pattern, i.e. inflectional reduction spread from singular to plural across time. Smits interpreted this as an attempt to maintain surface regularity:

> The marginal occurrence of verb stems in the plural may be related to the fact that in that context the generalization of monosyllabic and consonant-final stems affects surface regularity. Prototypical plural verb forms, namely, should be polysyllabic and end in *-en*. The use of verb stems in the 2nd and 3rd person singular, however, does not affect surface regularity, since prototypical singular forms are monosyllabic and end in a consonant.

Table 5.1 Use of Afrikaans forms according to verb classes in the *Corpus of Cape Dutch Correspondence.* The distribution is statistically significant at $p < 0.001$ (chi square).

	modal	**hebben**	**main**	***zijn***	**type-*gaan***
total	494	1296	1905	871	552
Afrikaans forms	295	746	1065	402	228
Afrikaans forms as % of total	60%	58%	56%	46%	41%

Table 5.2 Loss of verbal inflection for *schrijven* ('to write'), *leven* ('to live'), *regenen* ('to rain') and *beginnen* ('to begin') in the *Corpus of Cape Dutch Correspondence.* Frequency differences are not statistically significant.

	schrijve(n)/skryf	***leve(n)/leef***	***regent/reën***	***begint/begin***
total	48	15	20	41
stem forms	17	6	6	21
stem forms as % of total	35%	40%	30%	51%

Variation within the paradigm of main verbs is only marginally due to invariant inflected forms which compete (even in modern Afrikaans) with stem forms. According to Ponelis (1993: 393), such inflected forms were common until three generations ago, but their use has since been greatly reduced (although they still occur in non-standard varieties; cf. Van Rensburg 1984). Variation between stem forms and inflectional relic forms affects verbs with *f/v* or *d/t* voicing alternation (*skryf/skrywe*, 'write', *bid/bidde,* 'ask for'),[3] /x/ syncope (*kla/kla'e,* 'complain', from Dutch *klagen*), and verbs with stem final resonant (*reën/reënt,* 'rain', *begin/begint,* 'begin'). Use of these forms is lexically constrained and influenced by social and contextual variables such as region, age and speaker preference (Ponelis 1993: 391–393; Malherbe 1917: 55; Hoogenhout 1904: 13).

The data summarized in Table 5.2 shows that in the *Corpus of Cape Dutch Correspondence* stem forms were indeed less frequent in this sub-group of main verbs. The exception was *beginnen* which showed a close to average frequency for inflectional loss. However, lexical constraints also existed for verbs with stem-final voiceless obstruents (Table 5.3), which in modern Afrikaans occur only as stem forms. This suggests that loss of verbal inflection at the Cape might (in principle) be amenable to a lexical diffusionist explanation.

Although examples of the new system abound in the corpus, non-Afrikaans concord disagreement (as described for eighteenth century Cape Dutch Vernacular) still occurs (see examples (6)–(8)). However, in terms of

Table 5.3 Loss of verbal inflection for *maken* ('to make') and *spreken* ('to speak') in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	make(n)/maak	***spreke(n)/spreek***
total	45	22
stem forms	32	6
stem forms as % of total	71%	27%

quantity, such instances of 'misuse' of inflection (cf. Smits 1996: 43) are negligible (265 instances, i.e. 2% of all finite verb forms).

(6) **first person subject + second/third person verb:**
   a. ik vraag**t** ('I ask'; Wilhelmina January, 30/3/1897)
   b. ik denk**t** ('I think'; Cornelia Stanford, 6/8/1905)

(7) **singular subject + plural verb/infinitive:**
   a. dat de Doctor hunne voeten snyd**en** en reg draai ('that the doctor cuts their feet and twists them right'; B. J. Brümmer, 21/6/1900)
   b. ik moet**en** ('I must'; Mozes Mlimkulu, 1/12/1896)

(8) **plural subject + singular verb:**
   a. motor cars kom**t** ('motor cars come'; Wynanda Hoogenhout, 18/2/1907)
   b. ik en myn echtgenoote **ben** ('I and my spouse are'; M. J. Heroldt, circa 1892)

There is, however, a small group of four individuals (Beatrice de Waal, Wilhelmina January, Mozes Matodlana, Jacobus November) who show percentages of between 15%–30% of 'misuse' of inflection, including the frequent use of *-e(n)* with singular verbs which is otherwise extremely rare. The linguistic behaviour of this small group is best described in terms of hypercorrection, i.e. the overgeneralization of linguistic features which are perceived as prestigious, yet the writers' imperfect knowledge of the Dutch standard prevented their use of the correct forms. Three of these writers belong to the coloured community and their texts were produced in the context of the paternalistic, religious authority of the mission station (brethren-minister/missionary).

In the corpus, the new verbal paradigm is not yet fully crystallized for the set *gaan*, *slaan*, *doen*, *staan* and *zien* (see (9)–(11)). Variation between inflectional *-n* and *-t* is common, and forms ending in *-t* (e.g. *gaat*) were frequently generalized to the first person singular and the first/second and third person plural. Stem forms (e.g. *ga*), on the other hand, occurred rarely in the plural or

the second/third person singular (cf. also Changuion 1844: 121; Pannevis 1880 [1971]; 78, Mansvelt 1884 [1971]: 158).[4] Variation between *-n* and *-t* has been reported for non-standard Dutch and is still attested in Afrikaans for older speakers (Ponelis 1993: 391; Donaldson 1993: 218).

(9) a. maar ik **ziet** geen kans ('but I see no chance'; B. J. Brümmer, 26/4/1902)
b. ons **gaat** almal naar Stil Baai ('we are all going to Stilbaai'; J. F. W. Grosskopf, 8/12/1907)

(10) a. [de] paarden die men hier **zie** is groot maar te lang van rug ('[the] horses which one sees here are large but have a long back'; Louis Botha, 12/8/1909)
b. aanstaande week **ga** Danie Mev. Retief en kinderen halen ('next week Danie will fetch Mrs Retief and children'; Annie van Huyssteen, 23/7/1907)

(11) a. liefste Boykie, jy **zien** ik is by Annie ('dearest Boykie [term of endearment], you see I am at Annie's'; Johanna Brümmer, 10/5/1895)
b. als het mogelyk komt voor hy **gaan** ('if it is possible come before he leaves'; M. C. Breuris, 22/1/1906)

If one accepts the traditional dating for loss of person and number distinctions in the present tense paradigm to between 1750 and 1775, then the abundance of forms in the *Corpus of Cape Dutch Correspondence* is surprising. It is questionable whether such variation patterns can be explained simply as a result of dialect contact under diglossic conditions as argued, for example, by Scholtz and Raidt (i.e. contact between a relatively stable spoken norm of Cape Dutch Vernacular characterized by complete inflectional loss, on the one hand, and metropolitan standard Dutch with full inflectional paradigms, on the other hand). A similar line of thinking was suggested by N. M. Hoogenhout (1904: 13) in his early grammar of Afrikaans:

> In Gebeten usw. versucht man ausnahmslos seine Sprache, so gut oder so schlecht es geht, der niederländischen Schriftsprache anzupassen. Daher eine außerordentlich bunte Formenmischung.
>
> ['In prayer one tries without exception to accommodate one's language to the Dutch written norm. Thus a very colourful mixture of forms.']

In other words, the church and an increasingly elaborate school system are believed to have brought members of the Cape Dutch speech community into regular, if limited, contact with the norms of metropolitan standard Dutch,

leading to linguistic insecurity, random variability and *ad hoc* dialect mixture (via borrowings from the standard language) in formal situations.

However, this interpretation is disputable as misuse of inflection (that is, hypercorrection) is rare in the historical record. Trudgill (1996; cf. also Bailey 1980: 245) has suggested that in the absence of hypercorrection, the observed variation patterns can be interpreted as constituting 'a permanent and continuing feature' of the language or dialect. It appears that even around 1900 we are still observing a 'mediating system' (characterized by variable inflectional loss) located somewhere on a continuum between metropolitan Dutch and modern Afrikaans.

Revisiting Conradie (1979)

The apocope of [t] – An example of morphophonemic variation

Loss of verbal agreement interacted with the phonological process of *t*-apocope, characteristic not only of Cape Dutch Vernacular but also of early Dutch and today's Dutch dialects (especially Holland and Utrecht; cf. De Vries 1974; Goeman 1983 [1976]). *T*-apocope has been reported for upper class varieties of sixteenth and seventeenth century Dutch. Its prominence in Cape Dutch Vernacular suggests that Dutch prestige forms played a role in the early development of Afrikaans (Snyman 1979; Stoops 1980; Ponelis 1993: 154–155).

Within the verbal system *t*-apocope affected the finite verb (second/third person singular present tense) and the (regular and irregular) weak past participle. Three types of verbs can be distinguished in Dutch: weak and regular (*maken/gemaakt*, 'make/made'), weak and irregular (with stem mutation, *kopen/gekocht*, 'buy/bought'), and strong (with *-e(n)* inflection and ablaut, *schrijven/geschreven*, 'write/written'). Afrikaans has maintained the prefix *ge-* as a past participle marker, but lost the inflectional ending (*maak/gemaak*, 'make/made'). In both Dutch and Afrikaans the prefix *ge-* does not precede atonic verbal prefixes (thus *ontmoeten/ontmoet,* 'meet/met', *beloven/beloof(d),* 'promise/promised'). However, regularized forms such as *gebegin* ('begun') or *gevergeet* ('forgotten') have been reported for Griqua Afrikaans and other non-standard varieties (cf. Rademeyer 1938: 62–63; Loubser 1954: 102; Tiflin 1984: 62; Webb 1993: 163). Only few of such forms were found in the corpus:

(12) a. die wordt hem **geontnomen** ('they were taken away from him'; Anton April, 6/7/1913)

b. net nou uwe brief **geontvangen** ('just now received your letter'; B. J. Brümmer, 19/9/1907)

T-apocope in Cape Dutch Vernacular was a variable rule in the Labovian sense; that is, phonological and grammatical factors ([±voice] of the preceding segment and word class, i.e. finite verb, past participle or non-verbal item) furthered or inhibited the application of the rule (Scholtz 1963: 15–17; Conradie 1979: 117–123; Lubbe 1983). In his diachronic study of the past participle in Afrikaans, Conradie (1979: 92, 1981, 1994) analyzed the morphophonemic conditioning of *t*-apocope quantitatively by dividing his data (17 texts or text complexes covering the period from 1739–1851) according to the following categories:

i. third person singular present tense forms with stems ending in [–voice]
ii. past participles with stems ending in [–voice]
iii. non-verbal items ending in [–voice]
iv. third person singular present tense forms with stems ending in [+voice]
v. past participles with stems ending in [+voice]

The overall pattern of Conradie's analysis shows that *t*-apocope is more common after voiceless consonants (obstruents, i.e. fricatives, plosives and affricates) than after voiced consonants (resonants, i.e. laterals and nasals). It is also more frequent in the third person singular present tense than in the past participle and non-verbal items; the latter never showing *t*-apocope after [+voice]. Regarding the development of the morpho-phonemic variation pattern, Conradie identified three distinct stages on the basis of his analysis. The three stages are typologically rather than chronologically successive. The texts which conform to stage A were written between 1797 and 1838, those belonging to stage B between 1829 and 1851, and those belonging to stage C between 1830 and 1854.

Stage A The phonological factor supports *t*-apocope after [–voice] in all word classes, while the grammatical factor supports *t*-apocope in the third person singular present tense only.

Stage B Largely a continuation of stage A but the grammatical factor now also supports apocope for past participles. This leads to a small increase of *t*-apocope in the past participle especially after [+voice].

Stage C The phonological factor no longer plays a role for *t*-apocope in the verbal system (present tense and past participle). *T*-apocope is, however, still more common in the third person present tense than in the past participle.

Conradie (1979:29) suggests that at a later stage *t*-apocope within the past participle caught up with the present tense paradigm, thus leading to the regularization of *t*-apocope for inflectional [t]. Eventually loss of [t] after [–voice] would be generalized to non-verbal items. However, as yet non-verbal items in which voiced consonants precede [t] remain unaffected in (Standard) Afrikaans (but cf. Ponelis 1993:155 for historical and dialectal examples of *t*-apocope after resonants in non-verbal items).

Frequency of *t*-apocope was analyzed in the *Corpus of Cape Dutch Correspondence* based on the five categories identified by Conradie. A comparison of the scores given by Conradie for the *Kaapse Taalargief* (1712–1831) and his own text collection (1779–1854) with those for the *Corpus of Cape Dutch Correspondence* indicates a continuation of the phonologically and grammatically conditioned variation until after 1900 (Figure 5.2). This suggests, in opposition to the traditional view – which sees the process as completed before 1800 (cf. Raidt 1983:80–82, Ponelis 1993:155) – continuing variation or slow linguistic change. While frequency of *t*-apocope in the corpus lags behind Conradie's text collection (with the exception of past participles ending in [–voice]), apocope is considerably higher than in the *Kaapse Taalargief*. Most noticeable is the increase in the past participle after [+voice] for the 1880 to 1922 data.

When interpreting the evidence one should keep in mind that the three corpora differ diachronically and stylistically. As outlined in Chapter 1 and 3, the *Kaapse Taalargief* includes official letters and reports sent by field cornets to local magistrates. Conradie's corpus is heterogeneous, comprising official VOC reports, the Duminy and Trigardt diaries, individual letters from

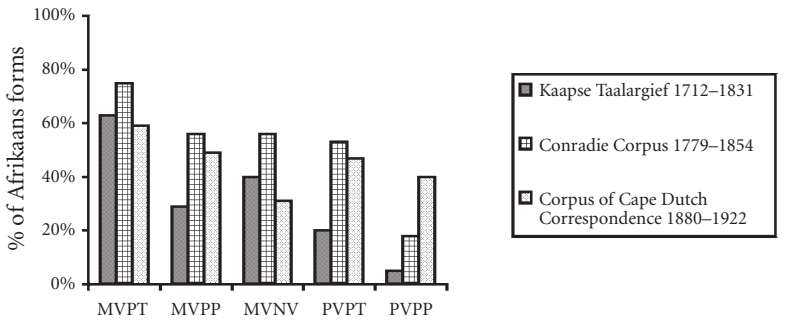


Figure 5.2 Apocope of [t]according to phonological and grammatical constraints: change in real time 1712–1922. MVPT = minus voice, present tense; MVPP = minus voice past participle; MVNV = minus voice nonverbal; PVPT = plus voice present tense; PVPP = plus voice past participle.

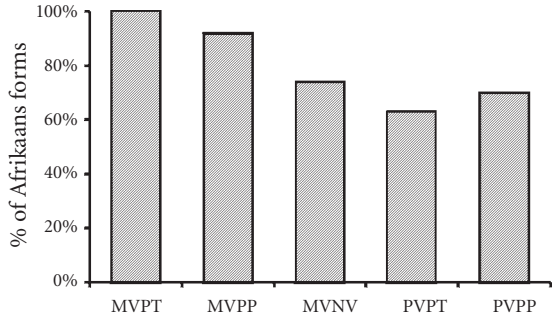


Figure 5.3 Apocope of [t] according to phonological and grammatical constraints: Johanna Brümmer (born 1877)

the *Kaapse Taalargief* and also examples of deliberate dialect writing (by De Lima, Meurant and Boniface; cf. Chapter 2), where *t*-apocope functioned as a sociolinguistic stereotype.

The graphs for individual writers indicate the existence of a finely graded continuum of *t*-apocope in the *Corpus of Cape Dutch Correspondence*: ranging from almost complete regularization (Figure 5.9) to relatively conservative patterns (Figure 5.8) where *t*-apocope was still limited to present tense forms ending in [–voice]. For some writers the grammatical constraint (present tense vs. past participle) had weakened, whereas the phonological constraint clearly retained its strength. This variation pattern was typical for a number of writers belonging to age group 2/class 2, i.e. young members of the emerging professional classes. An example is given in Figure 5.3.

The pattern illustrated in Figure 5.3 cannot be interpreted as 'following' Conradie's stage C which was already characterized by loss of the phonological constraint. Rather, it suggests the possibility that after Conradie's stage B the typological and/or diachronic development bifurcated into two distinct variation patterns:

i. pattern C[1] (Conradie's C) shows disappearance of the phonological constraint of [±] voice, but maintenance of the grammatical constraint (present tense vs. past participle);
ii. a second pattern, C[2], was characterized by the continuation of the phonological constraint and a weakening of the grammatical constraint.

A very advanced pattern, which could be termed stage D (if such a label were indeed necessary), is evident in Francois S. Malan's language use. Here

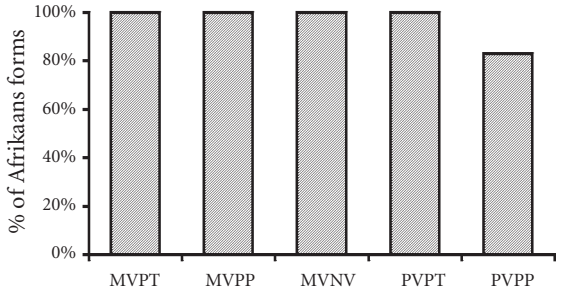


Figure 5.4 Apocope of [t] according to phonological and grammatical constraints: F. S. Malan (born 1871)

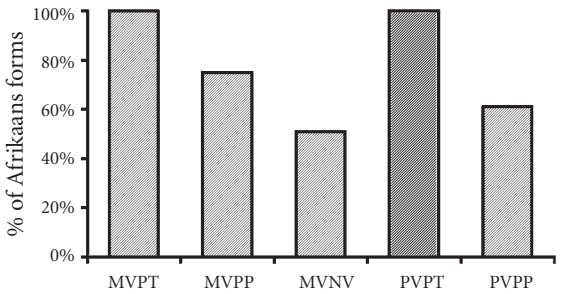


Figure 5.5 Apocope of [t] according to phonological and grammatical constraints: Petronella van Huyssteen (born 1854, white, class three)

t-apocope is nearly categorical (Figure 5.4; see Chapters 8 and 9 for a more detailed discussion of the letters of F. S. Malan).

An almost perfect example of C[1] is found in the linguistic behaviour of Petronella van Huyssteen, B.J. Brümmer and Eliabeth Eksteen (Figures 5.5, 5.6, 5.7; all three belong to age group 1 and class 3a, i.e. rural petty bourgeoisie): maintenance of the grammatical constraint but increasing loss of the phonological constraint. However, the three graphs differ strongly in terms of frequency of *t*-apocope: while Petronella van Huyssteen already shows categorical generalization of *t*-apocope in the present tense, B. J. Brümmer and Elisabeth Eksteen show more intermediate frequencies of *t*-apocope. Nonverbal items show an unusually high frequency of *t*-apocope in the letters of B. J. Brümmer. A conservative pattern exists in the letters of Louis Botha (also class 3a; Figure 5.8) where *t*-apocope is frequent only in the present tense after [–voice].

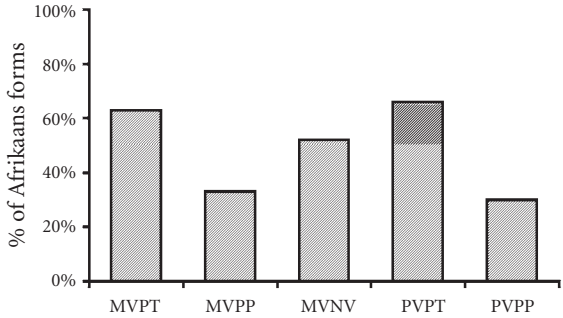


Figure 5.6 Apocope of [t] according to phonological and grammatical constraints: B. J. Brümmer (born 1841)

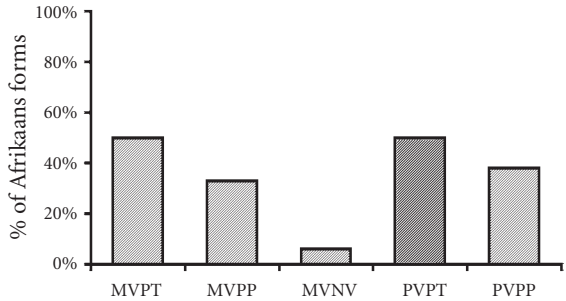


Figure 5.7 Apocope of [t] according to phonological and grammatical constraints: Elisabeth Eksteen (born 1860)

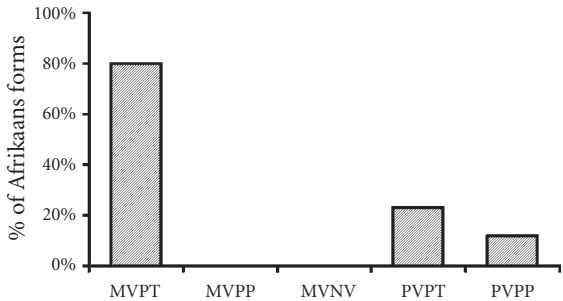


Figure 5.8 Apocope of [t] according to phonological and grammatical constraints: Louis Botha (born 1862)

Apocope of *-e(n)*

Unlike *t*-apocope, loss of word-final *-e(n)* – which within the verbal system affected the plural forms of the present tense, the infinitive and the past participle of strong verbs – was not the result of a phonological process. According to Scholtz (1963: 17–37), apocope of *-e(n)* gradually spread through the grammatical categories, leading to inflectional loss in the plural of the main verb by 1750, the infinitive around 1800, and finally the past participle (in predicative function) by the mid-nineteenth century (cf. also Raidt 1983: 124–127). During the last quarter of the nineteenth century a formal distinction arose between predicative and attributive past participles, with the latter maintaining the inflected form (Conradie 1979, 1982). The maintenance of the (original) inflected form in secondary functions (13) illustrates Kuryłowicz's fourth law of analogy: 'When a form undergoes differentiation as a consequence of a morphological change, the new form corresponds to its primary function and the original form is reserved for the secondary function' (quoted in Anderson 1992: 370).

(13) a. hy het die blompot **gebreek** ('he broke the flower pot')
b. die **gebroke** blompot ('the broken flower pot')

According to Conradie's (1979: 173) analysis of eighteenth and early nineteenth century documents, frequency of *e(n)*-apocope was roughly similar for infinitives and (predicative) past participles, while apocope in the present tense plural was far ahead of both (average for all texts: present tense 77%, infinitive 17% and past participle 19%). This frequency distribution suggests that the rate of change was roughly the same in the infinitive and the past participle. In the *Corpus of Cape Dutch Correspondence*, however, a (statistically significant) cline can be observed: verb (56%, $n = 1168$) > infinitive (40%, $n = 4129$) > past participle (32%, $n = 1104$). This supports Scholtz's original hypothesis of a three-tied grammatical constraint (Figure 5.9).

The past participle of strong verbs in Dutch is not only characterized by the inflectional ending *-e(n)*, but, for the majority of strong verbs, also by ablaut which was lost in Afrikaans (cf. Dutch *spreken/gesproken*, 'speak/spoken', Afrikaans *spreek/gespreek*). Psycholinguistic evidence suggests a fundamental processing difference between inflectional material and ablaut. Whereas inflections are relatively easy to process (as they are simply additions of linguistic material in previously empty positions), ablaut is the result of a substitution process in which the base vowel must be supressed in order to be replaced by the ablauted vowel (Berg 1998: 140–142). Processing difficulties of ablaut

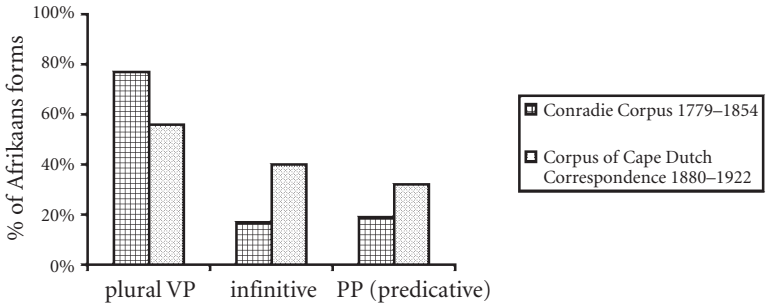


Figure 5.9 Apocope of *-e(n)* for plural verbs (present tense), infinitives and past participles change in real time, 1779–1922

substitutions are indicated, for example, by 'slips of the tongue' which frequently show overapplication of the base vowel, but only rarely hypercorrect ablaut vowels. Considering such processing difficulties as well as the fact that ablaut is not the primary marker of the past participle in strong verbs (which are unambiguously marked by the prefix *ge-* as well as the inflectional ending), the diachronic stability of ablaut in the Germanic group is surprising. It has been explained with reference to its high cue reliability (i.e. in the verbal system it is an unambigious marker of pastness), its occurance in high frequency verbs, and the fact that it is a phonologically straightforward process (cf. Berg 1998: 143). Den Besten (1989: 35) has argued that since ablaut is a relatively stable phenomenon in the diachrony of Dutch and its dialects, loss of ablaut in Afrikaans cannot be attributed to dialect levelling. Instead, he suggests gradual influence from a stable Cape Dutch Pidgin which had lost ablaut entirely. The assumption that ablaut was lost quickly in the basilectal contact variety is supported by evidence from non-native varieties of German in Namibia which also have a Khoe substrate (cf. Deumert 2003d). Non-native varieties of Namibian German have been attested from the beginning of the twentieth century and are still used by older speakers. Basilectal forms of these varieties show extensive loss of ablaut (which also extends to nominal derivations: *sprechen/gesprechen/die sprech*, 'to speak/spoken/the language', Standard German *sprechen/gesprochen/die sprache*), but maintenance of inflectional *-e(n)* (which in the usage of many speakers shows generalization to the weak past participle, e.g. *gearbeiten*, Standard German *gearbeitet*, 'worked').

Since regularization of the past participle involved loss of inflection as well as ablaut, it is important to clarify the relationship between the two processes, that is, whether inflectional endings and ablaut were lost at the same time, or

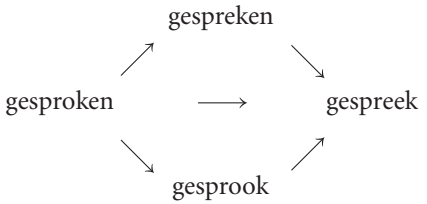


Figure 5.10 Possible lines of development for the regularization of the past participle of strong verbs

whether one was lost prior to the other (i.e. whether there was an intermediate stage; Scholtz 1963: 35). The possible stages of the regularization process are summarized in Figure 5.10.

Conradie's (1981) analysis indicates that the path of development was different for those strong verbs with a back vowel as ablaut vowel in the past participle (group I; <i>, <e> → <o> and <ē>, <ui>, <ie> → <ō>), and those with a front vowel (group II; <ij>, <i> → <ē>). In the corpus 259 examples of group I past participles, and 231 examples of group II past participles were found and tabulated according to the following categories:

		group I	group II
A	both ablaut and *-e(n)* maintained	*gesproken*	*geschreven*
B	ablaut lost, but *-e(n)* maintained	*gespreken*	*geschrijven*
C	ablaut maintained, but *-e(n)* lost	*gesprook*	*geschreef*
D	both ablaut and *-e(n)* lost	*gespreek*	*geschrijf*

The quantitative analysis of the corpus data (Tables 5.4 and 5.5) supports Conradie's interpretation. Regularization followed different paths in group I and II, and verbs with a front ablaut vowel in the past participle (group II) regularized earlier than those with a back ablaut vowel. Only group II verbs show a significant percentage of past participles which have lost the ablaut vowel but maintained the inflectional ending (which agrees with Conradie's analysis). This suggests that regularization involved an intermediate step only within

Table 5.4 Regularization of the past participle, group I (<i>, <e> → <o> and <ē>, <ui>, <ie> → <ō>). *N* = 259 in the *Corpus of Cape Dutch Correspondence*.

	A type: *gesproken*	B type: *gespreken*	C type: *gesprook*	D type: *gespreek*
no. of tokens	213	2	1	43
% of total	82%	0.7%	0.3%	17%

Table 5.5 Regularization of the past participle, group II (<ij>, <i> → <ē>). *N* = 231 in *the Corpus of Cape Dutch Correspondence.*

	A type: *geschreven*	B type: *geschrijven*	C type: *geschreef*	D type: *geschrijf*
no. of tokens	126	28	4	73
% of total	54.5%	12%	1.5%	32%

Table 5.6 Regularization of the past participle for *schrijven* ('write') and *krijgen* ('get') in the *Corpus of Cape Dutch Correspondence.* The distribution is statistically significant at $p < 0.001$ (chi square).

	A type: *geschreven*	B type: *geschrijven*	C type: *geschreef*	D type: *geschrijf*
schrijven (n = 104)				
no. of tokens	60	27	1	16
% of total	58%	26%	1%	15%
krijgen (n = 78)				
no. of tokens	35	1	2	40
% of total	45%	1%	3%	51%

group II. Verbs belonging to group I possibly maintained the ablaut vowel when group II verbs lost it, but (in analogy with forms from group II) at a later time lost both ablaut and inflectional ending simultaneously. The diachronic stability of back vowels may be due to an iconic constraint as the relationship between back vowels and pastness/distance has been described as non-arbitrary (cf. Louden 2000; Abraham & Conradie 2001: 40). The early loss of front vowels was possibly supported by their lack of association with pastness/distance.

Frequency of reduction was also lexically constrained (see Table 5.6). The distribution for the verb *schrijven* ('write'; group II), for example, shows only 15% for the loss of both ablaut and *-e(n)*. Loss of ablaut but maintenance of the inflectional ending occurred in about a quarter of cases. This contrasts with the frequency pattern for *krijgen* ('get'): just over half of the tokens in the corpus have lost ablaut as well as inflectional marking, and there is only one instance of category B (loss of ablaut, maintenance of *-e(n)*).

Conradie's quantitative and variationist analysis is a convenient point of reference. A comparison of Conradie's results with the variation patterns around 1900 (i.e. roughly a century later) suggests continuing variation characterized by gradual morphological erosion and regularization, rather than completed language change by 1750 or even 1800. However, while the over-

all patterns remained stable, the relative frequencies increased. This suggests that language change continued, although possibly at a slower pace than traditionally assumed by historians of Afrikaans.

Tense marking

Past tense variation

Dutch has a grammaticalized distinction between preterite and perfect which was lost in Afrikaans. In Dutch the preterite of main verbs is expressed by verbal inflection, the perfect and pluperfect by periphrastic constructions (inflected auxiliary verb plus past participle). The preterite of regular verbs shows the ending *-te* (after voiceless segments) or *-de* (after voiced segments) in the singular. The plural endings are *-te(n)/-de(n)*.[5] The preterite of strong irregular verbs is marked by a change in the stem vowel (ablaut) as well as inflectional *-e(n)* in the plural. Weak, irregular verbs inflect according to the regular paradigm, and show stem mutation (Geerts et al. 1984: 430; Booij & Van Santen 1995: 71–72).

(14) **Preterite**

I. Regular verbs

koken / ik kookte/ wij kookten, 'to cook/ I cooked/ we cooked'

horen/ ik hoorde/ wij hoorden, 'to hear/ I heard/ we heard'

II. Irregular verbs

a. weak irregular

kopen / ik kocht/ wij kochten, 'to buy/ I bought/ we bought'

b. strong irregular

zingen/ ik zong/ wij zongen, 'to sing/ I sang/ we sang'

Present Perfect

ik heb gekookt/ gekocht/ gezongen, 'I have cooked/ bought/ sung'

Past Perfect (Pluperfect)

ik had gekookt/ gekocht/ gezongen, 'I had cooked/ bought/ sung'

The semantic distinction between preterite and perfect is difficult to pin down and is subject to considerable variation in the West Germanic language group. A rough guide to the semantics of tense and aspect can be found in Reichenbach's (1947: 287–298) analysis of tense which distinguishes three relevant types of time: time of speaking, reference time (the time established

within the discourse/text), and event time (the time of the event which is being described).

The preterite is commonly used to describe past situations or actions which took place prior to the present (time of speaking); reference time and event time coincide (15). Regarding the internal contour of the action (aspect) the preterite indicates duration, habituality and/or continuity.

(15) a. De kinderen speelden beneden. ('The children played downstairs.')
b. Iedere dag ging hij naar de bibliotheek. ('Every day he went to the library.')

The perfect expresses semantics of completion and anteriority, i.e. the perfect refers to a completed action (perfective aspect), which took place some time prior to the reference time established within the discourse. The present perfect (16) is used in cases where the reference time is in the present (and thus coincides with the time of speaking), while the action/situation described is located prior to the present.

(16) Ik lees niet graag, maar dat boek heb ik gelezen. ('I don't enjoy reading, but I have read this book.')

The past perfect (or pluperfect) is used in cases where the reference time itself is located in the past and the reported event occurred prior to the reference time (17).

(17) Hij sliep nadat hij de krant gelezen had. ('He slept after he had read the newspaper.')

In contemporary Dutch (as well as in German) the present perfect has usurped many of the semantic functions of the preterite (for more a detailed discussion cf. Bosker 1961; Comrie 1985; Abraham & Conradie 2001), and is used for the description of past events which have perfective aspect, but which do not fulfil the criterion of anteriority (18). In English preterite tense would be required in this situation (cf. Vogel 1997: 91–92).

(18) Ik ben naar binnen gegaan en heb de vrouw gezien. ('I went inside and saw the woman.')

Afrikaans has lost the forms of the preterite as well as the past perfect, and has retained the present perfect as the only marked past tense form (*ek het gesing*, 'I [past] sing'). There are few relics of preterite forms in Afrikaans. These are mostly auxiliaries (*kan*/*kon*, 'can/could', *mag*/*mog*, 'may/might', *moet*/*moes*, 'must/must', *sal*/*sou*, 'shall/should', *wil*/*wou*, 'want/would', *wees*/*was*, 'be/was').

The preterite forms of *weet* (*wis,* 'knew'), *dink* (*dag/dog,* 'thought') and *hê/het* (*had*, 'had') are used only rarely (cf. Ponelis 1993: 429ff.; Donaldson 1993: 222–224; also Abraham & Conradie 2001: 27–28).[6] The modal auxiliaries show regularization in some Western Cape varieties of Afrikaans (cf. Kotzé 1989: 260). In Griqua Afrikaans regularized auxiliary past tense forms are found (i.e. *het … gehet, is … geïs* and *was … gewas*; cf. Van Rensburg 1984).

According to the chronology established by Scholtz and Raidt the pluperfect is believed to have been lost in Afrikaans around 1775, the preterite followed in the first half of the nineteenth century (Scholtz 1963: 37ff.; Raidt 1983: 121ff.). According to Raidt (1983: 122) loss of the preterite started in the late eighteenth century and progressed slowly. The preterite is still well-entrenched in the Duminy diary (1797) and in Louis Trigardt's diary (1836–1838); the latter shows regularization of originally strong verbs (e.g. *wijsde* instead of *wees*, 'knew', *roepte* instead of *riep*, 'called'; cf. Smuts 1968: 23). Early Modern Dutch and contemporary Dutch dialects also show a preference for the preterite (63% in Early Modern Dutch and 75% in contemporary Leiden speech; cf. Ponelis 1993: 430–431). However, a clear preference for the perfect is found in the *Kaapse Taalargief* (cf. Ponelis 1993: 448–449) and in basilectal varieties (e.g. in the speech of Isaac Albach as reported in Trigardt's diary, in Abu Bakr Effendi's *Bayân-ud-în*, and in the diary of Hendrik Witbooi; cf. Abraham & Conradie 2001: 44–51). Regarding the explanation of the development, both evolutive language change and creolization have been suggested. Lass (1987b: 326, also 1997: 255–256), for example, has argued that loss of the grammaticalized past/perfect distinction in Afrikaans 'is actually part of a recurrent or cyclical process in Germanic (as well as other branches of IE [Indo-European])'. Thus, the past (aorist)/perfect distinction of Proto-Indo-European was lost in Proto-Germanic, but was reintroduced in Germanic times when a new grammaticalized perfect (AUX plus PP) was created. In Yiddish and Afrikaans (also reported for Upper German from as early as the fifteenth century), perfect and preterite have merged again: the preterite forms have been lost and the perfect has taken over the functions of the old preterite (for Yiddish cf. Weinreich 1992: 319–320; for German cf. Von Polenz 1991: 199 and Erben 1980: 97-98; cf. also Visser 1992: 2–12). Ponelis (1993: 439; also De Kleine 1997), on the other hand, although acknowledging possible language-internal causes, has described the loss of the preterite in Afrikaans primarily as a result of contact-induced restructuring, i.e. 'another aspect of verbal deflection resulting from verbal invariance under creolisation' (cf. also Abraham & Conradie 2001).

Table 5.7 Present perfect/preterite ratio in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	main verbs	***hebben/zijn***
total	2344	1162
no. of perfect forms	1766	126
perfect forms as % of total	75%	11%

In the *Corpus of Cape Dutch Correspondence* there is a clear preference for the perfect (75%), i.e. a past tense ratio of roughly 3:1 in favour of the perfect (Table 5.7). The development was, however, different for the auxiliaries *hebben/zijn*. Both auxiliaries occur only rarely as periphrastic present perfects in the historical record (cf. also Du Toit 1976 [1897]: 25). In modern Afrikaans only *was* is still commonly used (and varies with *is gewees*) while *had* (rather than *het gehad*) is rare even in writing (Donaldson 1993: 236–240).

Unlike in the case of inflectional loss where the behaviour of individuals shows a broad variation continuum with two clearly identifiable poles (i.e. the norms of metropolitan Dutch and the emerging norm of Afrikaans; Figure 5.1), the majority of writers in the *Corpus of Cape Dutch Correspondence* make extensive use of perfect forms.

The distribution pattern summarized in Figure 5.11 indicates that the change in the past tense system had by 1900 penetrated upper mesolectal and acrolectal varieties, with only few individuals (less than one-fifth of the subsample) maintaining a past tense ratio in favour of the preterite. Although the process was far advanced by 1900, the question of whether the preterite should form part of the Afrikaans standard norm was still highly controversial in the

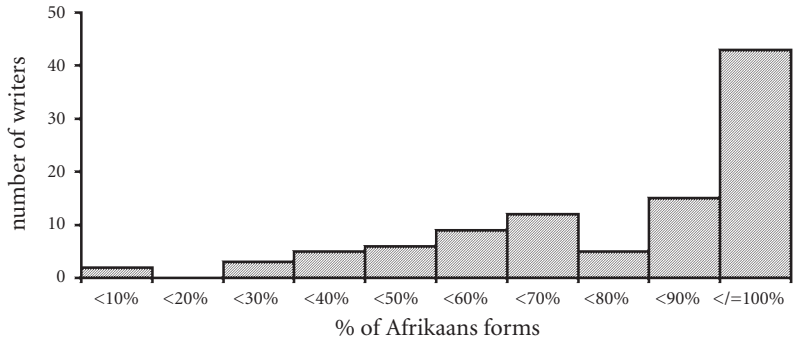


Figure 5.11 Histogram summarizing the use of present perfect by individual writers ($n = 100$); main verbs only

early twentieth century (Steyn 1989). It appears that, irrespectively of frequencies of usage, the preterite still functioned as a prestige marker at the Cape (cf. also Scholtz 1963: 45–46).

Examples of the past perfect occur in the corpus, but are rare. However, one individual (Cornelius Coopman; born 1856, coloured) stands out in that he generalizes past participle forms with *had* as an invariant marker of past tense (19). A similar usage has been reported for Griqua Afrikaans (cf. Rademeyer 1938: 63–64; cf. also Elffers 1903: 18), where the auxiliaries *het* and *had* are used interchangeably. Another (albeit sporadic) characteristic of Griqua Afrikaans, namely the omission of the tense auxiliaries *hebben/zijn* (see Roberge 2002c: 93), is absent from the corpus.

(19) die zelfde avond was myn Heer Fouchei ook in die vergadering en daar **had** myn Fouchei woorde **uitgelaat** wat glaat niet goed genoeg is van een predikant niet [,] die avond **had** myn Heer Fouchei aan my als ouderling **gezegt** daar in die vergadering hy trust my niet ver £5-0-0 in die week niet [,] ik wet niet wat meen die Leraar daarby niet om ver my zoo een woordt te zegt niet en nog meer dan **had** myn Heer Fouchei nog dingen in die Straat **gepraat** wat geheel en al niet Goed is niet [,] myn Heer Fouchei **had** in die straat **gezegt** hy **had** hem klaar **gemaak** ver my ver Corneulis Coopman en toen **had** ik die zaak voor die groot kerkraads **gehad** maar die groot kerkraad **had** my tot hier nog geen antwoordt **gegeeft** niet.
(Cornelius Coopman, 18/9/1909)

['the same evening was my Mr. Fouché also in the meeting and there my Fouché had let words out which are just not good enough for a clergyman [,] that evening my Mr. Fouché had said to me as a church elder there in the meeting that he doesn't trust me for £5-0-0 a week [,] I don't know what the teacher means with this to say to me such a word and more then my Mr. Fouché had said things in the street which are altogether not good [,] my Mr. Fouché had said in the street he had made himself clear to me to Cornelius Coopman and then I have had the affair before the great church council but the great church council had not yet given me an answer.']

Tense and aspect adverbials

In Afrikaans tense adverbials are used to specify the temporal and aspectual dimensions of an utterance (Donaldson 1993: 228ff.). The tense adverb *al* ('already'), for example, is used to indicate anteriority. In cases such as (20) past perfect would be required in Dutch.[7]

(20) Toe ek hom gesien het, het hy die brief **al** geskryf. ('When I saw him, he had already written the letter.')

Adverbs are also used to specify perfective aspect. In the historical record the Dutch form of the past participle of *doen* (*gedaan*, 'done', rather than the Afrikaans variant *gedoen*) occurs as an indicator of completive aspect (often in combination with *al*; cf. Roberge 1994b: 77–78). The oldest examples for *(al) gedaan* as a perfective marker (resultive perfect) are recorded in the eighteenth century Khoe and slave data, thus suggesting the possibility of a basilectal origin of this construction (21).

(21) Die gift **al gedaan** dood, wie kan hy meer wat schaden? ('The poison has already lost its power, who can it now harm?'; Khoe-speaker eighteenth century, cf. Roberge 1994b: 78)

Completive aspect can also be expressed by the adverb *(al) klaar* (Dutch 'ready, finished'; Afrikaans 'ready, already') which, according to Den Besten (1989), continues the pattern of *(al) gedaan*. The adverbs *al* and *alreeds* (both meaning 'already') can also be used in this function.

A well-established past tense marker in modern Afrikaans is *toe(n)* which can function both as an adverbial ('then') and a complementizer ('when'). *Toe(n)* (22) can be used with present tense verbs in Afrikaans, a usage already attested for eighteenth century Cape Dutch Vernacular (cf. Scholtz 1972: 30). In Dutch, on the other hand, *toen* requires the use of past tense (Donaldson 1981: 105, 122; Geerts et al. 1984: 462, 652).

(22) **Toe** hy daar aankom, **toe** groet die ou omie hom. ('When he arrived there, the old guy greeted him.'; Donaldson 1993: 228)

The use of the unmarked verb (present tense) to describe past events/situations is widespread in modern Afrikaans, and also occurs in the *Corpus of Cape Dutch Correspondence*. Example (23) comes from a letter (23/2/1894) by Rudolph Baalie (born 1864), describing a somewhat unpleasant journey in an over-booked train (bold: preterite; caps: present perfect; underlined: present tense).

(23) Tot hierheen **waren** wij tezamen in êên carriage, toen **moesten** wy ons verdelen, de vrouw **ging** met de kleinere kinders in een carriage waar reeds 2 vrouwen en 1 kind in **waren** en ik **ging** met de grooteren in eene carriage waar 1 police met 2 kaffer-convicts in **waren**. Die 2 vrouwen, die van Johannesburg **kwamen**, en lekker **sliepen waren** natuurlijk zeer boos om zoo gestoord te worden, en **wouden** ook niet opstaan. Zoo **moest** de arme

> Marie met de kinders … de nacht doorbrengen, in eene zeer onvriendelijke gezelschap. Met mij **ging** het nog erger, de police **joeg** mij weder uit, ik **liet** de kinders zitten en ga naar die station-master, en zeg hem: hij HEEFT GEZEGT ik moet hierin en de police jaagt mij uit. Hij **kwam** zamen en vraagt de police: waarom moet die man uit? Hij zegt: Ik **was** de eerste hierin en ik ben met convicts hier, ik wil hier geene andere menschen in hebben. Goed [,] zegt die station master, Ik moet alle menschen die tickets hebben plaats geven, er is geene andere plaats meer dan hier. Ga in! zegt hij voor mij en als hij U weder uitjaagt, then you wire back, and I shall teach him.
>
> ['Up to here we were together in a carriage, then we had to separate, the wife went with the smaller children to a carriage where there were already two women and a child, and I went with the bigger ones to a carriage where there was one police [officer] and two black convicts. The two women, who came from Johannesburg and who were sleeping nicely, were naturally very angry to be disturbed like this and didn't want to get up. Thus poor Marie had to spent the night with the children in very unfriendly company. For me it was even worse, the police chased me out again, I let the children sit and go to the station master and say to him, he has said that I must go there and the police chases me out. He came with me and asks the police: why must the man go out? He says: I was first here and I am here with convicts, I don't want to have other people in here. Good [,] says the station master, I must give places to all people who have tickets, and there is no other place anywhere but here. Go in! he says to me, and if he chases you out again, then you wire back and I shall teach him.']

In this extract the switch to the present tense functions as a discourse marker, separating the anecdote (which includes several examples of direct speech) from the rather straightforward description of the train journey. A similar pattern of variation exists in Aletta Schabort's (born 1871) memoir of her experiences at Merebank Camp (Durban) during the Anglo-Boer War: passages involving reported speech and descriptions typically favour a switch to the present tense.

Discourse marking through tense variation is also found in the above cited letter by Cornelius Coopman. In the letter, variation between *toen* plus present tense and *had* plus past participle is used to structure the temporal organization of the overall narrative: the background to Coopman's current problems (his conflict with Reverend Fouché; example (19)) is described using *had* plus past

participle, the more immediate consequences of the original conflict (Coopman's complaint to the church council, and Fouché's demand that Coopman and his family leave the church) are narrated using *toen* plus present tense (24).

(24) en tyd daarna toen gaan ik en ik gaan vrag ouderling van die groot kerk wat van die zaak gewordt hadt [,] toen zegt die ouderling voor dit verder schreef moet ik eens naar myn Heer Fouchei gaan [,] toen neem ik een van myn Diakens en ik gaan na myn Heer Fouchei en toen zegt myn Her Fouchei ik moet uitgaan uit die kerk met al my familie

['and a while later then I go and I go ask the [church] elder of the great church what happened about the business[,] then says the elder before it writes further I must go to my Mr. Fouché[,] then I take one of my deacons and I go to my Mr. Fouché and then says my Fouché I must leave the church with all my family']

The existence of non-verbal aspect markers and the widespread use of the historical present in Afrikaans have been compared to the tense systems of creole languages (cf. Den Besten 1989; Bruyn & Veenstra 1993; De Kleine 1997):

> Whereas [Dutch] employs the historical present as a stylistic device, the use of a present tense form of the verb in a past tense context in [Afrikaans] looks more like a pattern known in some of the creole languages: tense, once established in discourse, may be left unmarked. (Bruyn & Veenstra 1993: 40)

Extensive usage of the historical present is, however, also possible in English where present tense is frequently used as a stylistic norm when narrating comic stories and anecdotes. Tense variation is also used for discourse marking and narrative organization (cf. Wolfson 1979 and Leith 1995; see also the comments made by Francken 1900). The use of the historical present in Afrikaans, although possibly more widespread than in other Germanic languages, is thus by no means 'un-Germanic' (as suggested by De Kleine 1997: 302; note also that tense marking can be obligatory in Afrikaans, cf. Conradie 1998: 37; also Holm 1988/89 II: 347).

The use of adverbs to specify aspectual and temporal information is also too common in non-creole languages to be taken as an unambiguous indicator of creole origin (cf., for example, Crystal 1966 on the use of tense adverbials and unmarked verbs in English, and Cook 1992 on adverbial aspect marking in German). Furthermore, elicited information from Afrikaans speakers suggests that, rather than a clear and stable system of tense and aspect formatives (as known for creole languages), the use of tense/aspect adverbs in Afrikaans is highly context-dependent and subject to variation. Note, however,

that the situation is different in some non-standard varieties of Afrikaans. Griqua Afrikaans, for example, has a system of relatively stable pre-verbal markers, such as the durative marker *lê* ('lie') and the future marker *loop* ('walk'; Rademeyer 1938: 78–79; Roberge 1994c: 58ff.).

In the *Corpus of Cape Dutch Correspondence* there are two examples (25) in which *(al) gedaan* has been lexicalized as a (predicative) adjective with the meaning 'exhausted, finished, used up'. According to Roberge (1995: 83), this usage formed the semantic basis for its use as a perfective marker.

(25) sommige zoorten pruimen is **al reeds gedaan** en de andere soorten zyn nu ryp ('some types of plums are already finished and other types are now ripe'; B. J. Brümmer, 6/01/1901)

ik het naderhand Bob in my kamer laat slaap … want Bob was te **gedaan** ('afterwards I let Bob sleep in my room … because Bob was just too exhausted'; Johanna Brümmer, 15/2/1895)

There are a number of examples in the data in which *klaar* has been lexicalized as a predicative adjective meaning 'exhausted, finished' (26). This lexicalization along the same lines as reported for *gedaan* can be seen as corroborative evidence for Roberge's (1995: 84) argument that perfective *klaar* 'continues the pattern *(al) gedaan* in the Cape Dutch pidgin' (cf. also Den Besten 1989: 238).

(26) ***(al) klaar* (perfective)**

a. de ossen is almal **klaar** gery ('the oxen are all exhausted from the ride'; J. G. Du Preez, 11/8/1894)
b. zo velen oude leeraren die **klaar** gewerk is ('so many old teachers who are exhausted/finished from their work'; M. A. M. Oosthuizen, 3/1/1906)
c. ik heef voor een week terug **klaar** geskeer ('I finished shearing a week ago'; Eddie Smith, circa 1900)
d. ek geloof u Taalbond eksamens vragen zijn **al klaar** uit geskreven ('I believe your questions for the Language Union exam are already written out'; Bettie Smit, 20/5/1917)

Examples of *al* and *al reeds* (perfective and anterior) as well as *toe(n)* (past) also occur in the corpus (examples (27) to (29)).

(27) ***al* (anterior and perfective)**

a. ek het hulle ook **al** geseh dat hulle geen Afrikaans met oupa moet praat nie ('I have already told them that they mustn't speak Afrikaans with grandpa'; Hester Van der Bijl, 10/11/1916)

b. U het zeker **al** gehoor dat ik nou skool hou ('You have probably already heard that I am now teaching'; Leon Malan, 31/10/1917)

(28) ***al reeds* (anterior and perfective)**
als te zamen heb ik **al reeds** £4 ontvangen van de Heer Smith ('altogether I have already received £4 from Mr Smith'; Mozes Mlimkulu, 1/12/1896)

(29) ***toe(n)* (plus present tense)**
a. **toen** ik het geld zien huil ik hart ('when I see the money I cry heavily'; Martha Hillebrand, 13/10/1903)
b. gister aand **toen** ik hier aankom was prof Viljoen net uit ('Yesterday evening when I arrive Prof Viljoen was just out'; F. S. Malan, 7/6/1896)
c. we konden niet denken wat zij wou hebben + **toe** ze zij ik wil water ('we couldn't think what she wanted to have + then she says I want water'; Jacomina van Huyssteen, 14/4/1908)

However, compared to modern Afrikaans the use of tense adverbials was still limited in the corpus. This is clearly visible in the case of *toe(n)* as a past tense marker. 62 examples of *toe(n)* plus present tense were found in the corpus data (out of a total of 318 instances of *toe(n)*, i.e. 20%). This contrasts with Ponelis' (1993: 432–434) quantitative analysis of *toe(n)* in an extract from the *Lewende Bybel* ('Living Bible'; published 1983), which indicates a strong preference for the use of the historic present (in 25 cases out of 38, *toe* was followed by the present tense, i.e. 66%). It is thus possible that the grammaticalization of *toe(n)* as a past tense marker is, at least for acro- and mesolectal varieties, a development of the twentieth century since today *toe(n)* is 'most usually' used with present tense (Donaldson 1993: 228; cf. also Raidt 1983: 122). That *toen* plus present tense was prevalent in basilectal varieties from around 1900 is indicated by Van Ginneken (1928: 211).

Nominal gender agreement

Dutch has a two-gender system which distinguishes common gender (*de*-nouns; comprising both historically masculine and feminine words and constituting roughly two-thirds of Dutch nouns) and neuter gender (*het*-nouns) in the singular (for details cf. Geerts et al. 1984: 41–52). The old Indo-European three-gender system is maintained only in the Flemish dialects (Dekeyser 1980: 106–108). The Dutch system of *de*- and *het*-nouns is difficult to acquire for non-native speakers as it lacks overt morphological or semantic motivation. It is, however, of central importance for the morphosyntactic structure of

Dutch: the gender of the head noun determines the form of definite articles and demonstratives, relative pronouns, anaphoric and deictic pronouns, as well as the inflection of attributive adjectives. Grammatical gender has been lost in Afrikaans where the invariant definite article *die* is used with all nouns. Gender irregularities appear in the historical documents from the seventeenth century onwards (Scholtz 1963: 125ff., 1972: 39). Raidt has dated the loss of gender in the colonists' vernacular to approximately 1750 (Raidt 1983: 149, 1991: 208; cf. also Scholtz 1963: 125), while Ponelis has argued that gender was still relatively intact in the second half of the eighteenth century (Ponelis 1993: 172).[8]

As agreement is the way in which gender is realized syntactically, loss of gender can be defined as the partial or complete loss of agreement (Corbett 1991: 105, 315). The quantitative analysis of loss of gender in the *Corpus of Cape Dutch Correspondence* takes lack of agreement in two different types of determiners as a diagnostic feature for loss of grammatical gender: the definite article (Dutch *de/het*) and the demonstrative article (Dutch *deze, die/dit, dat*). To combine definite articles and demonstratives was necessary since in Afrikaans the single form *die* functions as both article and demonstrative, and it was often impossible to determine from the written text alone whether a particular example of *die* represented a definite article or a demonstrative.

In the *Corpus of Cape Dutch Correspondence* 55% of neuter singular nouns were preceded by common gender determiners (Table 5.8; $n = 60$). The opposite pattern, i.e. neuter article or demonstrative with common gender noun, was rare (less than 1%). As in the case of verbal agreement the behaviour of writers (Figure 5.12) shows two clearly identifiable poles (i.e. norm of metropolitan Dutch and the emerging norm of Afrikaans). Between the poles the distribution is continuous (45% of individuals are found in the middle field of the distribution).

Neuter nouns used with a common determiner include high-frequency nouns such as *geld* ('money'), *jaar* ('year'), *veld* ('field'), *dorp* ('village') and *huis* ('house'; Table 5.8). High frequency is generally seen as a conservative

Table 5.8 Gender regularization in the *Corpus of Cape Dutch Correspondence*. Frequency differences for these nouns are not statistically significant.

	all neuter nouns	***huis***	***geld***	***jaar***	***veld***	***dorp***
total	1166	57	30	20	18	21
common determiner *de/die*	637	36	18	10	7	6
de/die as % of total	55%	63%	60%	50%	39%	29%

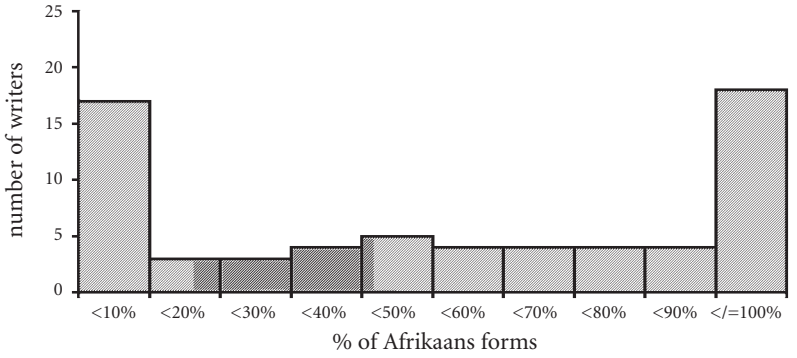


Figure 5.12 Histogram summarizing the degree of gender regularization by individual writers ($n = 60$)

force which can inhibit grammatical reduction (Stemberger & MacWhinney 1986; Smits 1993; Bybee & Hopper 2001). Loss of gender in these lexical items suggests that the reduction process was far advanced in nineteenth century Cape Dutch Vernacular.

In Dutch (as well as in other Germanic languages) there exist a number of suffixes which function as morphological gender markers. These include *-heid* and *-ij* for common gender (*de schoonheid*, 'the beauty', *de boekerij,* 'the library') and the diminutive *-je* (with the variants *-kje*, *-pje*, *-tje* and *-etje*) for neuter gender (*het kastje*, 'the [small] box'; cf. Donaldson 1981: 27–32; Geerts et al. 1984: 42–46). The diminutive ending had clearly lost its function as a marker of neuter gender in Cape Dutch Vernacular: over two-thirds of the diminutives in the corpus occur with a common determiner (64 out of 92, i.e. 70%). Thus, diminutives occur more frequently with a common determiner than neuter nouns in general.[9] The process of gender regularization in the diminutive appears to be independent of the semantic gender of the noun. Carolina Leipoldt, for example, uses *de Jannetje* [+human, +male] ('Johnny'), *de kannetje* [–human] ('the jar'), *de tuitje* [–human] ('the nozzle'), but *het meisje* [+human, +female] ('the girl').

Meisje is an interesting case, as it constitutes an example of what Corbett (1991: 225) has termed 'hybrid nouns', i.e. nouns in which semantic meaning and grammatical gender conflict with each other, leading to variation in gender agreement. Such variation is non-random and follows an agreement hierarchy (Corbett 1991: 226):

The Agreement Hierarchy
attributive < (predicative) < relative pronoun < personal pronoun
(predicative does not require agreement in West Germanic)

The further right an element is placed in the hierarchy, the more likely it is that agreement will be motivated by semantic factors. For example, with regard to the German noun *Mädchen* ('girl') grammatical agreement is common in attributive agreement targets (*das/dieses kluge Mädchen*, 'the/this clever girl') and relative pronouns (*das kluge Mädchen, das…*, 'the clever girl who…'), while (anaphoric) personal pronouns typically show semantic agreement (*das kluge Mädchen, das … sie ist nett,* 'the clever girl, who … she is nice'). The situation in Dutch is similar, and neuter nouns referring to male or female persons often take personal pronouns reflecting the semantic gender of the antecedent (30).

(30) Waar is dat **meisje**? ('Where is the girl?')
Ik weet het niet, ik heb **haar** niet gezien. ('I don't know, I haven't seen her.')

An important characteristic which distinguishes the (anaphoric) personal pronoun from other agreement targets is the fact that it can be separated from its antecedent, often across sentence boundaries. Its referential function is, furthermore, twofold: it refers to the grammatical antecedent as well as to the referent denoted by the antecedental noun. In cases where the natural gender of the referent disagrees with the grammatical gender of the referential noun, anaphoric pronouns are at the forefront of the conflict between the semantic and grammatical identity of hybrid nouns. They are therefore of special importance for the description of changes in gender systems (Corbett 1991:247). The use of hybrid nouns with semantically motivated (anaphoric) personal pronouns is widespread in the corpus, and frequently occurs in texts which otherwise conform closely to the norms of metropolitan Dutch (31).

(31) a. **'t Meisje** van Algie laat me zooveel aan haar denken. **Die** begint ook nou met loopen ('Algie's little girl makes me think about her so much. She also begins to walk now'; Elsie van Huyssteen, 2/3/1907)
b. hij drukte **het dochtertje** ettelike malen aan zijn boezem, kuste **haar** + ging ('He pressed his little daughter several times to his chest, kissed her + left'; Algie van Huyssteen, 27/7/1909)
c. en pas myn kind mooi op[,] wees een moeder over **hem** ('and look after my child[,] be a mother to him'; J. Henderson, c. 1900)

Loss of gender agreement with hybrid nouns is also found in determiners (32) and relativizers (33; for a detailed discussion of the Afrikaans relativizer see Chapter 6):

(32) a. als ik in de kerk trouwde waar **de meisje** behoord ('when I married in the church to which the girl belongs'; P. J. Eksteen, 15/3/1899)
b. en hoewel **de liewe kind** nu niet zo wel is, maar in bed met pijn in de linker zijde ('and although the dear child is now not so well, but in bed with pain in the left side'; Anna De Vries, 11/6/1906)

(33) a. een kleine **meisje** van omtrent 15 jaren **die** gaarne na U komen wil ('a little girl of about 15 years who would willingly come to you'; Carolina Leipoldt, 4/9/1880)
b. dat ik tot noch geen kind heb **die** voor gesondheid van huis moet gaan ('that I haven't yet had a child who had to leave home because of health'; M. Basson, 22/6/1904)

Gradual loss of the gender distinction is a common fate of extraterritorial varieties of Dutch: it has been reported for New Netherland Dutch and East Indian Dutch, and the gender distinction is said to be 'weak' in Surinamese Dutch and Antillean Dutch (Van Marle 1995). Reduction of the gender system is, furthermore, typical for the diachronic development of West Germanic languages, where it is one of the processes leading to the gradual 'loss of older West Germanic structures or contrasts' (Lass 1987b: 323). Dutch itself, which has lost the old tripartite gender system, has moved in the direction of reduction. However, only English and Afrikaans have lost gender entirely.

The attributive adjective inflection

The Dutch attributive adjective inflection is governed by the categories of gender, definiteness and number (34). Singular neuter head nouns favour -ø (unless preceded by a definite determiner), common gender and plural nouns the inflectional ending *-e* (pronounced as a schwa). In Dutch it is thus the inflected adjective which constitutes the 'general case' (Van Marle 1995: 284; also De Rooij 1980: 112), while the undeclined form is limited to one type of environment: [+neuter, –plural, –definite] (Booij & Van Santen 1995: 78).

(34) a. de jong-e vrouw ('the young woman')
het jong-e meisje ('the young girl')
b. (een) jong-e vrouw ('a young woman')
(een) jong-ø meisje ('a young girl')

c. de jong-e vrouwen ('the young women')
de jong-e meisjes ('the young girls')

Adjectives denoting materials and those ending in *-a*, *-o*, *-e*, *-i*, *-y* and *-en* (including past participles used as adjectives) are not inflected (e.g. *de nylon regenjas*, 'the nylon rain coat'; *het lila handdoek*, 'the purple towel'; *een geschreven brief,* 'a written letter'; exception: *naë bloodverwanten*, 'close blood relatives'). Adjectives ending in *-er* are inflected unless they are derived from a geographical place name (e.g. *Edammer kaas,* 'Edam cheese'; Geerts et al. 1984: 321–322; Haeseryn et al. 1997). The system of adjectival inflection in Dutch is, furthermore, subject to social, stylistic and geographical variation. Thus, uninflected forms of the adjective can occur in definite noun phrases, especially in constructions which form a semantic unit: *het doctoraal examen* ('the doctoral exam'), *het bijvoeglijk naamwoord* ('adjective'), *de waarnemend burgermeester* ('the acting mayor'; Geerts et al. 1984: 328; Odijk 1992: 198–199). Neuter nouns preceded by a possessive pronoun also favour the uninflected form of the adjective in many varieties of spoken Dutch (*haar ziek kind,* 'her sick child', *mijn nieuw boek,* 'my new book'; De Rooij 1980: 110, 113–119; Smits 1996: 67). In indefinite common gender noun phrases where the head noun refers to a person, the adjective can occur uninflected when used figuratively and expressing an intrinsic quality (35); a strategy 'exploited by Dutch speakers and writers for expression of a wide variety of subtle semantic nuances' (Shetter 1994: 60).

(35) een groot pianste ('a great pianist'; referring to her musical skills)
een grote pianiste ('a tall pianist')[10]

With the gradual erosion of grammatical gender from 1700, a central trigger for the adjective inflection (i.e. gender of head noun) was lost. The eighteenth century sources show largely random variation (for examples cf. Scholtz 1963: 129ff.). However, instead of simply regularizing the surface exponents of an oblique inflectional paradigm (i.e. always *-ø* or always *-e*), the adjective paradigm was restructured in Afrikaans in such a way that the inflection of the attributive adjective is no longer determined by morphosyntactic criteria, but by the form of the adjective itself (lexicalization). In modern Afrikaans, polysyllabic and morphologically complex adjectives (with the exception of those ending in *-er* or *-el*) take the suffix *-e* categorically. Monosyllabic adjectives are uninflected unless their (original Dutch) attributive form is strongly different from the simplex as a result of diachronic phonological processes such as apocope of /t/, lenition of /d/ or voicing alteration (*sag*/*'n sagte bed*, 'soft/a soft bed', *goed*/*'n goeie boek,* 'good/a good book', *blind* [blïnt]/*'n blinde* [blïnde]

kind, 'blind/a blind child'; cf. Raidt 1983: 142–147; Donaldson 1993: 163–169). Afrikaans thus maintained not only the statistically dominant form of the inflected adjective (from the perspective of metropolitan Dutch), but also the special case of the uninflected adjective. As in Dutch where obligatory adjective inflection can be dropped to express semantic distinctions, non-inflecting adjectives can occur inflected in Afrikaans (36).

(36) die arm man ('the poor (penniless) man')
die arme man ('the poor man, the man to be pitied'; Ponelis 1993: 365)

Inflection can also be re-introduced to increase the degree of formality and is maintained in fixed expressions (as in *die Verre Ooste*, 'the Far East'). Attributive adjective inflection is still subject to variation in contemporary Afrikaans and even native speakers often show insecurity about which forms are accepted as 'correct' (Carstens 1991: 81). According to Ponelis, the inflection of monosyllabic simplexes ending in consonant plus *s* is in flux: some are categorically uninflected (*spits*, 'pointed', *vars*, 'fresh', *pers*, 'violet'), some categorically inflected (*snaaks*, 'funny', *slinks*, 'clandestine', as well as those which are derived from a noun, as in *die Kaapse klimaat*,the Cape climate'), while others show variable inflection (e.g. *fluks*, 'smart', *preuts*, 'prim', *rats*, 'quick'; Ponelis 1993: 366). Regarding the group of monosyllabic adjectives ending in *-g*, there is disagreement among grammarians. While Donaldson (1993: 166) classifies monosyllabic adjectives ending in nasal plus *g* as inflecting, both Ponelis (1993: 366) and Raidt (1968: 107–108, 1983: 144) identify such adjectives (for example, *streng*, 'strict', or *bang*, 'anxious') as uninflecting.

Uninflected forms of monosyllabic adjectives appear in the historical record from the beginning of the eighteenth century, mostly in high frequency lexemes such as *klein* ('small'), *groot* ('big') and *oud* ('old'). Adjectives ending in *-er* also occur uninflected by the eighteenth century (Ponelis 1993: 367f.). Raidt (1968: 172–173) has suggested that the extensive of uninflected attributive adjectives in Afrikaans is was supported by patterns of adjectival inflection in upper-class varieties of seventeenth century Dutch. These varieties showed use of the uninflected adjective in additional syntactic and phonological environments (i.e. with indefinite determiners and masculine personal names in the nominative). Van Marle (1995: 291), however, has argued against assuming a direct link between the seventeenth century input and the current Afrikaans adjective inflection:

> In my view it is very unlikely … that there is a direct link between the prominent position of the undeclined adjective in these varieties of overseas Dutch

> [Afrikaans and Old New York Dutch, the latter of which has generalized the uninflected adjective] and the fact that in seventeenth century Dutch the undeclined adjective was more popular than in the present-day language. In the first place, there can be no doubt that in seventeenth century Dutch, too, it was the declined adjective which represented the general case … Secondly, the popularity of the undeclined adjective in seventeenth century Dutch was primarily a characteristic of the cultivated language, and it seems very doubtful whether this latter variety of seventeenth century Dutch has been very influential in relation to the rise and development of Afrikaans and Old New York Dutch.

Lass (1990) has discussed the Afrikaans adjective inflection as an example of exaptation (that is, the co-optation of historical surface structures to new functions), and described the development of a type of adjective declension based on morphological and morpho-phonemic criteria as 'a conceptually novel inflection-type, with no real Germanic precedents' (Lass 1990: 91; cf. also Van Marle 1995: 92: 'As a matter of fact, I am not aware of any other language with a system of adjectival declension which is even reminiscent of Afrikaans'). However, although the Afrikaans adjective declension is a paradigmatic example of what Ferguson (1995: 190) has called an 'elegance innovation' – that is, an innovation 'that introduces a neat, pleasingly designed pattern that "catches on" chiefly by virtue of its appeal to the human system-perceiving and system-constructing capacities' – it is not entirely without Germanic antecedent: in this case a marginal (but nevertheless stable) tendency within Germanic to allow for a special group of categorically undeclining adjectives. Thus, as noted above, the adjective declension of metropolitan Dutch shows morphophonological (and even semantic) conditioning by excluding certain types of adjectives (those denoting materials, those ending in vowels or *-e(n)*, polysyllabic adjectives) from the declension paradigm. In Swedish, adjectives ending in *-s,* *-e,* and *-a* are not inflected (e.g. *en bra bok*, 'a good book'); Norwegian (*Bokmål*) does not inflect adjectives ending in a stressed vowel, unstressed *-e* or *-s,* and German has a small class of uninflected adjectives, including *rosa* ('pink'), *prima* ('awesome') and *lila* ('lilac'; e.g. *das rosa Bändchen*, 'the pink ribbon').[11] In Frisian certain monosyllabic adjectives such as *âld* ('old'), *jong* ('young'), *lyts* ('small') and *nij* ('new') remain uninflected in cases where noun and adjective 'enter into a fixed collocation' (Hoekstra & Tiersma 1994: 513, as in *de jong faam*, 'the young woman'). However, while the general principle (i.e. to exclude certain types of adjectives from the inflectional paradigm based on phonological, morphological or semantic criteria) is not unknown in West and North Germanic languages, only in Afrikaans has this marginal feature been trans-

formed into a complex paradigm which is truly 'unique in Germanic' (Lass 1990: 95).

The attributive adjectives in the *Corpus of Cape Dutch Correspondence* were grouped into four categories for the variationist analysis:

i. monosyllabic adjectives (categorically non-inflecting in Afrikaans; n = 566);
ii. monosyllabic adjectives (categorically inflecting in Afrikaans; n = 382);
iii. polysyllabic adjectives (categorically inflecting in Afrikaans; n = 712);
iv. polysyllabic adjectives ending in *-er* (categorically non-inflecting in Afrikaans; n = 106).[12]

To investigate the extent to which the new system of the Afrikaans adjective inflection was already in place around 1900, only those environments where the Afrikaans inflections would differ from the Dutch norm were analyzed. The quantitative data summarized in Table 5.9 show that in particular those polysyllabic adjectives which are categorically inflecting in modern Afrikaans conform closely to the new system. The class of categorically uninflecting adjectives shows a considerably lower adherence to the emerging norm.

The overall result, however, once again conceals differences between lexemes. In the new group of categorically non-inflecting monosyllabic adjectives five high-frequency adjectives (*arm*, *klein*, *groot*, *mooi* and *ou*) were analyzed separately. The results are summarized in Table 5.10.

Table 5.9 The emergence of the new system of adjectival declension in the *Corpus of Cape Dutch Correspondence.* The distribution is statistically significant at $p < 0.001$ (chi square).

	no. of Afrikaans forms	**total**	**Afrikaans forms as % of total**
monosyllabic adj. (categorically -ø in Afrikaans)	203	479	42%
monosyllabic adj. (categorically *-e* in Afrikaans; e.g. *goed-goeie*)	21	37	57%
polysyllabic adj. (categorically *-e* in Afrikaans)	47	67	70%
adj. ending in *-er* (categorically -ø in Afrikaans)	49	93	53%

Table 5.10 Frequency distribution for five monosyllabic adjectives. The distribution is significant at $p < 0.001$ (chi square).

	arm ('poor')	*groot* ('big')	*klein* ('little')	*ou* ('old')	*mooi* ('pretty')
(a) no. of cases without inflectional ending	6	36	22	70	26
(b) no. of loci where Dutch would have required an inflectional ending	58	96	47	93	30
(a) as percentage of (b)	10%	38%	47%	75%	87%

The low percentage of uninflecting forms for the adjective *arm* and the high frequency of inflected forms for the adjective *ou* are largely semantically motivated and illustrate the co-existence and partial overlap of competing adjective paradigms at the Cape. Almost nowhere in the corpus is *arm* used with the literal meaning 'penniless', but always in the transferred sense of 'to be pitied' (37) which in modern Afrikaans also triggers inflection (cf. example (36)).

(37) a. die arme mans ('the poor men'; Aletta Schabort, 1902)
b. die arme dier ('the poor animal'; Maggie Marren, 13/6/1915)
c. ons arme volk ('our poor nation'; P. G. Kuhn, 16/5/1900)

Ou, on the other hand, occurs uninflected even in the writings of those whose texts conform otherwise closely to the norms of standard Dutch (38). In these texts the uninflected variant is used as a term of endearment and marker of informality.

(38) a. ou broer ('dear old brother'; J. H. H. De Waal, 8/1/1902)
b. ou Katie ('dear old Katie'; Rijkie Louw, 4/8/1909)
c. de ou boeken ('the old books'; F. S. Malan 1/6/1893)

The modern Afrikaans adjective declension does not parallel the general tendency towards reduction and regularization evident in other aspects of the morphosyntactic system. In other words, while gender and the verbal system went down the road of inflectional reduction, the declension of the adjective shows a complex re-structuring of an obsolete inflectional system. Inflectional generalization, on the other hand, is a common feature of many extraterritorial varieties of Dutch, i.e. East Indian Dutch, Surinamese Dutch and Antillean Dutch (cf. Van Marle 1995). Inflectional generalization has also been reported for some varieties of Western Cape Afrikaans (Kotzé 1989: 258; Ponelis

1993:373; Hoogenhout 1904:7; a preference for inflected forms is also evident in the Duminy diary), and at least one writer in the *Corpus of Cape Dutch Correspondence* appears to have adopted inflectional generalization as a linguistic strategy: Carolina Leipoldt (39) uses inflected adjectives almost categorically (the only exception is *een groot salaries*, 'a large salary').

(39) a. een zoete meisje ('a sweet girl'; 2x)
b. een kleine meisje ('a little girl')
c. een blikke kannetje ('a shining jar')
d. een zoete kind ('a sweet child')
e. een mooije meisje ('a pretty girl')

Four of these examples involve a head noun which although grammatically neuter is semantically female (*meisje*), and this could have triggered the inflection of the adjective. However, when Leipoldt uses *meisje* with a definite article she uses the correct neuter article *het*. To investigate the possibility of a (counter-) tendency towards inflectional generalization involving those monosyllabic adjectives which are today categorically uninflected in Afrikaans, only environments where uninflected forms would have corresponded to the norms of Dutch were considered (n = 87). Twenty inflected adjectives were found (23%). Regularization was unsuccessfully attempted by S. J. du Toit in the first grammar of Afrikaans. Du Toit stated that Afrikaans categorically requires *-e* in attributive adjectives, while predicates occur without the suffix (1980 [1876]:14; also Elffers 1903:15). However, Du Toit frequently violated this rule himself in the text of the same grammar (writing, for example, *een groot letter*, 'a big [printing] type', and *een rou boer*, 'an inexperienced farmer').

Summary: Morphosyntactic standardization as a process of rule extension

Analogy (i.e. regularization of surface forms through the extension of existing rules) is a driving force of morphosyntactic change. Sweet (1900:23) distinguished the abrupt nature of morphological change from the gradual movement which characterizes many phonetic/phonological changes. However, although morphological change involves an element of a rather abrupt and discrete process of rule substitution (or loss) – whereas phonetic change is often structurally continuous – morphological change is nevertheless gradual and continuous in terms of its temporal (and social) diffusion, and the relative frequency of morphological substitution or loss increases slowly over

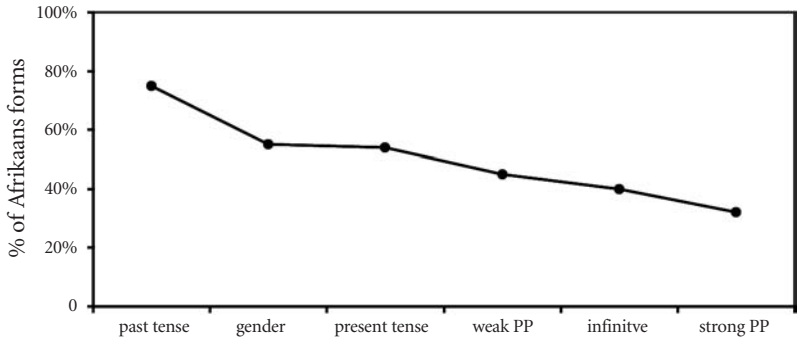


Figure 5.13 Morphosyntactic regularization in the *Corpus of Cape Dutch Correspondence* (frequency differences are statistically significant at 295.4275, df = 9, $p < 0.0001$)

time and/or across social groups. It is precisely this feature of gradual diffusion which makes not only the dating of morphological change difficult, but also the identification of its causes (Jespersen 1945: 169).

The variable concord phenomena found in the *Corpus of Cape Dutch Correspondence* (1880–1922) indicate that, in contrast to the existing datings, which assume completion of morphosyntactic regularization by 1800 (see Chapter 1), these changes were still in process in the (upper)mesolectal and acrolectal varieties around 1900. The 'principle of inflection' (although variable) thus remained productive until fairly late. Summarizing the relative frequencies for the different morphosyntactic variables, one observes a clear cline for the degree of morphological reduction in the corpus data (see Figure 5.13). Most advanced in the *Corpus of Cape Dutch Correspondence* is the generalization of the perfect as a marker of past tense (which, interestingly, is one of the few changes which Scholtz and Raidt have dated comparatively late). This is followed by the loss of nominal gender agreement, the regularization of the present tense paradigm, and finally the weak past participle, the infinitive and the strong participle.

The gradual and dynamic nature of morphosyntactic change and the existence of multiple norms (as illustrated most clearly in Figures 5.2 and 5.13) posed specific challenges to language standardization efforts at the Cape. When the first grammar of Afrikaans was published in 1876 its most remarkable feature was the complete linguistic uniformity and regularity of the language it propagated, thus clearly following the 'no variation'-principle of language standardization. The grammar (as well as other texts published by the GRA) established via processes of 'rule extension' a morphosyntactically uniform and

regular standard language which could be contrasted unequivocally with the norms of metropolitan Dutch. The complex Cape Dutch variation continuum was thus simplified and the gradual process of historical change accelerated. While morphosyntactic standardization was on the whole successful, some areas resisted the standardization efforts: the adjective inflection (which even today shows considerable variation), and a small set of lexical verbs which maintained an invariant inflected form (such as *skrywe, begint, gaat*) until quite recently (Ponelis 1993: 391–393).

Notes

1. The infinitive form *hê* is used in constructions where the infinitive follows present or imperfect modal verbs (as in *ek wil dit hê*, 'I want to have this'). *Het* is used as an auxiliary in compound tenses (as in *hy moet die bok geskiet hê*, 'he must have shot the buck'; Donaldson 1993: 239–240).

2. Cf. also Trudgill (1989: 228) on the increasing rate of endogenous innovations in contact situations; for a general argument in favour of endogenous explanations in the absence of clear evidence for contact see Lass and Wright (1986), Lass (1997: 207–209).

3. In Dutch, underlying voiced obstruents acquire the feature [–voice] word-finally.

4. Generalization of stem forms occurs in the letters of Louis Botha (7 out of 12) and Petronella van Huyssteen (3 out of 22). Both age group 1 and class 3a.

5. Since apocope of final *-n* is common in metropolitan Dutch plural preterite verbs are generally unmarked for plural in spoken Dutch.

6. The two preterite variants for *dink* originate from standard and dialectal Dutch respectively (standard *denken/ik dacht*, dialectal *denken/ik docht*; Donaldson 1993: 22, fn. 8).

7. Upper German, which also lost the past perfect has developed a double perfect as a substitute: *Zufällig* ***habe*** *ich an diesem Tag eine Forelle* ***gestohlen gehabt****, und der Fischer ist zornig zu uns gelaufen und hat geschrien* ('Incidentally I had stolen a trout that day and the fisherman ran to us angrily and shouted'; Erben 1980: 98). A similar construction exists in Afrikaans (*Hy* ***het*** *dit toe al* ***gedoen gehad***, 'He had already done it'), but is rarely used (Donaldson 1993: 233).

8. Case marking has disappeared in both modern Dutch and Afrikaans. There are, however, relics of case inflection in some Afrikaans expressions, such as *in my dag des lewens nie* ('not in my whole life'; genitive) or *in der haast* ('in haste'; dative; for more examples cf. Grayson 1962: 24ff.). In the corpus case marking occurs occasionally, but is rare.

9. The frequency difference between diminutives and other neuter nouns regarding loss of gender agreement in the determiner is statistically significant at $p < 0.01$ (chi square).

10. This usage appears to be typical for *nomina agentis* where it specifies the manner in which an action (playing the piano, writing, playing football, etc.) is carried out (Haeseryn et al. 1997). However, there appear to be restrictions on which nouns can occur in this

construction. Thus, while the phrase *een groot man* ('a great man') is acceptable, **een groot vrouw* ('a great woman') appears not to be possible (Wim Vandenbussche, personal communication; other non-permissable nouns include **vent*, **kerel*, **jongen*, cf. Odijk 1992: 198).

11. For Swedish see Ferguson (1995: 186), Holms and Hinnchliffe (1997: 46–47); for Norwegian see Strandskogen and Strandskogen (1986: 72); for German see Erben (1980: 72). However, in the case of German there is a tendency towards inflectional regularization and inflected variants (e.g. *das rosane Bändchen*) are not uncommon in the spoken and informal written language.

12. There were only four examples of polysyllabic adjectives ending in *-el.* These were excluded from the analysis. Adjectives which are categorically uninflecting in Dutch (see above) were also excluded.

Chapter 6

Morpholexical and syntactic variation

> Data! data! he cried impatiently, I can't make bricks without clay.
> Arthur Conan Doyle. *The Copper Beeches*

Personal pronouns

The pronoun system of Dutch differentiates between subject and object forms, as well as between stressed (strong) and unstressed (weak) forms (Table 6.1; for a summary of the etymological relations between strong and weak pronouns cf. Zwart 1997: 117–118). In Afrikaans the distinction between subject and object forms was maintained for singular pronouns, but was lost in the plural. The unstressed forms disappeared entirely (Table 6.2; for details cf. Donaldson 1993: 123ff.).

Raidt (1991: 213) described the loss of unstressed forms in a pragmatic-psychological fashion as the result of the 'neiging om nadruklik te praat' ('the tendency to speak with emphasis'). However, Zwart (1997: 199ff.) has shown that the unstressed Dutch pronouns are not merely phonologically reduced variants of the stressed pronouns. They are syntactic clitics which have

i. a specialized meaning that the stressed pronouns lack, i.e. *je* and *ze* (pl.) can be used to refer to the generic entity 'people', and *ze* (pl.) can be used to refer to both persons and things, and
ii. their syntactic status differs from that of the stressed pronouns, i.e. they can move to positions unavailable to stressed pronouns.

To describe the weak pronouns as special clitics means that their loss involved both lexical and syntactic change. Weak pronouns occur in the *Corpus of Cape Dutch Correspondence* but are rare (e.g. for the first person plural pronoun the ratio between *wij* and *we* is roughly 7:1). In modern Afrikaans weak pronouns have survived in some lexicalized forms, such as *dankie* ('thank you'; derived from *dank je*), or *ditsem* ('that's the stuff!'; derived from *dit is 'm*). According to Ponelis (1993: 196) these lexicalized forms 'might be taken as evidence of the

Table 6.1 The pronoun system of Dutch[1]

	subject pronouns		objects pronouns	
	stressed	**unstressed**	**stressed**	**unstressed**
First person singular	*ik (ikke)*	*'k*	*mij*	*me*
Second person singular	*jij*	*je*	*jou*	*je*
Third person singular	*hij, zij, het*	*'ie, ze, 't*	*hem, haar, het*	*'m, d'r, 't*
First person plural	*wij*	*we*	*ons*	–
Second person plural	*jullie*	*je*	*jullie*	*je*
Third person plural	*zij*	*ze*	*hun (hen)*	*ze*

Table 6.2 The pronoun system of Afrikaans

	subject	**object**
First person singular	*ek*	*my*
Second person singular	*jy*	*jou*
Third person singular	*hy, sy, dit*	*hom, haar, dit*
First person plural	*ons*	*ons*
Second person plural	*julle*	*julle*
Third person plural	*hulle*	*hulle*

fact that weak pronouns had gained a toehold in early Afrikaans'. Changuion (1844: 76) reports the use of weak pronouns for the nineteenth century, but describes them as socially and stylistically marked:

> *Je* behoort tot de taal des beschaafden omgangs … maar *jij* en *jou* is plat Holl. ('*Je* belongs to the language of educated conversation … but *jij* and *jou* is uneducated (*plat*) Dutch.')

The loss of the subject/object distinction in the Afrikaans pronoun system is similar to developments in creole languages which typically do not mark pronominal case (Ponelis 1993: 218). Case syncretism, however, is also a general characteristic of many West Germanic languages. The accusative/dative distinction, for example, has been lost in English and Dutch (cf. Sapir 1921: 163 who describes the levelling of the subject-object distinction as an example of drift). Nominativization of object forms has been reported for a number of languages: Upper German uses the dative first person pronoun *mir* as a subject form; the English second person (nominative) plural *you* is historically based on the dative/accusative form *ēow* (Lass 1997: 254); in some Dutch dialects the object forms *jou* and *hun* are used as subject pronoun (De Vooys 1953: 77; for further discussion cf. Scholtz 1963: 98–101). Thus, as in the case of morphological reduction, evolutive tendencies of internal Germanic language

change interacted with contact-induced restructuring in the development of the pronominal system. It should, however, be noted that with regard to several of the morpholexical and also syntactic variants, early attestations are prominent in the Khoe and slave Dutch data. This suggests that the locus of innovation was outside of the European settler community.

First person singular subject pronoun (*ek*)

The Afrikaans first person singular pronoun *ek* still varied in the written language with the Dutch form *ik* at the beginning of the twentieth century (Malherbe 1917: 74f.). 3652 instances of the first person singular form in the *Corpus of Cape Dutch Correspondence* included 858 examples of *ek* (i.e. 24%, see Table 6.5). Thus – as already suggested by Raidt's variationist study of nineteenth century dialect texts ([1992]: 303; cf. also Van Rensburg 1984: 148f.) – the variant *ek* was still rare at the turn of the century and most writers continued (at least) to write the Dutch form *ik*.

Third person singular subject/object pronouns

Afrikaans anaphoric pronouns reflect semantic gender differentiation in the third person singular and semantically male or female antecedents require the use of *hy* ('he') and *sy* ('she') respectively. This distinction has been lost in some non-standard varieties of Afrikaans where the masculine pronoun *hy* combines with female antecedents (Links 1989: 79f.; Van Rensburg 1989: 147; Ponelis 1993: 213). Roberge (1996: 140) has attributed this usage, which is entirely absent from the mesolectal/acrolectal historical data, to basilectal influence.

The loss of the Dutch third person singular neuter subject pronoun *het* and the rise of the Afrikaans form *dit* show the interaction of lexical and syntactic change in the development of the pronoun system. The syntactic position of the Dutch object pronoun *het*/*'t* is more restricted than the position of Afrikaans *dit*. The normal position for *het*/*'t* as an object pronoun is to the right of the V-cluster and to the left of the indirect object (1a), where it remains even if syntactic transformations are applied to the sentence (1b; cf. Seuren 1997: 230).

(1) a. Je hebt **het** haar gegeven. ('You have given it to her.')
b. Heb je **het** haar gegeven? ('Have you given it to her?')

In some varieties of Dutch movement of the pronoun to the right of the indirect object is possible for unstressed pronouns, thus emphasizing their status

as special clitics (2a/b). However, the position next to the verb (2a) is generally felt to be more standard (Zwart 1997: 123–124).

(2) a. Hij heeft **'t 'r** gegeven.
b. Hij heeft **'r 't** gegeven. ('He has given it to her.')

Het can occur sentence initially only as a dummy-subject (3a), and can function as a subject referring to a previous *de*-noun if the predicate following the copula is a noun (or an adjective used as a noun) preceded by an article (3b; Geerts et al. 1984: 169).

(3) a. **Het** regent. ('It's raining.')
b. Wie is die meneer? **Het** is de burgermeester. ('Who is this man? He is the mayor.')

The Afrikaans neuter pronoun *dit* is derived from the Dutch demonstrative pronoun (*dit*, 'this', alongside *dat*, 'that') and can occupy syntactic positions unavailable to *het* (4). While *het* (as a direct object) usually cliticizes to the verb, the demonstratives *dit* and *dat* can occur in non-clitic position, are stressable and can be used contrastively (Ponelis 1993: 209f.).

(4) a. *__Het__ gaf hij mij. **Dit** gaf hij mij. ('This he gave me.')
b. *Ik wil **het** hebben niet **het**. Ik wil **dit** hebben niet **dat**. ('I want this not that.')

Frequent use of non-clitic *het* in eighteenth century documents indicates the breakdown of the Dutch system and has been interpreted as 'masked *dit/dat*' (Ponelis 1993: 210; also Scholtz 1963: 57–58). In eighteenth century Cape Dutch Vernacular *het* competed not only with *dit,* but also with the demonstrative *dat*. From the early nineteenth century, however, use of *dit* increased and *dat* disappeared (Raidt 1983: 156; Ponelis 1993: 211–212). The analysis of the quantitative distribution of *het, dit* and *dat* in the *Corpus of Cape Dutch Correspondence* (as referential subjects and objects, following Ponelis' 1993: 210–211, quantitative analysis of the Barbier corpus) shows that *het* was still in a strong position in acrolectal registers, competing only with *dit*. Use of *dat* was negligible (Table 6.3).

The Afrikaans third person masculine object pronoun *hom* occurs only rarely in the *Corpus of Cape Dutch Correspondence* (15%, see Table 6.5; cf. also Malherbe 1917: 79 on continuing variation between the forms *hem* and *hom* in the early twentieth century).

Table 6.3 *Het, dit* and *dat* as referential subjects and objects in the *Corpus of Cape Dutch Correspondence*

	het	*dit*	*dat*	**total**
N	1272	793	31	2096
% of total	60%	39%	1%	

First person plural subject pronoun (*ons*)

Early examples of nominative *ons* are prominent in the Khoe-Dutch and slave Dutch data (cf. Scholtz 1963: 93–94). It was reported for the settlers' varieties of Cape Dutch Vernacular first by Otto Friederich Mentzel who visited the Cape in the 1730s (see also Chapter 1). In the account of his stay, he made the following, frequently cited, remark:

> The language of the country is just as far from being pure Dutch, as that of the German farmers is from pure German. The men have a broad accent and the women folk use certain expressions that are sometimes really ridiculous. For instance, if one were to ask them whether they have no Bible, the reply is: 'Onz heeft geen Bybel'; which means 'Us has no Bible'. If one were then to ask them: 'How many 'Onze' (ounces) in a pound?' they would blush.
>
> (Mentzel 1944 [1787]: 157–158)

This commentary suggests not only that nominative *ons* was used by the settler population, but also that speakers recognized the variant as 'incorrect' and 'non-standard'. Nominative *ons* occurs in the transcription (5) of an utterance by Hanna Wagenaar in 1772 (CJ 1103/225), which is generally remarkable for its similarity to modern Afrikaans (containing also the first attestation of *hulle* and shows reduction of the verbal system; cf. Scholtz 1963: 112; Ponelis 1993: 217).

(5) het **Hulle** dan Overal geloop waar **ons** geloop het ('have They then walked Everywhere where we have walked').

Scholtz and Raidt have argued that since nominative *ons* is known in older forms of dialectal Dutch and is still used in the Zeeland dialects, it should be interpreted not as a result of contact-induced change, but as a 'Germanic vulgarism' (Raidt 1983: 155; Scholtz 1963: 94–95). However, speakers of Zeeland varieties had little impact on the composition of the early society at the Cape, and nominative *ons* is not attested in Holland varieties of Dutch which featured prominently in the superstrate base of Afrikaans (see Chapter 1). Direct dialect origin is therefore unlikely (Ponelis 1993: 218).

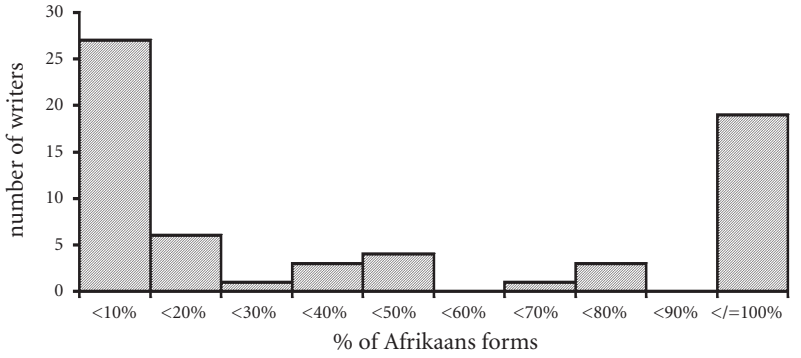


Figure 6.1 Histogram summarizing the use of nominative *ons* by individual writers ($n = 65$)

The comparatively late attestation of nominative *ons* in the settlers' vernacular contrast with the abundance of early examples for nominative *ons* in the Khoe and slave data, and Ponelis has, therefore, interpreted Afrikaans *ons* as an 'interlectally restructured form' of Dutch *wij*/*we* (Ponelis 1993: 32) – a 'linguistic innovation among Khoe and slave interlectals' (ibid.: 218; see also Valkhoff 1966: 222). While most Afrikaans linguists assume that nominative *ons* was fairly common by 1750 (Raidt 1983: 155, 1991: 213; Ponelis 1993: 218), Hoogenhout (1904: 9) still reported variation between *ons* and *wij* in subject position among older people at the beginning of the twentieth century.

In the *Corpus of Cape Dutch Correspondence*, *ons* replaces *wij* as a first person subject plural pronoun in over one-third of cases (36%; Table 6.5). The use of nominative *ons* by individual writers is summarized in Figure 6.1.

In contrast to loss of inflection and gender marking (Chapter 5), conformity to the norms of Dutch is (with over 40%) decidedly more pronounced in the pronominal system; the middle field of the distribution is small and discontinuous, containing less than one-third of writers (most of these writers are still oriented towards the traditional Dutch norms and show comparatively low frequencies of *ons* in their texts). However, many of the writers who do not (or rarely) use nominative *ons* nevertheless show extensive reduction in the verbal system. That is, although use of *ons* generally triggers the uninflected form of the verb, *wij* occurs frequently with uninflected verb forms in the corpus. Co-occurence restrictions are also reported in metalinguistic commentaries:

> De eenige Afrikaners, die *ik* in plaats van *ek* in een gewoon gesprek gebruiken, bezigen *wy zyn* of *wy is* in plaats van *ons is*.

['The only Afrikaner who use *ik* instead of *ek* in a normal conversation, also use *wy zyn* or *wy is* in place of *ons is*.'; *De Goede Hoop*, November 1906]

Third person plural subject/object pronouns (*hulle*)

The third person plural subject pronoun *hulle* is a variant of the Dutch object form *hunlui/hunlie(den)*. Its history thus reflects the same grammatical process as *ons*: nominativization. At the Cape, *hulle* competed as a subject form with the more 'genteel' variant *zulle/sulle* (derived from *sylui*, cf. Ponelis 1993: 219). Johanna Duminy, for example, uses *zulle/sulle* in her more acrolectal diary (1797) in variation with the standard forms (*zij/sij* and *haarlui*). Use of nominative *hulle* increased only after 1810 (Scholtz 1963: 96); however, variation with *zulle/sulle* continued until the early twentieth centuries (Malherbe 1917: 80).

Hulle as an object form competed with *haarle* (with the variants *haarly* and *haarlui*). *Haarle* was the dominant form throughout the eighteenth and early nineteenth centuries, and use of objective *hulle* increased only after 1840 (Scholtz 1963: 97). Like *zulle/sulle*, *haarle* was still used by older people at the beginning of the twentieth century (Malherbe 1917: 80).

In the *Corpus of Cape Dutch Correspondence* the form *hulle* (with the variants *hul* and *helle*) is widespread in subject position (49%) but less common in object position (32%; Table 6.4), where *hun* and its variants are still dominant. The greater prominence of *hulle* in subject position and the general maintenance of a subject/object distinction in the plural pronominal paradigm agrees with the patterns of variation reported for eighteenth and early nineteenth century Cape Dutch Vernacular (cf. Scholtz 1963: 99; Ponelis 1993: 220).

The variant *zulle/sulle* is rare in the corpus and only thirteen occurrences were found. Particularly interesting is the use of *zulle* in a letter by J. H. De Villiers to S. J. Du Toit (22/4/1891). At the beginning of the letter De Villiers explicitly announces as his aim to write in the kind of language he uses in or-

Table 6.4 *Hulle* as subject and object pronoun in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	subject	**object**
total of third person pl. pronouns	352	160
hulle	172	51
hulle as % of total	49%	32%

dinary conversation (note also his corrections of inflected Dutch forms such as 'spelt', and non-standard variation between 'i' and 'e' which occurs in many of the letters in the corpus):

> Ik wil nou probeer om in myn brief dieselfde taal te gebruik als ik ~~en~~ in gewone gesprek zou gebruik.
>
> ['I will now try to use in my letter the same language which I use in ordinary conversation.']

With regard to morphosyntax the language used in the letter is close to the emerging standard of Afrikaans. Yet, De Villiers' only example of a third person plural pronoun (subject) is *zulle*. He, furthermore, criticizes Du Toit's use of nominative *ons:*

> Bij voorbeeld, u ~~spelt~~ spel altyd 'veel' als 'veul' en zeg 'ons' waar in gewone gesprek 'wy' gezegd wordt.
>
> ['For example, you write always 'many' as *veul* and say *ons* where in ordinary conversations one would use *wy*.']

In Afrikaans *hulle* can also be used as a suffix to a noun when referring to a collective of people. Thus, depending on context *Piet-hulle* (Piet-3rd plural) can mean: 'Piet and his wife, Piet and his family, Piet and his friends' (Donaldson 1993:136). A similar construction, using the suffix *-goed*, occurs in Griqua Afrikaans (*Piet-goed*). This structure appears to be a reanalysis of the Dutch derivational morpheme *-goed* (e.g. *speelgoed*, 'toys, play things', *rookgoed*, 'smokables, smokes'), which shows surface similarity to the Nama plural suffix *-goe*. The use of *-hulle* as an associative plural marker constitutes a further step in the reanalysis chain (Nienaber 1994c).[2]

In the corpus data the use of *-hulle* as an associative plural was limited to three writers: Wynanda Hoogenhout, Rachel Steyn and A. J. Z. ('Janie') Van Huyssteen (6). All three belonged to the generation born in or after 1865 (age group 2; for a more detailed discussion of the letters of Wynanda Hoogenhout see Chapters 7 and 9). While the letters of Wynanda Hoogenhout and Rachel Steyn conform closely to the emerging standard of Afrikaans (e.g. loss of verbal inflection and gender, use of Afrikaans pronoun variants and the brace negation), Janie Van Huyssteen's letter is written in passable Dutch and shows only a slight tendency towards inflectional loss in the verbal and nominal system.

(6) a. **Mnr. Meyer hulle** behoort onder Charlie Meynhard zij gemeente ('Mr. Meyer and his family belong to the congregation of Charlie Meynard'; Wynanda Hoogenhout, 15/4/1909).

b. **Mrs. B. hulle** was ook zo goed fir my ('Mrs. B and her family were also so good to me'; Wynanda Hoogenhout, 28/4/1909).
c. Ek gaat dan somaar saam moet **Imker hulle** op reis ('I will then go travelling with Imker and his family'; Wynanda Hoogenhout, 10/6/1911).
d. **Nico hulle** ze passagie is al met die Kenilworth uitgeneem ('the [ship] passage of Nico and his family is already booked with the Kenilworth'; Rachel Steyn, 19/10/1919).
e. Zijn **Pieter hulle** al thuis? ('Are Peter and his family already home?'; Janie Van Huyssteen, 31/12/1907).

Attributive possessive pronouns

The Afrikaans attributive possessive pronouns are (except for the third person singular masculine/neuter) identical with the object forms of the personal pronouns (Table 6.5; cf. Geerts et al. 1984: 201–205 for general overview of the Dutch system, Donaldson 1993: 123, 138f. for Afrikaans).

There were important 'crossovers' in the variation patterns for object pronouns and possessive pronouns in seventeenth century Dutch (Ponelis 1993: 204; also De Vooys 1953: 75), and variation between *mij* and *mijn* occured in both grammatical contexts. However, the variant ending in *-n* was generally stigmatized as a social marker, and the *Statenvertalers* working on the Dutch version of the Bible rejected the variant *mijn* as too low and colloquial: 'nunquam myn, ut vulgus hic loquitur' ('never use *myn* as lower class people do'; cited in De Waal 1992: 124). At the Cape, the Dutch low prestige *n*-variant dominated in both grammatical contexts during the first half of the eighteenth century (see Table 6.6).

An even clearer preference for *mijn* (in both functions) existed at the Cape during the early nineteenth century. Forms without *-n* showed an increase

Table 6.5 Attribute possessive pronouns in Dutch and Afrikaans

	Dutch		Afrikaans
	stressed	**unstressed**	
first person singular	*mijn*	*m'n, me*	*my*
second person singular	*jouw*	*je*	*jou*
Third person singular	*zijn, haar*	*z'n, d'r*	*sy, haar*
first person plural	*ons/onze*		*ons*
second person plural	*jullie*		*julle*
third person plural	*hun*	*'d, d'r*	*hulle*

Table 6.6 *Mijn/mij* (in objective and attributive function) at the Cape 1710–1750 (from Ponelis 1993: 205)

	object	**attributive possessive**
total	124	77
mijn	74	46
mij	50	31
mijn as % of total	60%	60%

in object position only from about 1840 (cf. Ponelis 1993: 205), and *mijn* remained an important variant of the attributive possessive until the early twentieth century. It also continued to be used as an object pronoun (ibid.: 226–227; also Malherbe 1917: 82).

While the *n*-less form *mij* dominates in object position, for example, in the dialect texts of Meurant, Boniface, and Zwaartman (see Chapter 2), the publications of the *Genootskap van Regte Afrikaners* (*Die Patriot* and *Ons Klyntij*) maintain a preference for objective *mijn* (73%; objective *mijn* also occurs in Abu Bakr's *Bayân-ud-în* 1869; cf. Van Selms 1979; Ponelis 1993: 205; also Lätti 1978: 114). Variation continued until the early twentieth century, but overt attitudes towards the use of *-n* object forms remained negative. Thus Hoogenhout (1904: 9) remarked in his grammar: 'Akkusative Singular *myn* für *my* ist nicht nachzuahmen' ('Accusative singular *myn* for *my* is not to be imitated').

Objective *mijn* is rare in the *Corpus of Cape Dutch Correspondence* (6%). Its use is limited to a small group of speakers (all older generation; age group 1, class 3a) who have generalized it (B. J. Brümmer, D. J. LeRoux, and H. H. Pienaar). The *n*-form, however, is still used frequently as a possessive. The percentage of *n*- forms is even higher for the third person singular possessive *zij(n)/sij(n)* (cf. also Ueckerman 1987: 15; Ponelis 1993: 227). In the plural, use of possessive *hulle* (third person plural) is relatively widespread: it is found in 38 out of 94 cases (i.e. 40%; Table 6.8).

Summary of pronoun use in the corpus

The frequency patterns for the different pronoun variables in the corpus are in line with the results from other historical studies; that is, very low frequencies are found for *ek* and *hom*, the third person singular neuter variant *dit* is dominant over *dat* but still in competition with *het*, and the third person plural form *hulle* is more prominent in subject than in object position. The overall frequencies of Afrikaans forms are, however, considerably lower than for the morphosyntactic variables (Chapter 5), and metalinguistic comments (such as

Table 6.7 *Mij* and *zij/sij* as possessive pronouns in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	first person singular possessive: *mij(n)*	**third person singular possessive: *zij(n)***
total	940	463
forms with *n*-apocope	262	70
n-apocope as % of total	28%	15%

Table 6.8 Summary: The subject/object pronoun variables in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	ek	***dit***	***hom***	***ons***	***hulle* (subj.)**	***hulle* (obj.)**	***hulle* (poss.)**	***my* (poss.)**	***sy* (poss.)**
Total	3652	2096	451	1236	352	160	94	940	463
no. of Afrikaans forms	858	793	69	448	172	51	38	262	70
Afrikaans forms as % of total	24%	39%	15%	36%	49%	32%	40%	28%	15%

in the above cited letter by De Villiers) suggest that incorporation of these variants into the emerging Afrikaans standard was not uncontroversial. The results for the six pronoun variables are summarized in Table 6.8.

The relativizer

In Dutch the form of the relative pronoun in direct relative clauses is determined by the gender of the noun phrase antecedent: *die* is used for common gender nouns and plurals, *dat* (and its colloquial and dialectal variant *wat*) for neuter singular nouns. In standard Dutch the form *wat* occurs regularly after indefinite pronouns (*dat is iets wat ik al wist*, 'that is something that I already know'), and in clause-linked relatives where no direct antecedent exists (as in *ik kan niet krijgen wat ik nodig he*, 'I cannot get what I need'). In Afrikaans only a single relativizer (*wat*) is used (7).

(7)

Dutch	Afrikaans	
de man **die** Piet heet	die man **wat** Piet heet	'the man who is called Piet'
het meisje **dat** Engels praat	die meisie **wat** Engels praat	'the girl who speaks English'
de meisjes **die** Engels praten	die meisies **wat** Engels praat	'the girls who speak English'

Middle Dutch did not have a separate class of relative pronouns, but used in the typical West-Germanic fashion demonstratives (D-forms: *die, dat, daar, diens*) and interrogatives (W-forms: *wie, wat, waar, wiens, welke;* Burridge 1992: 243). Middle Dutch interrogative pronouns occurred mainly in prepositional relatives (as is also the case in standard Dutch: *de vrouw met wie ik gisteren praatte*, 'the woman with whom I spoke yesterday'), and soon dominated after indefinite pronouns, although D-forms still occurred in this position (Ponelis 1986: 53–54). Use of *wat* with neuter singular antecedents was not uncommon in nineteenth century Dutch. It is also attested for dialectal Dutch where it competes with the D-form *dat* (Sassen 1983; Geerts et al. 1984: 250–251). Use of W-forms (interrogatives) as relativizers is widespread in other West Germanic languages (English, Yiddish, southern German dialects, Frisian as well as the Dutch creoles Negerhollands and Berbice Dutch; cf. Kouwenberg 1994: 236; Den Besten 1996: 16–17).

It is necessary to distinguish clearly between relative pronouns and relative markers (or complementizers; cf. the overview in Lightfoot 1979: 314). Unlike relative pronouns, relative markers are invariant (i.e. they do not vary according to the co-referential noun phrase), cannot serve as prepositional objects and lead to preposition stranding. English *that,* for example, is not a pronoun but a complementizer: **this is the university at that she works* (but periphrasis is possible: *this is the university [that] she works at*). Similarly Frisian *wat*, Bavarian *wos*, Yiddish *vos*, Berbice Dutch *wati* and Afrikaans *wat* are best described as relative markers or complementizers (as argued in detail by Den Besten 1996). Like English *that* Afrikaans *wat* is not only invariant, but also cannot be the object of a preposition (8a). Afrikaans has maintained the (Dutch) construction of 'preposition plus *wie*' (with human antecedent) and '*waar* plus preposition' (with non-human antecedent) to introduce prepositional relatives. However, in cases of preposition stranding *wat* is obligatory (8b; Den Besten 1996: 12; 8b).

(8) a. die vrouw **met wie** ik gister gepraat het — 'the woman with whom I spoke yesterday'
die problem **waarvan** jy gepraat het — 'the problem about which I have spoken'
b. die vrouw **wat** ek gister **mee** gepraat het
die problem **wat** jy **van** gepraat het

That the introduction and grammaticalization of an invariant relative marker is not a necessary or direct consequence of the loss of gender can be seen from the case of Yiddish where a three-gender system has been maintained, but the invariant relative marker *vos* (9) is used in addition to the inflected relative pronoun *velkh* (Lowenstamm 1977; Weinreich 1992 [1949]: 197, 306; Jacobs, Prince, & Van der Auwera 1994: 416).

(9) der mentsh **vos** zitst baym tish iz mayn tate — 'the man who sits at the table is my uncle'
dos bukh **vos** ligt oyfn tish iz interesant — 'the book which lies on the table is interesting'

Vos cannot be the object of a preposition unless the clause contains a resumptive pronoun which then functions as the head: * *der mentsh* ***mit vos*** *ikh hob geredt* ('the man with whom I spoke'), but *der mentsh* ***vos mit im*** *hob ikh geredt*; or unless the antecedent is nonhuman *di teme* ***af vos*** *er redt iz* … ('the topic about which he speaks is …'; Lowenstamm 1977; Jacobs, Prince, & Van der Auwera 1994: 416; Weinreich 1992[1949]: 198). Invariant relativizers are widespread in North Germanic: e.g. Danish (*som*), Icelandic (*sem*), Norwegian (*som*) and Swedish (*som*). An invariant relative marker is also used in Pennsylvanian German (*as*, 'that').

Although the development of an invariant relativizer is certainly not a direct consequence of gender loss, the disappearance of nominal gender distinctions in Cape Dutch Vernacular probably facilitated the generalization of *wat* to common gender and plural antecedents. This was argued by Ponelis (1993: 193):

> [G]ender helps to delimit the respective domain of D- and W-relatives (as common gender *die* vs. neuter *wat* in current Dutch), and neutralization of gender, as in English, Negerhollands and Afrikaans, does remove an obstacle in the path of W-relatives.

Ponelis (1987: 64–65, 1984) further suggested that generalization of the neuter relativizer to common gender and plural antecedents might have been influ-

enced by dialectal and colonial Dutch usage where occasional examples of *dat* (but not *wat*) as an invariant relative can be found.

However, Den Besten (1996: 23) has shown that the use of an invariant relativizer is not attested for Holland varieties of Dutch which featured prominently in the dialectal basis of Afrikaans. Direct dialect origin and continuation of superstrate structures is therefore unlikely. Within the substrate an invariant relative marker is attested for Creole Portuguese (*que;* cf. Ponelis 1993: 193), and Hesseling (1923: 121) has suggested that influence from Malay (which uses the invariant relative marker *yang*) might also have played a role.

In the *Corpus of Cape Dutch Correspondence* the relativizer *wat* occurs in about one-third of relative clauses (32%; Table 6.9). The relatively high percentage of *wat* contrasts with the eighteenth century Cape Dutch Vernacular data analyzed by Ponelis (1993: 193–194) where use of *wat* is still rare. Although *wat* is well established in the corpus, generalization of *die* as an invariant relativizer was clearly the dominant strategy. Frequencies for the neuter D-form *dat* were low (see Table 6.9).

Generalization of the common gender and plural relative pronoun *die* to neuter singular noun phrases (10) paralleled the development in the gender system where the common gender and plural article/demonstrative *de*/*die* was generalized to neuter singular nouns.

(10) a. die Geld **die** wij voor Mr. Hendrickse gecollecterd heeft ('the money which we collected for Mr. Hendrickse'; Aletta April, 27/1/1913)
b. een jong meisje **die** naar school gaat ('a young girl who goes to school'; Katie van Huyssteen, May 1907)

Variation between *die* and *wat* has also been described for Trigardt's diary (Smuts 1968) and the dialect writing tradition (Van Rensburg & Combrink 1984; Raidt 1994 [1992]; see also Table 2.1). It is possible that the choice between *die* and *wat* was stylistically motivated and, at least in part, a function of text register. This is suggested by the discussions surrounding the Bible translation, where *wat* was felt by many not to be appropriate when referring to God (Steyn 1989: 26).

The conventional dating of invariant *wat* to 1800 (Raidt 1983: 160; Ponelis 1987: 68a) appears rather early when considering the distribution patterns summarized in Table 6.9. According to Den Besten (1996) the prominence of *die* in the historical record suggests the existence of two competing systems until the late nineteenth and possibly early twentieth centuries, i.e. a simplified Dutch system which used *die* as a general relative pronoun and *waar/wie*

Table 6.9 Relative marking in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	die	*dat*	*wat*
a. common gender/plural antecendent			
N	225	6	105
% of total	67%	2%	31%
b. neuter gender antecedent			
N	41	18	30
% of total	46%	20%	34%
c. both			
N	266	24	135
% of total	63%	6%	32%

with prepositional relatives (acro-and mesolectal varieties), and a substratum system using only the invariant relative marker *wat* (basilectal varieties).[3]

The demonstrative pronouns *hierdie* and *daardie*

Modern standard Afrikaans has the demonstratives *hierdie*, *daardie,* emphatic *dié* and the independent demonstrative *dit* (Donaldson 1993: 142–145; cf. example (11)). In colloquial Afrikaans a third form, *doerdie* ('yonder'), is also used (Ponelis 1993: 169).

(11) a. **hierdie/daardie** boek is interessant ('this/that book is interesting')
b. **dié** book is interessant ('this book is interesting')
c. Het jy **dit** gelees? / Het jy **hierdie** gelees? ('Have you read this?')

The Dutch demonstratives *deze/dié* (preceding common gender singular and plural noun phrases) and *dat/dit* (preceding neuter singular noun phrases) are attested in our Cape Dutch sources up until the mid-eighteenth century (Scholtz 1963: 126–128). There is but one single case of *daardie* for the eighteenth century in a letter written by fieldcornet Floris Visser (1797). Visser reports an incident between a white colonist (*dese Scholtz*) and a black chief (*dadi Kaffer Capiteijn*). In the context of the letter the use of *dadi* is clearly pejorative, while *dese* appears to have been the unmarked form (Raidt 1993: 287–288). Examples of *hierdie* and *daardie* are rare even in the second half of the nineteenth century, and only one (albeit ambiguous) example (12) has been found so far for the 1860s.

(12) dan zal **dar die** dag bloet geloop het ('then that day blood would have run' or 'then blood would have run there that day'; *Het Volksblad* 14.01.1864, quoted in Raidt 1993:288)

Regarding the origin of the demonstratives, Roberge (1994b:71) has suggested a substrate-oriented explanation. It is likely that basilectal varieties of Cape Dutch Vernacular used deictic adverbs (*hier, daar*, and *doer*) to specify proximity and distance. Evidence for this come from contemporary Griqua Afrikaans where the deictic adverb alone can fulfil the function of a determiner (Roberge 1994b:72–73). The basilectal pattern converged with more mesolectal and acrolectal usage where *die* as the sole demonstrative could be pragmatically 'sharpened by means of a preceding deictic adverb' (ibid.).

Demonstratives consisting of a deictic locative plus definite article are, however, also common in the Germanic languages (cf. also Raidt 1993).[4] The constructions given in (13a, b) are well-established in the spoken language. The use of deictic adverbs has been grammaticalized in Yiddish (13c), Swedish (13d) and Afrikaans.

(13) a. German **hier das** Buch ist sehr interessant ('this book is very interesting')
b. English **these here** people are French (dialectal)
c. Yiddish **ot der** tsimer is groys ('this room is very big')
d. Swedish **den där/här** mannen ('these men')

The forms *hierdie* and *daardie* appear suddenly in the sources from around 1880 without any noticeable period of transition, and they are reported in most grammatical and lexicographical descriptions from the late nineteenth century as the common forms of the demonstrative (cf. Van der Merwe 1971). To explain the absence of *hierdie/daardie* from the pre-1880 historical data with the familiar argument of stigmatization leading to conscious avoidance (Ponelis 1993:169) is unconvincing: this should not have affected the dialect writing tradition, which consciously aimed at representing the colloquial language and did not generally avoid linguistic stereotypes and stigmatized expressions.

To account for the late occurrence of *hierdie* and *daardie* in the sources, Raidt (1993) draws on Henn-Memmesheimer's (1986) notion of 'habitualized' speech patterns (see introduction to this volume). That is, although these forms occurred only infrequently, they nevertheless constituted 'habitualized' norms which existed side by side with more frequent forms, especially emphatic *dié*:

> Our data indicate that the demonstrative *dié* was by far the strongest form before 1800, but that there were other, non-standard variants, i.e. the locative reinforced *hierdie* and *daardie* which had become habitualised patterns … [T]he meteoric rise of the new forms [after 1880] … reveals the potential vitality of unnoticed variants in the process of linguistic change. (Raidt 1993: 291)

However, to assume habitual, but marginal (i.e. infrequent) usage does not explain how and why these variants suddenly assumed generality and were selected into the emerging standard language. It is interesting in this context that Hoogenhout (1904: 10) described in his *Lehrbuch* a rather unstable system of demonstratives and reported the existence of several deictic demonstratives, most of which were not selected into the system of standard Afrikaans and did not crystallize as stable non-standard norms: *di, hier-di, di-een* ('the-one'), *deus-kant-sy* ('this-side-them'; all glossed as 'this'), *daar-di, di-anner-een* ('the-other-one'), *anner-kant-sy* ('the-other-side-them'; all glossed as 'that'; cf. also Van Ginneken 1928: 211 who lists *di-een* and *di-anner-een* as demonstratives). Hoogenhout's description suggests the existence of a highly variable system of deictic determiners which was simplified in the process of standardization.

Only few examples of *hierdie* and *daardie* appear in the corpus: there are nine occurrences of *hierdie* (14), and only six of *daardie* (15). Almost half of these are found in the letters of one writer: J. F. W. Grosskopf (born 1885), a young teacher, writer and enthusiastic Afrikaner nationalist.

(14) a. en denk **hier di** roos of bloem het mama geplant ('and think this rose or flower has been planted by mama'; B. J. Brümmer, 19/9/1907)
b. **hierie** selfde broer van my is sekretaris van hul Debatsvereniging ('this same brother of mine is secretary of their debating society'; J. F. W. Grosskopf, 1/12/1907)
c. om **hierie** brief te pos ('in order to post this letter'; J. Grosskopf, 26/3/1907)
d. ons het **hier di** laaste tyd vrot weer gehad ('we've had lousy weather this last while'; J. F. W. Grosskopf, 3/6/1907)
e. sover as ek kan sien sal ek **hierdie** week sewe volle ure hê ('as far as I can see I will have seven full hours this week'; J. F. W. Grosskopf, 8/12/1907)
f. **hierie** mej. Hofmeyr en haar maat ('this Miss Hofmeyr and her chum'; J. F. W. Grosskopf, 18/2/1909)
g. zy heeft aperkatie vir **hier die** school gemaak ('she has made an application for this school'; Elisabeth Eksteen, 1/1/1922)
h. **hierie** kaartje is van wilde beesten ('this postcard shows gnus'; Jacques Malan, 1910)

i. **hier die** plek is van die bezigheidplek van Pretoria ('this place is the marketplace of Pretoria'; Jacques Malan, 1910)

(15) a. die zeg **daardie** vuiles, en **daar die** vuiles is julle voor mannen ('he says those miserables a and those miserables are men for you'; Aletta Schabort, 1902)

b. die een wat julle **daarie** slag op Bloemfontein gesien het ('the one which you saw that time in Bloemfontein'; J. F. W. Grosskopf, 1/12/1907)

c. myn smaak ook … in **daardie** rigting gelop het ('my taste also … went in that direction'; Jan de Waal, 3/6/1910)

The negation

To express negation of a sentence or sentence constituent in Afrikaans a negator (or negative operator, *nie*-1) and an obligatory scope marker (or associated negative, *nie*-2) are used. The syntactic position of the negator is the same as that of the Dutch negator *niet*, i.e. within the verb phrase and not post-subject as in many pidgin/creole languages (Ponelis 1993: 461–462). The main negator is the negative particle *nie,* but as in Dutch negative adverbs (e.g. *geen*, 'not a, not any', *niemand*, 'nobody', *niks,* 'nothing') can also fill this position. The second negative element (*nie*-2) is clause final and only the lexical form *nie* can be used (16). Use of *nie*-2 is optional after clause-final negative adverbs (17). Its omission after clause-final *nie*-1 is obligatory (18). Colloquial varieties of Afrikaans usually assimilate *nie*-1 and *nie*-2 onto the preceding element and use the reduced enclitic form *-ie* (19).[5]

(16) ek praat **nie** graag **nie** ('I don't like talking')

(17) ek hoor **niemand** (nie) ('I don't hear anyone')

(18) ek ken hom **nie** ('I don't know him')

(19) ek wil-**ie** slaap-**ie** ('I don't want to sleep')

Nie-2 as a scope-marker can negate entire sentences (simplex and complex sentences; cf. 20a) and also sentence constituents (cf. 20b; Waher 1978: 57; Ponelis 1993: 455–457).[6]

(20) a. sentence negation
Wie besef dit **nie** goed **nie**? ('Who doesn't realize it well?')
Wie besef dit **nie** goed dat die stelsel verander moet word **nie.** ('Who

doesn't realize well that the system must be changed?'; cf. Ponelis 1993: 456–457)

b. constituent negation
Nie die man **nie** maar die vrouw was die skuldige. ('Not the man but the woman was the guilty one.'; Le Roux 1923: 167)

Regarding the origin of the brace negation, Dutch influence and language contact (substrate influence) have been advanced as explanations. Following Pauwels (1958), Dutch dialect influence was emphasized by Scholtz and Raidt who have argued that the Afrikaans negation originates from a facultative Dutch dialect pattern which grammaticalized at the Cape (cf. Scholtz 1980: 90–91; Raidt 1983: 190, 1991: 223). However, Den Besten (1985 and 1986) has shown that the syntax of the Afrikaans negation differs structurally from the negation used in dialectal Dutch. That is, in those Dutch (and also German) dialects for which double negation patterns have been reported, the second negator typically occurs clause internally and is thus clearly a 'phenomenon of the 'middle field' (or inner field)' (Den Besten 1986: 204). The second negator can occur sentence finally only:

i. if the V2-rule has removed the verb from the final position, or
ii. if the internal negator is situated close to the end of the clause.

Sentence-final *nie*-2 at the end of complex embedded clauses (such as in example (20a)) is not possible in dialectal Dutch. Furthermore, those Dutch dialects for which the use of a facultative double negation has been reported are southern dialects of Dutch (especially Brabant and Flanders). Southern dialects, however, played only a limited role in the development of Afrikaans (see Chapter 1). Bipartite negation patterns were used only sporadically in Early Modern Dutch (which, however, still shows occasional use of the Middle Dutch negation *en . . . niet*), and are not attested in other colonial varieties of Dutch (Roberge 2000: 119).

Waher (1994a), who accepts Den Besten's argument about the impossibility of a direct continuation of Dutch dialectal negation patterns, nevertheless maintains that Dutch dialect influence was important for the development of the Afrikaans negation. She argues that a general word order change took place in eighteenth century Cape Dutch Vernacular which resulted in the move of scope-bearing elements such as modal auxiliaries, adverbials and the negation particle to the left of the verbal complement, often leading to the replication of the moved element in the original position (21; cf. also Webb 1993 and Kotzé

1989 on leftward auxiliary movement in non-standard varieties of twentieth century Afrikaans).[7]

(21) a. dat een braaf soldat, aen een Lang discours, kan **niet** sich houden, van Oorlogen te spreeken ('that an upright soldier in a long conversation cannot stop himself from talking about wars'; E. Barbier 1739; Pheiffer 1980, appendix 67)

b. waaruit men supponeerde dat dat de redden was waarom zij **niet** met Pienaar **niet** mede naar boord wilde gaan ('from which one assumed that this was the reason why they did not want to go on board with Pienaar'; Van Rheenen's journal, 1793; Godée-Molsbergen vol. 2 1916:153)

The leftward movement of the negator and the resulting replication of the negative particle in the original position led to a situation in which a second sentence internal negative became a common feature of Cape Dutch Vernacular. Initially it probably functioned as an emphasis marker. The development was further supported by the existing Dutch (and also German) dialect structures which allowed for the placement of a second negative in the middle field. Subsequently, the sentence internal second negative was increasingly pushed to the right, i.e. became sentence final, lost its function as an emphasis marker, and grammaticalized into a scope marker since its previously variable position could lead to semantic ambiguity. The transition period of this development was, according to Waher, characterized by the occurrence of multiple sentence internal negative particles. Such triple (and even quadruple) negation patterns remained productive until the early twentieth century (Van Oordt 1913:83), and are still common in colloquial spoken Afrikaans (22; cf. Feinauer 1994; Rademeyer reports such constructions also for Griqua Afrikaans 1938:74).

(22) a. Of was dat, tusse de voornaamste teugstanders van die geseyde predikant **nie** altemit **nie** een afgesproke werk **nie**. ('Or was this between the most prominent opponents of the mentioned minister not perhaps a prearranged work'; Jantje Eenvoudig, 1864; quoted in Waher 1994a:107)

b. Hulle gee vir jou absoluut **geen** handleiding **nie** op universiteit **nie**. ('They give you absolutely no guidance at university'; Feinauer 1994:95)

c. **niemand** sal **nie** daarvan kry **nie** ('nobody will get anything of it'; Rademeyer 1938:74)

The Afrikaans negation is thus believed to have developed 'on the back' of available dialectal patterns of (sentence internal) multiple negation, but it is in itself an autochthonous innovation (Ponelis 1993: 478; cf. also Roberge's 2000: 129 assessment: 'Waher's hypothesis … posits an autochthonous innovation that converged with a metropolitan feature').

A rather different hypothesis was put forward by Nienaber (1965, 1994b) and Den Besten (1985, 1986), who have argued that the construction is the result of substratum influence from Khoe. In other words, Khoe speakers of Afrikaans developed this construction in analogy to their native language which has sentence-final negators. The basilectal construction was later borrowed into the colonists' vernacular.

Nama (as well as Korana) has three negative operators; two of these occur clause finally: (indicative) *tama* and (future) *tite*. The prohibitive negative *taa*, however, usually occupies the first position (for details on the Khoe negation see Ponelis 1993: 470–471; Roberge 2000). When learning Dutch the Khoe shifted the locus of negation to the left, but maintained their substrate pattern of a sentence final negative element which was later re-analyzed as a scope marker (Den Besten 1986: 221). Three main arguments (cf. Ponelis 1993) have, however, been brought forward against substratum influence from Khoe:

i. Nama negation is not strictly sentence final, i.e. the negator *tama* is often moved to the left, and (as noted above) the prohibitive negator *taa* is never sentence final;
ii. *nie*-2 is today facultative in the strongly Khoe-influenced varieties of Northwestern Afrikaans (Orange River Afrikaans) but obligatory in other varieties;
iii. and *nie*-2 is not attested in the historical corpus of Khoe and slave utterances.

Notwithstanding such facultative patterns in today's substrate communities, in the early nineteenth century use of *nie*-2 was clearly associated with non-European speakers of Cape Dutch Vernacular (cf. Nienaber 1994b). An interesting piece of evidence (23) comes from the early nineteenth century, when C. J. Foecee (Fouché) transcribed two verbatim utterances by his Khoe farmhand Stoffel in a report to the magistrate (KT 206, 1810).

(23) a. wat hij **niet** weet **niet** ('what he doesn't know')
b. hij … zeg dat ik hem van avoon **niet** moet los maak **niet** ('he says that I mustn't make him loose tonight')

Ponelis (1993:475) suggests that 'Foecee might be stigmatising Stoffel's speech by putting a vernacular stereotype in Stoffel's mouth'. However, he rejects the possibility that this evidence could point to a real sociolinguistic difference, since 'there is clear evidence soon after 1810 of ample use of *nie*-2 in the speech of whites' (cf. also Raidt 1983:190, 1991:223). This evidence, however, falls almost entirely into the category 'dialect literature' which, although useful for identifying the absence or presence of historical features, generally fails to capture the structures of linguistic variability (see Chapter 2). In 1830 the traveller Teenstra remarked with regard to the negation in Cape Dutch Vernacular 'deze herhaling van *niet* is hier zeer algemeen in gebruik' ('this repetition of *niet* is commonly used here'), and gives several examples of sentence final *nie*-2 in his linguistic representation of a *boerenvrouw* in Caledon.[8] However, as noted by Roberge (2000:143–144, fn. 38), it is not entirely clear that Teenstra's examples of *nie*-2 are meant to reflect the language use of a white woman:

> There are grounds to suspect that the boerenvrouw is actually a woman of color, a fact that Teenstra has encoded linguistically. Beyond the settler agricultural areas closer to Cape Town, the situation of a white farmer living with a mixed-race woman as mistress of the farm, though not routine, would not have been exceptional either.

Roberge (2000:145) has furthermore drawn attention to the fact that *nie*-2 occurs in Trigardt's diary (1837) in the reported speech of Isaak Albach, a Frenchman who lived at the colonial frontier and who was married to a coloured woman. Note that *nie*-2 is entirely absent from Johanna Duminy's acrolectal diary (Franken 1953), and from the main text of Trigardt's more mesolectal diary (Smuts 1968). Roberge concludes that use of *nie*-2 was closely connected to socioeconomic status and ethnicity:

> The most plausible sociolinguistic evaluation of final *nie*-2 at the time of its entry into the philological record is as a variable feature with a strong association with persons of color and with the lowest socioeconomic strata of Europeans, whose communication networks brought them and their children into intimate linguistic contact with the labor force and its children. (ibid.)

Moreover, use of *nie*-2 was also a function of register and style. That negation is a central pragmatic features of conversational interactions was argued, for example, by Cheshire (1998 [1995]:134):

> I want to stress that negation generally … has an interactional role in ensuring the coherence of the emerging discourse. In other words, it can link the current turn to the previous one, by negating a presupposition that has just been expressed, whilst simultaneously ensuring that the interlocutors have a shared

> orientation to the topic that they are pursuing, so that the subsequent turn is felicitous. This cohesive function is perhaps one reason why negation occurs more frequently in spoken than in written discourse: Tottie (1991) found twice as much negation in her sample of spoken English as in the sample of written English.

The spoken and discourse-dependent nature of negation is reflected, for example, in Boniface's vernacular dialogues which show stylistically marked, multiple and emphatic negatives such as *Nee, niks geen drank, nie, niks* ('No, nothing, no drink, no, nothing') which are unlikely to have survived in conventional written prose. That *nie*-2 originates from discourse-dependent structures 'for which appropriate contexts would not readily present themselves in writing' has also been argued by Roberge (2000: 120). Not only is negation generally more frequent in the spoken language, but multiple negation strategies involving emphasis marking will be typical of spoken rather than written discourse or, more precisely, of spoken discourse in heterogeneous language and dialect contact situations where communicative success is precarious and participants revert to the use of salient agreement and negation markers to express unambiguously their intentions and desires. The locus of the innovation and re-analysis which gave rise to *nie*-2 is thus found the multilingual substrate community, where sentence-final *nie* functioned initially as an emphatic tag negation (Roberge 2000: 148), and was used to strengthen the negative force of a statement in spoken interaction. The emphasizing function of *nie*-2 was lost when it became grammaticalized as a scope marker (i.e. as a functional category) and today *nie*-2 is as a rule unstressed.

In a re-assessment of the history of *nie*-2, Raidt (1994 [1992]: 305) remarked cautionary that in the light of contemporary patterns of variable *nie*-2 usage in colloquial and dialectal varieties of Afrikaans, historical linguists might have been over-confident regarding their early dating of *nie*-2.

> Dit wil amper voorkom of die negasie patroon van die sogenaamde 'dubbele ontkenning' tog nie so vas en sonder uitsondering was nie. … Hoewel die ontkenning met eind-*nie* as tweede negasiepartikel voor 1875 al sekerlik die normale negasiepatroon was, kon ontkenningsvariante sonder die eind-*nie* ook gebruik word.
>
> ['It almost seems as if the negation pattern of the so-called 'double negation' was after all not so stable and without exception. … Although the negation with end-*nie* as the second negator was before 1875 definitely the normal negation pattern, negation variants without the end-*nie* could also be used.']

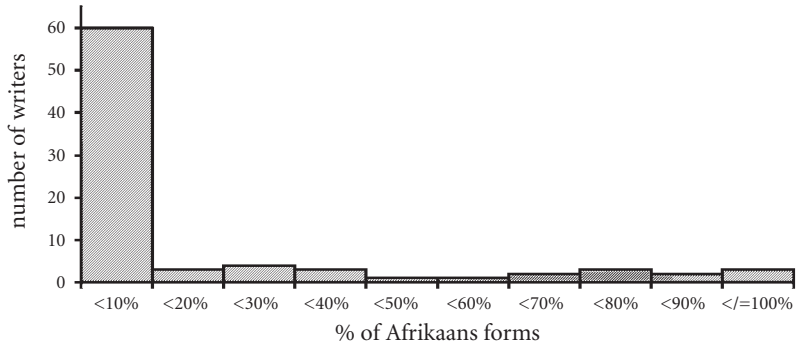


Figure 6.2 Histogram summarizing the use of *nie*-2 by individual writers ($n = 85$)

Leaving open for the time being the question which negation pattern was the 'normal' one (in the sense of quantitatively dominant), pushing the stabilization of *nie*-2 as an obligatory scope marker into the second half of the nineteenth century entails the possibility that obligatory *nie*-2 might – at least in part – be a standardization feature, an interpretation which is supported by the continuing existence of variable multiple negation patterns as well as lack of *nie*-2 in non-standard dialects of Afrikaans (cf. Ferguson 1995: 179 for a similar argument regarding third person singular *-s* in standard English as a standardization feature). Standardization of *nie*-2 not only involved grammatical rule extension (i.e. the identification of *nie*-2 as an obligatory feature), but also 'stylistic extension' as *nie*-2 was a prominent feature only of spoken discourse and, moreover, associated with low socioeconomic status and substrate speakers. *Nie*-2 was identified as a central feature of Afrikaans in the first grammar from 1876 ('Nes in Frans het ons een dubble ontkenning in ons tweemal *nie*', 'As in French, we have a double negation in our twice *nie*'); yet as in the case of the adjective inflection this prescriptive rule was not followed conistently by the grammarian Du Toit himself (cf. Roberge 2000: 165).

Examples of *nie*-2 are rare in the *Corpus of Cape Dutch Correspondence* and only 21% of negative clauses show the Afrikaans brace negation (363 out of 1746 negative clauses). The distribution across individuals writers (Figure 6.2) is highly skewed with only about a third of individuals employing the feature at all. Use of *nie*-2 was not yet categorical for any writer.

Individuals who employ *nie*-2 use it in agreement with the grammatical rules of Afrikaans. Among those writers who use *nie*-2 relatively frequently there is a tendency to omit the second negative in cases where several constituents would intervene between the two negators, especially if the position

of *nie*-2 would be at the end of a complex subordinate clause. It might be that such constructions are partially constrained by the V-final rule in subordinate clauses, which would be upset by final *nie*-2. Examples of (Germanic) negative spread also occur in the data (24), as well as examples of *nie*-3 (25).

(24) a. hy heef **nooit geen** vout gemaak ('he hasn't made a mistake'; Martha C. Hillebrand, 13/10/1903)
b. de schoolkamer zoo klein is dat hy **onmogelyk niet** voor 23 kinderen kan school houden ('the classroom is so small that he cannot possibly hold school for 23 children'; Elizabeth Malan, 9/5/1893)

(25) a. ik glo **niet** dat hulle iets zal doen **nie** voor dat de Vrystaat zyn volksraad zit **nie** ('I don't believe that they will do anything before the Freestate's parliament sits'; D. C. De Wet, 24/11/1888)
b. was **nie** lang genoeg **nie** om die Afrikaanse beweging van alle kante te bespreek **nie**, en makaar syn standpunt te verstaan **nie** ('[it] was not long enough to talk about the Afrikaans movement from all sides and to understand each other's standpoint'; F. S. Malan, circa 1895)

The inclusion of *nie*-2 into the emerging standard language is a somewhat puzzling selection decision. *Nie*-2 was associated with the substrate community, had little 'overt' prestige and was used only infrequently in mesolectal and acrolectal varieties. Moreover, codification of pragmatically emphatic forms is unusual for most standard languages (Stein 1997). *Nie*-2 thus seems an unlikely candidate for inclusion into the Afrikaans standard norm. Yet, it was precisely its socially and stylistically marked identity which contributed to the selection of *nie*-2. In other words, sociolinguistic distinctiveness of the new colonial standard was achieved through 'hyperdemotic' selection decisions which iconically reflected the symbolic identity of Afrikaans as the language of 'the people'. This social symbolism was intricately connected to the positive stereotype of the honest, down-to-earth and common-sense South African *Boer*, which had been established in the popular dialect literature from the 1850s, and whose language (*Boere-Afrikaans*) was the focus of the GRA's propagandistic activities (see Chapter 2; cf. also Irvine 2001 on iconization as the creation of a seemingly natural connection between structural properties of a language and social characteristics of its speakers). At the same time the regularity and exceptionless structures codified in the first grammar distinguished the new norm clearly and unambigiously from the variable and mixed varieties spoken by people of colour and/or those of low socioeconomic status.[9]

In sum, while *nie*-2 is likely to be a relatively old discourse-based, pragmatic feature which emerged in the context of inter-ethnic communication

in the eighteenth century (Roberge 2000), generalization of the structure as an obligatory scope marker occurred relatively late and interacted with early language standardization in the context of Afrikaner identity politics.

The infinitive clause

In Dutch there are three types of infinitive clauses: bare infinitives (26), constructions with the infinitival marker *te* (27), and constructions including the complementizer *om* as well as the marker *te*. Infinitives with *om… te* are obligatory only in cases where the infinitive phrase either describes the head noun further (predicative nominative, as in example (28a)), or clarifies the aim or intent of an action (that is, constitutes an adverbial of purpose as in example (28b); Geerts et al. 1984: 716–717, 789–790, 824; Robbers 1997: 47–51). Bare infinitives can occur in subject position, they can also follow modal verbs and some copula verbs (as in *zij hoorde mij komen*, 'she heard me coming'). In all other cases *te*-infinitives are acceptable in Dutch.

(26) Slapen is gezond. ('Sleeping is healthy.')

(27) Laura was blij met haar moeder **te** reizen. ('Laura was pleased to travel with her mother.')

(28) a. Het is een meisje **om te** zoenen. ('The girl is worth kissing.')
b. Ik vraag het niet uit nieuwsgierigheid maar **om** je **te** helpen. ('I am not asking out of curiosity but in order to help you.')

In spoken Dutch *om… te* infinitives frequently replace *te* infinitives. The distribution of *om… te* and *te* infinitives varies stylistically and use of *om… te* is widespread in informal registers (De Vooys 1953: 163–164; Donaldson 1981: 158; Geerts et al. 1984: 790, Ponelis 1993: 292). *Om… te* infinitives replace *te* infinitives also frequently in Frisian (Hoekstra & Tiersma 1994: 525) and extensive use of *um… zu* is common in spoken, informal varieties of German. Afrikaans has generalized the *om… te* construction and infinitives with *te* are largely obsolete in spoken Afrikaans. Bare infinitives in subject position can be paraphrased with *om… te* in Afrikaans (see (29)).

(29) Slaap is gesond. – **Om te** slaap is gesond.
Laura was bly **om** met haar moeder **te** reis.

Te-infinitives have survived in idiomatic expressions (for examples cf. Raidt 1983: 177f.), and the modals *behoor(t)* ('should/ought to'), *hoef* ('(not) to

need') and *durf* ('to dare') as well as the copula usually combine with *te* (Ponelis 1979: 432; Donaldson 1993: 274–278; Robbers 1997: 89). In more formal styles the verbs *skyn* ('appear'), *blyk* ('appear/be obvious'), *meen* ('mean') and *wens* ('wens') are frequently followed by *te* infinitives (Ponelis 1979: 248). However, it appears that the use of *te* is decreasing in modern Afrikaans. Today, *te*-infinitives are generally perceived as archaic and formal (Robbers 1997: 214).

The basic syntactic structure of the *om… te* infinitive in both Dutch and Afrikaans can be summarized as: *om* + (X) + *te* + (infinitive). The position of (X) can remain empty or can be filled with a noun phrase, prepositional phrase or adverb. In cases where the verb has a separable prefix, the particle *te* is positioned between prefix and stem (*opstaan*, 'stand up' → *om op te staan* 'to stand up'). Non-standard dialects of Afrikaans show considerable variation in the structure of the infinitival complement. This has been described in detail for Griqua Afrikaans. The different non-standard structures are summarized in (30) (from Webb 1993: 168–169). Standard structures are given in brackets.

(30) **om te** werk soek (**om** werk **te** soek), 'to find work'
om te hier **te** kom (**om** hier **te** kom), 'to come here'
om die waarheid sê (**om** die waarheid **te** sê), 'to speak the truth'
om die bees **te** vasmaak (**om** die bees **vas te** maak), 'to tie up the animal'

Roberge (1994b: 87–89) has suggested that these non-standard infinitives originate from basilectal infinitive constructions which only used the complementizer *om*.[10] Mesolectal varieties, which were under the influence of the Cape Dutch acrolect, added the particle *te* in different syntactic positions (as exemplified in the Griqua data). The generalization of *om… te* infinitives in acrolectal Cape Dutch Vernacular, on the other hand, can be linked to inherited metropolitan structures.

1200 infinitive clauses where *om… te* was not obligatory in Dutch were identified in the *Corpus of Cape Dutch Correspondence* (predicative nominatives and adverbials of purpose were excluded from the analysis). In just over half of these infinitive clauses writers used the complementizer *om* in combination with *te* (618; 52%). Compared to the distribution patterns for the other variables, the lack of polarization and overall continuity in Figure 6.3 is remarkable, indicating that the majority of individuals made moderate to frequent use of *om… te* infinitives in environments where (spoken) Dutch would have allowed *te* as well as *om… te*. Thus, on the whole the construction is common in the data without having yet been entirely generalized, and *te* infinitives still occur.

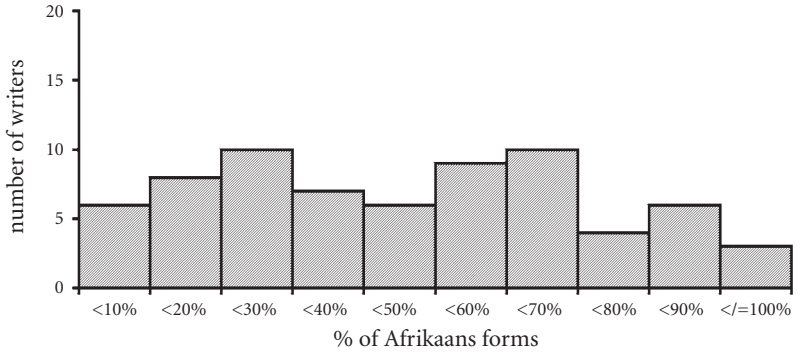


Figure 6.3 Generalization of the long infinitive with *om… te* by individual writers (*n* = 74)

Te-infinitives are categorical in the corpus after the modals *hoeven/hoef* and *behooren/behoor(t)* (cf. the examples given in (31)), which is also the case in modern Afrikaans and Dutch (Donaldson 1981: 152, 154; Geerts et al. 1984: 535–536; Donaldson 1993: 247–249).

(31) a. en ek **hoef** nie **te** sé dat dit heerlik is ('and I don't need to say that this is wonderful'; N. J. De Wet, 12/5/1895)
b. Krause **behoort** 'n First Class gekregen **te** hebben ('Krause should have received a First Class'; F. S. Malan, 1/6/1893)

The verb *begin* shows variation patterns which are partially in line with Afrikaans. In the corpus, *begin* is generally followed by *te* + (V) (32) or a bare infinitive (33). Only one example of *begin* plus *om… te* was found (34). In modern Afrikaans *begin* is usually followed by a bare infinitive, but can also be followed by *om… te* (Donaldson 1993: 274, 277). In Dutch, on the other hand, *beginnen* is generally followed by (*om*) *te*. Use of a bare infinitive is only possible if the infinitive follows *beginnen* directly (*Ik moet beginnen werken*, 'I must begin working', **Ze beginnen vandaag werken*, 'they begin today working'; cf. Van den Toorn 1982: 193; Geerts et al. 1984: 585).

(32) ***begin*** + **te** + (**V**)
a. Zaterdag het dit weer **begint te** reënt ('Saturday it began raining again'; Johanna Brümmer, 29/3/1895)
b. de vogels **begen** al **te** eet er aan ('the birds are already beginning to eat it'; Hester Vermaak, 19/2/1906)
c. **begon** de menschen **te** huilen ('people began to cry'; W. C. Cronje, 19/7/1905)

(33) ***begin*** + (V)

a. de tuin **begint** prachtig **worden** ('the garden is beginning to become lovely'; B. J. Brümmer, 19/9/1907)

b. het ek dadelik **begin** geld bij mekaar **maak** ('I immediately began getting money together'; Lelon Malan, 31/10/1917)

c. ek **begin** al so stadig die motor **nasien** ('I am beginning slowly to check the car'; Andries Hauptfleisch, 7/9/1921)

(34) ***begin*** + ***om te*** + (V)
wil ik by de Einde van deze maand **begin om te** snij ('I want to begin cutting at the end of this month'; Eddie Smith, c. 1900)

The complementizer *om* also occurs as a causal conjunction in the corpus (35), replacing the standard form *omdat* ('because'). Use of *om* in this function has been described for earlier Cape Dutch Vernacular texts (Roberge 1994b: 87). It was branded by Malherbe (1917: 104) as colloquial and strongly discouraged.

(35) pa moet voor hem een wijnig geld stuur **om** hij niets heeft ('father must send a bit of money for him because he has nothing'; Petronella van Huyssteen, circa 1906)

Non-standard infinitives of the type illustrated in (32) do not occur in the *Corpus of Cape Dutch Correspondence*, yet were found in a letter (36) written by a native English speaker who, according to his own testimony, had just begun to learn Dutch (SAL-MSC1-34-9). The letter was not part of the main corpus.

(36) a. ik heeft ooit probeerd **om te** oplos ('I have also tried to figure out')

b. **om te** zien zoovele arme kinderen ('to see so many poor children')

Objective *vir*

The use of *vir* (Dutch *voor*, 'for') as an object marker is a typical feature especially of spoken Afrikaans (cf. Molnárfi 1999). *Vir* is obligatory with prepositional objects and in bi-transitive constructions where the direct object is followed by the indirect object (37a). In such cases *vir* can be replaced by the historically correct Dutch preposition (e.g. *aan*, 'at', or *naar*, 'to') in formal registers. *Vir* is facultative in bi-transitive constructions with the indirect object preceding the direct object. It is also facultative before direct objects in general (37b; see Carstens 1991: 59–60; Donaldson 1993: 342–343; Ponelis 1993: 265–266; Raidt 1994 [1969]).

(37) **a. obligatory objective *vir***

Dutch	Afrikaans
Ik luister **naar** Oom Jan.	Ek luister **vir** (na) Oom Jan. ('I listen to uncle Jan.')
Hij heeft het boek **aan** zijn broer gegeven.	Hy het die boek **vir** (aan) sy broer gegee. ('He has given the book to his brother.')

b. Facultative use of *vir*

Dutch	Afrikaans
Hij heeft zijn broeder het boek gegeven.	Hy het (**vir**) sy broer die boek gegee. ('He has given the book to his brother.')
Ik zie Nelsie.	Ek sien (**vir**) Nelsie. ('I see Nelsie.')

Use of *vir* is optional with proper nouns and pronouns that are scrambled (i.e. that appear to the left of adverbs, as in *Hulle het (vir) Sarie gister gesien*, 'They saw Sarie yesterday'), but obligatory when they are not scrambled (*Hulle het gister vir Sarie gesien/*Hulle het gister Sarie gesien*; Ponelis 1993: 268). The use of facultative *vir* in contemporary Afrikaans is governed by the nature of the noun,[11] animacy (inanimate nouns do not usually take *vir*),[12] emotivity and degree of formality (emotional involvement favours use of *vir*; Ponelis 1979: 202–203, 1993: 266ff.). There is some debate surrounding the question as to whether facultative *vir* should be interpreted as a preposition, or whether it is more adequately described as a case prefix or (direct) object marker (Den Besten 1981, 2000; Markey 1982; Molnárfi 1997; see Ponelis 1993: 269–271 for an argument in favour of the preposition status of *vir*).

There are some marginal (Dutch and German) dialectal structures which suggest that objective *vir* is not 'entirely without a Germanic antecedent' (Ponelis 1993: 271; cf. also Raidt 1994 [1976]; LeRoux 1923: 140ff.). Examples of Dutch *voor* exist in bi-transitive constructions (indirect object following direct object: *mijn oupa heeft een trui voor me gebreid*, 'my grandfather knitted me a jersey') as well as with prepositional objects (*ik zal voor hem betalen*, 'I will pay for him'), and in Early Modern Dutch the preposition *voor* occurred in many contexts where modern Dutch would require different prepositions (for a detailed discussion and examples cf. Raidt 1994 [1976]). However, the evidence is scarce for those environments where Afrikaans allows facultative *vir,* especially preceding direct objects.[13] Since the construction is also not attested for colonial Dutch, Afrikaans historical linguists have generally interpreted objective *vir* as an extraterritorial innovation triggered not so much by the Germanic input but by substrate influence from Creole Portuguese (Raidt 1994 [1976]).

The similarities between the Creole Portuguese construction *per* plus personal object and the Afrikaans construction with *vir* were first noted by Hugo Schuchardt (1885; cf; also Hesseling 1899, 1923; Raidt 1994 [1976]). In Creole Portuguese (38) both direct and indirect objects (including personal pronouns, proper names and substantive phrases) can be preceded by the preposition *per* (and its variants).

(38) nos atja ponta **per** eli
WE CAN SHOOT FOR HIM (Raidt 1994 [1976]: 133)

The substrate hypothesis is supported by the fact that early attestations of objective *vir* occur predominantly in the speech of slaves most of whom had at least some knowledge of Creole Portuguese and/or Malay (for examples cf. Raidt 1994 [1976]; also Ponelis 1993: 273f.; Den Besten 2000, also draws attention to the use of object markers in Malay). It is commonly assumed that objective *vir* was a well entrenched feature of Cape Dutch Vernacular speech by the beginning of the nineteenth century, and Changuion (1844) described the widespread use of the construction as 'very noticeable' in the mid-nineteenth century:

> Zeer opmerkelijk is het gebruik van overal het voorzl. *voor* tusschen het wkw. en den 3. of 4. naamv. in te voegen: ken je nie voor mijn nie? Kent gij mij niet? Zie je voor die paard? Ziet jij dat paard? Ik geef voor jou wat moois. Ik zeg voor hem: as jij sla voor die kind, ik zal voor jou bijkom. Ik zeg hem, als jij het kind slaat, zal ik je weten te vinden. (Changuion 1844: 130)
>
> ['Very noticeable is the habit of inserting particularly the preposition *voor* ['for'] between the verb and the third or fourth nominal case: don't you know for me? Don't you know me? Do you see for the horse? Do you see the horse? I give for you something nice. I tell for him: if you hit for the child, I will come past for you. I tell him, if you hit the child, I will know where to find you'.]

The last example given by Changuion suggests the possibility that objective *vir* was a social marker; the image of a child-beating thug threatened with violence in return is certainly suggestive, and according to Raidt (1994 [1969]: 114, 1994 [1976]: 140) the use of objective *vir* was indeed socially stratified. This interpretation is also supported by the discussions surrounding the translation of the Bible into Afrikaans, where the use of objective *vir* was a highly contested issue (Raidt 1994 [1976]: 141ff.). Frequent use of *vir* was typical for many vernacular texts in the late nineteenth century. However, important differences existed between individual authors: S. J. du Toit, for example, used *vir* almost excessively in his writings; others made little use of the construction, often limiting it to

the reporting of direct speech, especially when representing people of colour (see Raidt ibid.).

Altogether 237 instances of objective *vir* (with the variants *voor* and *ver*) were found in the *Corpus of Cape Dutch Correspondence* and 65 individuals (43% of the sample) made use of the feature. A tendency to use objective *vir* more frequently before proper nouns than before other nominal categories is noticeable in the corpus (39).

(39) a. groet de kinderkies en **ver** Frans ('greet the children and Frans'; B. J. Brümmer, 28/11/1907)
b. beste groete aan almaal en ook **vir** Katie van Winnie ('many greetings to everyone and also to Katie from Winnie'; Wynanda Hoogenhout, 13/3/1911)

Frequent (27 instances) was the construction *jammer vir iemand wees* ('to be sorry for someone, to pity someone'), which was reported by Changuion (1844: xiv) and remarked upon by Scholtz (1965: 93) in his analysis of the Hendrik Kok dialogues as 'eienardig' ('strange'). The use of the copula *lijken/lyk* ('look, appear') plus *vir*, followed by what Ponelis has termed an 'experiencer NP', is also noteworthy (as in *dit lyk vir my* 'this appears to me'; cf. Ponelis 1993: 275–276; Raidt 1994 [1969]: 95; Donaldson 1993: 342).

The quantitative analysis presented in Table 6.10 and Table 6.11 indicates that use of *vir* was at least partially dependent on the lexical verb involved; lexical restrictions are also attested for contemporary Afrikaans, where certain transitive verbs preferably combine with *vir* (cf. Raidt 1994 [1969]; Carstens 1991: 60). It is surprising that *vir* occurs with such low frequency for *geven/gee* ('give'), *vragen/vra* ('ask') and *zeggen/sê* ('say'); these verbs typically combine with *vir* in modern Afrikaans.

Table 6.10 Use of objective *vir* with different lexical verbs preceding direct and indirect objects in the *Corpus of Cape Dutch Correspondence*. The distribution is statistically significant at $p < 0.001$ (chi square).

	total	**no. of objective *vir***	**% of objective *vir***
sturen/stuur ('send')	82	34	42%
schrijven/skryf ('write')	124	20	16%
zeggen/sê ('say')	113	12	11%
vragen/vra ('ask')	67	7	11%
vertellen/vertel ('tell')	34	3	9%
geven/gee ('give')	133	8	5%

Table 6.11 Use of objective *vir* with different lexical verbs preceding direct objects in the *Corpus of Cape Dutch Corresponence*. The distribution is statistically significant at $p < 0.001$.

	total	**no. of objecive *vir***	**% of objective *vir***
lijken/lyk ('look, appear')	19	12	63%
zien/sien ('see')	133	17	13%
ontmoeten/ontmoet ('meet')	30	3	10%

The periphrastic possessive

Apart from the use of possessive pronouns, three other constructions are available in Dutch to indicate possession: the sigmatic genitive (*-s*), the *van* construction and a periphrastic construction with the pronouns *zijn, z'n/haar,d'r/hun, ze* (which vary according to gender and number of the possessor; Geerts et al. 1984: 53; Haeseryn et al. 1997). The three constructions are illustrated in examples (40) to (42). In spoken Dutch the weak form of the pronoun is typically used, although strong forms have made inroads into spoken Dutch (cf. Norde 1998).

(40) sigmatic genitive — moeder**s** huis ('mother's house')

(41) *van* construction — het boek **van** mijn vriendin ('the book of my (girl)friend')

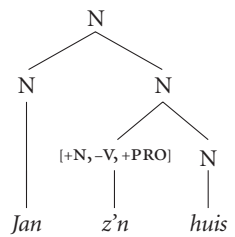
(42) *z'n/d'r* construction — mijn vriendin **d'r/haar** boek ('my (girl)friend her book')
mijn vriend **z'n/zijn** boek ('my friend his book')

While the sigmatic genitive had already disappeared by the eighteenth century (Scholtz 1963: 108), the *van* construction is still (albeit rarely) used in modern Afrikaans (Donaldson 1993: 99). A periphrastic possessive similar in surface structure to the Dutch *d'r/z'n* construction is, however, highly productive in Afrikaans. The Afrikaans construction (43) uses an invariant form *se* which can follow not only animate nouns (as in Dutch), but also inanimate nouns, personal pronouns and other noun phrases, including expressions of measurement and time (cf. Donaldson 1993: 98–100; Ponelis 1993: 231–239; Roberge 1996: 123–125; for a generative account see Oosthuizen and Waher 1994). Periphrastic possessives with *se* are perfectly acceptable in spoken and written Afrikaans (Donaldson 1993: 98; Ponelis 1993: 241), while the *z'n/d'r* construction is marked as colloquial in Dutch (Geerts et al. 1984: 208–209).

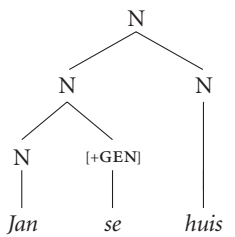
(43)	a. moeder **se** fiets	(Dt. moeder d'r/haar fiets)	'mother's bike'
	b. die huis **se** tuin	(Dt. *het huis z'n/zijn tuin)	'the garden of the house'
	c. hy **se** boek	(Dt. * hij z'n/zijn boek)	'his book'
	d. vyf kilo **se** aartappels	(Dt. vyf kilo aartappels)	'five kilo potatoes'
	e. tien uur **se** kant	(Dt. tegen tien uur)	'around ten o'clock'

The use of personal pronouns as possessors, although not part of the system of standard Afrikaans, is widespread in non-standard varieties of Western Cape Afrikaans (Ponelis 1993: 231) and Griqua Afrikaans (Rademeyer 1938: 66–67; cf. also Malherbe 1917: 83). Afrikaans, furthermore, allows for the insertion of postmodifying clauses between possessor and possessive pronoun (as in *die man met die mooi huis se kind*, 'the child of the man with the nice house') and the construction is (potentially infinitely) recursive (as in *die man se dogter se vriend se boek…* 'the book of the friend of the man's daughter…'). The Afrikaans periphrastic possessive with *se* is not only more productive than the Dutch *z'n/d'r* construction, but also differs structurally from the Dutch pattern. The invariant form *se* is best described a free possessive particle [+Gen] which cliticizes to the preceding noun phrase. The possessor plus *se* thus forms the determiner which specifies the head noun (Le Roux 1923: 89; Den Besten 1978: 31–38). Donaldson (1993: 99, fn. 54), for example, notes the assimilation of *se* to the possessor in sentences such as *die Bloemfontein se stadrad* ('the Bloemfontein city council'), yielding a pronounciation almost indistinguishable from die adjectival construction *die Bloemfonteinse stadrad*.

Dutch



Afrikaans



Originally Afrikaans invariant *se* had been interpreted as a spontaneous reanalysis of the Dutch pronominal construction which was triggered by loss of gender agreement in the nominal system (Scholtz 1963: 107–108). However, it is today generally accepted that the development of the structure owes much to language contact and convergence, and strikingly similar constructions have been reported for Creole Portuguese, Malay and Nama (LeRoux 1923: 97–98; Valkhoff 1966: 17–18, 227–229; Den Besten 1978: 31; Ponelis 1993: 225).[14] All three languages employ a possessive which is periphrastic in structure, contains

a free and invariable possessive particle, shows the same order of elements, and allows noun phrases and personal pronouns as possessors (44)–(46).

(44) Creole Portuguese

me **sua** nomi	minya **sua**
I POSS name	I POSS
'my name'	'mine'

(45) Malay

Dodi **punya** rumah	saya **punya** rumah
Dodi POSS house	I POSS house
'Dodi's house'	'my house'

(46) Nama

khoeb **di** haab	sada **di** haab
man POSS horse	we POSS horse
'the man's horse'	'our horse'

(examples from Ponelis 1993: 244–245)

The Creole Portuguese possessive (*sua*, weakened to *su*) furthermore allows for the expression of temporal relations in ways similar to Afrikaans (Ponelis 1993: 246). In the case of Nama, however, the genitive particle *di* is not obligatory and direct substrate influence is thus less likely. Roberge (1996: 134, 141) has suggested that the facultative particleless possessive structure in Nama supported the juxtaposition of possessor and possessed noun in some basilectal varieties. He cites the following construction from a letter written by 'a Griqua' – who would have been fully or partially bilingual in Khoe and Cape Dutch Vernacular (see Chapter 1) – to the *South African Commercial Advisor* (1830): *in ander menschen land* ('in other people's land'). Similar constructions were found by Luijks (1998) in the letters of the Nama leader Jan Jonker Afrikaner: *God woorte* ('God's words'), *my Vader blads* ('my father's place'). Juxtaposition of possessor and possessed noun also occurs in the Pidgin German of Namibia (Deumert 2003d) which has a Nama substrate: *ihr bruder sachen* ('her brother's things'), *der schwarze schule schulmeister* ('the teacher of the black school').

The lexical form *se* has been interpreted as a reduced allomorph of the Dutch pronoun *zij(n)*. Ponelis (1993: 235), for example, explains *se* as a result of atonic reduction and nasal deletion (*sein* → *sen* → *se*). Ponelis suggested that *se* was part of the spoken language in the eighteenth century (also Scholtz 1963: 109). Roberge (1996), however, has argued that to explain *se* as an allomorph of Early Modern Dutch *zijn* could be 'barking up the wrong tree' be-

cause *se* is so very rare in the historical sources (where *sen* and *syn* are the common forms). Moreover, use of masculine possessive pronouns with feminine or plural antecedents was rare until 1850, and Dutch gendered forms are attested until the early twentieth century (Malherbe 1917: 83; Le Roux 1923: 88–89; Bouman & Pienaar 1925: 76; Scholtz 1963: 108; cf. also Elffers 1903: 12, who gives *s'n* as a generalized singular form, masculine and feminine, and *hulle* as the plural form; for a detailed discussion of the diachronic development cf. Roberge 1996: 135ff.; also Ponelis 1993: 232–237). It is thus likely that the origin of invariant *se* is found outside of the acrolectal Dutch system. In other words, *se* is a substrate feature and as such probably derived from the reduced Creole Portuguese possessive formative (i.e. *su (su(w)a → su → se*). The use of the borrowed invariant particle was initially limited to basilectal slave varieties (Khoe varieties as noted above probably used a simple juxtaposition structure to express possession), but later converged with the Dutch-based periphrastic possessive of the European settlers which used gendered emphatic forms (*syn* and *haar*, as well as *hulle* for the plural).

Over one-third of the writers in the *Corpus of Cape Dutch Correspondence* make use of periphrastic possessives in their writings (38% of the sample; 52 writers). Of the 148 periphrastic constructions in the corpus, 55 had a female or plural possessor. About half of these combined with a variant of *zyn* (masculine pronoun) or *se* (possessive particle; (47)–(50)). In the majority of cases the nominal possessors were animate and human. Non-animate nouns occur as possessors but are rare. Only three examples were found in which a pronoun fills the slot of the possessor (49).

(47) **Possessor: NP** [+**female**]

a. twee van Eliza **zyn** broeders ('two of Eliza's brothers'; Johan Brümmer, 12/7/1900)

b. Hester **zijn** schoolgeld ('Hester's school money'; H. H. Pienaar, 17/5/1905)

(48) **Possessor: NP** [+**plural**]

a. zyn kenders **zen** gesontheid ('his children's health'; B. J. Brümmer, 14/10/1904)

b. die boontjies **ze** kos ('the edible part of beans'; Bettie Smit, 23/11/1917)

(49) **Possessor: NP** [+**pronoun**]

a. met hun allen **zen** boerderij ('with their farming'; B. J. Brümmer, 26/4/1902)

b. mekaar **zyn** standpunt ('each other's point of view'; F. S. Malan, c. 1895)
c. mekaar **se** gesondheid ('each other's health'; J. Grosskopf, 26/3/1907)

(50) **Possessor: NP [–animate]**
a. die zoo **z'n** postkaartjes ('the postcards of the zoo'; Jacques Malan, circa 1910)
b. Queenstown **se** kerk ('Queenstown's church'; Hlela, 29/9/1906)
c. Barkley **zijn** distrek ('the district of Barkley'; E. Smith, c. 1900)

It appears that there existed stylistic constraints regarding variation between the invariant particle and the use of gendered and plural pronouns. Hester van der Bijl, for example, uses mostly invariant *se* (spelled *sê*) in her rather chatty vernacular letters (*my ma sê soetste kind*, 'my mother's sweetest child'; *sy Pa sê ouders,* 'my father's parents'). Yet when writing within the same corpus of letters about the recent death of her sister-in-law Johanna Hoogenhout, she uses the correct Dutch pronominal form (*Johanna haar dood*, 'Johanna's death'; also *Mijnheer Roussouw zijn dood*, 'Mr. Roussouw's death'). Rachie Malan (51) uses the construction with *se* in a letter (8/5/1909) which otherwise conforms entirely to the norms of metropolitan Dutch (i.e. other linguistic markers of Afrikaans such as inflectional reduction, loss of gender, nominative *ons*, *nie*-2 or objective *vir* are absent from her letter). The construction occurs in the context of a short anecdotal description of the monolingual limitations of English speakers who cannot make themselves understood on a German ship:

(51) het wekt 'n mens **ze** lachspieren op ('it wakes up your laughing muscles')

Summary: Comparing distribution patterns

Individual writers in the *Corpus of Cape Dutch Correspondence* make use of the linguistic features which are now part of the Afrikaans standard norm to differing degrees. With regard to loss of inflection and gender, language use is polarized and two well-defined norms (metropolitan Dutch and the emerging standard of Afrikaans) are discernible in the corpus data. At the same time, the distribution is continuous and many writers are found in the middle field of variable usage. The pronoun variables, on the other hand, show less continuity in the middle field and more pronounced adherence to the norms of metropolitan Dutch. The distribution is skewed for *nie*-2 (with the majority of individuals not making any use of the Afrikaans brace negation, positively skewed). A skewed distribution was also found for the loss of the preterite (with

the majority of individuals making extensive use of the perfect as a generalized marker of pastness, negatively skewed). Lack of polarization is noticeable in the distribution pattern for the *om... te* infinitive, which is characterized by a general and continuous distribution across the middle field.

An important question is whether and how such distribution patterns across speakers (or writers) in apparent time relate to the diffusion of a linguistic change in real time. In other words: how do the distributions summarized in the histograms relate to S-curve graphs which describe the spread of innovations across time (linguistic and others; cf. Hakkarainen 1983; Mahajan & Peterson 1985)? The S-curve model predicts slow spread in the initial stages of a change, followed by a period of (rapid) acceleration in which the change catches hold, and finally the trajectory of the diffusion slows down before it is completed.

Chambers and Trudgill (1980: 178–180) have argued that the S-curve model can also be applied to the description of change in apparent time by plotting speakers (rather than dates) on the horizontal axis (the vertical axis remains unchanged, indicating the percentage or frequency of usage of a linguistic variant). It is then hypothesized that those speakers who show low percentages 'will move rapidly into the middle areas and into the high percentages in the course of time' (Chambers & Trudgill 1980: 180). Similarly Wardhaugh (1992: 212–213) suggested that an apparent time curve which closely approaches the ideal S-curve can be interpreted as depicting a change in progress which has reached its mid-point. To link diffusion across speakers to language change across time is supported by the theoretical reasoning which underlies the S-curve model: changes in speaker behaviour are the result of interpersonal communication within a group; that is, contact between individuals is seen as the driving force in language change (cf. also Nettle 1999, Chapter 3), and diffusion is the result of gradual adoption through imitation (Hakkarainen 1983).

Figure 6.4 plots the data for loss of gender and *nie*-2 by arranging writers on a continuum according to the degree to which they conform to the Afrikaans norm. In an ideal S-curve the proportions should be roughly: 37.5% of writers $\leq$ 20%, 25% of writers 21–80%, 37.5% of writers 81–100% (Bailey 1973: 77).

As illustrated in Table 6.12, in the case of gender regularization we have a strong middle field of variable usage which might indicate a linguistic change which is fast approaching completion. Only few speakers still maintain a bipartite gender system. Following Trudgill and Chambers (ibid.), it can be argued that those writers who show low and variable percentages will soon move

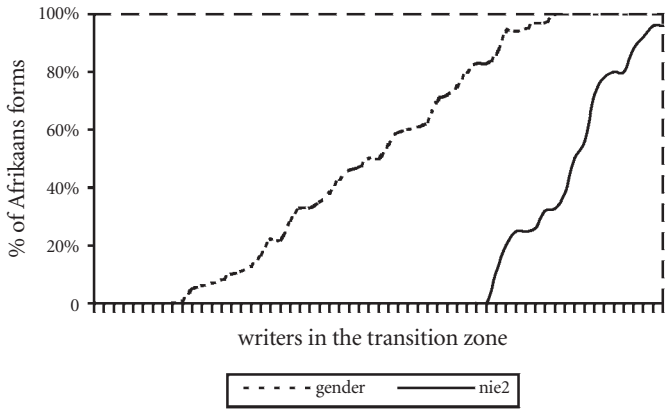


Figure 6.4 Diffusion of gender regularization and *nie*-2 across speakers (apparent time) in the *Corpus of Cape Dutch Correspondence*

into the higher percentages. Alternatively, however, the distribution could also reflect a pattern of stable variation and multiple vernacular norms. *Nie*-2 is probably best described as an incipient linguistic change; around 1900 only few speakers make use of this feature in their written language. However, while the S-curve model has been useful for description and also prediction of long-term processes of diachronic change, it relevance of the study of change in apparent time is less well-established and such synchronic patterns thus need to be interpreted carefully (cf. Mahajan & Peterson 1985: 25, 41–43; Devitt 1989: 48).

Table 6.12 Distribution of individual writers in the *Corpus of Cape Dutch Correspondence*

	'ideal' S-curve	**gender regularization**	***nie*-2**
0–20%	37.5%	30%	72%
21–80%	25%	37%	18%
81–100%	37.5%	33%	6%

Notes

1. The general object pronoun for the third person singular is *hun*. *Hen* is used in more formal registers as a direct object and after prepositions (Donaldson 1981:53–55; Geerts et al. 1984:171).

2. There appears to be a possible additional substrate source in Malay as well: *mama dong* (mother-3rd plural; Den Besten 2000:960).

3. Substrate systems also made use of zero marking since in Khoe the relative pronoun is sentence final (Carla Luijks, personal communication).

4. Cf. LeRoux (1923:67), Scholtz (1963:138–139); for the use of deictic adverbs in standard German cf. Erben (1980:230); for dialectal German cf. Henn-Memmesheimer (1986:186–189); for Yiddish cf. Weinreich (1992 [1949]:191); for English cf. Trudgill (1990:79–80); for Swedish cf. McClean (1947:108–109), Holmes and Hinchliffe (1997:41, 70–71). A similar form (*da di*) is also attested for Negerhollands, cf. Hesseling (1905:100).

5. Cf. Donaldson (1993:401ff.) and Ponelis (1993:453ff.); for generative account of the Afrikaans negation cf. Waher (1978), Robbers (1992) and Oosthuizen (1998).

6. Note in this context that in some varieties of Afrikaans repetition of syntactic constituents also affects particles such as *al* (*ek is al moeg al* 'I am already tired'), prepositions such as *in* (*hy woon in hierdie huis in* 'he lives in this house') as well as verbs (*sing sal sy sing* 'She will [definitely] sing'; cf. Boonzaier 1989:282–283 for examples).

7. Leftward movement of negative markers has been reported for many languages, including pidgins/creoles (Horn 1989). Note also that in modern Afrikaans, unlike Dutch, the preferred position for *nie*-1 is towards the left of the verbal complement: *zij hadden naar mijn verhaal* ***niet*** *geluisterd* (Dutch); *hulle het* ***nie*** *na my verhaal geluister* ***nie*** (Afrikaans; 'they haven't listen to my story').

8. Interestingly, the slave in Teenstra's dialogue uses *nie*-1 in post-subject position: *hij niet spreek* ('he not speak'). Post-subject negation occurs also in Malay Afrikaans texts such as Abu Bakr's *Bayân-ud-în* (cf. Ponelis 1981:76).

9. Cf. also Roberge (2003:26): 'The idea was to promulgate a set of linguistic norms that were neither too strongly associated with people of colour nor emblematic of the urban elite'.

10. Infinitives with the structure *om* + (X) + (infinitive) are also attested for Early Dutch and the southern Dutch dialects (Ponelis 1993:292; LeRoux 1923:104).

11. Hierarchy: proper nouns (almost always) > pronouns (frequently) > nouns (seldom).

12. The use of *vir* plus inanimate object pronouns has, however, been described for some varieties of Western Cape Afrikaans (Kotzé 1983:256f.).

13. According to Schuchardt (cited in LeRoux 1923:141) the construction *er sagte für mich* ('he said to me') is attested for the dialect of Thuringia. In the Cologne urban vernacular similar constructions are possible (e.g. *gib das für mich* 'give that to me'; Peter Deumert,

personal communication). However, it is likely that this is a more recent development which was not part of the seventeenth century Middle German dialect input at the Cape (see Chapter 1).

14. Some dialects in Flanders also make use of an invariant possessor (*se, sen* if the following word begins with a vowel; cf. Taeldeman 1994); however, since Holland dialects were dominant in the historical base of Afrikaans, direct dialect influence is unlikely.

Chapter 7

The Cape Dutch variety spectrum

Clusters, continua and patterns of language alternation

> Statistics can only tell us what is normal. What is natural – the goal of linguistic theory – must be determined on the basis of patterns.
>
> C.-J. Bailey (1980). *Conceptualizing dialects*

Identifying lects in the data

While in Chapters 5 and 6 the distribution patterns of individual linguistic variables were considered separately, the final (multivariate) analysis describes patterns of co-variation between variables. It was argued in Chapter 4 that numerical clustering methods such as cluster analysis, multidimensional scaling and principal components analysis (PCA) are useful tools for the grouping of entities (in this case the idiolects of individual writers), and allow for the identification of large-scale linguistic varieties in terms of similarity and co-occurrence. Ten linguistic variables, all of which are of diagnostic value for the description of Afrikaans as an independent linguistic system, were included in the multivariate analysis:

i. loss of neuter gender, understood as loss of agreement between article/demonstrative and neuter noun (variable GENDER);
ii. loss of person and number distinction in the present tense, main verbs only (variable VERB);
iii. loss of inflection in the infinitive (variable INF);
iv. loss of inflectional endings in the past participle, loss of *-t* and *-en* considered separately (variables PP_T and PP_EN);
v. loss of the preterite and generalization of the perfect, main verbs only (variable TENSE);
vi. use of the Afrikaans pronoun variants, combined scores for the use of *ons, hulle, dit, my/sy* (possessive) and relative *wat* (variable PRONOUN);

vii. generalization of the *om... te* infinitive (variable OMTE);
viii. use of *nie-2* (variable NIE2);
ix. use of objective *vir* per 1000 words (variable VIR).

Except for VIR all other variables were quantified as percentages. The decision to use combined scores for the pronoun system was motivated by pragmatic considerations, i.e. only few individuals had scores for all the different pronominal subcategories. Altogether thirty-five individuals had sufficiently high scores on all ten variables (see Chapter 5) and were included in the multivariate analysis.

The results from hierarchical cluster analysis (Ward's method and average linkage; distance measure: Euclidean) indicate that in terms of their linguistic performance three main linguistic groups can be distinguished (see Figure 7.1 and 7.2 for the results from Ward's method).

The three-cluster solution was confirmed using complete linkage, average linkage and *k*-means clustering. Cluster membership was found to be stable with four exceptions (Andries Hauptfleisch, Carolina Leipoldt, Martha Basson and Louis Botha). Using multidimensional scaling the relative location of the individuals in the two-dimensional space was computed (Figure 7.3; each

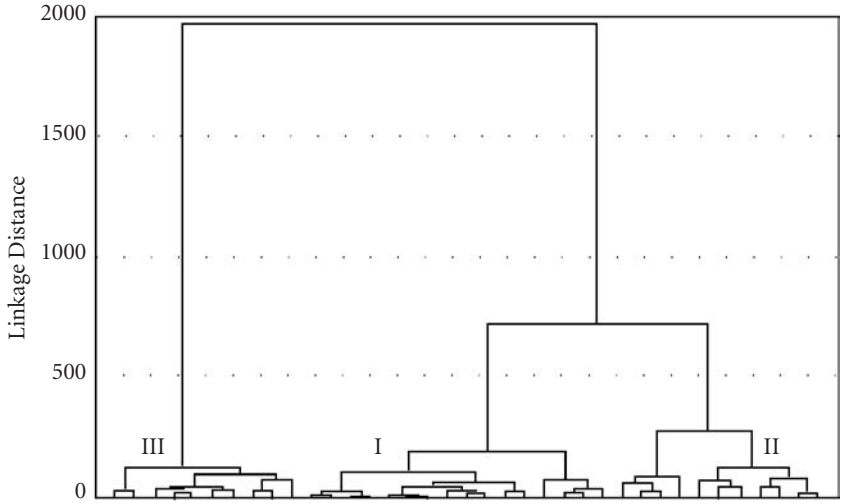


Figure 7.1 Tree diagram (Ward's method, Euclidean distances) for the variables GENDER, VERB, INF, PP_T, PP_EN, TENSE, PRONOUN, OMTE, NIE2 and VIR. Individual writers are located on the x-axis. Three main groups (I to III) were identified by the cluster algorithm.

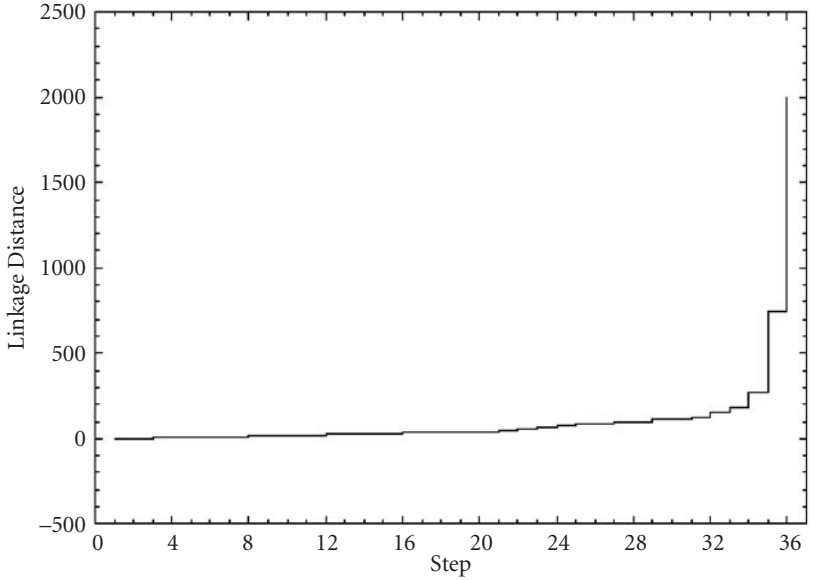


Figure 7.2 Graph of amalgamation schedule for Ward's method

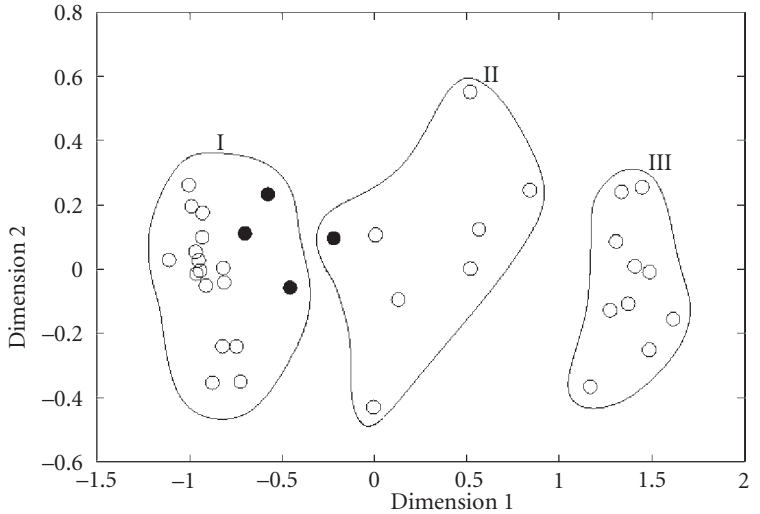


Figure 7.3 Multidimensional scaling (Guttman Lingoes) of the distance matrix for cluster I–III; based on Euclidean distances. Variables: GENDER, VERB, INF, PP_T, PP_EN, TENSE, PRONOUN, OMTE, NIE2 and VIR. Alienation = 0.04, stress = 0.04. Individuals with variable cluster membership across different clustering methods (linkage algorithms) are indicated by a solid circle.

writer is represented by a circle). Following Berruto (1987: 265) who described linguistic varieties as 'points of density in a continuum' ('Verdichtungspunkte in einem Kontinuum'), the three clusters can be seen as indicating the existence of three linguistic codes or varieties.

Dimension one in the scatter plot (Figure 7.3) represents the progression from Dutch to Afrikaans, and individuals located furthest to the left of the plot (cluster I) have largely maintained the Dutch system. The linguistic behaviour of individuals furthest to the right (cluster III), on the other hand, conforms closely to the system of Afrikaans. This interpretation was supported by the results from multiple regression analysis. Briefly, the idea is to use the coordinates of the configuration as independent variables and the linguistic variables as dependent variables. Multiple regression analysis then analyzes the relationship (correlation) between independent variables (the two axes of the multidimensional scaling plot) and each dependent variable (the ten linguistic variables). Correlation coefficients can then be interpreted as relating systematically to the position of data points in the two-dimensional configuration (cf. Kruskal & Wish 1978; Table 7.1).

All ten variables correlate highly and positively with dimension one. Dimension two is characterized by two weak (positive) correlations (TENSE, PP_EN), and one moderate (negative) correlation on the variable OMTE. In other words, individuals with high scores on dimension two show relatively low scores on the variable OMTE and moderately high scores on the variables TENSE and PP_EN.

Table 7.1 Multiple regression (beta correlation coefficient). Correlations are significant at $p < 0.01$.

	dimension 1	**dimension 2**
GENDER	0.942	–0.08
VERB	0.962	0.033
INF	0.987	0.013
PP_T	0.970	–0.01
PP_EN	0.916	0.277
TENSE	0.705	0.295
PRONOUN	0.951	–0.03
OMTE	0.651	–0.42
NIE2	0.842	–0.20
VIR	0.687	–0.19

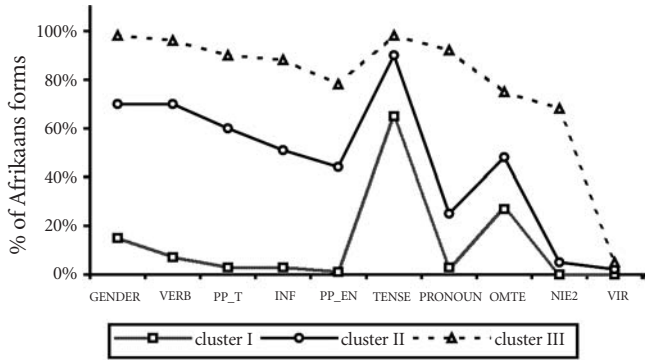


Figure 7.4 Plot of means for the three cluster solutions (GENDER, VERB, INF, PP_T, PP_EN, TENSE, PRONOUN, OMTE, NIE2 and VIR, based on the results of k-means clustering)

The linguistic behaviour of the writers in the three clusters can be compared by plotting the cluster means for each of the ten variables (Figure 7.4). The structure of the plot of means shows what Labov has termed a 'complex pattern of regular relations' (1991 [1972]:124), i.e. writers in the three clusters differ in their linguistic behaviour quantitatively, but the overall pattern is symmetrical. This is most clearly visible for the morphosyntactic variables where all three groups show roughly the following hierarchy of morphosyntactic reduction:

GENDER $\geq$ VERB > PP_T $\geq$ INF > PP_EN

Variation on the variables TENSE and OMTE also increases linearly from cluster I to cluster III. However, the frequency differences between the three clusters are much less pronounced. This suggests that around 1900 change in these two variables had already diffused into the higher lects (which also agrees with the item-based analysis in Chapters 5 and 6). Variation of the morpholexical pronoun variables and the negation, on the other hand, is discontinuous and high frequencies are only found in cluster III. In other words, individuals who with respect to their use of the morphosyntactic variables can be classified as an intermediate group, bridging the gap between cluster I and III, move closer to the writers in cluster I with regard to their behaviour on the morpholexical variables and NIE2. A gradual (rather than discontinuous) pattern of usage across clusters is also characteristic of objective *vir*.

The clusters identified by the clustering algorithm are first and foremost descriptive constructs which do not necessarily have a reality in the minds of

Table 7.2 Cluster means variables as well as the corpus mean. Except for *vir* all means are given as percentages.

	cluster I	cluster II	cluster III	corpus
GENDER	15	70	98	51
VERB	7	70	96	46
PP_T	3	60	90	42
INF	3	51	88	38
PP_EN	1	44	78	29
TENSE	65	90	98	78
PRONOUN	3	25	92	28
OMTE	37	48	75	48
NIE2	0	5	68	16
VIR	0.2	2	6	2

speakers or the world (cf. McEntegart & Le Page 1982:115–116). The visual representation of the cluster formation in Figure 7.3 suggests that cluster III fulfils one of the criteria usually seen as characteristic of natural classifications, i.e. cluster III is well separated in the two-dimensional space. The boundary between cluster I and II, however, is blurred: four individuals show variable cluster membership, and these individuals can be said to have degrees of membership in both clusters. Case instability and the general dispersion of writers within all three clusters suggests that the clusters show properties of fuzzy sets: they describe linguistic groups which are characterized by graded membership and ill-defined boundaries (cf. Zadeh 1965; also McNeill & Freiberger 1993).

Turning to the degree of within-cluster cohesion (another criterion for naturalness), the statistical measure of variance provides a useful index of the degree to which the individual scores within a cluster deviate from the cluster average. A small variance indicates that most of the scores lie close to the average, a large variance represents scores that are widely scattered.

The information presented in Table 7.3 shows that the three clusters can be ordered in terms of their within-cluster variance: variance cluster II > variance cluster III ≥ variance cluster I.

The variance values indicate that the different variables contribute to differing degrees to the variation within the data set as a whole as well as within clusters. Considering the entire corpus, loss of the preterite tense (TENSE), the generalization of the *om... te* infinitive (OMTE) and the double negation (NIE2) contribute least to the variation in the data. For the variables TENSE and OMTE this is due to their relatively broad diffusion; for the variable NIE2, this is a con-

Table 7.3 Variance ($\Sigma(x–m)^2/ n–1$) for the three clusters and the corpus (x = any score in a data set, m = mean). The comparatively low variance index for the variable VIR is due to its different quantification (i.e. per 1000 words, not percentages).

	cluster I	cluster II	cluster III	corpus
GENDER	255	471	6	1530
VERB	200	137	9	1575
PP_T	42	768	57	1694
INF	23	214	53	1397
PP_EN	5	1107	517	1267
TENSE	283	449	8	592
PRONOUN	17	171	57	1254
OMTE	575	413	223	847
NIE2	0	89	523	881
VIR	0.1	5	12	12
Σ	1400.1	3824	1465	11049

sequence of lack of use by the majority of individuals. The three clusters can be described as follows:

i. Cluster I is best described as acrolectal, i.e. the individuals grouped in this cluster conform relatively closely to the norms of metropolitan Dutch. Variation is most pronounced for the infinitive clause and past marking. Variation in the strong past participle, which is substantial in the other two clusters, is marginal. There is some variation for the morphosyntactic variables, which suggests that morphological reduction had begun to affect varieties close to the Dutch norm. While acrolectal varieties generally maintain the distinction between finite and infinite verbs and also inflectional marking of the past participle, they show increasing loss of gender and loss of person and number distinction of the verb.
ii. As is visible in the scatterplot and confirmed by the variance values, the intermediate cluster II is highly variable. Again, the degree of within-cluster variation is not evenly distributed across the different linguistic variables and variation is particularly prominent for past participle regularization. Within-cluster variation of NIE2 is low as only one group member (Hester Vermaak) uses the double negation at all. Individuals grouped together in this cluster are not only intermediate, but, as is visible in the scatter plot, their linguistic usage positions them either closer to the new norms of Afrikaans (cluster III) or to the traditional standard of Dutch (cluster I). The broad dispersion of writers within this cluster, as well as the existence of a fuzzy boundary with cluster I, suggests the possibility of a pattern of

continuous variation, bridging the gap between the poles from Dutch to Afrikaans.

iii. Finally, (early) Afrikaans (cluster III) is a relatively focused variety. Variation is prominent only in the strong past participle which had not yet been completely regularized, and in the structure of the negative clause where single, double and triple negative markings co-exist. The generalization of the *om. . . te* infinitive is also still variable, with some individuals still making frequent use of the *te* infinitive. Cluster III can be described as an emerging standard.

The visual representation in the scatterplot (Figure 7.3) shows interesting differences in the internal structure of the two varieties: cluster I shows a clear cluster core surrounded by peripheral cluster members. Cluster III, on the other hand, does not show a cluster core; however, the data points (although dispersed) are relatively close together (i.e. the cluster is loose but shows high density). One may hypothesize that disintegration of the acrolectal Cape Dutch norm will continue in cluster I, while the emerging standard of cluster III will gradually become more focused (in the sense of Le Page & Tabouret-Keller 1985).

Returning to the notion of co-occurrence as a diagnostic criterion for the identification of linguistic varieties, the correlation matrix given in Table 7.4 shows that most of the ten variables are highly and positively correlated, i.e. high (low) values on the variable VERB go with high (low) values on the variables INF, PP_T, and so forth. More moderate correlations exist for the variables TENSE, OMTE as well as VIR and NIE2.

Table 7.4 Correlation matrix for the ten variables. Correlations are significant at $p < 0.05$ (Pearson r).

	GENDER	VERB	INF	PP_T	PP_EN	TENSE	PRONOUN	OMTE	NIE2	VIR
GENDER	1									
VERB	0.94	1								
INF	0.92	0.95	1							
PP_T	0.92	0.95	0.96	1						
PP_EN	0.85	0.85	0.91	0.87	1					
TENSE	0.55	0.73	0.70	0.69	0.65	1				
PRONOUN	0.86	0.87	0.94	0.88	0.86	0.64	1			
OMTE	0.60	0.56	0.63	0.62	0.54	0.40	0.63	1		
NIE2	0.72	0.73	0.81	0.76	0.71	0.56	0.90	0.56	1	
VIR	0.65	0.67	0.68	0.69	0.51	0.53	0.69	0.46	0.60	1

The existence of strong correlations between variables was confirmed by the results from principal components analysis. Based on the interpretation of the scree plot (Figure 7.5), two factors with eigenvalues of 7.7 and 0.6 were maintained. Together the two factors account for 80% of variance in the data.

The first factor which shows high positive loadings on all ten variables, accounts for over three-quarters of data variance. The second factor is weakly characterized by a relatively high, negative loading on OMTE and a moderate, positive loading on TENSE. Thus, the two variables which are most widely diffused across the sample, are again shown to occupy a special position. The weak loadings on factor II as well as the structure of the scree plot (Figure 7.5) suggest that variation in the *Corpus of Cape Dutch Correspondence* is largely unidimensional: from cluster I to cluster III there is a linear increase in the frequency of Afrikaans forms on all ten variables.

The item-based analysis presented in Chapters 5 and 6 suggested that the diffusion of the morpholexical variables was less comprehensive and continuous than that of the morphosyntactic variables. Different distribution patterns were also indicated by discontinuities in the plot of cluster means (Figure 7.4). To investigate these differences in more detail, cluster analysis and multidimensional scaling were performed on sub-sets of the data: (i) including only the morphosyntactic variables, and (ii) including only the morpholexical variables. Considering morphosyntactic and morpholexical variables separately is of interest in the context of on-going discussions about modularity in linguistics, i.e.

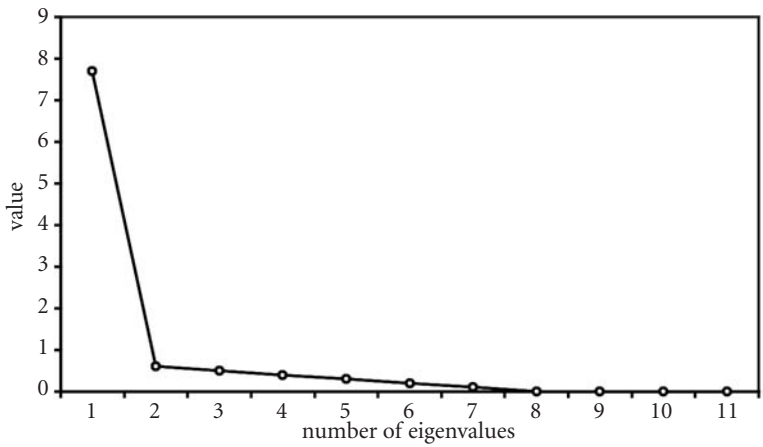


Figure 7.5 Scree plot

Table 7.5 Principal components analysis, factor loadings for factor 1 and factor 2

	factor 1	**factor 2**
GENDER	0.93	–0.05
VERB	0.96	0.13
INF	0.98	0.03
PP_T	0.97	0.04
PP_EN	0.90	0.08
TENSE	0.74	0.45
PRONOUN	0.95	–0.06
OMTE	0.68	–0.62
NIE2	0.85	–0.11
VIR	0.73	0.04
% of variance accounted for	76.5%	6.4%
eigenvalue	7.7	0.6

the idea that parts of grammar from sub-systems and that different sub-systems might behave differently in language change (cf. Ferguson 1995: 191–192).

Morphosyntactic variation

Considering the verbal system (i.e. variables VERB, INF, PP_T, PP_EN) separately, a three cluster solution was again indicated by the graph of amalgamation schedule (Ward's method, and average linkage; distance measure: Euclidean; Figure 7.6). However, the structure of the solution differs in important ways from the results discussed above. Firstly, the intermediate cluster II shifts to the right, towards the emerging standard of Afrikaans (cluster III) and there is now some overlap between clusters II and III (cf. the variable cluster membership of four writers: Engela Furter, Eddie Smith, Elizabeth Steyn and Johanna Brümmer). The gap between cluster I and II, on the other hand, is slightly more pronounced, and only one individual (Andries Hauptfleisch) shows variable cluster membership. The means for the three clusters again indicate a symmetric hierarchy, i.e. VERB > PP_T≥ INF > PP_EN.

The general fuzziness of the solution suggests that in this case the algorithm imposed a cluster structure on a situation which is probably more accurately described as a continuum – that is, a finely graded 'spectrum of usage' (DeCamp 1971b: 36) which is 'not amenable to either mono- or bisystemic treatment' (Bickerton 1973: 641). DeCamp (1971a, 1971b) has pointed to the implicational relations which often characterize linguistic continua, and following DeCamp's analysis of seven Jamaican speech samples, Guttman scalogram analysis is today commonly used for the description of implicational

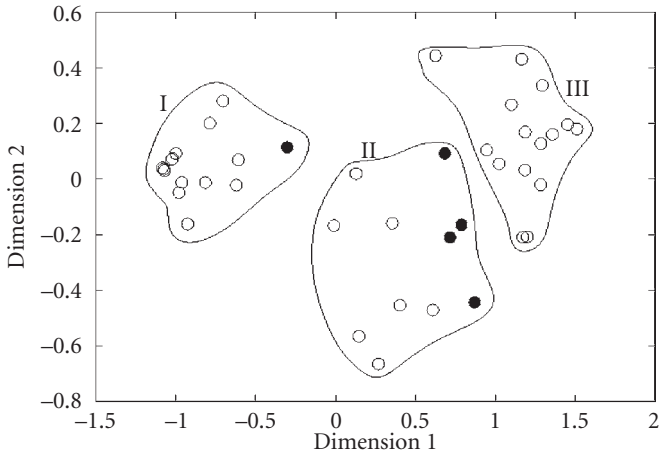


Figure 7.6 Multidimensional scaling (Guttman Lingoes) of the distances matrix for cluster I–III; based on Euclidean distances. Variables: VERB, INF, PP_T, PP_EN. Alienation 0.03, stress 0.03. Individuals with variable cluster membership across different clustering methods (linkage algorithms) are indicated by a solid circle.

continua (for the methodology see Guttman 1944 and Togerson 1958: 307–331). Early implicational scales were mainly based on binary data (i.e. a certain linguistic feature was classified as being either absent or present). However, as linguistic data usually do not come in such an all-or-nothing way, Bickerton (1973: 642; cf. also Fasold 1970: 553–554) suggested the use of three-place scales which allow for the incorporation of variable structures. A perfect continuum using a three-place-scale is illustrated in Table 7.6. The implicational relations which exist in the continuum can be summarized as follows: the presence of property D in the speech of an individual means that the speaker will also use properties A to C; if D is absent but C is present then properties A and B will also be present, and so forth. In other words, D implies C implies B implies A. Perfect continua are, however, not to be expected in empirical research, and DeCamp (1971b: 35) therefore likened the continuum model as a theoretical construct to the perfect shapes which underlie the study of geometry:

> You do not obtain a square by carefully measuring thousands of floor tiles and then averaging the measurements, for a square has four equal sides by *a priori* definition, not by empirical measurement. However, geometry is useful to us precisely because there are shapes in our real world which, though irregular, are similar enough to the ideal geometrical shapes that they can be described by surveyors and navigators as if they really were circles, squares, and triangles.

Table 7.6 Three-place implicational scale. X = categorical usage, V = variable usage, O = no usage (adapted from Romaine 1982: 171).

	ENVIRONMENT				
	A	**B**	**C**	**D**	
lect 1	X	X	X	X	system I
lect 2	X	X	X	V	
lect 3	X	X	V	V	
lect 4	X	V	V	V	
lect 5	V	V	V	V	
lect 6	V	V	V	O	
lect 7	V	V	O	O	
lect 8	V	O	O	O	
lect 9	O	O	O	O	system II

The systematic ordering of idiolects portrayed in implicational scales has been interpreted as a 'synchronic [reflection] of diachronic processes, depicting the acquisition or diffusion of features as waves spreading across social (speaker-to-speaker) and linguistic (environment-to-environment) space' (Rickford 1987b: 155; cf. also Bailey 1973: 67–109). In other words, implicational scales are believed to illustrate the path of diachronic change from one system to another, and to preserve 'diachronic changes in a synchronic state' (Bickerton 1975: 17). The data for the variables VERB, INF, PP_T and PP_EN was scaled using a three-place scale with the following thresholds:

O 0–25%, i.e. largely corresponding to the norms of metropolitan Dutch
V 26–74%, i.e. variable usage
X 75–100%, i.e. largely corresponding to the emerging norms of Afrikaans

Following DeCamp (1971b: 34) these three categories are (at least in some sense) meant 'to model the idealized competence of the persons involved in those speech acts'. It is of course questionable whether the setting of such ultimately arbitrary thresholds is theoretically acceptable (cf. Romaine 1982: 178, fn. 19); however, relying on similarly arbitrary thresholds Romaine (ibid.: 170–173) used implicational scaling successfully to illustrate the syntactic diffusion of relative deletion in Scots English.[1]

Table 7.7 shows the scaling of the individual writers depicted in Figure 7.6. A key concept in scalogram analysis is reproducibility and in a perfect scale the linguistic behaviour of any individual can be reproduced from the knowledge of its rank position alone. However, as noted above, perfect scales are not to be expected in real-life data and a coefficient of reproducibility (CR) is calculated

Table 7.7 Implicational analysis including the variables VERB, INF, PP_T, PP_EN. Deviations indicated by *, CR = 0.93. Empty cells were included as deviations (cf. Romaine 1982: 176).

	VERB	PP_T	INF	PP_EN	
N = 8	X	X	X	X	Afrikaans
N = 1	X	X	V*	X	
N = 1	V*	X	X	X	
N = 6	X	X	X	V	
N = 3	X	X	V	V	
N = 0	X	V	V	V	
N = 3	X	V	V	O*	
N = 1	V	V	V	X*	
N = 1	V	V	V	V	
N = 2	V	V	V	O	
N = 1	V	O*	V	O	
N = 0	V	V	O	O	
N = 4	V	O	O	O	
N = 19	O	O	O	O	Dutch

to measure the degree of deviation from a perfect scale (CR = 1 – Σ errors/ Σ possible errors). Scales with a CR > 0.85 can be considered as approximations to a perfect scale, and a CR ≥ 0.93 approaches the 0.05 level of significance (Guttman 1944: 140; Dunn-Rakin 1983: 105–107).

The implicational analysis shows the same hierarchical ordering of the four verbal variables as was indicated in the plot of cluster means (cf. Figure 7.4): individuals show highest frequencies of regularization in the verbal system, followed by the weak participle and the infinitive, and finally the strong participle. In other words, the hierarchical ordering VERB > PP_T > INF >PP_EN is not an artefact of group analysis but is reflected in the individual lects of writers. The implicational relations in the data show that the intermediate lects grouped in cluster II are not the result of unstructured dialect mixture and *ad hoc* borrowings from metropolitan Dutch: linguistic variation at the Cape was structured and systematic. The scale supports an interpretation of morphosyntactic variation at the Cape as a finely graded linguistic continuum.

The implicational relations hold when including loss of gender as a fifth morphosyntactic variable (Table 7.8). The CR is, however, lower when counting empty cells as deviations (CR = 0.86). However, since the scale is based on fewer individuals (N = 40), it is possible that the empty cells would be filled if more data were available.

Table 7.8 Implicational analysis including the variables GENDER, VERB, INF, PP_T AND PP_EN. Deviations indicated by *, CR (empty cells counted as deviant) = 0.86, CR (empty cells excluded) = 0.95.

	GENDER	VERB	PP_T	INF	PP_EN	
N = 7	X	X	X	X	X	Afrikaans
N = 4	X	X	X	X	V	
N = 1	X	X	X	V	V	
N = 2	V*	X	X	V	V	
N = 1	V*	X	X	V	X*	
N = 0	X	X	V	V	V	
N = 1	X	V	V	V	X*	
N = 0	X	V	V	V	V	
N = 2	X	V	V	V	O*	
N = 1	V	V	V	V	V	
N = 1	V	X*	V	V	O	
N = 0	V	V	V	V	O	
N = 1	V	V	O*	V	O	
N = 0	V	V	V	O	O	
N = 2	V	V	O	O	O	
N = 3	V	O	O	O	O	
N = 1	O	V*	O	O	O	
N = 14	O	O	O	O	O	Dutch

Morpholexical variation

Multivariate analysis was performed including only the pronoun variables *ons, dit, my/sy* (possessive) and *hulle* (subject and object combined). The results differ in important ways from those discussed for the morphosyntactic variables. The analysis indicates a clear two cluster solution (distance measure Euclidean, linkage rules: Ward's method, complete and average linkage). Cluster membership was entirely stable across the various linkage rules.[2]

Unlike in the previous analysis, no intermediate variety (cluster II) is indicated and the distribution is unambiguously bi-systemic: a large gap exists between two well-defined clusters. Dimension 1 can easily be identified as leading from Dutch (cluster I) to the new standard of Afrikaans (cluster III). Dimension 2 is necessary only for the placement of F. S. Malan (co-ordinates: 0.6 and 1.3), who conforms to the new standard of Afrikaans with one exception: he still makes regular use of possessives ending in a nasal (*myn/syn*). Individuals who were previously classified as belonging to cluster II are now grouped as (peripheral) members of cluster I. In other words, membership of cluster III remained stable, and cluster II collapsed with cluster I. Moreover,

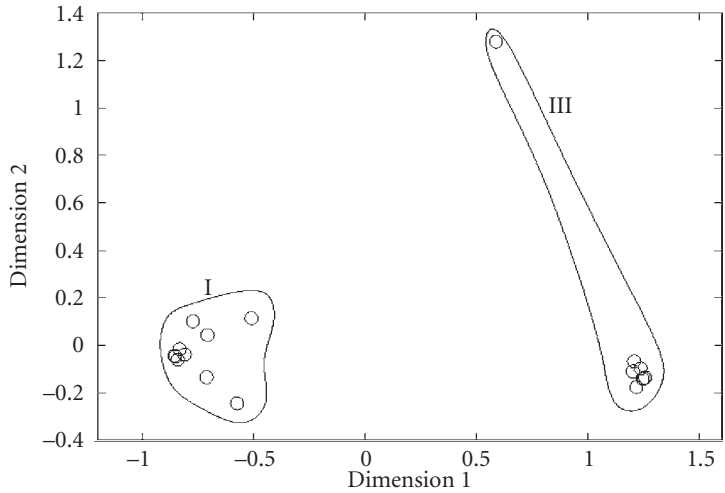


Figure 7.7 Multidimensional scaling (Guttman-Lingoes) of the distance matrix for cluster I and III; Euclidean distances. Variables: *ons*, *dit, my/sy, hulle*. Alienation = 0.01, stress 0.008.

the new standard of Afrikaans now shows a clearly defined cluster core and within-group variation is (with the exception of F. S. Malan) minimal.

Summary of results of the multivariate analysis

The overall results of the multivariate analysis can be summarized as follows:

i. All ten variables are highly and positively correlated, and the results from PCA suggest that over two-thirds of the variation in the *Corpus of Cape Dutch Correspondence* can be summarized by a single, linear factor which can be described as [±Afrikaans].
ii. Morphological regularization and reduction (verbal and nominal) is continuous, implicational and gradually increasing from cluster I to III.
iii. Variation on the variables TENSE and OMTE also increases linearly from cluster I to III; however, frequency differences between the three clusters are much less pronounced than on the other variables, indicating that by the turn of the century change on these two variables had affected most varieties within the Cape Dutch speech community. This interpretation is supported by the item-based analysis presented in Chapters 6 and 7.
iv. Variation of the morpholexical variables (pronoun system) is bi-systemic. This reflects the existence of two distinct and well-defined norm systems

in the spoken language. Most of the writers were still oriented towards the Dutch norm and used Afrikaans variants sparingly.

The distribution of the three clusters in the two-dimensional space is evocative of the concepts of acrolect, mesolect and basilect. However, it will be shown in the next section that the social composition of cluster III is at odds the social characteristics typically reported for basilectal varieties.

Examining the social dimensions

So far the discussion has focused on the patterns of linguistic variation and has not considered the social location of the writers. While general correlation analysis of the data (using multiple regression) yielded inconclusive results, an interesting social pattern was found to structure the three cluster solution described above (see Table 7.9).

Unfortunately the actual numbers in Table 7.9 are marginal. Since the reliability of the χ^2 test depends on sufficiently large expected frequencies (generally 5 or more), testing for statistical significance is problematic. Using the χ^2-test under conditions of low expected frequencies increases the likelihood of what is known as a type 1 error (α error), i.e. the erroneous rejection of the null-hypothesis (that is, the assumption that the distribution in the sample does not exist in the population from which the sample is drawn; Woods, Fletcher, & Hughes 1986: 145). At the same time type 2 errors (β error; erroneous rejection of the alternative hypothesis) are generally likely when the sample size is very small (ibid.: 127, 130). The results were thus not tested for statistical significance. Similarly, the indicated percentage values should only be read as data summaries which allow comparison across clusters; they cannot not be interpreted as being accurate to within one part in a hundred.

Table 7.9 shows that socially the subset of individuals included in the multivariate analysis is more biased than the overall corpus: it only includes speakers belonging to classes 2 (the emerging professional class) and 3 (the traditional petty bourgeoisie). The distribution is also skewed with regard to age; that is, only one-third belonged to age group 1 (born before 1865), two-thirds belonged to age group 2 (born after 1865).

Cluster III (the emerging standard of Afrikaans) is characterized by a dominance of age group 2 and the new professional class. Note that the two class 3 individuals included in cluster III are both young (age group 2), non-rural and relatively well educated. Both were classified as class 3b, i.e. they were indepen-

Table 7.9 Social characteristics of the three cluster solution

		sex		age group		social class	
		F	**M**	**1**	**2**	**2**	**3**
all	N	19	18	13	24	19	18
	%	51%	49%	35%	65%	51%	49%
cluster I	N	9	10	7	12	11	8
	%	47%	53%	38%	62%	58%	42%
cluster II	N	5	3	5	3	0	8
	%	63%	37%	63%	37%	0%	100%
cluster III	N	5	5	1	9	8	2
	%	50%	50%	10%	90%	80%	20%

dent small-scale entrepreneurs with at least moderate education (i.e. several years of schooling).

The intermediate cluster II is entirely composed of members of class 3 (almost exclusively class 3a, i.e. rural petty bourgeoisie). The cluster also contains a relatively high percentage of older speakers (age group 1) – especially when compared to their under-representation in the sub-sample as a whole. With regard to sex, cluster II is characterized by a slight predominance of women, and considering that this in-between cluster shows both progressive and conservative features, the distribution agrees with Raidt's (1995:138) assessment of the role of women in linguistic change at the Cape: 'women in the Cape Colony played a double role in linguistic change – both progressive and conservative'. Finally, cluster I (the old Dutch standard) does not show any particular bias, but reflects roughly the social distribution of the sub-sample.

The social patterning across clusters I to III holds when including only subsets of the linguistic variables, such as the verbal variables (VERB, INF, PP_T and PP_EN) discussed in the previous section. However, additional tendencies are noticeable in the social distribution: the morphosyntactic variables which show continuity across clusters are also less categorically structured on the social dimensions (Table 7.10). The intermediate cluster II is again characterized by a dominance of age group 1 and class 3. The higher percentage of writers belonging to class 3 and age group 1 in cluster III is a result of excluding the more dichotomously patterned morpholexical pronoun variables and *nie*-2. These variables can thus be interpreted as markers of the language use of younger and educated writers.

Although the social correlates of variation across the three clusters are not categorical and depend to some extent on the linguistic variables involved, the overall tendency is stable. That is, while the linguistic patterns across the clus-

Table 7.10 Social characteristics of the three cluster solution for morphosyntactic variables. Based on four linguistic variables: VERB, INF, PP_T and PP_EN.

		sex		age group		social class	
		F	**M**	**1**	**2**	**2**	**3**
all	N	25	25	19	31	25	25
	%	50%	50%	38%	62%	50%	50%
cluster I	N	11	13	8	16	14	10
	%	46%	54%	33%	67%	58%	42%
cluster II	N	5	6	6	5	2	9
	%	45%	55%	55%	45%	18%	82%
cluster III	N	9	6	5	10	9	6
	%	60%	40%	33%	66%	60%	40%

ters are linear (i.e. gradual increase in the frequency of Afrikaans forms, see Figure 7.8), the social patterns are curvilinear with respect to age and class as well as sex. The different patterns of relationships are summarized in Figures 7.9 to 7.11.

The three clusters can now be described sociolinguistically:

> Cluster I
> Acrolectal Dutch; relatively low level of linguistic variation; both age group 1 and 2, as well as class 2 and 3 [–Afrikaans, –variation, +age group 1/2, +class 2/3]
>
> Cluster II
> Mesolectal Cape Dutch Vernacular; no clear systemic identity; high level of within-cluster variation; dominated by age group 1 and class 3 [± Afrikaans, +variation, +age group 1, –age group 2, –class 2, +class 3, +female]
>
> Cluster III
> 'Afrikaans'; relatively low level of linguistic variation; dominated by age group 2 and class 3 [+Afrikaans, –variation, –age group 1, +age group 2; +class 2, –class 3]

The terminology of acro-, meso- and basilect has proved useful for the description of linguistic variation in extraterritorial settings. However, while most linguists would comfortably identify cluster I as acrolectal (i.e. more or less corresponding to metropolitan norms; probably best described as an extraterritorial dialect of Dutch), and cluster II as mesolectal, labelling cluster III basilectal does not seem appropriate since the concept of a basilect not only refers to the linguistic entity furthest removed from the acrolect (in terms of

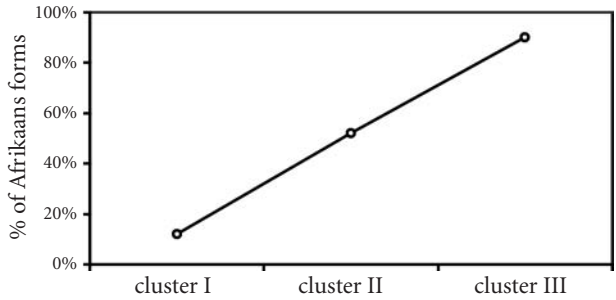


Figure 7.8 Distribution of Afrikaans forms across the three clusters

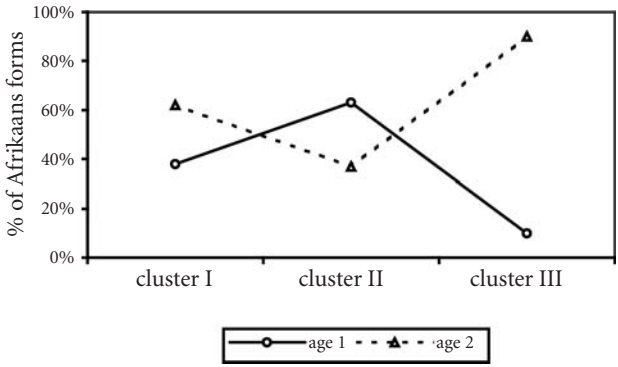


Figure 7.9 Distribution of social groups across clusters: age

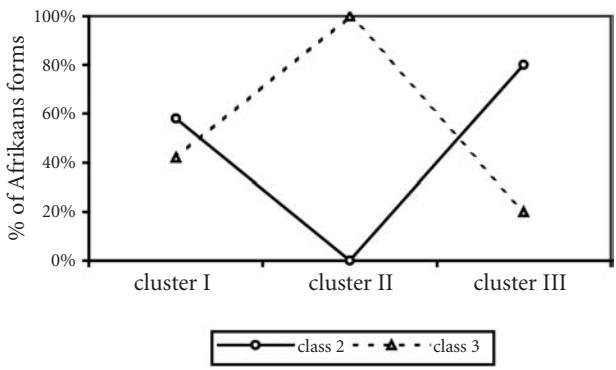


Figure 7.10 Distribution of social groups across clusters: class

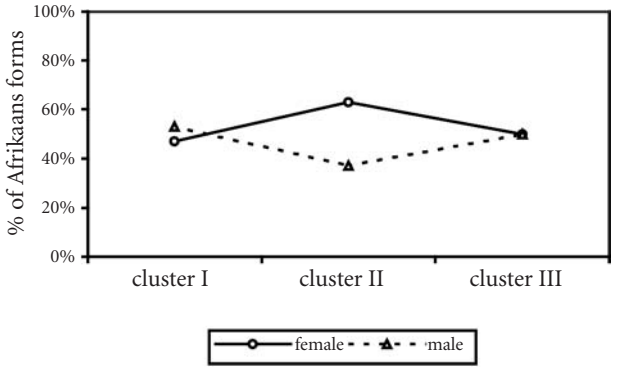


Figure 7.11 Distribution of social groups across clusters: sex

linguistic structure), but also signals its social location and identity as a stigmatized low-prestige variety. Considering the social composition of cluster III, especially the dominance of the 'new' (highly educated) professional class, it appears more appropriate to label this cluster 'elevated basilect' or 'emerging non-conservative standard'. Thus, although situated at the opposing pole in terms of its linguistic identity, cluster III exhibits social characteristics typical for acrolectal varieties: that is, it is a relatively uniform variety and is used by educated writers.

Linguistic patterns in the dialect writing tradition

It was suggested in Chapter 2 that dialect writing contributed to the standardization of Afrikaans by providing prospective writers with a 'typological' conception of what constituted 'the vernacular'. To further investigate the position of the vernacular literature in the spectrum of linguistic variation at the Cape, a sample of nine dialect texts was included in the multivariate analysis (Figure 7.12; the arrangement of writers in the scatter plot differs from Figure 7.3 as a result of the inclusion of additional cases). The sample covers the time period 1828 to 1909.

1. Letters by *Mietje* (author: S. De Lima, *De Kaapsche Courant of Verzamelaar* 1/10/1828, 8/10/1828; reprinted in Nienaber 1971: 25–30).
2. Dialogues with Hendrik Kok (author: C. E. Boniface, *Zuid-Afrikaan* 6/8/1830, 13/8/1830; reprinted in Nienaber 1971: 36–44).

3. *Zamenspraak tusschen Klaas Waarzegger en Jan Twyfelaar, No 1* (author: L. Meurant 1861).
4. *Benigna van Groenekloof of Mambre* (anonymous, 1873).[3]
5. *Boerenbrieven* by Samuel Zwaartman (author: H. W. A. Cooper, *Het Volksblad* 12/1/1870, 20/1/1870; reprinted in Nienaber 1942: 5–9).
6. *Geskiedenis van Josef ,* Chapters I–II (author: C. P. Hoogenhout, 1873).
7. *Manifes van die Genootskap van Regte Afrikaners* (*Afrikaanse Patriot* 15/1/1876).
8. *Voor Volk en Taal*, scenes 5/6 (author: M. Brink, 1905).
9. *Straatpraatjes* (author: A. Abdurahman, *APO Newspaper* 5/6/1909, 19/6/1909: reprinted in Adhikari 1996: 26–34).

Except for De Lima and Brink, all dialect texts are grouped with the emerging standard of Afrikaans (cluster III). Nienaber (1971: 24) has described De Lima's *Mietje*-letters as an example of what he termed *Hoog-Afrikaans* ('High Afrikaans'); a local variety which was meant to represent the language used by the upper classes in Cape Town. From the perspective of the metropolitan Dutch standard, De Lima's language use is marked by various non-standard forms: gender marking of the noun is absent, loss of verbal inflection is com-

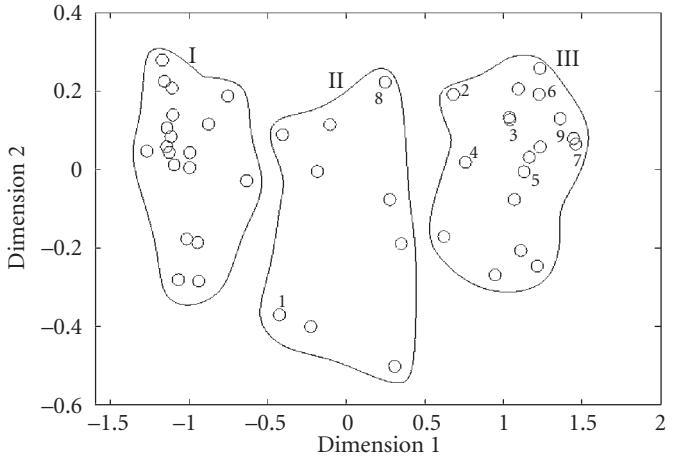


Figure 7.12 Multidimensional scaling (Guttman Lingoes) of the distance matrix for cluster I–III; data basis: *Corpus of Cape Dutch Correspondence* and nine vernacular texts. Variables: GENDER, VERB, INF, PP_T, PP_EN, TENSE, NIE2 and VIR. Vernacular texts are indicated as follows: De Lima 1, Boniface 2, Meurant 3, Benigna 4, Zwaartman 5, Josef 6, Manifes 7, Brink 8, Straatpraatjes 9.

plete for modal verbs, but variable for auxiliaries, main verbs and weak past participles; infinitives and strong past participles are, however, inflected. With respect to the more dichotomously patterned pronoun variables De Lima still follows the Dutch system. He uses only *wy* in the first person plural, *zulle* and *zy* for the third person plural, and although he uses *dit* as a third person singular pronoun and *wat* as a relative (with non-animate antecedent only), examples of this are rare. No instances of *nie*-2 exist in the *Mietje*-letters.

Brink, who explicitly aimed at writing what he called Cape Dutch (see Chapter 2), is further advanced towards Afrikaans than De Lima. His language use conforms to the Afrikaans system for auxiliary and modal verbs; however, he frequently marks inflection on main verbs (roughly one-third of main verbs are inflected). Brink also shows a tendency to maintain inflection in the past participle. He has generalized the use of *ons* (instead of *wy*) and *dit* (instead of *het*), but shows variation for the third person plural pronoun (*hulle* ~ *zulle)*. Examples of *nie*-2 occur but are infrequent (22% of negative sentences show *nie*-2). Of the dialect texts grouped in cluster III, the anonymously published novel Benigna (4) and Boniface's Hendrik Kok dialogues (2) appear to be more peripheral: they show a still variable pronoun system, variation in the use of *nie*-2 and incomplete regularization in the past participle.

The incorporation of the sample of dialect texts into the analysis gives empirical support to some hypotheses formulated in earlier chapters:

i. There were clear typological continuities between the dialect literature and the early Afrikaans standard norm (see Chapter 2). It is likely that dialect writing contributed to the standardization of Afrikaans by offering prospective writers, from the 1860s onwards, a clear model of how to write 'Afrikaans'.
ii. The linguistic characteristics of the texts belonging to the dialect writing 'tradition' and the private letters of young and educated writers can be described with reference to what Britto (1986:232) has called a 'Pseudo-H'. That is, they constitute a relatively stable linguistic variety typically found in eye dialects where it is clearly identifiable as the Low variety (thus commanding local prestige), but not stigmatized since it is used by the most educated writers. It is essentially a socially acceptable, albeit imperfect and stylized representation of the Low variety which shows a high degree of uniformity and conventionality, and thus conforms to the structural requirements of a standard language.
iii. The concept of simplified Cape Dutch (*eenvoudig Kaaps Hollands*) – which was advocated by writers and academics such as Brink, Viljoen, Mansvelt

and D'Arbez as an alternative to the Afrikaans standard propagated by the *Genootskap van Regte Afrikaners* (see Chapter 2) – was not a language-political fiction, but was based on a linguistic reality: that is, the variable structures of the acrolectal and mesolectal varieties identified in the *Corpus of Cape Dutch Correspondence*. To characterize such variable linguistic usage simply as a random and ungrammatical mixture of Afrikaans and Dutch under diglossic conditions fails to recognize the structured nature of these intermediate forms.[4]

iv. Cognitive psychologists investigating the principles underlying the formation of categories have shown that human categorizing usually proceeds from a prototype or core which represents the clearest case(s) of the category. The category-core is surrounded by non-prototype members which show varying degrees of similarity to the prototype. Rosch and Mervis (1975) have argued that proximity scaling is an important tool for the study of categories: items that are located closely to one another can be interpreted as conceptual prototypes, with prototypicality diminishing the further one moves from the core (cf. also Kempton 1978:53 who argued that sharply bound categories appear as tight clusters in a graph). Considering the multidimensional scaling plot given in Figure 7.3, it is evident that clusters I and III, which have been described as relatively focused, show substantial item proximity. This probably facilitated their categorization in folk taxonomies (see Chapter 2). Cluster II, on the other hand, is loose and its boundaries are 'fuzzy' (i.e. show some overlap with cluster I and also III, cf. Figures 7.3 and 7.6). It is thus not easily amenable to categorizing. The diffuse and ambiguous nature of many forms of Cape Dutch Vernacular prevented its inclusion into well-defined, prototypical folk taxonomies, and supported its the pejorative evaluation as a *tussending* (in-between-thing).[5]

Standardization and diglossia revisited

The low level of within-cluster variance for the emerging standard of Afrikaans (cluster III) indicates that by 1900 (i.e. before the onset of academy-organized language planning/standardization) Afrikaans had developed into a relatively focused written variety. The social composition of cluster III suggests that the agents of this focusing process were predominantly found among the emerging professional class, who also played a central role in early Afrikaner nationalism and contributed to the vernacular writing tradition from the 1850s. Alongside

this new code developed a puristic language ideology which, as argued in Chapter 2, identified linguistic influence from Dutch as artificial and non-authentic. However, only those who had ample exposure to the Dutch standard (usually via the education system) were in a position to recognize, and thus avoid or enhance, Dutch features in their linguistic usage (as they had a clear typological and grammatical conception of the norms of metropolitan Dutch). That it was the élite who first showed fluent usage of a standardized Low variety free from the influence of the High variety, has also been noted for Greece by Frangoudaki (1992: 370):

> definitely since the 1950s, fluent usage of an elaborate D [Dhimotiki, L variety] free of K [Katharevousa, H variety] mixtures placed speakers in the highest social ranking, identifying them as members of an intellectual elite.

Kazazis (1992) noted that hyperdemoticism was most dominant in written styles, while the spoken language of the Greek élite contained considerably fewer demotic forms (on hypervulgarisms, i.e. the excessive use of Low forms, as typical for non-casual styles cf. also Janda & Auger 1992). Contemporary linguistic commentary and the increasing number of excuses for 'incorrect' Afrikaans in private letters from about 1910 suggest that the production of pure (i.e. non-Dutchified) Afrikaans was the result of conscious efforts by the writers (cf., for example, the documents quoted in Truter 1997: 96–97; in the late nineteenth century similar apologies were habitually made for writing 'incorrect' Dutch). Malherbe's (1917: 44) observations show a clear rejection of any influence from Dutch and a strong awareness of the norms of the new standard (my emphasis):

> Ook van die kant van Nederlands dreig ons 'n gevaar. Ons weet hoe moelik dit vir 'n openbare spreker was wat gewoon geraak het aan *sijn soort Nederlands* om 'n Afrikaanse anspraak te lewer. *Hy verval onbewus in Nederlandse vorm.* Dit geld ook van schrijwe … Eers in die laaste jare *na veel studie en bewuste strewe*, kan 'n mens 'n groot verbetering opmerk. Dit gaan hand aan hand met 'n toenemende hoogskatting van ons taal wat op sijn beurt met jaloerse nouwgesetheid 'n so suiwer moontlike gebruik eis. Die "Brandwag"-artiekels vergelijk vandag pragtig met dié van vroeër jaargange. Tog word nog so gedurig foute gemaak.
>
> ['There is also danger from the Dutch side. We know how difficult it was for a public speaker who had got used to *his kind of Dutch* to deliver an Afrikaans speech. *He falls unknowingly into the Dutch form*. This is also true for writing … Only in recent years *after much study and conscious endeavour*, can one notice a great improvement. This goes hand in hand with an increasing es-

> teem for our language which in turn demands with jealous conscientiousness a use as pure as possible. The *Brandwag* articles compare today splendidly with those from earlier years. However, mistakes are still being made.']

Several writers in the corpus were found to make alternate use of both acrolectal Afrikaans and acrolectal Dutch in their writing. Following DeBose (1992: 161), a text or section of a text was considered to be in Afrikaans or Dutch 'to the extent that it is grammatically compatible with an autonomous description' of one or the other. The individual letters of these writers could be quite unambiguously classified as representing one or the other variety, that is, either the letters showed verbal inflection and nominal gender in accordance with the norms of Dutch, Dutch pronoun forms and the syntax of Dutch negation; or they showed complete loss of nominal gender and verbal inflection, use of Afrikaans pronoun variants and the brace negation. Variety choice was usually determined by interlocutor (typically: Afrikaans to friends and siblings, Dutch to parents) and also varied over time (see Chapters 8 and 9 for a more detailed discussion of individual writers).

It appears that these writers maintained a relatively clear-cut code differentiation through the use of strict and well-defined co-occurrence restrictions. Non-diglossic situations (such as the formal/informal distinction in English), on the other hand, show more continuous variation patterns and variants are not ordered into clearly defined H/L doublets (Ferguson 1959: 334; Hudson 2002: 16). In Figure 7.13 the relative position of the different letter-corpora is indicated for two bi-dialectal writers: Francois Malan (A) and Wynanda Hoogenhout (B): [1] Dutch and [2] Afrikaans.

The clear-cut distinction these writers maintain between the two codes suggests something one could describe as 'diglossic competence'. When Karl Krumbacher used the term *diglossie* in his book *Das problem der modernen griechischen schriftsprache* ('The problem of the modern Greek written language'; 1902) he supplemented it with many descriptive synonyms, such as *doppelgesicht* ('double-face'), *doppelkopfigkeit* ('double-headedness') and *doppelsprachigkeit* ('double-lingualism'), which express clearly the individual and psychological correlates of diglossia, i.e. the co-existence of two grammars within the individual (or alternatively the existence of one grammar and a set of transformation rules; on the possibility of transformation rules in diglossia cf. Wexler 1971).

The presence of such diglossic or bi-dialectal writers in the corpus again raises the question of diglossia as a sociolinguistic model for the description of the Cape speech community. While the functional and acquisitional as-

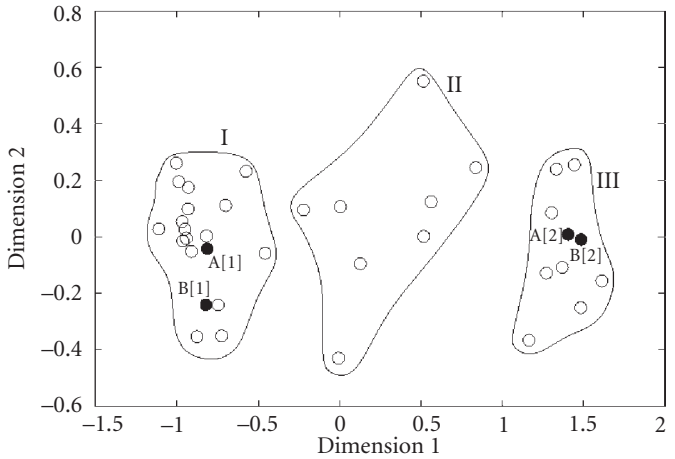


Figure 7.13 Multidimensional scaling (see Figure 7.3). Position of the bi-dialectal speakers F. S. Malan (A[1], A[2]) and Wynanda Hoogenhout (B[1], B[2]).

pects of diglossia (i.e. domain complementarity and acquisition of the High variety as a second language) have featured prominently in the study of diglossic societies and are seen as principal determinants of diglossia (e.g. Hudson 2002), Ferguson's original paper also contains several remarks on the structural relationships between codes in diglossia. A central criterion of diglossia is structural discontinuity between the Low and High varieties. Although Ferguson has emphasized the bi-systemic nature of diglossia, he acknowledges the possibility that additional varieties might develop (1959: 332). However, these varieties are typically situated within the fundamental linguistic duality which is the defining characteristic of diglossia (cf. Ferguson 1991: 226: 'in the diglossia case the analyst finds two poles in terms of which the intermediate varieties can be described; there is no third pole'; also Britto 1986: 15ff.). Similarly Fishman (1980: 4) remarks that the structural gap between the different systems or linguistic codes needs to be such 'that without schooling the written/formal-spoken cannot be understood (otherwise every dialect/standard situation in the world would qualify within this rubric)'; and Wexler (1971: 336) notes that 'diglossia is not meant to refer to any condition of multiple norms, but specifically to that condition where there is a broad structural gap' correlating with patterns of functional complementarity (cf. also Keller 1982: 73; Winford 1985: 349; Hudson 2002: 11–12 for a critical evaluation of the linguistic aspects of diglossia).

However, at closer inspection even in the defining cases of diglossia continuity between the systems appeares to be norm rather than the exception. With regard to Greek diglossia, Pappageotes and Macris (1964: 58) make direct reference to the possibility of a continuum situation when they state: 'The distinction between the Demotic and the Puristic is not absolute except at the two ends of the scale'. And Daltos (1980: 84; also Alexiou 1982) emphasizes that while High and Low varieties exist as social constructs in the mind of the speakers, the analysis of actual language data shows High and Low to be idealized codes which are imposed on a linguistic continuum in which stylistic and social constraints interact (cf. also Note 5).[6] Structural discreteness was not taken to be a requirement of diglossia by De Silva (1974) when describing Sinhala diglossia as a continuum, and more recently Schiffman (1993/1997) defined Tamil diglossia as an example of what he calls 'heterogeneous diglossia':

> Between the styles [High and Low] there are shadings from one style into another, i.e. it is possible to write modern literary Tamil with an archaic lexicon, but with nonarchaic grammar; it is also possible to swing in the other direction and to make modern literary Tamil more spoken, or to make educated colloquial more literary in flavour, or more nonstandardly colloquial. In any event, though linguistic cultures think of diglossia as either-or, it is often a gradient cline, with one variant shading into another. (Schiffman 1997: 210)

In other words, while speakers (and linguistic cultures, i.e. speech communities) perceive and describe linguistic variation in bi-systemic terms of High and Low, the situation is different 'in the field': the linguistic systems employed by different speakers can often only be loosely grouped, and individual lects overlap and show a finely graded continuum of usage. The question, however, remains whether such scenarios of linguistic continuity can and should still be referred to as representing cases of diglossia.

It was noted by Paolillo (1997) that the description of linguistic variability in diglossic (and other) language situations has often been impressionistic, based either on folk taxonomies or unspecific linguistic intuitions. Paolillo thus emphasizes the need for an empirical, variationist approach to the study of diglossia:

> Although the different models of diglossia disagree as to whether they regard H and L as distinct, no one appears to have developed an empirical procedure to test for distinctness in diglossia, or indeed any other language situation. Instead, diglossia research typically reports impressionistic judgements about the relatedness or divergence of varieties used in a given linguistic situation ...

> Claims about one, two, three or uncountable numbers of canonical varieties are simply asserted without little further justification.

Multivariate clustering techniques are useful descriptive tools for investigating the distinctiveness of the codes used in diglossic and other linguistic situations. The multivariate analysis presented in the first part of this chapter shows that bi-systemicity and discontinuity was a defining feature of morpholexical and aspects of syntactic variation (i.e. *nie-2*). Morphosyntactic variation, on the other hand, was continuous and non-discrete (i.e. non-diglossic).

In diglossic speech communities, all language behaviour is subsumed under the linguistic duality of diglossia. Varieties which show both High and Low features do not normally constitute a category of their own. Their linguistic identity is, according to Ferguson, 'unstable' and 'embarrassingly random' since they are the result of *ad hoc* interference from one discrete grammatical system on the other (i.e. High with interference from Low, or Low with interference from High; cf. Ferguson 1963:77). Britto (1986:316) refers to such intermediate forms as 'hybrids', that is, mixed varieties whose 'features are random and whose diasystemic identity is unspecifiable'; an intermediate variety 'cannot be identified as either H or L … is stigmatized (since it reflects the language user's incompetence)'. These varieties are thus seen as an unsuccessful attempt to produce either the High or the Low norm (and thus fall within the broad spectrum of High or Low rather than constituting an alternative norm). They are reminiscent of Haugen's notion of schizoglossia, i.e. the *ad hoc* production of linguistic utterances which correspond neither to the High (standard language) nor the Low variety, and whose in-between status fills their users (who generally have had some exposure to the standard via the education system) with feelings of inferiority and insecurity (cf. Haugen 1972:148; Martinet 1961:154 described similar forms in the context of what he calls *la situation patoisante;* cf. also Labov 1991 [1972]:213).

To assume that influences from metropolitan standard Dutch were random, and to discount the variation patterns in the written record as a writer's unsuccessful attempt to conform to the norms of Dutch, has been the position of traditional Afrikaans historical linguistics (cf. Chapter 1). However, linguistic variation in cluster II, which is 'fuzzy' in terms of its systemic status – that is, it cannot be said to form a linguistic system of its own, but rather bridges the gap between cluster I and III – is neither random nor unstructured. Cluster II thus differs qualitatively from the *ad hoc* mixed and intermediate varieties which have been described by Ferguson and others for diglossic situations. The fact that the overall pattern of the three clusters shows symmetrical re-

lations similar to those reported in social dialect studies further supports the interpretation of the Cape Dutch speech community as dialectal rather than classically diglossic. As a social dialect cluster II is not merely the transient result of mass education and limited exposure to the standard, but constitutes a highly variable, yet structured variety in the repertoire of the Cape Dutch speech community.

In order to investigate the sociolinguistic situation at the Cape and the linguistic strategies of individual speakers in more detail the following section looks at patterns of code-mixing/switching in the corpus data. Sociolinguists have argued that qualitatively different strategies of bi-dialectal and bilingual language behaviour correspond to different types of diglossia and language contact (cf. Meyers-Scotton 1986; Cochran 1997; also Muysken 1997, 2000). Investigating the patterns of Dutch/Afrikaans variation from a code-mixing/switching perspective can, therefore, be seen as complementing the quantitative analysis. In the following sections the hybrid expression 'code-mixing/switching' will be used as a cover term in cases where a more specific categorization of the type of bilingual speech is impossible. Although some linguists (e.g. Muysken 2001; McCormick 2002) maintain a relatively clear-cut distinction between 'code-mixing' (involving single-item switches and structural convergence) and 'code-switching' (involving alternations of more than one word, often inter-sentential), the two processes cannot always be clearly separated in the analysis of individual texts or speech events, and terminological precision is often impossible (cf. Auer 1999).

Afrikaans-Dutch code-mixing/switching

Code-mixing/switching between Dutch and Afrikaans occurs throughout the corpus data. Switches into Afrikaans are particularly common in cases where direct speech is reported ('quotational code-switching', Romaine 1995: 175; cf. also Gumperz 1982: 75–77). This type of (inter-sentential) code-switching also characterizes literary texts such as *Benigna* (1873, see above) and Lub's short-stories (1908). Quotational code-switching is also used when representing the speech of children (1; letter from Johanna Brümmer to F. S. Malan, 14/5/1901). Switch into Afrikaans in italics.

(1) Nu toen ik hem vroeg wat ik aan u moet schryven zeide hy: '*Ze ver pappie ik leer my lees*'.

['Now when I asked him what I should write to you he said: "*Say to daddy that I learn my lesson.*"']

Inter-sentential code-switching into Afrikaans also occurs in letters (or parts of letters) addressed to children. The letters by Maria Malan to her sister-in-law Katie are all in passable Dutch; however, on several occasions she shifts into Afrikaans when providing post-scripts or comments intended for Katie Malan's children (2).

(2) Wat doen de kleinen? Wat zegt Vrede van zyn briefje? zal hy weer schryven? *zé ver hem en Rechtie, hulle moet ver oupa kom help werke, arme oupa heeft zoo veel te doen!* Nu liefste, myn tyd is beperkt. Pa wacht om de brieven te posten. (29/5/1906)

['How are the little ones doing? What does Vrede say about his letter? Will he write again? *Tell him and Rechtie that they must come and help Grandpa with the work, poor Grandpa has so much to do*! Now my dearest, my time is limited, Dad is waiting to post the letter.']

Narrating anecdotes can also trigger a switch into Afrikaans, as evident in a letter by Francois Malan to Johanna Brümmer (10/4/1892; Afrikaans in italics, English in small caps). Francois Malan writes about his recent travels in the colony and refers to a mutual acquaintance who he met on the train (3).

(3) Ik zal u nu niet lastig vallen met een breedvoerig verhaal van mijn reis, doch als gij er belang in stelt zal ik later zulks doen. Overall werd ik getroffen door de vriendelijkheid de menschen [,] behalve op Pt. Elizabeth heb ik nooit iets voor kost uitgegeven. Bij Mr. Wilkie onder andere bij wien ik een avond heb geslapen (in 't district van bedford) waren de meisje en jongetje de volgende dag zoo gek naar mij dat zij huilden om mee te gaan naar de statie. Kost had ik altijd in overvloed en als de anderen kort schoten komen zij maar by Malen eten.
Daar zit een jonge dopper langers een ou baas in die trein. Hulle praat oôr die lokaties, die Turks-veige, die vooruitzichte en die boerderij. Kom ek gé jou af als jij raai dat de jong – een stellenboscher is. Dit is oom voor en oom na en oom is net spijt als oom moet uit klim. Ja dit is die zelftde jong dopper wat drie maande gelede te lui was om een perd te huur en naar Herschel te rij.
NEVER MIND, HE WILL GET THE OPPORTUNITY SOME DAY AGAIN.
Graaf-Reinet lijkt net zooals Worcester en Uitenhage is een fraai dorp. Maar wacht ik zal eerst uitvinden of gij er van houdt dat ik deze dingen schrijf.

['I won't bore you with a detailed narrative of my travel, but if you are interested I will do so later. Everywhere I am impressed by the friendliness of people, except for Port Elizabeth I haven't yet spent anything on food.

> At Mr. Wilkie, for example, where I slept one evening (in the district of Bedford), the girl and the boy were the next day so besotted with me that they cried to come with me to the station. Food I always had in abundance and when the others have little they just come and eat at Malan.
> *A young Dopper* [*member of the Dutch Reformed Church*] *sits next to an old man in the train. They talk about the locations, the prickly pears, the future and about farming. Come, I'll leave you alone if you guess that the boy – is from Stellenbosch. It is uncle before and uncle after and uncle is just sorry when uncle must climb out. Yes, this is the same young Dopper who three months ago was too lazy to hire a horse and ride to Herschel.* NEVER MIND, HE WILL GET THE OPPORTUNITY SOME DAY AGAIN.
> Graaf-Reinet looks just like Worcester and Uitenhage and is a beautiful village. But wait, I will first find out whether you like it that I write these things.']

It appears that this strategy was not limited to members of the Cape Dutch speech community, but was also employed by English-speaking South Africans. Thus Schreiner (1992 [1923]: 79) remarks:

> It is often complained by persons lately from England, that when the English South African has a joke to make, or comic story to tell, he lapses into the Taal, which is not understood by the newcomer; the truth being that it is the use of the Taal which transforms an ordinary sentence into a joke, and makes the simplest story irresistibly comic.

Much research on code-switching/-mixing makes use of the concept of a base or matrix language which is seen as providing the morphosyntactic frame for the utterances and which is usually identified 'on the basis of relative frequency of morphemes' (cf. Myers-Scotton 1993a: 486–487; 1993b: 46ff.). However, it is not always possible to identify the matrix language of a given utterance unambiguously (cf. Giacalone Ramat 1995: 55; Gardner-Chloros 1995: 71; Muysken 1995: 181–182). This is illustrated in (4) which comes from a letter by F. S. Malan (1/6/1893). The text shows frequent (multi-constituent) switches between Dutch, Afrikaans and English at clause boundaries (Afrikaans italicised, English in small caps).

(4) *Kalfoog is deur! Die results is uit!* Ik weet niet of dit nou juist zoo danig aangenaam nieuws zal wezen, maar myn hart kan ook bly wezen. Of ik het verdiend heb weet ik niet maar het is toch daarom lekker. Ik wil *Joe haar hart ook 'n beetje wakker scheed.* Denk er aan, *ou Kalfoog het oek 'n ou examentje gemaak.* Ik had gehoopt om in Class II te komen doch daar er geen Class I is komt het niet zoo groot op aan. …

> Pas 'n brief van Hendrik Theron van London ontvangen. Onder andere schryft die stonkerd: "*Arie, hoe better zal die kleinding ze hartje klop in Afrika als zy dit hoor. Ja, ni, dan droom zy zomaar weer lăt zy ver hom zoo in zyn Toga zien staan*." waar hy er aan komt weet ik niet. Het schynt my of hy meer weet dan ik. wien kan hy bedoelen? Wiens hart zal om *ou Kalfoog klop?* "Fishing again" you will say. Mag ik zulks hopen? Joe dear?
>
> [rest of the letter in Dutch] (1/6/1893)

> ['*Kalfoog* ['calf eyes' – term of endearment] *is through! The results are out!* I don't know if this will just now be such very good news, but my heart can also be happy. Whether I deserved this, I don't know but it is nevertheless great. I will *also stir up Joe's heart a little bit.* Think about it, dear old Kalfoog has also done a dear old exam. I had hoped to get into Class II but since there is no Class I it doesn't matter that much …
> Just received a letter from Hendrik Theron from London. Among other things the trouble maker writes: "*Oh, how much better will the heart of the little thing in Africa beat as she hears this. Yes, no, then she dreams just again that she sees him standing there in his toga*." How he arrives at this I don't know [?] It seems as if he knows more than I do. Who can he mean? Whose heart will for *old Kalfoog beat*? "Fishing again" you will say. May I hope for such a thing? Joe dear?']

It is difficult to pinpoint individual switching sites (i.e. the actual point of switching) in bi-dialectal language behaviour as utterances and texts usually contain several 'homophonous diamorphs' (Clyne 1987: 754; also Clyne 2003: 164ff.), i.e. morphemes which are identical in both codes (such as *is* for the third person singular as in *Kalfoog is deur!*). The typographical marking of code variation by italicization should therefore be seen only as a general interpretative guideline.

The switches in example (4) involve both grammar and lexicon, are multi-constituent and are typically located at clause boundaries. Muysken (1997: 361, 2000: 96ff.) has discussed such patterns of code-switching/mixing under the term 'alternation', that is, a type of bilingual/bi-dialectal speech characterized by regular switches involving complex structures (typically several constituents in a row). Alternation is 'a strategy of mixing, in which the two languages present in the clause remain relatively separate' (ibid. 2000: 96), and can be distinguished from 'insertion' (i.e. the embedding of individual linguistic constituents into a clearly defined matrix language; 'embedded language islands' according the Myers-Scotton's terminology) and 'congruent lexicalization' (i.e. the continuous mixing of material from two languages, with no matrix lan-

guage identifiable). Both insertion and congruent lexicalization will be discussed in more detail below.

Example (4) also shows possible instances of triggering, i.e. cases where 'an item of ambiguous affiliation (that is, one that belongs to the speaker's two systems) triggers off a switch' (Clyne 1987: 754–755 and 2003: 162–163; cf. also Myers-Scotton 1993: 491; Giacalone Ramat 1995: 59). In this case the word *lekker* belongs to both lexical systems but has a wider semantic range in Afrikaans. In the extract from the letter *lekker* triggers a clause with an uninflected verb phrase. Note that the possessive construction *Joe haar hart* (rather than standard Afrikaans *Joe se hart*) is typical for Francois Malan's use of the periphrastic possessive in his Afrikaans texts (while he avoids the periphrastic construction altogether when writing Dutch). A second trigger is the use of an English expression (reported as direct speech) which triggers a switch to English, and the sentence is then finished according to the morphosyntactic rules of English.

In those letters with Afrikaans as a matrix language (cluster III) switches into Dutch are typically related to topics connected to those social and communicative domains (education and religion) in which Dutch still functioned as the norm up until the early twentieth century (5). In texts with Dutch as the matrix language, switches into Afrikaans were frequently triggered by topics connected to the private and intimate domain of the family (6). Again switching occurs typically at clause boundaries or the periphery of sentences (Afrikaans italicized).

(5) *Mij zendeling vriendin en ek het net al zwaar gevoel over die tekort*, nog gister ontving ik een brief van haar, waar in zij mij aansporen om tog nog meer en meer te bidden voor *die Werk*.
(letter continues in Afrikaans, Bettie Smit, 23/11/1917)

['*My mission (girl)friend and I felt very sad about the shortfall*, just yesterday I received a letter from her in which they encourage me to pray more and more for *the Work*'; letter continues in Afrikaans]

(6) wij zijn zeer gelukkig en tevreden. *Winnie maak zulk lekker kos* en is zoo handig in al hare huis werkjes dat ik waarlik verwonderd ben, en er is dus geen gevaar dat ik eenige neiging zou hebben om naar 't hotel terug te gaan. (letter continues in Dutch, Andries Hauptfleisch 11/6/1914)

['We are happy and content. *Winnie cooks such delicious food* and is so handy in all her housework that I am truly surprised, and there is thus no danger that I would have any inclination to go back to the hotel.']

The social meaning of Afrikaans as a language of intimacy and familiarity is stated almost programmatically in a letter by Hester Van der Bijl (born c. 1885) to her future parents-in-law (7). However, at the same time Afrikaans also seems to have carried connotations of disrespect and lack of education (8).

(7) Weet u wat? Ek virlang regtag om u te zien … Daarom schrijf ek ok sort van Afrikaans omdat dit vir my is, net of ek u al lang ken. (8/11/circa 1905)

['Do you know what? I really long to see you … Therefore I too write a type of Afrikaans because it is for me as if have already known you for a long time.']

(8) Ek kan niet H. Hollansch schrijf, ek meen nie goed nie … zoo schrijve ek lievers mij sort Afrikaans, maar dit lijk toch zoo 'disrespectful' om aan u zoo te schrijf.

['I cannot write High Dutch, I mean not well … therefore I prefer to write my type of Afrikaans, but it seems just so 'disrespectful' to write to you in this way.']

That alternation between codes was not limited to the written register, but also occurred in spoken language interaction, is suggested by an anecdote narrated by the same writer (9; letter dated 10/11/1916). Hester van der Bijl's son Pieter (born 1908) is shown to have both Dutch and Afrikaans at his disposal, alternating between the two in the course of the same situation (switch into Dutch in italics).

(9) Laatste Maandag was hier 'n klein Hollander meisje, en sij moest help 'n huisie van petrol blikke bouw, dit was in die middag, ek was in die kamer, maar kon alles hoor deur die oop venster, Pieter was eigentlik die bouwer en sij en Klasie die aan draërs, dinge het miskien te stadig na sij smaak gegaan, toe hoor ek hem uitroep na die klein meisie, met 'n opregte hollandse uitspraak, '*Antonia, breng twee binnen tegelijk dan is het wel spoedig af gedaan*' – dit was te mooi om te hoor, en gewoonlik praat hij sij beste met haar, Op 'n ander keer was het weer 'heh! daar zit nog wat leuke sousjes in de pan.'

['Last Monday there was a little Dutch girl here, she had to help to build a house out of petrol tins, this was at lunchtime, I was inside, but could hear everything through the open window. Pieter was actually the builder and she and Klaas the carriers, things went maybe to slowly for his taste, then I hear him call the little girl with a proper Dutch pronunciation: "*Antonia, bring two inside together then this is quickly done*" – this was too beautiful

to hear, and normally he speaks his best with her. At another time it was again "Hey! there are still funny juices in the pan."']

Muysken (2000: 8) has suggested that strategies of alternation, i.e. regular multi-constituent switches, are 'particularly frequent in stable communities with a tradition of language separation'. Speakers who show language alternation typically have a strong norm awareness and perceive the two languages or varieties as well-defined and clearly distinct linguistic entities (Muysken 200: 249). Cochran (1997: 49–50) has argued that the existence of such alternation patterns is a diagnostic feature of (classical) diglossic communities. In the corpus, strategies of alternation are found only in clusters I and III. In addition, members of clusters I and III also show examples of insertion, i.e. single-item switches (isolated constituent insertions) within a clearly identifiable and statistically dominant matrix language.

While those individuals who can be described as bi-dialectal (cf. Figure 7.13) show code-mixing/switching patterns of alternation (i.e. multi-constituent switches, typically at clause boundaries or inter-sentential switches) and insertion (single-item switches), other writers also mix elements of the two systems – however, they never conform to any one system for a longer stretch of discourse. Muysken (1997: 362) termed this latter type of mixing/switching 'congruent lexicalization', and described it as 'a situation where the two languages share a grammatical structure which can be filled lexically with elements from either language'. The speaker's awareness of linguistic norms is typically diffuse and the languages or varieties involved are not seen as being in overt competition, opposition or conflict. Unlike alternation, congruent lexicalization does not involve multi-constituent switches and switching sites are not located at clause boundaries. According to Muysken congruent lexicalization is 'akin to … monolingual variation' (ibid., also 2000: 123–126) and style shifting, and can be described quantitatively using Labov's notion of the linguistic variable. Congruent lexicalization as a type of bilingual speech also supports structural convergence between languages (and dialects; e.g. Gumperz & Wilson 1971; Louden 1994; McCormick 2002; cf. also Chapter 1 on 'convergent hybridization' in the history of Afrikaans).

Examples (10) and (11) illustrate patterns of congruent lexicalization in the corpus. Switching occurs within constituent boundaries, for example, within the verb phrase *(ik heef gekregen*, 'I have received'), as well as between noun phrase and verb phrase where the noun phrase is the head determining inflectional marking (*wij verwacht*, 'we expect', *het gaan*, 'it goes'). Afrikaans items are in italics in (10) and (11). Code-mixing within constituents is rare in

the texts produced by bi-dialectal writers, but is characteristic of those writers belonging to cluster II.

(10) ik *heef* nu tog een brief van Edmund Smith uit Barkley gekregen [,] het *gaan* met hem en Martha *reg* goed [,] ook met Johan en Barend gaat het *reg* goed (B. J. Brümmer, 7/12/1901)

['I have now received a letter from Edmund Smith from Barkley, it goes well with him and Martha, it also goes quite well with Johan and Barend']

(11) zaterdag *heeft ons* bazaar [,] dan *ver wacht* wij veel menschen [.] zij hebben *mos* een plan om hier een ge meenten te *stigt* [,] maar of zij het *zal* regcht krijgen weet ik niet (Petronella van Huyssteen, circa 1906)

['we are having a bazaar on Saturday, then we expect many people. They have indeed a plan to establish a congregation here, but I don't know if they will get this right']

The three-cluster solution, established on quantitative grounds using the construct of the linguistic variable, thus holds when approaching the variability in the data from a dialect switching perspective; the three clusters relate to different processes or patterns of code-mixing/switching.

Cluster I insertion and alternation
Cluster II congruent lexicalization
Cluster III insertion and alternation

Alternation and congruent lexicalization differ not only structurally, but probably also with regard to their sociolinguistic and neurolinguistic embedding. Alternation, as discussed above, correlates with social stability and code complementarity (i.e. is characteristic of a broadly diglossic scenario), while congruent lexicalization is believed to be typical for 'second generation immigrant groups, dialect/standard and post-creole continua' (Muysken 1997: 362) and can – over time – support the stabilization of a mixed code as the regular vernacular of a community (cf. McCormick's 2002, work on the District Six in Cape Town).

Although there is still considerable debate regarding the neuro-anatomical organization in bilinguals, some preliminary thoughts on the psycholinguistic correlates of these structural patterns should be permitted (for a recent overview and discussion cf. Clyne 2003: 193–214). Studies of bilingual aphasics support the hypothesis of the existence of two I-languages for bilinguals (see Paradis 1985: 22), and it appears that bilingual 'speakers can and often do keep their languages separately' (Muysken 2000: 277). Grosjean (1989, 1995,

2001) postulates two separately stored language networks in bilinguals and assumes that during code-mixing/switching (bilingual or bi-dialectal speech mode) both languages are activated, i.e. speakers have simultaneous access to both languages. Muysken (1997: 364) has suggested that different types of code-mixing/switching have distinct psycholinguistic correlates with respect to the 'level of monitoring in both languages, the triggering of a particular language by specific items, and the degree of separateness of storage and access system'. Thus, during alternation (cluster I and II) psycholinguistic activation probably switches from one language to the other, while during insertion the dominant language might be deactivated temporarily. Regarding congruent lexicalization (i.e. cluster II), however, it seems that the two (or more) languages or dialects do not constitute distinct codes with separate neuro-anatomical organization, but share a single processing and storage system.

Summary

The multivariate analysis of the corpus data discussed in this chapter shows that a complex sociolinguistic dialect continuum existed at the Cape up until the early twentieth century. Using hierarchical cluster analysis, three main sociolinguistic varieties were identified in the corpus: acrolectal Cape Dutch Vernacular, mesolectal Cape Dutch Vernacular and the emerging standard of early Afrikaans. These varieties formed overlapping rather than distinct systems, especially with regard to the structures of morphosyntactic variation.

A comparison with texts from the dialect writing tradition showed the new norm of Afrikaans to be typologically similar to the dialect literature. As regards the sociolinguistic structure underlying the threefold division of the data, a clear social patterning was found: mesolectal varieties were mainly used by older speakers belonging to the traditional petty bourgeoisie; acrolectal Cape Dutch vernacular and the emerging standard of Afrikaans, on the other hand, were found to be used predominantly by younger and highly educated speakers (belonging to what has been termed the new professional class). The latter made use of both Dutch and Afrikaans in their written texts and maintained a clear-cut distinction between the two norms, and their behaviour can be described as diglossic. This interpretation is also supported by the code-mixing/switching strategies of these writers (alternation and insertion). More mesolectal speakers, on the other hand, showed patterns of congruent lexicalization which are best described in Labovian, variationist terms.

Notes

1. Romaine (1982: 172) specified the following thresholds for her analysis of relative deletion in Middle Scots: O = 5% or less deletion, X = 5–25% deletion, 1 = 25% or more deletion.

2. The variables *ek* and *hom* as well as relative *wat* were included in a control analysis which was carried out to ensure the stability of the solution, and the two cluster solution was repeated. However, while two clusters were again clearly separated in the two-dimensional space they showed more within-cluster-variability. This is not surprising since these forms were less broadly diffused than the other pronoun variables.

3. As the narrative of the novel is in Dutch, only sections which reported direct speech were included into the analysis.

4. Ponelis (1996: 134), for example, described such intermediate linguistic usage as '[d]utchified Afrikaans or approximate Dutch … essentially the product of poorly educated Afrikaans-speakers struggling to write and speak formal Dutch'. To recognize the structured nature of these intermediate forms is not to say that borrowing (from the standard variety) played no role in the process, and growing exposure to the Dutch standard norms in the context of a gradually expanding school system supported the maintenance of Dutch variants.

5. It appears that mesolectal varieties are rarely recognized in folk taxonomies; cf. also Patrick (1999: 273) with respect to the situation in Jamaica: 'The belief that there are two stable and distinct codes, commonly labelled "English" and "Patwa", with nothing in between, is even more widely held by Jamaicans than it is by creolists. This folk position, which is the easiest one to elicit from JC [Jamaican Creole] speakers, does not explicitly recognize intermediate varieties, and confers no firm sense of alternative identity upon the mesolect'.

6. The nature of the Greek continuum is, however, qualitatively different from that of the Cape Dutch speech community at the Cape. In Greek, for example, inflected verbs cannot be assigned to either the High variety or the Low variety since there exist forms which are combinations of High and Low morphemes, e.g. *parusiástike* (L form) – *eparusiástike* – *eparusiásthike* – *eparusiásthi* (H form; 'he was present'; from Hawkins 1983: 226).

Part III

Establishing the norm

Chapter 8

Engels, Engels, alles Engels

Language contact, conflict and purism

> The educated are, consequently, exposing their Dutch vocabulary to what the Afrikaans Patriot might call the contagion of English, and when they are called upon to deal with subjects to which the Taal cannot do justice they will, whether consciously or mechanically, resort to English for the terms that must supply the shortage.
>
> A. J. Barnouw (1934). *Language and Race Problems in South Africa*

British colonial rule: 1806–1910

The advent of British rule in the late eighteenth century not only changed the political conditions at the Cape but also had far-reaching sociolinguistic consequences. The first British occupation took place in 1795. The colony was, however, returned to the Dutch for the years 1803 to 1806 after which it remained in British hands. In 1815 British control of the Cape was confirmed by the Congress of Vienna. From early on the Colonial Office advocated a policy aimed at assimilating the Dutch-speaking inhabitants into the English culture and language, and in 1822, under Lord Charles Somerset, a government proclamation granted official status to English only, thus reducing the use of Dutch to the religious and private sphere (including privately financed schools where Dutch remained the medium of instruction). Somerset's policy was, however, only implemented slowly and often took second place to more pressing political issues such as the consolidation of the eastern frontier or the politics of slavery before emancipation in 1838 (for a detailed overview of the anglicization policies at the Cape cf. Sturgis 1982). Although speakers of English were numerically in the minority during the nineteenth century (Watts 1976: 43–44), their numbers gradually increased through on-going immigration from the British Isles. English quickly gained ground in the urban centres of Cape Town, Kimberley and Port Elizabeth where the majority of immigrants

settled (Watts 1976: 46; cf. also Colquhoun 1906: 144–145; Schonken 1914: 193; Barbouw 1934: 38). McCormick (2002: 30) describes the growing dominance of English in the urban environment – with reference to Cape Town – as follows:

> English was the main medium of communication for Cape Town's growing national and international connections with commercial, manufacturing, and mining interests. This had a ripple effect on the city's inhabitants: many people needed to learn English for trade and employment purposes. Thus, it came to be used at work by speakers of Afrikaans, Yiddish, Portuguese, and other European languages, as well as by speakers of African languages from within and beyond the colony. Many of them learned English in the workplace. These adults then wished their children to become proficient in the language and kept them at school for as long as they could afford to. In some communities … this led to language shift. In other communities bilingualism became the norm … Whereas schools had been relatively insignificant as sites of language learning during the period of VOC rule, in the British colonial period schools and formal classes for adults were very important loci of language learning.

The historical record indicates little overt conflict between English and Dutch colonists during the first decades of the nineteenth century, and until the 1850s there was a general consensus that English would be the future language of the colony; however, first dissenting voices were heard in the periodical press (cf. the source material presented in Scholtz 1964: 27–28; also Steyn 1980: 124–131; Du Toit 1985: 210ff.; Zietsman 1992: 4–6). From the 1860s onwards social life in the colony was significantly anglicized and British social and linguistic norms came to dominate in Cape society. 'Englishness' developed into a major social symbol and determined 'what was right and acceptable' in the political and, in particular, the social and cultural life of the colony (Ross 1999a: 43).

Although there are no language shift statistics for the nineteenth century,[1] there is historical evidence for inter-generational language shift in traditionally Dutch/Afrikaans speaking families. The following comment comes from a letter by Elizabeth J. Malan (9/5/1893):

> oom Gawie (dokter) was verleden week hier, by ons. Hy was 14 dagen te Wellington, zyn klein zoontje was zamen met hem, hy spreekt alles engelsch
>
> ['Uncle Gawie (doctor) was last week here, with us. He spent 14 days in Wellington, his little son came with him, he speaks only English']

It is difficult to reconstruct to what extent English had invaded the informal domains in traditionally Dutch/Afrikaans-speaking families in the mid- to late nineteenth century. A cursory look at, for example, the private and official correspondence of ministers of the Dutch Reformed Church shows dominance

of English in most private documents (diaries and private letters), while the official correspondence on church matters was regularly conducted in Dutch (cf., for example, the papers by B. F. G. De La Bat, NGK/P-54/1, or T. C. B. Vlok, NGK/P-24). A case in point is Andrew Louw (born 1862) – son of Ds Andries Adrian Louw (born 1827) – who corresponded with his siblings and mother, as well as with his fiancée and later wife Francina Malan, exclusively in English. Writing to his father, however, he used Dutch only. The same pattern of language use (i.e. Dutch to the father, English to everyone else) is also characteristic of the private papers of his sister Maria Johanna.

The growing bilingualism gave rise to strategies of code-mixing/switching between Dutch/Afrikaans and English. Although code-mixing/switching research has by and large focused on contemporary immigrant communities and post-colonial societies, bilingual practices have also been described historically, for example, for Middle and Early Modern English (Schendl 1996; Wright 1998) and seventeenth-century German (Von Polenz 1991: 225–226). Mesthrie (1993) has argued that there is ample evidence (traveller's journals, poems, plays, descriptions of the education system as well as reported linguistic usage), which shows that code-mixing/-switching between Dutch/Afrikaans and English was a sociolinguistic reality at the Cape during the nineteenth century.

English-Dutch/Afrikaans code-mixing/switching

There are a number of literary examples of English-Dutch/Afrikaans code-mixing/switching for the nineteenth and early twentieth centuries, such as the popular poem/play *Kaatje Kekkelbek* written by A. G. Bain in 1838 (cf. Mesthrie 1993 for a discussion), Brink's play *Volk en Taal* (1905) and the short-stories by Lub (1908). These texts show code switching/-mixing as a regular, yet stigmatized sociolinguistic practice. While in *Kaatje Kekkelbek* the protagonist is the stereotypical *Hotnotsmeid* ('Hottentot girl'), later representations of code-switching/-mixing (such as Brink and Lub) identify young Afrikaner women (and to a lesser degree young men) as those typically engaged in such bilingual behaviour (cf. also Viljoen 1891: 4; De Vos 1891: 22–23; Te Winkel 1896: 371).

A distinction between situational and conversational (or metaphorical) code-switching was introduced by Blom and Gumperz (1972: 424–426). Situational switching is akin to style-shifting in that it is a direct response to the context in which the speech event occurs (generally change of setting or participants). Conversational code-switching, on the other hand, is not triggered by a change in the social parameters of the situation, but typically relates to

the discussion of particular topics or subject matters, though it can also occur without direct contextual motivation (cf. Gumperz 1982: 60ff.). Myers-Scotton (1986: 406) termed this latter type 'switching as an overall unmarked [i.e. socially expected] choice'. It is a type of bilingual speech characteristic of societies where the social identities of speakers are associated with both languages:

> When more than one social identity is salient for the rights and obligations balance which speakers wish to have in effect for the current speech event, and each identity is encoded in the particular speech community by a different linguistic variety, then those two or more codes constitute the unmarked choice. … In this case, each switch in itself has no special social significance; rather, it is the overall pattern of using two varieties which carries social meaning.
>
> (ibid.)

Both intra-sentential and inter-sentential code-mixing/switching occur in the *Corpus of Cape Dutch Correspondence*. Most of the examples in the data show single English lexemes which are incorporated into the morphological and syntactic structure of the matrix language. It is not entirely clear to what extent these lexemes were integrated into the Dutch/Afrikaans lexicon of the time (and are thus better described as stable borrowings), or whether they reflect a type of code-switching/mixing behaviour which Myers-Scotton (1993a: 486, 1993b: 77–78) had glossed 'ML [matrix language] + EL [embedded language] constituents' (Sankoff, Poplack, & Vanniarajan 1990 describe such single lexeme constituents as 'nonce borrowings'; Muysken refers to them 1997: 361 as 'insertions', see Chapter 8, and Clyne 1987: 740, 2003: 76–80 as 'transferences'). Today, several of these lexemes constitute well-established borrowings in colloquial Afrikaans and are seen as an integral part of the non-standard Afrikaans lexicon. Single lexeme switches in the data are occasionally marked by the writers themselves using inverted commas (7). This suggests that they were not yet fully integrated. The examples involve nouns ((1) and (2)), adjectives within a noun phrase (3), adjectives in predicate function (4), finite verbs (5), non-finite verbs in infinitival subordinate clauses (6), and infinitives as part of a verbal cluster (7). English lexemes are italicized.

(1) ik het een van myn menschen afgevaardig om na een *Propperty* te kyk die my angebode is ('I have sent out one of my people to look at a property which was offered to me'; F. S. Cillie, 1900)

(2) Zy is een zeer knappe *nurse* ('she is a very capable nurse'; Katie Beyers, 25/2/1906)

(3) want *3rd* [class] rijdt elke *common* mensch ('because every common person travels third class'; Rudolf Baalie, 23/2/1894)

(4) ons is vandaag ook *jollie* want het regen nu ('we are today also jolly because it's now raining'; P. A. Goosen, sen., 27/7/1909)

(5) Denk toch niet dat ik dit uit een boek *copy* ('don't think that I am copying this out of a book'; Sophie Hendrikz, circa 1900)

(6) dat ek jou hond aan Berth overhandig het om saam met sij oû hond Jock te *train* ('that I have given your dog to Berth to train together with his dog Jock'; F. P. Hoogenhout, 23/3/1912)

(7) Morre moet ik al weer gaan '*call*' en over morgen moet ik naar een '*wedding*' (engels) ('tomorrow I must "call" again, and the day after tomorrow I must go to a "wedding" (English)'; Hester Van der Bijl, 13/2/1906)

Turning to the conversational motivations for these single item switches, many of them relate to specific topics, such as transport technology (8), (9), employment (10), social life and entertainment (11), (12), sport (13), and education/school environment (14), (15). Code-mixing/switching is also used to reiterate information (16).

(8) de vrouw ging met de kleinere kinders in een *carriage* ('the wife went with the smaller children into a carriage'; Rudolf Baalie, 23/2/1894)

(9) ons het gegaan met Mr. Brink z'n *motor bike* en *side-car* ('we went with Mr. Brink's motor bike and side-car'; Wynanda Hoogenhout, 15/6/1914)

(10) zoo dat ik kan weet om mijne *notice* te geven ('so that I can know in order to give my notice'; Sophia Daniels, 23/9/1900)

(11) Hy het gestuur 'n pragtige 1/2 doz *china tea cups*, 6 silver lepeltjes – en 'n prachtige *stand* met *teapot* en di *spirit* lampie ('he sent a wonderful 1/2 doz. china tea cups, 6 silver spoons and a wonderful stand and the spirit lamp'; Sebastina Goosen, 19/4/1910)

(12) Ik heb Susie '*Chess*' leeren spelen en nu zal zy my '*painting lessons*' geven ('I taught Susie to play "chess", now she will give me "painting lessons"'; Johanna Hoogenhout, 26/10/1902)

(13) Hoe het jul '*Tennis*' afgekom? Ek speel nou op Mr. Theron zij *courts*, maar hul is vreeslik slecht ('How was your 'Tennis'? I play now on the courts of Mr. Theron, but they are awfully bad'; Mimmie Laubscher, 1/12/1909)

(14) Mr. de Kock zeg dat Jacob en Jannie en ik ge*pass*ed hebben ('Mr. de Kock says that Jacob and Jannie and I have passed'; Susie Goosen, 23/2/1909)

(15) waar is logies om al de menschen met de *ordinary* of gewone treinen [?] ('where is accommodation for all the people with the ordinary or common trains [?]'; C. J. Van der Merwe, 30/11/1900)

Examples (14) and (16) show code-mixing within word boundaries. Such mixed constituents are quite common in the corpus and occur mostly in the past participle (cf. also Te Winkel's, 1896:361, commentary in which he gives the following examples: *ons het ons baing geënjoy*, 'we have enjoyed ourselves tremendously', and *ons het gewalk*, 'we have walked').

(16) a. ik heb … nog een weinig by hem ge*order*d ('I ordered a little bit more from him'; J. J. De Villiers, 4/12/1882)
b. dat fir hulle almaal … spesiale eerste Klas sitplekke ge-'*book*' is ('that for all of them … special first class seats are booked'; J. W. F. Grosskopf, 21/10/1906)
c. Ek het vir hom en Ruth gezien en by hulle ge*supper* in 'n restaurant ('I have seen him and Ruth and have eaten with them in a restaurant'; F. S. Malan, 2/4/1896)

Code alternation in the corpus also involves longer constituents such as English idioms (17), as well as strings of material following English morphosyntactic rules ('embedded language islands' in Myers Scotton's terminology) as shown in example (18).

(17) *'By the way'* zeg hom, dat ik dikwijls over hem denk ('By the way, tell him that I am frequently thinking about him'; Jan de Waal, 8/1/1902)

(18) Ga in! zegt hy voor mij en als hij U weder uitjaagt, *then you wire back and I shall teach him* ('Go in! he says to me, and if he chases you out again, then you wire back and I shall teach him'; Rudolf Baalie, 23/2/1894)

Although there are several examples of English-Dutch/Afrikaans code alternation in the *Corpus of Cape Dutch Correspondence*, in terms of relative frequency the occurrence of code-switched material is mostly limited to one or two switches per text. There are, however, some interesting exceptions, most notably Johanna Brümmer (born 1877), the young fiancée of Francois Malan. In one letter in particular (10/5/1895) she quite regularly inserts English constituents into the Afrikaans frame. Example (19) is a characteristic extract from the letter.

(19) Van aand gaan ons *practice* by die pastorie – wil jy ni saam gaan ni? Connie Newnham zal dar wees, en dan zal ons net 'n *jolly* aand deurbring. Ik zal so *like* dat jy ver haar ontmoet. Zy is altyd zo *bright*, al is dit maar treurig in

hulle huis dan lag zy altyd. Ons het di *pledge* ge*sign* om ver Dr. Newnham aan te moedig om ver hom terug te hou van di drank. Ons draag now almal wit *ribbons* want ons behort aan di *association*. Een van di dage gaan ons Dr. Newnham vra om di *pledge* te *sign* en dan kry hy oek 'n '*white ribbon*'.

['Tonight we are going to practice at the parsonage – don't you want to come with us? Connie Newnham will be there, and then we will spend such a jolly evening. I would like so much for you to meet her. She is always so cheerful, even if it is sad in the house she always laughs. We have signed the pledge to encourage Dr. Newnham to hold back from the drink. We are all wearing white ribbons because we belong to the association. One of these days we will ask Dr. Newnham to sign the pledge and the he will also get a "white ribbon"'.]

The structure of (19) exemplifies Myers-Scotton's (1993:482) concept of code-mixing/switching as an unmarked choice, i.e. it occurs in a peer-group situation and is characterized by frequent intra-sentential switches. Inter-sentential switches are common in Johanna Brümmer's letters, where they often occur in a topic-comment structure. In example (20) the main topic is discussed in Afrikaans, followed by a metalinguistic commentary in English (in italics), then the switch to English is commented on in Afrikaans. The rest of the letter is entirely in Afrikaans (on code alternation in topic-comment constructions, cf. Romaine 1995:163).

(20) Van môre voor dat ik jou brief gekry het dach ik by myself dit is tyd dat ik weer ver Kalfoog in zy geliefde taal skryf. En toe herinner gy ver my dit in jou brief van môre. Dankie daarem (*I cannot help laughing at the easy way of spelling)*. Ah! Exkuus Kalfoog ik zal nie weer soo deur makaar praat nie.
(Johanna Brümmer, 7/12/1893)

['This morning before I received your letter I thought by myself that it is time that I write again to Kalfoog ['calf's eye', term of endearment] in his beloved language. And then you remind me about it in your letter this morning. Thank you for it (I cannot help laughing at the easy way of spelling). Ah! Sorry Kalfoog, I won't talk so mixed-up again.']

Use of inter-sentential code-switching in topic-comment constructions is also found in Francois Malan's (born 1871) letters to Johanna Brümmer (21).

(21) Ik denk dat ik u vroeger reeds geschreven heb dat mijn jongste broertje (omtrend twee jaren oud) een ongeluk gehad heeft en nu dikwijls trekkings krijgt. *Poor boy, it works on his nerves + brain + the Drs. don't*

think he will ever become well again. (F. S. Malan, 10/4/1892)

['I think that I wrote to you earlier that my youngest brother (about two years old) had an accident and he now frequently has convulsions. Poor boy, it works on his nerves + brain + the Drs. don't think he will ever become well again.']

Topic-comment code alternation is also a prominent feature in the letters of Mimmie Laubscher (born c. 1885 and a teacher at the prestigious Bloemhof school in Stellenbosch) to her fiancée Daniel Marais Hoogenhout. In example (22) and (23), the discourse topic is established in English. Afrikaans (in italics) occurs in the comment position.

(22) P.S. By the way I have not had a photo taken yet – *ik is zoo bang di glas zal bars* ('I am so scared the glass will break') – but I hope to have one taken soon, and will then let you have a copy (Mimmie Laubscher, 22/5/1909).

(23) [I]f you gave me full permission to kiss him do you think I would? *Dit zal erger als een 'seep-schottel' smaak, na die lekker laaste zoen van jou bij die trein die aand* ('This will taste worse than a soap dish, after the nice last kiss from you that evening at the train'; Mimmie Laubscher, 9/8/1909).

In several cases the postscript following a Dutch or Afrikaans letter is written in English (e.g. Maria Malan in the postscript of a Dutch letter to her brother Francois, 14/5/1892: 'Please write soon dear! Adieu Won't you be grand in going to Europe eh? Please excuse paper and writing as I have a truly horribly bad pen to write with'); or a reference to, for example, an English book or letter can trigger a complete switch to English half way through a Dutch letter (e.g. Rijkie Louw to Katie van Huyssteen, 4/8/1909).

Issues of language choice are also reflected in the overall linguistic behaviour of many of the writers. In some cases one sees a clear interlocutor effect as in the case of the Louw family letters (mentioned above), where Dutch was the language of choice only in correspondence with the father, while siblings communicated with one another in English. Interesting are also the overall language choice pattern in Johanna Brümmer's letters to her fiancée Francois Malan, who repeatedly urged her to write to him in Afrikaans. Johanna Brümmer, however, shows a clear preference for English and when she started writing in Afrikaans it was to please and appease him (see example (20)). Eventually she suggested the following compromise:

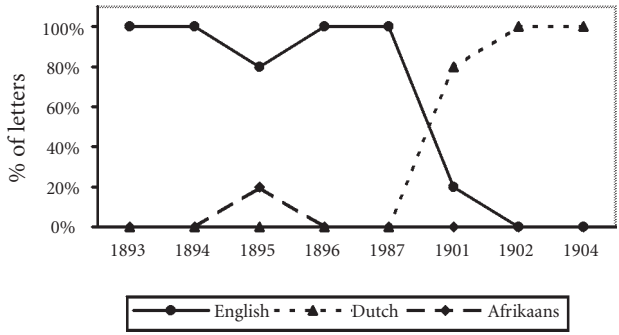


Figure 8.1 Code choices in Johanna Brümmer's letters to F. S. Malan, 1893–1904. *N* = 209 letters; in percentages.

> You don't mind my writing in English this week [?] I hope always to write in Afrikaans every other week. Are you satisfied with this arrangement? (8/3/1895)

Yet, judging from the quantitative overview of her language choices (illustrated in Figure 8.1), this agreement was on paper only, and did little to minimize the dominance of English in her early correspondence. English disappeared from her repertoire as a favoured code-choice after the Anglo-Boer War when it was replaced by Dutch.

Questions of national identity politics within the context of an anglicized society provided conflict potential for private relationships, and repeatedly Johanna Brümmer explained and justified her preference for English. Judging from her own testimonial (letter to F. S. Malan 23/11/1894), her choice of language appears more a result of habit than lack of identification with early Afrikaner nationalism.

> Kalfoog Mine, Your last letter made me think a great deal. From what I can understand you wish to know my feelings with regard to my own people the Africanders. Is that it? I know I have often given you to understand that I am prejudiced against them. My having done so was never in earnest. Perhaps you still remember Ben's way of talking about the 'boere' and at heart he is a true Patriot. I have unconsciously drifted into this habit by pretending to be above them – If any one else should do the same things [in] my presence it hurts me at once and I am not slow to resent it too. I have never once thought it a disgrace to own that I am a 'boere meisje' … I do not remember you telling me why you work for Afrikaans – But understood you wish to be of help and use to your country men. 'It may be only a means to an end' I do not understand … As to the language – I simply write English because I am used to it – As you

know we speak Afrikaans – In entertaining visitors it depends upon themselves which language we use. I had never seen Afrikaans on paper until you wrote me the first letter – So imagine how strange it must have seemed at first.

A clear preference for English is also evident in Mimmie Laubscher's letters (cf. also example (23) and (24) above) to Daniel Hoogenhout. Only two of her sixty-six letters to him are written completely in Afrikaans, yet in fifteen of the English letters she makes use of Afrikaans in topic-comment structures (examples (22), (23)) and in postscripts. Examples such as (24) reflect her playful use of linguistic resources (Afrikaans in italics), including a fanciful Latin construction.

(24) Chum I find it best not to worry about the future. Of course I know at times one is apt to do so, and cannot help it, but after all it won't mend matters. Even if as you say the present state does continue for 2 or 3 years, well then all we can do is to be satisfied. *Alles zal wel recht kom, net ek is jammer vir jou wat jij zal met een old-maid opgeskeep wees!* ['All will come right, I am just sorry for you because you will be stuck with an old-maid'] Cum multum amorem, Your own Mimmie. (2/09/1909)

Anglicisms

At the Cape, societal and individual bilingualism not only supported the development of what Muysken (2000) has called 'bilingual speech', i.e. code-mixing/switching patterns and practices, but also structural interference, calquing and lexical borrowing. So-called 'anglicisms' are widespread in modern Afrikaans and are also attested for the nineteenth and early twentieth centuries (cf. Pienaar 1931: 173–174; Nienaber 1936: 141; cf. Donaldson 1991: 18ff. for a critical discussion). Although this is not the place to deal with these contact phenomena in detail, some comments should be permitted, especially as this area has largely been neglected in diachronic studies of Afrikaans (cf. Donaldson 1991: 7). As already argued in Chapter 1, by placing the formation of Afrikaans in the late eighteenth century, Afrikaans historical linguists have traditionally interpreted the nineteenth century as a time of linguistic consolidation and continuity. This position has been challenged by Donaldson (1995: 223; cf. also Valkhoff 1966: 195), who suggested that with regard to language contact 'the linguistic transformation that would take place after the

British occupation of the Cape in 1795 was to be as great as, if not eventually greater than, all the changes that had taken place hitherto'.

Several of the loan translations listed by Swart and Prinsloo (1936: 1–70) and also Donaldson (1991: 170–177) were found in the corpus (examples (25)–(32)). This suggests that many of these loan translations are possibly not so much 'the result of spontaneous translation by the bilingual Afrikaner' (Donaldson 1991: 170), but have a certain historical depth and stability.

(25) **so ver (Engl. so far; up till now)**
Met de boerderij gaat het nog *zoo ver* alles goed ('so far it all still goes well with farming'; E. Smith, circa 1900)

(26) **heel dag (Engl. all day)**
a. en natuurlik ons zien hom *heel dag* ('and naturally we see him all day'; Catharina Beyers, circa 1918)
b. ik loop maar *heel dag* en de tuin ('I just walk all day in the garden'; B. J. Brümmer, 19/9/1907)
c. want ik word *heel dag* daarna gevraagd ('because I am asked for it all day'; C. J. Van der Merwe, 30/11/1900)

(27) **skaam voel (Engl. to feel ashamed)**
ik *voel* altijd zo *skaam* fir my briewe die ek aan u schrijf ('I always feel so ashamed for my letters which I write to you'; Hester Van der Bijl, 13/2/1906)

(28) **een van die dae (Engl. one of these days)**
a. *een van die dage* gaan ons Dr. Newnham vra ('one of these days we will go and ask Dr. Newnham': Johanna Brümmer, 10/5/1895)
b. *een van die dagen* kom ik weer thuis! ('one of these days I will come home again!'; Wynanda Hoogenhout, 21/9/1905)

(29) **lyn (Engl. line)**
a. net een paar *lynen* om u te vragen ('just a few lines to ask you'; P. J. Le Roux, 18/2/1908)
b. in haas net gou 'n paar *lijntijes* ('in haste just quickly a few lines'; Hendrina Viljoen, 13/5/1919)

(30) **verlang vir (Engl. to long for)**
en kan part dae vreselik *vir* die plaas *verlang* ('and some days I can long for the farm very much'; Rachel I. Steyn, 4/12/1919)

(31) **lyk (Engl. to like)**
a. als u nie *lykt* dat ik van hier een order moet stuur ('if you don't like that I must send an order from here'; Engela Furter, 5/10/1909)

b. Ina sal dit baing *lijk* hier ('Ina will like it here very much'; Wynanda Hoogenhout, 7/2/1908)

(32) **bly (Engl. to stay)**

a. van dat ik in de Pastorie *blijft* ('since I stay in the parsonage'; Maria Cronje, 30/7/1906)

b. soals u weet *bly* ons nou in die Kaap ('as you know we now stay at the Cape'; Susie Baartman, 6/3/1920)

Common in the *Corpus of Cape Dutch Correspondence* is also the use of *meen(en)* ('to mean') for Dutch/Afrikaans *bedoel(en)* and *beteken(en)* (Swart & Prinsloo 1936: 30; DeKlerk & Bosch 1998: 47), and the emphatic reflexive extension *-self* (33), which has been generalized at the Cape under English influence to cases where no emphasis is implied (Ponelis 1979: 83; Raidt 1983: 137; Donaldson 1991: 273).

(33) a. hartelyke groete van Martha en *myself* aan u allen ('many greetings from Martha and myself to you all'; E. Smith, circa 1900)

b. Theron, Watermeyer en *jouself* ni te vergeet ni ('Theron, Watermeyer and yourself not to forget'; N. J. De Wet, June 1895)

Language conflict and language purism: *Moenie jou languages mix nie*

Code-mixing practices as well as interference from English were noted critically in the 1840s by Changuion (1844) in his Dutch grammar, and in the metalinguistic discourse of the late nineteenth century bilingual speech was heavily stigmatized and ridiculed. Mansvelt (*Zuid-Afrikaan* 1/5/1890) described such bilingual practices ironically as *Mixpickles Afrikaans*, and an anonymous article published in the 1890s in *Ons Land* (25/10/1895) discusses these practices at length (switches into English in italics):

> Laat mij beginnen door te zeggen, dat in wat ik nu ga doen het volstrekt niet noodig is voor mij, om op mij *imagination* te gaan *draw* ver *examples*, integendeel dit is maar net om te *pick* en te *choose* uit wat een mens dagelijks hoor in die *conversations* van diegene mit wie hij in aanraking kom … Hoewel gij al dadelijk moet *genotice* hebben, dat ik geen *advocate* ben van zulk een *language*, wil ik nu ronduit zeggen, dat my *object* is om allen, die *guilty* moet *plead* aan 't gebruik van Kaapsch, so *thoroughly disgusted* met zichzelven te maak, dat hulle in de *future* liever hulle mond zal hou, als hulle dan nie een fatsoenlijker *language* kan *employ* … *True*, baing van ons oudere menschen, die met

> moeite 'Di Patriot' kunnen lezen, zullen wat ik hier geschreven heb, glad en al niet verstaan, maar dit neemt hoegenaamd niet weg dat wat ik schrijf waar is. Zoo, en niet anders spreekt een groot deel van onze menschen, vooral het jongere geslacht, de spes patriae.
>
> ['Let me begin by saying that in what I am now going to do, it is absolutely not necessary to *draw* on my *imagination* for *examples*, on the contrary this is just to *pick* and to *choose* from what one hears daily in *conversations* with those with whom one comes into contact … Although you must already have *notice*d that I am not an *advocate* of such a *language*, I will now say openly that it is my aim [*object*] to make all of those who must *plead guilty* to the use of Kaapsch so *thoroughly disgusted* with themselves, that they will rather keep quiet in *future* if they cannot *employ* a more decent *language* … *True*, many of our older people, who can read "The Patriot" only with effort, will not at all understand what I have written here, but this does not alter the fact that what I write is true. A great number of our people, in particular the younger generation, the spes patriae, speak this way and not otherwise.']

The evidence from the *Corpus of Cape Dutch Correspondence* gives some support to the nineteenth century stereotype that code-mixing was most prominently used by the urban-based younger generation (age group 2). Code alternation is particularly prominent in the letters of young women, Johanna Brümmer and Mimmie Laubscher being principal examples (see above); however, it is not absent from letters written by young men (such as Francois Malan, Gideon Goosen and J. F. W. Grosskopf). Age has been identified as an important extralinguistic variable for code-mixing/switching in general. Bilingual speakers are found primarily to 'engage in code-mixing as adolescents; when they have turned into "responsible" adults they keep their languages more apart. Adolescent code-mixing has indeed been reported for many communities' (Muysken 2000: 227).

These playful peer-group based code alternation practices stand in sharp contrast to the anti-English attitudes and polemics which characterized the Afrikaner nationalist agenda. Anti-English propagandistic texts such as C. P. Hoogenhout's popular poem *Fooruitgang* ('Progress'; written in the early 1870s) give voice to perceptions of linguistic conflict and competition which were prominent motifs in the nationalist discourse.

> Engels! Engels! alles Engels! Engels wat jy sien en hoor;
> In ons skole, in ons kerke, word ons moedertaal vermoer.
> …
> Wi hom ni lat anglisere word geskolde en gesmaad.
> Tot in Frystaat en Transfaal al, oweral diselfde kwaad.

'Dis vooruitgang!' roep die skreeuwers, 'dis beskawing wat nou kom!'
Di wat dit ni will gelowe, die is *ouderwets en dom*.

['English! English! All is English! English what you see and hear;
In our schools, in our churches our mother tongue is killed.

...
Those who don't let themselves be anglicized are scolded and abused.
Even into the Free State and Transvaal, it's everywhere the same evil.
"This is progress", the loud-mouths call, "this is education that is coming now!"
Those who don't believe this, they are old-fashioned and stupid.']

The vilification of the English language and culture as *bastardtaal* ('mixed language'), *brabbeltaal* ('jabber language'), or *rooitaal* ('red language'), found expression in numerous texts and poems, and became part of Cape Dutch popular culture. The following extract from a poem, which was published in the *Zuid-Afrikaansche Tijdschrift* in February 1893, reflects the often playful nature of these popular diatribes.

Toen die duivel, altijd besig	When the devil, always busy
Eens een uurtjie ledig had	Once had a spare hour
En hij rustend van zijn streken,	And he, resting from his endeavours,
Peizend in zijn Helstoel zat;	Sat pensive on his hell-chair;
Kwam die plan in zijn gedachte	Came a plan to his mind
Om een nieuwe taal te brouw	To brew a new language
Die voorzeker bij de menschen	Which would definitely among the people
Strijd en twist verwekken zou.	Rouse strife and quarrel.
Want jul weet mos, twist en ruzie	As you know well, quarrel and dispute
Geef dien duivel steeds pleizier	Always give pleasure to the devil
Daarom riep hij tot zijn knechten:	Therefore he called to his servants:
'Breng een groote kookpot hier!	'Bring a big cooking pot here!
En begin die vuur te stoken,	And begin to heat up the fire,
Net zoo vinnig as jul kan!	As quickly as you can!
Want ek gaat een kostje koken	Because I will cook a meal
Voor mijn vriend, den Engelsman!'	For my friend, the Englishman!'

From the 1930s – that is, after the recognition of Afrikaans as one of the official languages of the Union of South Africa – anti-English attitudes proliferated

and a strongly puristic language movement turned its attention not only to Afrikaans lexical and syntactic constructions – which were believed to be the result of English influence (cf. Donaldson 1991: 64ff. on *anglisismejagtery* 'anglicism hunt') – but also advised members of the emerging Afrikaans speech community to avoid the use of English in all official and private conversations. Purist lexical elaboration frequently took place through what has been called *Vernederlandsing* ('Dutchification'; cf. Uys 1983); that is, extensive borrowing from the metropolitan Dutch norm (cf. also Donaldson 1991: 159). *The Handhaaf en Bou!* ('Maintain and Build!') Congress of 1929 issued a strict list of recommendations to members of the just founded *Federasie van Afrikaanse Kultuurverenigingе* ('Federation of Afrikaans Cultural Organizations') for the use of Afrikaans in everyday life (cited in Pienaar 1946: 374):

> Lede sal uitsluitlik Afrikaans besig: (a) As hulle telefoniese oproepe maak … (b) As hulle met *alle* regerings-, prowinsiale, municipale of bankbeamptes te doen het … (c) As 'n Afrikaner-amptenaar met sy kollega in die diens praat. (Gebruik Afrikaanse vakterminologie.) (d) As hulle aan sport deelneem. … (e) In die sosiale omgang. (Hou u taal suiwer.) (f) As hulle vreemdelinge aanspreek … (g) As hulle met dokters, prokureurs en ander professionele persone in gesprek is. (h) As hulle handelsake verrig
>
> ['Members should use only Afrikaans: (a) When they make telephone calls … (b) When they deal with all government, provincial, municipal or bank employees … (c) When an Afrikaans civil servant talks to his colleagues during work hours. (Use Afrikaans specialist terminologies.) (d) When you participate in sport … (e) In social interactions. (Keep your language clean.) (f) When you talk to strangers … (g) When you talk to doctors, attorneys and other professional people. (h) When you conduct business']

English-Afrikaans code-mixing/switching became increasingly associated with coloured speakers, and within the Afrikaner community such practices were seen as a marker of low social status (cf. Barnouw 1934: 29, 39–40; Donaldson 1991: 35). Le Roux (1944: 42), for example, maintained in the 1940s that code-mixing/switching is typical only of the Malay Afrikaans 'dialect'. He cites the following *diensmeide-sinnetjie* ('housemaid's sentence') which locates this variety as a marker of a specific socioeconomic and ethnic group: 'En *besides*, ek moet nog *clean towels* in die *gentlemen* se *room* gaan sit' ('And besides, I still need to put clean towels in the gentlemen's room').

Afrikaans language purism served to consolidate *suiwer* ('pure') Afrikaans as the national language of educated white Afrikaners. The bilingual linguistic practices of other ethnic and/or social groups, on the other hand, were deni-

grated as *gemixte taal* ('mixed language'). This further supported the marginalization of especially coloured and black speakers of Afrikaans, which was in line with their political exclusion under apartheid. McCormick's (1989, 1995, 2002) work on language use of the coloured community of Cape Town's District Six has shown that the District Six dialect of Afrikaans diverges from the standard norm in particular with regard to the lexicon, which contains a great number of regular English loans. Extensive code-mixing/-switching remains a salient feature of the community's language use (cf. also Carstens 1991: 274–275; Ponelis 1993: 115).[2]

While most early language activists favoured a clear separation of languages, others saw individual bilingualism and language contact as an intrinsic part of South African identity. Support of societal and individual bilingualism was expressed at the first and second Afrikaans language congresses which took place in 1896 and 1897. The first bilingual English-Afrikaans dictionary published in 1902/1904 also reflects a liberal attitude to the effects of language contact and includes a great number of assimilated English loans (such as, for example, the colloquial form *auntie* instead of the Dutch-based variant *tante* or *tannie*).

Francois Malan's early diaries are an interesting sociolinguistic document regarding the relationship of English, Dutch and Afrikaans (see above and also Chapter 9 for a discussion of his letters). Francois Malan uses the three languages in strict and systematic alternation; that is, if one diary entry is written in English the next one will be written in either Dutch or Afrikaans. Initially (diary 1892) the alternation is only between Dutch and English, but in 1893 when he was studying in Edinburgh, Afrikaans joined the repertoire and gradually replaced Dutch. After his return to South Africa, his entries are again in Dutch and English. From 1906 he uses only Dutch. In his correspondence Malan expresses a tolerant and liberal attitude to the question of 'anglicisms' and argues that the development of Afrikaans should be non-purist in orientation: English had become too strong a force at the Cape to make its purist exclusion a viable project (for further discussions on Malan's language use see Chapter 9; cf. also the above discussion of Johanna Brümmer's letters).

> Dit is nou eenmaal 'n feit – goed of sleg – dat die Engelse taal 'n vaste grond het in di verbeelding van d'n menigte mense by di Kaap, onder di vind jy 'n menigte egte Afrikaanse families en soos jy self weet vooral onder di meisiis … die mense daar oer sleg te maak, hulle te beskuldig van karakterloosheid of wat oek al, sal niks help ni. Daar net één maniir waarop die kwaad kan verhelp worde en dit is: om di fondamente so breed te lé dat di mense – of dit hulle bedoeling is of ni – onwillekeurig sal werk om di groote werk tot stand

> te bring. As jy 'n stroom siin ankom dan kan jy dit ni verhinder maar as jy kan ver jou so inrig dat jy di water kan opvang of weg ly, so that in plaas van 'n verspoeling en verwoesting te weeg te bring, dit jou tot nut en diins is. So moet ons nog om di Engelse stroom te gebruik en ni teen te staan.
>
> (Francois Malan to W. Viljoen 4/11/1894)

> ['This is now just a fact – good or bad – that the English language has a firm grounding in the imagination of many people at the Cape, among them you find many Afrikaans families and, as you know yourself, especially among the girls … to malign people because of it, to accuse them of lack of character and what have you, will not help. There is just one way in which the evil can be remedied and this is: to make the foundation [of Afrikaans] so broad that the people – whether they intend to or not – will involuntary work to achieve the great work. If you see a stream coming then you cannot stop it, but you can arrange things for you in such a way that you can catch the water or lead it away, so that instead of causing a wash-away and devastation, it is now for your benefit and service. In this way we must still use the English stream and not oppose it.']

However, such liberal attitudes did not agree with the popular and political *Zeitgeist* which – as discussed in Chapter 2 – favoured clear-cut linguistic and cultural boundaries. Linguistic diversity and 'polyglottism' were seen as incommensurate with 'peace and quietness', with political stability and economic prosperity (Alderson 1909: 11). The Anglo-Boer War (1899–1902) fundamentally changed the relationship between Afrikaners and English South Africans, and created bitter conflicts and political schisms. Avoidance of English loans in the emerging standard of Afrikaans now became a question of national pride and Afrikaner ethnic identity.

Notes

1. The category 'language' was only from 1918 included in the census.

2. Only in the 1980s when an inclusive, politically-oriented language movement (the so-called *Alternatiewe Afrikaans Beweging*) emerged, were coloured and black speakers (and their language practices) recognized as full members of the Afrikaans speech community (cf. February 1991: 128ff.; see also Van den Heever 1987).

Chapter 9

Social networks and the diffusion of standard Afrikaans*

> A slow advance in the beginning, followed by rapid and uniformly accelerated progress, followed again by progress that continues to slacken until it finally stops: these are the three ages of … invention … if taken as a guide by the statistician and the sociologist, [they] would save many illusions.
>
> Gabriel Tarde (1903). *The Laws of Imitation*

Networks, modernization and nationalism

The diffusion aspects of standardization were emphasized in Ferguson's (1988) paper *Standardization as a form of language spread*. Linguistic variants belonging to the standard norm increase over time across language functions (registers and communicative domains), and across communicative networks (social groups and geographical spaces). The classic sociological attribute approach which identifies the users of the new national standard language as young and educated individuals (Chapter 7) provides results which are not unlike those described for other standard language histories (cf., for example, Van de Craen & Willemyns 1988). In order to acquire a more detailed and dynamic understanding of how the new standard norms diffused across individuals and social groups, the methodology of social network analysis is useful.

The application of social network analysis to sociolinguistic research was pioneered by Lesley Milroy in her variationist study of Belfast vernacular (Milroy 1980), and in their work on linguistic change and language maintenance Jim and Lesley Milroy (e.g. Milroy & Milroy 1985b) have argued convincingly that the structures of a society's network relations can advance or obstruct the diffusion of linguistic changes. The network model of language change and maintenance can be summarized as follows:

i. Frequency of contact in close-knit community networks, which are characterized by a large number of strong ties, supports the development of highly focused group norms and to promote language maintenance. That

is, the more regularly and intimately individuals interact with one another, the more similar their behaviour will become (norm emergence). In addition, in-group pressures to conform to the norms of the community impede the use and diffusion of new linguistic variants (norm maintenance through sanctions). Close-knit, strong-tie networks can therefore act as a conservative force counteracting the spread of innovations.

ii. Groups (or societies) with infrequently interacting group members, on the other hand, will exhibit more diffuse or 'unfocused' linguistic and social practices. Moreover, since weak-tie networks lack a linguistic norm of their own, the adoption of new linguistic variants or other forms of behaviour by individuals does not violate existing community norms. Weak-tie networks are therefore generally susceptible to innovation and change.

Thus, structural properties of interpersonal networks (such as density; the number of ties in a network) and interactional properties (such as the frequency and intensity of interpersonal relationships) influence the development and maintenance of linguistic and cultural norms. The connectivity of networks is not only a factor in the explanation of behavioural differences and similarities between groups of actors, but network ties are also channels through which resources, influence and knowledge about new modes of behaviour flow or diffuse. Following Granovetter's (1973, 1982) work on the cross-cutting 'strength of weak ties', weak ties are believed to play a central role in diffusion processes: they connect diverse local networks into larger social systems; that is, they serve as 'bridges' between localized, cohesive networks. Although strong ties can also act as bridges, actors generally have fewer strong ties than weak ties and are thus more likely to get new information from a weak tie.

Nevalainen's (2000) application of social network analysis to the Early Modern English period is an example of a macro-historical approach, describing the general changes and characteristics of a society's network structure without attempting to reconstruct individual networks with any degree of quantitative or empirical detail. Briefly, in the Early Modern English period on-going migration to London can be said to have supported an increase in weak ties. As a result, linguistic and other behavioural innovations could diffuse easily and quickly across social groups. This macro-sociohistorical process correlates with the observation that linguistic change proceeded faster in London than in the more rural, close-knit areas of the North of England and East Anglia.

One could argue that at the Cape from the mid-nineteenth century processes of urbanization, immigration, economic diversification and improved

transport technologies supported the development of weak ties within the previously close-knit, localized and largely rural society. These changes in network structures facilitated the diffusion of social and linguistic innovations, and supported the emergence of a linguistic code which could function adequately in these weak-tie contexts. In other words, a general supra-local and supra-social variety was now required for communication at the level of the national community and came to replace the highly localized vernacular norms of mesolectal Cape Dutch in most formal domains (cf. also Santa Ana & Parodi's 1998 model of nested speech communities discussed in Chapter 3). One could also hypothesize – still at the macro-level – that while urbanization and social mobility created weak multi-tie networks which supported the speedy diffusion of new variants, Afrikaner nationalism established the conditions for linguistic and cultural norm development by creating a special kind of symbolic social cohesion.

Anderson (1991) has described nations as symbolic or 'imagined communities' whose members, although not necessarily directly linked, develop a strong feeling of belonging to an imagined, homogenous socio-cultural whole. Improved technologies of print reproduction facilitated the distribution of texts and images which established the nation as a point of reference and identity across large geographical areas, as well as across diverse social groups (cf. also Delanty & O'Mahony 2002: 18 on nationalism as a 'form of knowledge by which members of society come to "know" their society'). Nationalist discourse thus transforms the weak ties which characterize modern societies, and which impede the establishment of new norms, into symbolically strong ties, facilitating both the diffusion of new forms of behaviour (and thus linguistic and cultural change), as well as the emergence and focusing of new national traditions or linguistic standards (i.e. norm emergence through symbolic or 'imagined' social cohesion).

The network model can also be adopted as a descriptive and explanatory concept for the analysis of the early dialect writing tradition. Network theory assumes that innovators are marginal network members. They are therefore not subjected to the norm-enforcing mechanisms of close-knit social groups and are free to introduce new linguistic variants and linguistic practices without risking community sanctions. The early writers of 'Afrikaans', who published their work in pamphlets as well as the periodical press, were clearly marginal members of the Cape speech community: De Lima, Boniface and Abu Bakr Effendi were first-generation immigrants, and both Meurant and Cooper were from English-speaking backgrounds (Chapter 2; also Schoeman 1998: 226–227). First-generation immigrants were also among the leading

members of the GRA (e.g. C. P. Hoogenhout and Arnoldus Pannevis who both arrived at the Cape from the Netherlands as young adults).

Reconstructing historical social networks

Network theorists typically distinguish between 'ego network analysis' and 'whole network analysis' (Erickson 1997:150–151; Wetherell 1998:127–128). The whole network approach aims at describing all relations or ties between actors of a given group or sub-group (the guiding question is: 'who knows whom?'). The connections between individuals are represented in the form of a square matrix or sociogram (Figure 9.1). The methods and concepts of graph theory are then used to tease out prominent structural patterns in the network. Whole network analysis has been used in sociolinguistics, for example, by Labov in his work on Harlem gangs (1972) and in Labov and Harris's (1986) study of language variation in Philadelphia. More recently Paolillo (1999) used whole network analysis in his study of internet communication and Eckert (2000) in her study of adolescent speech in a Midwestern high school. Analysts of ego networks, on the other hand, take the individual as the unit of analysis and describe in detail the ties reported by each individual and the range of their connectivity (the guiding question is: 'who does X know?'). In Milroy's (1980, 1987) Belfast study the range of an individual's network was interpreted as part of his or her social profile (alongside the familiar categories of social class, age and gender), and was then correlated with the individual's language use. Similarly Milroy and Li Wei (1995) established on the basis of the proportion of Chinese ties in personal networks a so-called 'ethnic index score' for each individual in their study of the Tyneside Chinese community, and correlated the ethnic network score with the individual's patterns of language use and code-switching practices.

All writers included in the final multivariate analysis (discussed in Chapter 7) were also included in the network analysis. As described in detail in Chapter 7 the varieties used by these writers are: acrolectal Cape Dutch, the emerging standard of Afrikaans and varieties of mesolectal Cape Dutch. The description of the network relationships between writers follows Bax's (2000) proposal to distinguish clearly between network patterns or structures (functional aspects of the network, i.e. existence or absence of relationships between actors), and attitudinal factors (emotional aspects of the network, i.e. the intensity and meaningfulness of these relationships). Although the fragmentary nature of the archival record is a general problem for historical network analy-

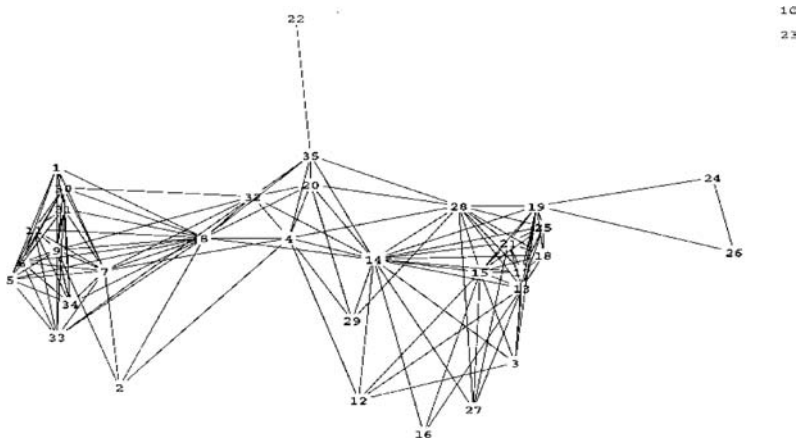


Figure 9.1 Sociogram of the functional relations among 35 speakers/writers (Krackplot 2.0; multidimensional scaling)

sis (as well as for historical sociology in general), and there is always the danger of uncovering incomplete information, the specification of the functional aspects of relationships (i.e. presence or absence of ties) is often possible on the basis of careful archival work (cf. Wetherell 1998 for a discussion of historical social network analysis). In this case, birth and death notices, information from the voters' lists, marriage registers, estate records, commercial trade directories and membership lists for political/religious organizations provided sufficient information to allow for the description of the overall network structure with regard to family connections, residential patterns, and political and organizational (e.g. religious) alliances. Additional information about patterns of interaction was provided in the letters themselves.

The connections between individuals can be represented in the form of a network sociogram (Figure 9.1). The numbers stand for individuals, the lines represent their functional ties; they do not yet contain any information about the attitudinal and emotional aspects of the relationships (the program Krackplot 2.0 was used for the creation of the network graph; cf. Krackhardt, Lundberg, & O'Rourke 1993).

The visual impression is one of two groups situated towards the right and the left of the diagram. These groups, which appear to be quite densely connected, are linked indirectly through several actors who occupy the middle field in the sociogram. In addition there are a number of peripheral network members, situated towards the edges of the diagram. Two actors (10 and 23) had no ties with anyone in the network.

Table 9.1 Cliques in the network. *Peripheral sets and individuals:* church circles (13, 3; ties of organizational affiliation), community of Albertinia (26, 24; ties of neighbourhood), individuals (2, 16, 22, 27). Numbers in bold indicate individuals with double clique membership.

network members	functional description of the cliques
A: 1, 5, 6, **7**, **8**, 9, 11, 30, 31, 33, 34	Hoogenhout: ties of kinship and neighbourhood
B: 12, **14**, 15, 17, 18, 19, 21, 25, **28**	Malan: ties of kinship and neighbourhood
C: 4, **7**, **8**, **14**, 20, **28**, 32, 35	Afrikaner nationalists: ties of political/organizational association

Larger social networks can be described structurally as amalgamations of densely connected sub-groups or clusters of varying sizes. Social network analysts have introduced a variety of methods to identify groups within a given network structure, and have thus formalized our intuitive understanding of social groups in mathematical terms. The graph theoretical definition of a clique was used for the analysis of the network relations in Figure 9.1: a clique is a maximal complete sub-graph, that is, any number of actors who have all possible ties present between themselves (Wasserman & Faust 1994: 254ff.). The analysis was conducted using the network software UCINET IV (Borgatti, Everett & Freeman 1996), and three densely connected, partially overlapping cliques were found to exist in the diagram (see Figure 9.2 and Table 9.1):

a. two groups linked by functional ties of neighbourhood and marriage/kinship. These two groups are indicated here by the name of the dominant family within the network (Hoogenhout and Malan respectively); and
b. the organizational and political network connections of Afrikaner nationalists.

In addition, we can identify a periphery of individuals who are only loosely connected to these cliques. The periphery includes: the community of the small town of Albertina (actors 26 and 24), members of the politically and linguistically conservative church circles (actors 12 and 3) and several individuals who, judging from their social profile, were deeply embedded in rural and localized networks (actors 2, 16, 22, 27 and 29; on the concept of a network periphery cf. Everett & Borgatti 2000; also Eckert 2000: 205–210).

Since tie-strength (that is, the interactional and emotional meaning of a tie) has been shown to be an important variable for the explanation of norm formation and linguistic change, an attempt was made to specify the intensity of the functional ties represented in Figures 9.1 and 9.2. The classification of the emotional aspects of relationships is difficult not only for historical studies

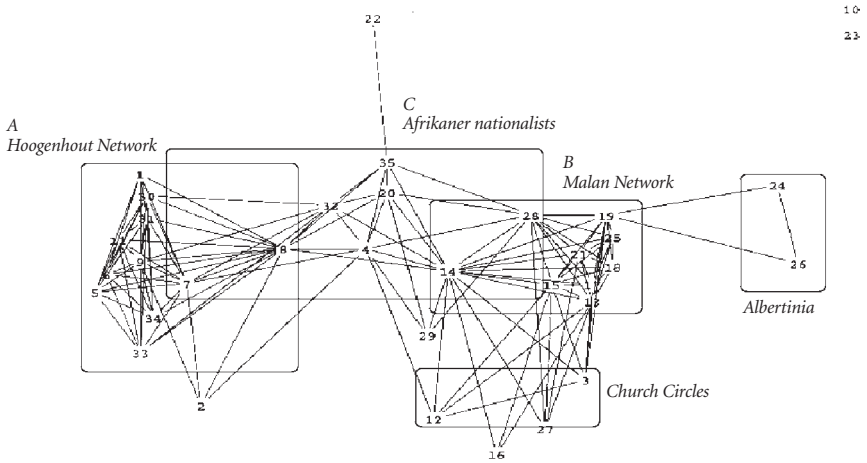


Figure 9.2 Cliques in the network

as feelings in general resist mathematical quantification. Consequently, there is a strong subjective and ethnographic dimension to the experience of relationships and 'people may have different ideas of what constitutes friendship' (Bax 2000: 283). The reconstruction of the interactional content of ties is particularly difficult for historical network analysis where one relies on fragmentary evidence, and often lacks the relevant ethnographic information which infuses ties with meanings of intimacy and reciprocity (Fitzmaurice 2000a/b).

In the case of the *Corpus of Cape Dutch Correspondence*, substantial information and extensive comments about personal relationships existed in the letters themselves. This information allowed a broad assessment of two salient aspects of these interpersonal ties, namely, mutual confiding (as an indicator of intimacy), and frequency of interaction and reciprocity (cf. Granovetter 1973). Weak ties can be conceptualized as infrequently used, non-intimate connections, while strong ties are intimate, frequently used ties located in reciprocal exchange networks.[1] The network structure of three sub-graphs can be described as follows:

i. The Hoogenhout network, that is clique A, is geographically homogenous. It is characterized by numerous strong ties and a clear sense of regional and communal identity. Kinship relations within the extended Hoogenhout family were regular and intimate, as were their relations with their neighbours and friends. Although there was geographical and social mobility, close relationships were maintained by those who (for rea-

sons of work/marriage) had left the original neighbourhood network near Wellington in the Western Cape. At the same time, this mobility supported the development of weak ties outside of the closely-knit local network and exposed individuals to innovations from different directions.

ii. The Malan network is particularistic rather than embracing. Most members of the network were deeply rooted in their respective local communities (Paarl/Stellenbosch, Western Cape; Lady Grey, Eastern Cape) and had only few (weak) ties outside of these closely-knit, strongly localized networks (speech locales in the terminology of Santa Ana & Parodi 1998). The dominance of strong ties supported the maintenance of non-standard local vernacular variants and norms (i.e. mesolectal Cape Dutch Vernacular). Unlike in the Hoogenhout network, socially and geographically mobile network members of the Malan network did not maintain close and regular contact with the rural, local community from which they originated.

iii. The ties of the Afrikaner nationalist network are mainly weak (infrequent, non-intimate) and limited to the organizational structures of Afrikaner politics.

In terms of linguistic innovativeness – that is, their time of adoption of the new standard of Afrikaans – the three groups differ strongly from one another.[2] The Afrikaner nationalists were the first to use the new standard norm in their private correspondence. Members of the Hoogenhout network followed during the first decade of the twentieth century; members of the Malan network and the periphery only adopted the new standard after 1920 when it had been officially recognized by the cultural and political authorities. Within the Malan network, letters written in non-standard, mesolectal Cape Dutch Vernacular can still be found in the 1930s.

The different groups can be described with reference to the adopter categories introduced in Rogers' standard textbook *Diffusion of Innovation* (1995: 257–279): early adopters, early majority and late majority (Table 9.2). These categories reflect the observation that individuals in a social system do not adopt an innovation simultaneously but in an 'over-time sequence' (ibid.: 252; cf. also the discussion of the S-curve model in Chapter 6). The authors of the dialect literature (from the 1830s onwards) and the members of the first language society (mid-1870s) can be described as linguistic innovators.

The language spread process summarized in Table 9.2 describes not only diffusion across groups of individuals, but also across language functions; that is, from the domain of public texts (dialect literature, newspaper articles and political pamphlets) to the use of Afrikaans in private correspondence.

Table 9.2 Adopter categories and the diffusion of Afrikaans

social networks	adopter category	adoption time	nature of ties
dialect literature/GRA	innovators	from 1830s	mostly weak ties
Afrikaner nationalists	early adopters	late 1880s/early 1890s	mostly weak ties
Hoogenhout network	early majority	1900–1910	combination of strong and weak ties
Malan network and periphery	late majority	1920s/1930s	mostly strong ties; few weak ties

The early adopters, who made use of Afrikaans in their private correspondence from the late 1880s, were actively involved in the Afrikaner nationalist movement and its organizational structures. Only after 1900 did the new norm enter other, more localized networks; probably through actors who functioned as bridges or links between the Afrikaner nationalist circles and their respective local friendship and kinship networks. Examples of such bridges in the sociogram (Figure 9.2) can be found in the overlapping areas between groups, that is, actors 7 and 8, on the one hand, and actors 14 and 28, on the other hand.

The successful adoption of Afrikaans as the new standard norm within the Hoogenhout network can be explained as a result of the network's social and geographical homogeneity and its general strong-tie identity. This supported behavioural convergence once the innovation had entered the network. Characteristic of the Hoogenhout network is also a mix of strong and weak ties; this supported exposure to innovation as well as norm development.[3] In the Malan network (and the periphery), on the other hand, the existence of strong local community networks supported the maintenance of mesolectal Cape Dutch Vernacular. The lack of weak ties which characterized the Malan network also meant limited exposure to innovations. Maintenance of Dutch as the standard variety of choice within the network was further supported by the close association of the network with members of the linguistically conservative church circles. This interpretation is in line with current network and diffusion theory (cf. Roger 1995) as well as with the Milroys' network-based model of linguistic change and maintenance (Milroy & Milroy 1985b, 1992).

The strong-tie/weak-tie dichotomy is, however, less useful when describing the network relationships of the most innovative group: the Afrikaner nationalists. How should one explain the emergence of Afrikaans as a communicative norm in Afrikaner nationalist circles? As noted above, these nationalist networks were characterized by weak ties of restricted and non-intimate interaction, and, according to traditional network theory, such weak-tie networks impede the development of focused group or community norms. In order to

solve this puzzle it is necessary to side-step the conventional dichotomy of strong- and weak-tie networks which has dominated much network theorizing, and to specify more clearly the nature of the relationships which characterize the network structures of social movements.

Ties of coalition and cooperation: The Afrikaner nationalists

Fitzmaurice's (2000) notion of coalitions is useful for the description and analysis of the Afrikaner nationalist network. Coalitions are social groups formed 'in order to achieve particular goals or to pursue a particular common agenda' (ibid.: 266). The ties between members of such coalitions are 'targeted', that is, they are strategic 'connections formed as motivated by a limited set of social wants' (ibid.: 274); they are, 'temporary alliance[s] of distinct parties for a limited purpose' (Boissevain 1974: 171). Fitzmaurice (2000) describes (in the context of the standardization history of English) the relationship between Addison and Steele as a paradigmatic example of a coalition:

> The alliance contracted between Addison and Steele in the *Spectator* enterprise is a good example of such a connection; they join forces for the specific purpose of co-authoring a daily periodical. Although their friendship (a separate matter) endures after the lifetime of the periodical, their *Spectator* alliance is dissolved in 1714, after the publication of the collected edition … [T]he ties contracted in the formation of a coalition or an alliance are not straightforwardly weak in the way in which 'weak' ties have usually been understood in social network analysis. … *the kinds of ties involved are of a highly restricted kind; it is clear that they are consciously contracted in order to create the best possible conditions for good outcomes, whether economic, social or personal.*
>
> (ibid. 275, 276; my emphasis)

Targeted ties of coalition – although not strong in the conventional kinship-and friendship-based sense – have been shown to lead to language behaviour that resembles that of close-knit networks; that is, they support the development of group-based norms through processes of linguistic and cultural focusing. Coalition-based networks, characterized by the weak, yet targeted ties of mutual endeavour, rather than the strong, intimate ties of mutual confiding, define an important network type in standardization histories, most of which have been shaped to varying degrees by conscious linguistic and political activism.

Fitzmaurice's notion of coalitions can be linked to Eckert's (2000; cf. also Milroy & Gordon 2003: 118–119) discussion of social networks as 'communities of practice', that is, voluntary groups or associations in which people cre-

ate and maintain social meaning, practices and identities – beliefs, languages, shared patterns of behaviour – through their joint participation in purposeful, strategic activities and enterprises. Eckert (2000: 35) captures the special nature of these interaction networks in the context of her study of adolescent language use in an American high school with the example of a garage band:

> A community of practice is an aggregate of people who come together around some enterprise … That endeavor develops a life of its own as local practices develop around it, transforming the enterprise, the activity, and knowledge. The practices that emerge as a rock 'n roll band works together include such things as the choice of songs the band plays, the kind of music, a view of its place in the wider landscape of music, an attitude towards other kinds of music, the band's 'sound' and the contribution of each instrument to that sound, ways of dressing, ways of getting and choosing gigs, ways of developing new songs and rehearsing, ways of behaving and talking in encounters with band members and when representing the band.

Again, the nature of the interpersonal connections which characterize communities of practice cannot be described adequately in terms of the conventional strong/weak tie dichotomy. The ties which connect individuals in their activities are neither strong nor weak, rather they resemble the targeted, goal-directed ties as discussed by Fitzmaurice.

Both coalitions and communities of practice are different from networks in the sense that they are 'about' something. They are purposeful alliances of individuals in which linguistic choices are made and enacted, attitudes are developed and maintained, and ways of in-group and out-group representation are created through strategic social participation. In the history of Afrikaans two communities of practice (or coalitions) can be identified: the writers of the dialect literature (who established Afrikaans as a typological norm and way of speaking; see Chapter 2), and the various formal and informal networks of Afrikaner nationalism (which established Afrikaans as the preferred linguistic norm in a wide range of domains, including private correspondence). That the purposeful use of Afrikaans for in-group correspondence was a common practice among Afrikaner nationalists is indicated by a comment made by D. C. De Wet in a letter to E. Esselen (24/11/1884):

> Ik hoop jy neem my nie kwalyk dat ik jou in die Patriot taal, ons taal, geschreywe het nie, ik zien dat jy en Hofmeyr zoo correspondeer, en ik schryf ook graag zoo, gewoonlyk schryf ik aan Bondmannen in die taal.
>
> ['I hope you don't mind that I have written to you in the Patriot language, our language, I see that you and Hofmeyr correspond in this way, and I also enjoy

> writing in this fashion, normally I write to members of the Bond [*Afrikaner Bond*] in *die taal* [i.e. the Afrikaans language].']

Not only did the Genootskap decide to use Afrikaans as the language of their proceedings (see Nienaber 1974), but Viljoen (1891: 3–4) also reports that Afrikaans was used in meetings of the *Debats- and Christelijke Jongenvereeniging* and among students. J. W. F. Grosskopf describes in a letter (4/11/1908) to Imker Hoogenhout the use of Afrikaans as a kind of playful and ritualized in-group code by South African students in Edinburgh. Grosskopf referred to this way of speaking as *Edinburgse Afrikaner-slang*.

> Dit is net Stellenbosch of di Kaap, met diselfde bleddie Textbook en eksamen sisteem, en met onbeperkte billiart, whiskey en poeier. Hulle praat wel deur di bank Afrikaans daar. Maar, God bewaar ons, hoe? En ook net om van die Skotsmanne nie verstaan te word nie.
>
> ['This is just Stellenbosch or the Cape with the same bloody textbook and exam system, and with unlimited billiard, whiskey and powder. All of them speak Afrikaans there. But, God help us, how? And also only as not to be understood by the Scotsmen.']

The use of Afrikaans thus developed into an overt marker for the expression of national identity and in-group membership within these network structures. The writer and journalist Jan De Waal (1932: 20–21) described in his autobiography in anecdotal style how changes in language use and choice were diffused through cultural organizations and communities of practice.

> Ongeveer in 1890, onderwyl ek student aan die Normaalkollege was … het ek aangesluit by die debatsvereeniging wat daaraan verbind was, en gevind dat al die studente, waar hulle nie hul toesprake in Engels afsteek nie, dit in Hoog-Hollands probeer doen het. Hier het my ingebore opstandige gees hom laat geld. Alhoewel ik self maar in die algemeen 'n slaaf van die mode was, het dit my daarem as eenmaal te erg voorgekom om onder mede Afrikaner-studente by voorkeur die taal van Holland te gebruik; so het ek sommer uit die staanspoor uit weggeval in Afrikaans, in sy ongekunstelde vorm, al sou ek my met minder moeite in Hoog-Hollands kon uitgedruk het as die meeste van die ander. 'n Paar vrinde het my effens skeef aangekyk, 'n paar het smalend geglimlag en een het 'n beledigende aanmerking gemaak, maar ander het dit goedsmoeds geduld, *en weldra het ek volgelinge gekry.*
>
> ['In about 1890, while a student at the Normaalkollege, I joined the debating society that was attached to it and found that all the students, if they didn't give their speeches in English, would try to do so in High Dutch. Here my inborn rebellious spirit asserted itself. Although I myself was in general a slave

> of fashion, it seemed to me to be going to far to prefer using the language of the Netherlands among Afrikaner students; so right from the start I just fired away in Afrikaans, in its natural form, even though I could have expressed myself with less effort in High Dutch than most of the others. Some friends looked at me a little askance, some smiled contemptuously, and one made an insulting comment, but others accepted it cheerfully, *and soon I had followers*'; my emphasis.]

The self-conscious adoption of such semiotically-marked linguistic practices was not limited to within-group communication but also affected the linguistic choices of writers in other domains, and with interlocutors who did not belong the Afrikaner nationalist networks. Francois Malan's letters to his fiancée Johanna Brümmer (Figure 9.3; cf. also Figure 8.1 for Johanna Brümmer's letters to Francois Malan) show a clear pattern of code-choice: that is, increasing use of Afrikaans from 1893, peaking in 1894; during the same time period Malan wrote few letters in Dutch and English. From 1895 his use of Dutch and English increased again, but English disappeared from his linguistic repertoire with the beginning of the Anglo-Boer War. After 1897 there are no longer entire letters written in Afrikaans. Malan is number 14 in the sociogram (Figure 9.2).

The exploratory changes in Malan's private language use can be interpreted as a reflex of his political alliances or coalitions within the broader context of Afrikaner nationalism and South African politics. As a student in Edinburgh (cf. the above mentioned comment by Grosskopf) Malan identified strongly with the more radical wing of Afrikaner nationalism and followed other young nationalists in his early and enthusiastic adoption of the emerging standard of Afrikaans for his private correspondence (cf. also Chapter 8 on his diary lan-

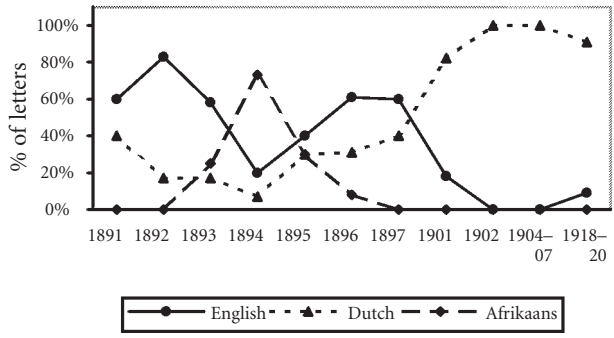


Figure 9.3 Code choices in Francois Malan's letters to Johanna Brümmer, 1891–1920. $N = 631$; in percentages.

guage use). In the late 1890s when working as an editor for the Dutch-language paper *Ons Land* Malan appears to have sided with the language-politically more conservative groups within the Afrikaner nationalist movement, represented by the wealthy wine farmers of the Boland region. His use of Dutch increased gradually and after the Anglo-Boer War Dutch became his main language of personal communication. Malan expressed his opposition to English hegemony and his Afrikaner identity not only linguistically (through use of Afrikaans and Dutch respectively), but also through symbolic clothing practices in the Cape Parliament (Malan was appointed as cabinet minister in 1908). In June 1909 the satirical column *Straatpraatjes* commented ironically on his conspicuous clothing style:

> Mr. Malan lijk meer boerachtiger. Stoffel se die wijnboere het hom goet gegee di anner dag, en hom gese dat hij hom te veel ophou moet Dr. Jamieson en die aner Engelse. Nau het hij weer sij das verkeert vas gemaak en een weie broek angetrek, so dat hij weer soes e boer kan lijk. (Adhikari 1996: 27)
>
> ['Mr. Malan looks more farmer-like. Stoffel says that the wine farmers gave him trouble the other day and said to him that he spends too much time with Dr. Jamieson and other Englishmen. Now he has again fastened his tie the wrong way and put on wide trousers so that he can again look like a farmer.']

Malan's wide-legged trousers in particular were considered 'rustic' and 'Boer-like' in cosmopolitan English-speaking Cape Town, and clearly distinguished him from urban 'Englishness' which was represented symbolically by the *harrebol-kuiljies* of the other parliamentary members (*hardebolkeiltjie*, 'bowler hat'; cf. Adhikari 1996: 36; on the symbolic reflection of urban and rural identities though clothing practices in the Cape Colony see also Ross 1999a: 88).

The rise of Afrikaans: *Daar buite in die bloue lug*[4]

A period of alternation between the old and new standard norm, as indicated in Malan's code choices, is generally observable in the linguistic behaviour of the early adopters and the early majority, all of whom were found to go through a time in which they used both standard Dutch and Afrikaans in their private correspondence. 'Trialability' of an innovation (that is, 'the degree to which an innovation may be experimented with on a limited basis'; Rogers 1995: 243–244) is important for a successful diffusion process. The code choices of Wynanda Hoogenhout (born 1888) between 1905 and 1921 also illustrate a transition period: both norms were employed to varying degrees and with

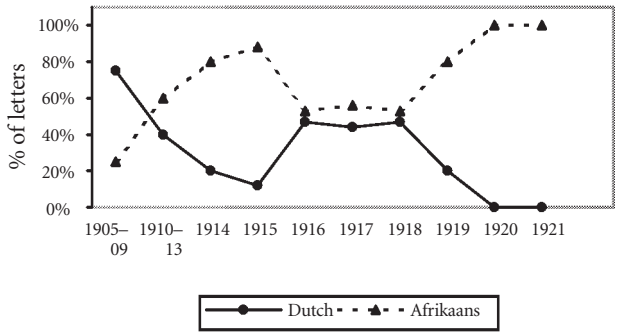


Figure 9.4 Code choices in Wynanda Hoogenhout's letters, 1905–1921. $N = 155$, in percentages.

different interlocutors (generally Afrikaans to peers and Dutch to the parent generation; Figure 9.4).

There is no evidence that such an extended transition period existed for the late majority. Here the change from Dutch to Afrikaans appears to have been more abrupt, possibly as a result of the fact that a 'critical mass' of users had already been reached and the new standard now abounded in everyday life, leading to increased exposure as well as creating opportunities for its use (Rogers 1995: 313–330).

Afrikaner nationalism gained large-scale popular support across all social groups in the years after 1910 (Hofmeyr 1987), and a new strong language movement vigorously propagated the new standard norm in several newly founded Afrikaans-medium periodicals (1910 *Die Brandwag,* 1914 *Ons Moedertaal,* 1916 *Die Huisgenoot*). As a result, Afrikaans became an accepted code choice in the periodical press as well as in non-fictional texts, literary texts and poetry, and the publication of books in Afrikaans increased dramatically (Figure 9.5).

At the same time, language political changes took place: in 1914 Afrikaans was introduced as medium of instruction in primary schools, in 1915/17 the first spelling rules were published by the *South African Academy*, in 1918 Afrikaans was introduced in the civil service and at some universities, and in 1919 it was officially adopted by the church. Students in particular were active in the language movement. Their activities included, for example, the organization of several *taalfeste* ('language festivals') which contributed to the ritual elevation of Afrikaans as a symbol of Afrikaner *volkseenheid* ('national unity'; cf. Deumert 2003c). Those involved in the *taalstryd* – the language

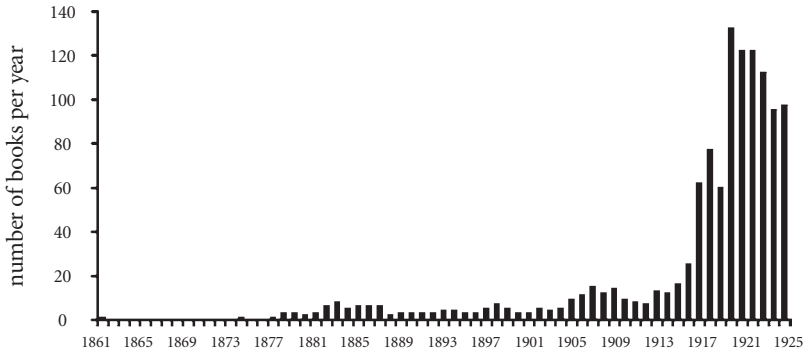


Figure 9.5 Publication of books in Afrikaans, 1861–1925. Number of books per year (based on Nienaber 1948).

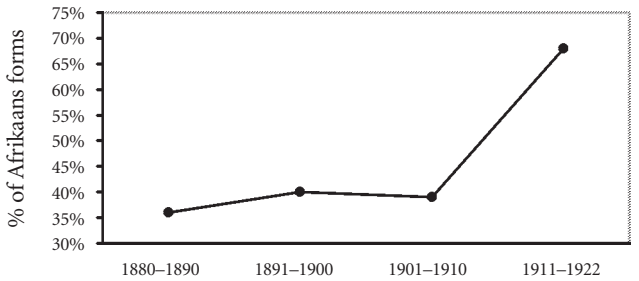


Figure 9.6 Percentage of Afrikaans forms in the *Corpus of Cape Dutch Correspondence* across time

struggle – had little doubt that the movement was now irreversible and had developed its own momentum. W. M. R. Malherbe, the editor of *Die Brandwag*, wrote in 1916 to C. P. Hoogenhout, one of the founders of the first Afrikaans language society:

> Wat moet dit uw hart tog bly maak om te sien, hoe wonderlik snell die beweging voor Afrikaans vorder in die laatste tyd. Dis byna ongelooflik, dit maak 'n mens byna bang, dat dit nou met een weer al te vinnig gaat. (MSC 1-37-2)
>
> ['How it must please you to see how incredibly fast the movement for Afrikaans is progressing in the last little while. This is almost unbelievable, it makes one almost afraid that everything is suddenly happening so quickly.']

Malherbe's observation is supported by the real-time diffusion of Afrikaans variants across the *Corpus of Cape Dutch Correspondence*. When grouping the corpus texts into four ten-year periods, the percentage of Afrikaans forms (all

variables combined) shows a step increase only after 1910, i.e. once the new standard had been adopted by the early majority.

Summary

The discussion in this chapter complements the sociolinguistic analysis of Chapter 7, which described the social profile (age, gender, class) of the early users of the new standard norm. Patterns of association and interaction played an important role at the Cape and influenced the speed of the diffusion process. The results agree with current network theory: late or delayed adoption of Afrikaans as a new norm was common for those who were deeply integrated into their respective local communities and had no or few ties to the innovative centre of Afrikaner nationalism, where the new linguistic practices crystallized during the late 1880s and early 1890s as a marker of national identity. On the other hand, a combination of strong and weak ties, as described for the Hoogenhout network, supported both access to the innovation as well as the gradual diffusion and establishment of Afrikaans as the new norm within the network. The ties within the innovative network of Afrikaner nationalism are best described in terms of mutual endeavour; such ties are characteristic of strategic coalitions and communities of practice, including language movements.

For many of the bi-dialectal speakers (discussed in Chapter 7; cf. Figure 7.13), the change from Dutch to Afrikaans was a case of a clear-cut language shift, that is, from a predominantly Dutch system to a predominantly Afrikaans system (cf. Figure 7.13). The adoption process was different for mesolectal speakers whose vernacular norms were variable and conformed neither to the norms of Dutch nor to those of Afrikaans. Their 'route to the written standard variety' (Elspaß 2002:48) is less well documented and probably involved various processes of partial accommodation and variety shifting. Letters written in mesolectal Cape Dutch Vernacular can be found as late as the 1930s and in 1932 Hester Van Huyssteen (a member of the Malan network) apologized in a letter to her step-daughter Katie Malan (née Van Huyssteen): 'Lees tog maar al myn fouten reg hoor!' ('But just read all of my mistakes right, OK!'; 6/8/1932, MSB 853-5-1) – thus acknowledging her lack of proficiency in the accepted written standard.

Notes

* An earlier version of this chapter was published under the title 'Standardization and Social Networks: The Emergence and Diffusion of Standard Afrikaans' in *Standardization. Studies from the Germanic Languages*, N. McLelland and A. Linn (Eds., 2002).

1. Granovetter defined tie strength as follows (1973:1361): '[T]he strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie.' Tie strength is a continuous variable; however, for purposes of analysis Granovetter (and most network analysts) treat it as a discrete variable.

2. Rogers (1995:261) defined innovativeness as follows: 'the degree to which an individual or other unit of adoption is relatively earlier in adopting new ideas than other members of a social system'.

3. The Hoogenhout kinship and neighbourhood network was directly linked to the first language society, the *Genootskap van Regte Afrikaners*, through C. P. Hoogenhout (born 1843) who was one of the founding members of the society. However, it seems that this connection had little effect on informal linguistic practices within the network. Although Hoogenhout was an active member and a well-known writer of Afrikaans poems and stories, he did not use Afrikaans in his private correspondence and diaries. His children wrote to him mostly in Dutch (and he pedantically corrected orthographic and grammatical mistakes in their letters).

4. Langenhoven described in 1914 the spread of Afrikaans as 'n snellstygende boog … nie net vinnig opgande nie, maar opgande na 'n vinnige vermeerderende rede' ('a steeply mounting curve, not only rising sharply, but doing so at a rapidly increasing ratio', i.e. a hyperbola), which would soon reach a point 'daar buite in die bloue lug van Bloemfontein' ('out there in the blue sky of Bloemfontein'; *Versamelde Werke*, vol. 10).

Epilogue

Language standardization and language change

Historiography is a difficult genre: it is based on evidence which is always incomplete (and often inadequate), and only through conjecture and interpretation can the historian link the scattered pieces of evidence into a coherent narrative. That this is true also for language history was emphasized by Lass (1997: 4–5), whose notion of myth largely agrees with the way the term narrative is used here, that is, 'a myth in the widest sense is a story or image that structures some epistemic field'. Lass continues:

> The histories of languages (as objects available to or made by linguists) are, like all histories, myths. We do have documents for portions of many of our histories; but even these are subject, like scripture, to exegesis: we don't know what they mean (the less, the older they are). We do however tell (and believe) stories about them, not just the documents but the languages they supposedly reflect. (ibid.)

Thus, rather than a solution or definite answer to the question *wie ist es nun eigentlich gewesen?* ('what was it like exactly?', to rephrase Leopold von Ranke's well-known statement), historical knowledge is best seen as 'a balance of probabilities' (Stanford 1994: 112), probabilities which can amount to near certainty – yet a certainty with no other base than the interpretation of an always incomplete record. To emphasize the tentative nature of historical knowledge is not, however, to open the doors to a post-modern 'anything goes'-relativism, since our interpretations and conjectures can always be measured against the criteria of empirical responsibility, coherence and plausibility (Lass 1997). As a case study in historical sociolinguistics this monograph partakes in all of these epistemological problems and the reconstruction of the history of early standard Afrikaans is thus best presented in the form of hypotheses.

Hypothesis I: An argument for slow and gradual change

The patterns of variation described for the *Corpus of Cape Dutch Correspondence* indicate that for the acrolectal and (upper) mesolectal varieties the process of language change was slow and gradual, continuing, in many cases (as shown in detail in Chapters 5 and 6), phonological, grammatical and lexical patterns described for the mid- to late eighteenth century. This challenges the conventional dating of the emergence of Afrikaans to c. 1775 and suggests that language change at the Cape proceeded at various speeds in the different social, ethnic and geographical varieties.

The acrolectal and (upper) mesolectal varieties described in this study developed in a complex process which involved evolutive language change, endogenous innovation and contact-induced restructuring. Since the Cape colonial society was not strictly segregated until the late nineteenth century, contact between speakers of different varieties of Cape Dutch Vernacular (some of which were strongly marked by substratum influence and the universals of untutored second language acquisition) was an important mechanism of language change, leading to the transfer of substratum features and convergent hybridization.

Hypothesis II: Against diglossia

A central argument of this study is that language change at the Cape did not result in the development of a rigid diglossic structure with Dutch (and its acrolectal varieties) as the High variety and the basilectal Cape Dutch dialect as the Low variety, each being used in discrete and functionally separate domains. Rather, linguistic stratification (social, geographic and ethnic) was complex and a wide range of varieties and ways of speaking developed at the Cape.

Rejecting the conventional interpretation of the Cape speech community as classically diglossic also means that the written data should not be seen as reflecting a different type of language (the High variety), which never occurred in spoken interaction, but rather as reflecting a different style (oriented towards the formal side of the stylistic continuum). In other words, while under diglossic conditions the relationship between spoken and written language is discontinuous, continuity along an informal-formal continuum is characteristic of social dialects. The interpretation of the Cape Dutch speech community as dialectal rather than diglossic is supported by the patterns of structured heterogeneity which were found to exist in the *Corpus of Cape Dutch Cor-*

respondence – linguistic patterns which are typical for standard-with-dialects scenarios (see Chapters 5, 6 and 7).

To emphasize continuity and to reject the traditional diglossic interpretation is reminiscent of debates in Romance historical linguistics regarding the diachronic delineation of Latin and Proto-Romance (see Chapter 2). According to the traditional model, linguists have assumed a diglossic situation with post-classical Latin fulfilling High functions in society and Proto-Romance fulfilling Low functions. This theory was challenged by Wright (1982, 1991) who proposed what Lodge (1993) has called the 'one-norm-theory'. The parallels to the debates in Afrikaans historical linguistics are striking:

> For Wright, Latin and Proto-Romance constituted not two discrete varieties with little in between, but, for a good deal longer than was previously thought, a single stylistic continuum embracing the language of formal educated usage at one end and uneducated informal speech at the other (the 'one-norm-theory'): mutual intelligibility between speakers of this amalgam of varieties continued to be possible (according to Wright) for much longer than is allowed for in the 'two-norm-theory' through normal processes of linguistic accommodation … and through the application of 'automatic conversion rules' …, just as it can be maintained across the diverse varieties of English currently in use in various parts of the world. (Lodge 1993:90)

The hypothesis that Afrikaans existed as a relatively uniform spoken vernacular from about 1800 but that diglossia prevented the structures of the spoken language from surfacing in the written record is epistemologically suspect. By classifying texts which conformed neither to the emerging norms of Afrikaans nor to the norms of continental Dutch as 'approximate Dutch' (i.e. the theoretically uninteresting *ad hoc* result of linguistic insecurity and schizoglossia) Afrikaans historical linguists have engaged in an unconfessed historical purism, and have, as a result, neglected the detailed analysis of a vast number of documents which reflect the structured but highly variable nature of nineteenth century Cape Dutch Vernacular.

Although the notion of diglossia fails to capture the complexities and layers of linguistic variation at the Cape, diglossic tendencies nevertheless existed. In the corpus a small number of individuals were found to use Afrikaans and Dutch in strict alternation, and the language of their letters could be identified unambiguously as representing either one code or the other. It seems that within this sociolinguistic group a dual norm system was indeed present and the two-norm-theory might thus be applicable within limits. Those who engaged in such patterns of code alternation (across and within texts) were young (born after 1865) and generally well educated, which agrees with Wright's

(1991) observation that diglossia is not a naturally developing characteristic of speech communities, but emerges as a result of more or less conscious engineering and is reproduced via the education system (see Chapter 2).[1]

Hypothesis III: Shaping the linguistic market

As an exercise in historical sociolinguistics this study was concerned not only with the description of linguistic variation, but also with the notion of speaker agency (i.e. what speakers do) and speaker attitudes (i.e. what speakers believe and what they say they believe). The concepts of focusing and projection, which were introduced to sociolinguistic theorizing by Le Page and Tabouret-Keller (1985), are useful for a broad, sociolinguistic understanding of standardization. The two concepts are based on an explicitly context- and speaker-oriented approach to language history: it is through a multitude of individual acts of accommodation that linguistic regularity norms emerge and acquire meaning (see introduction). In this context, cultural and linguistic nationalism is a powerful social movement, which has transformed – through the development of new discourses and practices – the sociolinguistic make-up of many societies, creating novel forms of symbolic societal integration (i.e. 'imagined communities' in the context of supra-local communicative networks) and disintegration (i.e. nationally distinct cultural and linguistic identities). The development of a clear-cut conceptual division between a Low and a High variety was initiated from the 1850s with the dialect writing, but only crystallized in the 1870s in tandem with the early standardization efforts and the rise of Afrikaner nationalism. The discursive creation of diglossia supported a clear sense of cultural and linguistic autonomy (*vis à vis* Dutch), and created the attitudinal conditions for successful vernacular standardization.

As in many other language histories, vernacular elevation as a language-political goal did not refer to the acceptance of the realities of sociolinguistic diversity and variability as the new norm, but involved the strategic simplification of a complex variety space, the replacement of the old High variety (Dutch) and the creation of a new uniform linguistic norm (Afrikaans). In his discussion of the 'avant-garde of acculturation', Joseph (1987: 52, 106) argued that the new High variety is always created on the basis of the 'normalized' model provided by the old High variety. It therefore reflects the linguistic structures of the heterogeneous Low varieties only within limits:

> [T]hose who would undertake the standardization of L seem to betray the L culture … They 'sell out' L for a position within H culture that is certain to elevate their own personal prestige … Thus, although the work of the L cultural avant-garde in adapting a synecdochic dialect for standard language functions is typically motivated by a desire to promote the instruction and acculturation of the L community at large, *the very adaption they perform alters the L language – by standardizing it – such that it is no longer the language of the mass of the L community.* (ibid.; my emphasis)

If the socio-historical analysis presented in this study is plausible, it follows that the adoption of Afrikaans as a new standard norm led to different processes of variety shifting, i.e. the bi-dialectal speakers shifted from one well-defined and comparatively uniform (i.e. standardized) linguistic system to another, whereas others shifted (probably gradually) from the variable structures of non-standard mesolectal Cape Dutch Vernacular to the highly regularized new standard language which came to replace the old Dutch standard in education and formal registers. However, many older speakers – who moved in highly localized networks – maintained their variable non-standard norm, and mesolectal Cape Dutch Vernacular can be found in private letters up until the 1930s (Chapter 9).

Hypothesis IV: The role of the 'middle classes' and the standard as a social symbol

The social make-up of the early standard users, and the network analysis presented in Chapters 7 and 9 illustrate the path through which the Afrikaans standard language diffused, i.e. via the socially (and geographically) mobile new professional class (cf. Ross 1999a: 47: 'intellectuals, often lawyers, schoolteachers, clergymen and so forth.'). The link between standard language usage, education, urbanization and an intermediate social identity or status is a recurring feature of many otherwise quite distinct standard language histories. With regard to the standardization of Dutch in Flanders Van de Craen and Willemyns (1988: 51) refer to the 'challenging role of a rising middle class vis-à-vis the [French-speaking] power elite', and in the context of French language history Bourdieu (1991: 47) describes the 'local bourgeoisie of priests, doctors or teachers, who owed their position to the mastery of the instruments of expression and had everything to gain from … linguistic unification'. Similarly, Mattheier (1991: 42) has argued for German that the adoption of the standard norm by the educated middle-classes was not only a function of access and exposure

(via the education system and popular print culture), but that the standard language itself developed into a symbol of bourgeois social identity:

> Das Bildungbürgertum hat im 19. Jahrhundert … die Durchsetzung und Ausbreitung der Standardsprache … veranlaßt und geprägt … Bildungsbürgertum und Standardsprache stehen also in einem gegenseitigen Wirkungs- und Verwertungsverhältnis.
>
> ['In the nineteenth century the educated bourgeoisie was instrumental in the establishment and diffusion of the standard language … the educated bourgeoisie and the standard language thus stand in a relationship of mutual influence and utilization.']

The central role played in standardization histories by those who achieve social distinction not by descent but through education deserves closer and comparative investigation. It should be approached from a broad perspective of social change taking into account the societal transformations which are part and parcel of the modernization process, and which contributed to the rise of standard languages (i.e. migration, urbanization, public education, industrialization as well as new forms of public participation, political democratization and administrative bureaucratization; see also Deumert & Vandenbussche 2003).

Hypothesis V: Language standards as 'focused' clusters of idiolects

The statistical analysis presented in this volume is based on the hypothesis that the early stages of standardization are reflected in a gradual increase of linguistic uniformity (i.e. the emergence of focused regularity norms) and can be investigated with the methods of variationist sociolinguistics (cf. also Ferguson 1988; Devitt 1989; Lippi-Green 1994). Multivariate clustering techniques are useful tools for modelling linguistic varieties and norm emergence on the basis of frequency data. They are potentially applicable to a wide range of sociolinguistic and historical research projects.

As regards the process of linguistic standardization, the patterns of variation discussed in Chapter 7 suggest that established standard varieties and newly focusing varieties have different structures of speaker membership. Established standard varieties show a *Kernlandschaft* or nucleus of individuals who are very similar in their linguistic behaviour. Variation is largely the result of so-called outliers or residuals, i.e. individuals who show more marginal adherence to standard norms, but who themselves do not cluster. Newly focused or focusing varieties combine variability with increasing uniformity, i.e.

individuals show relatively uniform linguistic behaviour without yet constituting a tight nucleus of very similar cases (cf. Figure 7.3 cluster I and III). In addition, the principal components analysis of the corpus data indicates that during standardization, linguistic variation – at least in written and more formal domains of communication – is stream-lined and increasingly (yet not exclusively) ordered along a single standard/non-standard dimension (cf. also Bourdieu's 1991 discussion of the 'unification of the linguistic market'). Whether this is a general phenomenon of early standardization, and whether this simplification of the sociolinguistic space is maintained diachronically (i.e. even during processes of re- and de-standardization) is open to further investigation.

Using the history of Afrikaans as a case study and approaching the historical material in the *Corpus of Cape Dutch Correspondence* from a range of quantitative and qualitative sociolinguistic perspectives (i.e. classical variation analysis, multidimensional clustering techniques, discourse analysis, code-mixing/switching and social network analysis), this book has shown that as linguists we cannot ignore the question of historically-situated speaker agency in our explanations of language change. Ideally, language-historical studies should be based on detailed sociolinguistic analysis and the careful ethnographic reconstruction of the beliefs, intentions and social identities which speakers wish to express through language, and which they enact in social practices. Although we have to accept that a sociolinguistic approach to language history is not always possible on the basis of the surviving historical record (which is often too fragmentary for meaningful variation analysis and ethnographic description), a sociolinguistic, speaker-oriented approach to language change should always function as 'a methodological guideline and heuristic aid' (Lass 1976: 225; Deumert 2003a).

Histories of language standardization – in which gradual and cumulative processes of emergent focusing interact with deliberate speaker invention, the goal-directed manipulation of linguistic resources, and discursive rationalization (as expressed in the Milroys' notion of a 'standard language ideology') – are of particular interest to the field of historical sociolinguistics. Deliberate language change propagated by individuals or group action (language societies, language academies, etc.) interacted in particular from about 1800 (i.e. with the emergence of language-oriented nationalism) with on-going processes of sociolinguistic change, and supported within modern speech communities the formation of relatively uniform, written and increasingly prescriptive standard norms. Although such deliberate changes did not necessarily or immediately

affect the majority of the speech community, the emergence of social groups whose casual spoken language became a variety close to the new standard norm has become an important aspect of language change in the twentieth century.

Note

1. Ponelis' recent critique of the one-norm theory remains doubtful (cf. Ponelis 2002:96). Ponelis remarks correctly that the period discussed in this study 'falls within the present generation's consciousness of spoken Afrikaans' and is thus open to direct critique. He refers to a recording he made of a woman born in 1886 whose speech 'was practically identical to her children (my contemporaries) and my own'. Based on this evidence Ponelis maintains that varieties close to modern Afrikaans must indeed have constituted the normal spoken language around 1900, and that the linguistic patterns of the written record were an artefact of education, reflecting an imperfect and transient approximation of High Dutch. Ponelis' observation, however, does not actually address the argument of this study: the variable structures of Cape Dutch Vernacular were characteristic primarily of speakers/writers who were born *before* 1865 and who belonged to the traditional petty bourgeoisie (Chapter 7). For those born later (i.e. Ponelis' informant) the norms of the emerging Afrikaans standard already formed a point of normative orientation, leading to the gradual emergence, formation and focussing of a new linguistic habitus and, ultimately, to the replacement of acrolectal and mesolectal varieties of Cape Dutch. The woman's close approximation to modern Afrikaans testifies to the success of the diffusion process in the early twentieth century. The *Corpus of Cape Dutch Correspondence* provides ample evidence for such language shift processes in the post-1865 generation (see Chapters 7 and 9).

Appendix

The *Corpus of Cape Dutch Correspondence*

CA = The State's Archives Depot, Cape Town
G = Archives of the Moravian Church, Genadendal
NGK = Archives of the Nederlands Gereformeerde Kerk, Cape Town
SAL = South African Library Manuscript Collection, Cape Town
UCT = Manuscript Collection of the University of Cape Town

April, Aletta	letters to C. P. Hoogenhout (27/1/1913, 19/9/1913, 25/10/1918, two letters c. 1918) [SAL-MSC 1-26-4/ 1-41-8]
April, Anton	letter to Mr. Schmidt (6/7/1913) [G-Vorsteheramt Briefe 1913–1914]
Baalie, Fredrik	letter to P. O. Henning (3/12/1896) [G-Briewe Onderwysers aan Superintendent 1850–1900]
Baalie, Rudolf	letters to P. O. Henning (23/2/1894, 30/10/1897) [G-Briewe Onderwysers aan Superintendent 1850–1900]
Baartman Susie	letters to C. P. Hoogenhout (18/7/1895, 30/2/1902, 22/8/19, 1/3/1920) [SAL-MSC 1-25-5]
*Basson, Mary	letters to C. P. Hoogenhout (22/6/1904, 28/6/1905, 3/9/1906, 20/4/1907) [SAL-MSC 1-10-4/ 1-28-1/ 1-41-1/ 1-27-2]
*Beyers, Catharina E.	letters to C. P. Hoogenhout (20/1/1906, 25/2/1906, one letter circa 1918) [SAL-MSC 1-25-7]
Blake, Ebenzer J.	letters to C. P. Hoogenhout (30/4/1909, 9/8/1909) [SAL-MSC 1-41-1/1-27-3]
Bonthuis, Jan	letters to C. P. Hoogenhout (26/4/1909, 3/5/1909) [SAL-MSC 1-41-1]
Bosman, Louisa	letter to C. P. Hoogenhout (26/1/1900) [SAL-MSC 1-41-8]
*Botha, Louis	letters to F. S. Malan (12/8/1909), to Mrs. C. C. De Villiers (4/4/1903), to E. Esselen (7/8/1909), to J. C. Smuts (4/8/1902) [CA-A 583-2; UCT-B32-2/ UCT-BC 94-D3.4/ UCT-BCZA-77\104\3]

Breuris, Frans M. C.	letters to Katie Van Huyssteen (22/1/1906, 28/12/1905), to David Malan (11/12/1905) [SAL-MSB 853-4-1/ 853-1-3/ 852-1-13]
*Brümmer, Barend J.	letters to his daughter Johanna Brümmer (12/4/1900, 21/6/1900, 23/5/1901, 4/8/1901, 7/12/1901, 26/4/1902, 14/10/1904, 1/11/1904, 19/9/1907, Oct. 1907, 28/11/1907), to his wife Gerberg Wilhelmina (25/12/1901), to his son N. J. Brümmer (6/1/1901) [CA-A 583-85/ 583-3/ 583-7]
*Brümmer Johanna	letters to her fiancée/husband F. S. Malan (7/12/1893, 16/11/1894, 15/2/1895, 15/3/1895, 10/5/1895, 29/3/1895, 14/5/1901, May 1901), to her sister-in-law Katie Van Huyssteen (22/12/1905) [CA-A 583-8; SAL-MSB 853-4-1]
Brümmer, Johann J.	letter to his sister Johanna Brümmer (12/7/1900) [CA-A 583-85]
Brümmer, Nikolaas J.	letters to F. S. Malan (12/11/1894, 8/4/1895) [CA-A 583-2]
Burger, Jacobus A.	letter to Katie Van Huyssteen (1/1/1906) [SAL-MSB 853-4-3]
Cillie, Francois S.	letters to his parents (10/6/1896), to his-brother-in law C. P. Hoogenhout (15/4/1889, 1900, 4/5/1909, 31/5/1910) [SAL-MSC 1-22-3/ MSC 1-22-3/ MSC 1- 26-6/ MSC 1-28-1]
Coopman, Cornelius	letter to the Sendings-Commissie (18/9/1909) [SK-S 5-2-57-1]
Cronje, Daniel C.	letter to Katie Van Huyssteen (7/6/1906) [SAL-MSB 853-4-1]
Cronje, Hendrik S.	letter to Katie Van Huyssteen (15/6/1906) [SAL-MSB 853-1-5]
Cronje, Maria	letters to Katie Van Huyssteen (2/7/1906, 30/7/1906) [SAL-MSB 853-4-1/ MSB 853-1-5]
Cronje, Wessels C.	letter to David Malan (28/7/1905) [SAL-MSB 852-1-13]
Dampies, Piet	letter to the Sendings-Commissie (30/8/1891) [SK-S 5-21-26]
Daniels, Sophie	letter to P. O. Henning (23/9/1900) [G-Briewe Onderwysers aan Superintendent 1850–1900]
De Villiers, J. De W.	letter to C. P. Hoogenhout (11/3/1894) [SAL-MSC 1-29-1]
De Villiers, J. H.	letter to S. J. Du Toit (22/4/1891) [SAL-MSC 7-8-4]
De Villiers, Jan J.	letter to D. J. Du Toit (4/12/1882) [SAL-MSC 1-34-9]
De Vos, Helena Catherina	letter to her niece Sophie Hendrikz (23/8/1899) [SAL-MSB 892-1-4]
*De Vries, Anna	letters to Katie Van Huyssteen (13/6/1906), to C. P. Hoogenhout (6/6/1899, 3/12/1901, 11/6/1906, 10/9/1906) [SAL-MSB 853-1-5; MSC 1-24-3]

De Waal, Beatrice	letter to C. P. Hoogenhout (21/12/1908) [SAL-MSC 1-29-5]
*De Waal, Jan H. H.	letters to F. S. Malan 28/10/1901, 8/1/1902), to C. P. Hoogenhout (3/6/1910) [CA-A 583-2; SAL-MSC 1-36-4]
*De Wet, D. C.	letter to E. Esselen (24/11/1888) [UCT-BC 94-D3.20]
De Wet, Nicolaas J.	letters to F. S. Malan (1894, 21/5/1895, June 1895) [CA-A 583-2]
Dippenaar, Johannes	letter to Sophie Hendrikz (22/12/1900) [SAL-MSB 892-1-3]
Du Preez, Johannes G.	letter to his brother N. J. Du Preez (11/8/1894) [SAL-MSB 162-1]
Du Preez, P. E.	letter to his brother N. J. Du Preez (24/8/1891) [SAL-MSB 162-1]
*Eksteen, Elizabeth J. H.	letters to her daughter Rachel Steyn, née Eksteen (31/8/1919, 1/1/1922) [SAL-MSC 33-3-1/ MSC 33-5-1]
Eksteen, P. J.	letter to P. O. Henning (15/3/1899) [G-Briewe Onderwysers aan Superintendent 1850–1900]
Es(s)op, Napie	letters to C. P. Hoogenhout (23/4/1919, 5/7/1919) [SAL-MSC 1-26-4]
Furter, Engela J. M.	letters to C. P. Hoogenhout (5/10/1909, 30/8/1909, 24/11/1909, 17/9/1909) [SAL-MSC 1-26-3]
*Goosen, Gideon J. C.	letters to Imker Hoogenhout (4/2/1909, 19/6/1909, 22/10/1909), to C. P. Hoogenhout (13/11/1913) [SAL-MSC 1-17-5/ MSC 1-17-4/ MSC 1-27-3]
Goosen, Jan A. B.	letters to Imker Hoogenhout (10/10/1908, 25/4/1909, 2/8/1909) [SAL-MSC 1-17-5/ MSC 1-17-4]
*Goosen, Pieter A. jun.	letters to Imker Hoogenhout (24/5/1909), to C. P. Hoogenhout (6/5/1909, 18/10/1911) [SAL-MSC 1-27-2/ MSC 1-17-5/ MSC 1-27-3]
Goosen, Pieter A. sen.	letters to Imker Hoogenhout (27/7/1909), to C. P. Hoogenhout (2/2/1912) [SAL-MSC 1-17-4/ MSC 1-27-4]
*Goosen, Sebastina	letters to her parents-in-law C. P. Hoogenhout and C. M. Hoogenhout (17/6/1905, 7/4/1910, 19/4/1910), to her husband Imker Hoogenhout (20/3/1912) [SAL-MSC 1-10-4/ MSC 1-14-2/ MSC 1-14-5/ MSC 1-17-1]
Goosen, S. J. F. sen.	letter to C. P. Hoogenhout (1913) [SAL-MSC 1-22-3]
Goosen, S. J. F. jun.	letters to Imker Hoogenhout (23/2/1909, 24/10/1909), to C. P. Hoogenhout (8/4/1911, 14/4/1912, 12/5/918) [SAL-MSC 1-29-3/ MSC 1-27-4/ MSC 1-27-3/ MSC 1-27-5/ MSC 1-27-5]

*Grosskopf, J. F. W.	letters to Imker Hoogenhout (12/8/1906, 21/10/1906, 26/3/1907, 3/6/1907, 13/9/1907, 1/12/1907, 8/12/1907, 4/11/1908, 18/2/1909) [SAL-MSC 1-17-2]
*Hauptfleisch, Andries P.	letters to his father-in-law C. P. Hoogenhout (4/3/1914, 11/6/1914, 9/7/1915, 19/2/1918, 1909, 31/7/1918, 7/9/1921, 26/09/1921) [SAL-MSC 1-20-1/ MSC 1-20-5/ MSC 1-20-4/ MSC 1-20-2]
Hatting, Benjamin	letter to the Sendings-Commissie (18/7/1886) [SK-5-2-126]
Henderson, James	letter to Marie Koopmans-DeWet (c. 1900) [CA-A 583-89]
Hendrikz, Sophie	letter to her cousin 'Johannes' (c. 1900) [SAL-MSB 892-1-3]
Heroldt, Michael J.	letter to his grandchild 'Annie' (28/9/1891) and biographical narrative (circa 1900) [SAL-MSB 229-1/ MSB 229-2]
Hillebrand, Martha C.	letter to her grandchild Katie Van Huyssteen (13/10/1903) [SAL-MSB 853-1-2]
Hlela, Johannes	letters to the Moravian mission in Genadendal (22/8/1906, 20/9/1906) [G-Buchkasse Briefe 1905–1907]
Hoogenhout, Dan. M.	letters to Mimmie Laubscher (2/11/1909), to his parents (17/6/1909, 24/10/1909) [SAL-MSC 1-19-4/ MSC 1-19-1]
*Hoogenhout, Francois P.	letter to his brother Imker (23/3/1912), to his parents (16/2/1910, 20/9/1911, 8/6/1912) [SAL-MSC 1-11-5/ MSC 1-11-1/ MSC 1-11-2]
Hoogenhout, Johanna	letters to her parents (5/7/1893, 6/2/1896, 2/1/1897, 26/5/1899, 29/4/1902, 26/10/1902, 4/11/1902, 11/06/1905) [SAL-MSC 1-10-2]
*Hoogenhout, Nicolaas M.	letters to his parents (10/3/1897, 29/4/1897, 14/4/1905, 20/5/1906, 24/6/1910) [SAL-MSC 1-5-2/ MSC 1-5-3]
Hoogenhout, Petrus I.	letters to his parents (24/5/1902, 8/5/1908, 23/5/1908, 25/5/1908), to his brother Imker (23/2/1896) [SAL-MSC 1-17-1/ MSC 1-17-5/ MSC 1-14-3]
*Hoogenhout, Wynanda	letters to her parents (21/9/1905, 8/12/1905, 11/10/1905, 17/10/1905, 18/2/1907, 8/3/1907, 7/2/1908, 15/4/1909, 14/5/1911, 27/5/1911, 10/6/1911, 24/6/1911, 12/6/1913, circa 1914, 15/6/1914, 16/7/1914, 11/6/1915, 12/10/1915, 23/6/1916, 23/8/1919, 9/12/1916, 15/1/1917), letters to her brother Imker (28/4/1909, 3/11/1909) [SAL-MSC 1-20-4/ MSC 1-20-5/ MSC 1-20-1/ MSC 1-20-2/ MSC 1-27-1/ MSC 1-27-2/ MSC 1-17-4/ MSC 1-20-3]

January, Wilhelmina	letter to P. O. Henning (30/3/1897) [G-Briewe Onderwysers aan Superintendent 1850–1900]
Koch J. A.	letter to Katie Van Huyssteen (27/7/1909) [SAL-MSB 853-1-8]
Kotze, T. N.	letter to F. S. Malan (1/11/1904) [CA-A 583- 4]
Krige, Leopold J.	letter to his sister Isabella Smuts (née Krige) (7/3/1907) [UCT-BCZA- 77/107/14]
Kuhn, Pieter G.	letter to F. S. Malan (16/5/1900) [CA-A 583-4]
Laubscher, Mimmie	letters to D. M. Hoogenhout (Afrikaans/Dutch postscripts and notes in the following letters 1909, 1909, 19/1/1909, 18/4/1909, 29/4/1909, 6/5/1909, 15/5/1909, 22/5/1909, 15/5/1909, 21/6/1909, 26/6/1909, 29/6/1909, 3/8/1909, 9/8/1909, 12/8/1909, 2/9/1909, full Afrikaans letters 1/12/1909, 16/2/1910) [SAL-MSC 1-19-2/ MSC 1-19-3/ MSC 1-19-4]
*Leipoldt, Carolina J. M.	letters to her son Christian Friedrich (31/7/1880, 21/8/1880, 4/9/1880, 2/10/1880, 27/11/1880, 11/12/1880, 19/3/1881, 21/5/1881, 13/8/1881, 23/6/1883, 21/7/1883, 27/9/1883, 5/11/1884, 22/11/1884) [UCT-BC 94 C.263/ BC 94 C2.64-68/ BC 94 C2.70/ BC 94 C2.72/ BC 94 C.2 75/ BC 94 C2.83/ BC 94 C2.91/ BC 94 C2.92/ BC 94 C.2-94]
Le Roux, D. J.	letters to C. P. Hoogenhout (15/11/1886, 17/11/1886) [SAL-MSC 1-41-1]
*Le Roux, Petrus J.	letters to his step-father C. P. Hoogenhout (22/3/1905, 1/5/1908, 18/2/1908, 6/6/1908, 16/8/1909, 6/4/1912, 16/11/1913, three letters without date) [SAL-MSC 1-22-1/ MSC 1-29-1]
Louw, Andries A. sen.	letters to his son Andries (29/3/1886, 1/4/1891, 11/11/1894) [NGK-P 52-1-2]
*Louw, Andries A. jun.	letter to his father A. A. Louw, sen. (28/2/1887) [NGK-P 52-1-2]
Louw, Francina S.	letter to her father-in-law A. A. Louw, sen. (14/3/1899) [NGK-P 52-1-2]
Louw, Maria J.	letter to her father, A. A. Louw, sen. (27/5/1901) [SAL-MSB 57]
Louw, Rijkie Hester	letter to Katie Van Huyssteen (4/8/1909) [SAL-MSB 853-1-8]
Louwrens, Anna S.	letter to Katie Van Huyssteen (16/3/1905) [SAL-MSB 853-1-3]
Lourens, Johannes F.	letter to Daniel Malan (28/12/1905) [SAL-MSB 853-5-1]

Malan, Daniel G.	letters to his daughter-in-law Katie Van Huyssteen (29/5/1906, 22/6/1906, 14/12/1906, 25/8/1907) [SAL-MSB 853-1-5/ MSB 853-1-6]
*Malan, Elizabeth J.	letters to her son F. S. Malan (9/5/1893, circa 1900), to her daughter-in-law Johanna Brümmer (16/12/1893), to her daughter-in-law Katie Van Huyssteen (17/12/1905, 28/5/1906, 22/6/1906, 20/8/1907, no date), to her grandchildren Vrede, Rechtie and David Malan (25/7/1909) [CA-A 583-7/ A 583-85; SAL-MSB 853-1-3/ MSB 853-1-5/ MSB 853-1-6/ MSB 853-1-8]
*Malan, Francois S.	extracts from diaries (1892–1895), letter to W. Viljoen (transcript of the letter in his diary 4/10/1894), letters to his fiancée/wife Johanna Brümmer (10/4/1892, 1/6/1893, 31/10/1894, 28/3/1895, 2/11/1894, 7/4/1895, 2/4/1896, 7/6/1896, 27/12/1906), letter to his sister-in-law Katie Van Huyssteen (20/1/1906) [CA-A 583-16/ A 583-17/ A 583-18/ A-583-83/ A583-84; SAL-MSB 853-4-1]
Malan, Jacques	postcard to his cousin Vrede Malan (circa 1910) [SAL-MSB 853]
Malan, Lelon	letter to C. P. Hoogenhout (31/10/1917) [SAL-MSC 1-29-1]
*Malan, Maria C.	letters to her brother F. S. Malan (14/5/1892, 23/4/1893), to her sister-in-law Katie Van Huyssteen (22/12/1907, 29/5/1906), to her nephews Vrede, Rechter and David (4/11/1906) [CA-A 583-7; SAL-MSB 853-4-1/ MSB 853-1-5/ MSB 853-1-5]
Malherbe, Cornelia	letter to Mr. Murray (29/8/1916) [NGK-S 5-2-113]
*Marais, F. P.	to C. P. Hoogenhout (2/3/1895) [SAL-MSC 1-22-2]
Marais, Jacobus P. Z	letters to C. P. Hoogenhout (3/8/1908, 4/8/1908, 1/11/1918, 5/11/1918, 9/11/1918) [SAL-MSC 1-22-3/ MSC 1-22-4]
Marais, Helena E.	letters to her brother-in-law C. P. Hoogenhout and her sister C. M. Hoogenhout (23/1/1896, 10/2/1896) [SAL-MSC 1-22-3/ MSC 1-22-4]
Marais, Rachel	letter to C. P. Hoogenhout (8/5/1909) [SAL-MSC 1-27-3]
Marais, Stephanus J.	letter to C. Rayner (5/6/1901) [SAL-MSB 744-1-39]
Marren, 'Maggie'	letters to C. P. Hoogenhout (24/3/1915, 13/6/1915) [SAL-MSC 1-27-5]
Matodlana, Mozes	letter to P. O. Henning (5/7/1898) [G-Briewe Onderwysers aan Superintendent 1850–1900]

Mlimkulu, Mozes	letter to P. O. Henning (1/12/1896) [G-Briewe Onderwysers aan Superindentent 1850–1900]
Muller, Hermanus Steyn	letter to Katie Van Huyssteen and David Malan (18/5/1905) [SAL-MSB 853-1-3]
November, Jacobus	letter to C. P. Hoogenhout (13/2/1912) [SAL-MSC 1-26-4]
Odendaal, Johanna	letter to Katie Van Huyssteen (2/2/1906) [SAL-MSB 853-4-1]
Oosthuizen, M. A. M.	letters to Katie Van Huyssteen (3/1/1906, 27/12/1905) [SAL-MSB 853-4-1/ 853-1-3]
Pauw, Anna	letter to C. P. Hoogenhout (17/6/1917) [SAL-MSC 1-29-5]
Pienaar, Hendrik H.	letter to David Malan (17/5/1905) [SAL-MSC 583-1-3]
Roos, Lenie	letter to Marie Koopmans-DeWet (29/1/1901) [CA-A 583-89]
Roussouw, Aletta M.	letter to her brother-law C. P. Hoogenhout and her sister C. M. Hoogenhout (4/4/1904) [SAL-MSC 1-29-1]
Roux, Katie	letter to J. W. C. Hofmeyr (27/12/1919) [NGK-S 5-2-18]
*Schabort, Elizabeth A. S.	narrative of her experiences before and during detention at Merebank Camp, Durban [SAL-MSB 437-1-1]
Schrade, Hester	letter to J. W. C. Hofmeyr (12/2/1920) [NGK-S 5-2-56-1]
Small, Nicolaas	letter to C. P. Hoogenhout (1/7/1898), to Wynanda Hoogenhout (27/6/1898) [SAL-MSC 1-26-4/ MSC 1-21-6]
*Smit, Elizabeth ('Bettie')	letters to C. P. Hoogenhout (18/10/1911, 13/2/1917, 20/5/1917, 23/11/1917, 24/3/1918, 28/7/1918, 2/11/1919, 8/2/1920, 15/8/1920, 19/12/1920) [SAL-MSC 1-22-4/ MSC 1-25-3]
*Smith, Edmund T.	letters to his father-in-law B. J. Brümmer (circa 1900, 27/10/1904) [CA-A 583-3/ A 583-85]
Smuts, Jacobus A. ('Koos')	letter to his brother J. C. Smuts (11/5/1902) [UCT-BCZA 77-104-102]
Smuts, J. C.	letter to his brother J. A. Smuts (4/1/1902) [UCT-BCZA-77-104-108]
Smuts, Isabella M. ('Isie')	letters to Rachel I. Steyn (9/9/1912, 13/11/1912) [SAL-MSC 33-17-2]
Stanford, Cornelia	letters to Katie Van Huyssteen (6/8/1905, 9/6/1906) [SAL-MSB 853-1-3/ MSB 853-4-1]
Steyn, Anna	letter to Katie Van Huyssteen (16/3/1906) [SAL-MSB 853-4-3]
Steyn, Daniel	letter to Katie Van Huyssteen (2/6/1906) [SAL-MSB 853-1-5]

*Steyn, Elizabeth G.	letters to Katie Van Huyssteen (2/6/1906, 16/3/1905) [SAL-MSB 853-1-5/ MSB 853-1-3]
*Steyn, Rachel I.	letters to her daughter-in-law Rachel Eksteen (4/6/1919, 19/10/1919, 4/12/1919) [SAL-MSC 33-17-1]
Uys, Susan	letter to Katie Van Huyssteen (8/1/1906) [SAL-MSB 853-4-3]
*Van der Bijl, Hester	letters to her-parents-in-law C. P. Hoogenhout and C. M. Hoogenhout (13/2/1906, 8/11/1906, 18/11/1906, 21/2/1907, 10/3/1908, 17/5/1908, 10/11/1916) [SAL-MSC 1-5-2/ MSC 1-5-4/ MSC 1-5-1]
Van der Merwe, C. J.	letters to D. S. De Villiers (26/11/1900, 30/11/1900, 3/12/1900, 4/12/1900) [SAL-MSB 510-1-1]
Van Huyssteen, Anna Chr.	letter to her sister-in-law Katie Van Huyssteen (23/7/1907) [SAL-MSB 583-1-6]
Van Huyssteen, A. J. Z. ('Janie')	letter to her sister Katie Van Huyssteen (31/12/1907) [SAL-MSB 853-1-6]
Van Huyssteen, Daniel P. S.	letters to his sister Katie Van Huyssteen (17/1/1906, 16/6/1906, 30/6/1909) [SAL-MSB 853-1-5/ MSB 853-1-5]
*Van Huyssteen, Eliza J.	letters to her sister Katie Van Huyssteen (2/3/1907, 16/4/1908) [SAL-MSB 853-1-6]
*Van Huyssteen, E. C. P.	letters to her daughter Katie Van Huyssteen (10 letter 1905/06; not dated), to her daughter Jacomina Van Huyssteen (1906) [SAL-MSB 853-1-5/ MSB 853-1-10/ MSB 853-1-3]
Van Huyssteen, Hendrik	letters to his daughter Katie Van Huyssteen (9/1/1906, 3/10/1908), to his grandson David (14/10/1909) [SAL-MSB 853-1-5/ MSB 85-1-3/ MSB 853-1-8]
*Van Huyssteen, J. P. J. W. ('Algie')	letters to his sister Katie Van Huyssteen (20/5/1907, 27/7/1909) [SAL-MSB 853-1-6/ MSB 853-1-8]
Van Huyssteen, Jacomina H.	letters to her sister Katie Van Huyssteen (14/4/1908, 5/11/1909) [SAL-MSB 583-1-6/ MSB 853-1-8]
*Van Huyssteen, M. C. ('Katie')	letters to her parents Hendrik and E. C. P. Van Huyssteen (21/4/1906, 6/8/1906, May 1907) [SAL-MSB 853-1-1]
Van Huyssteen, Petrus H. H.	letters to his sister Katie Van Huyssteen (21/2/1906, Dec. 1905) [SAL-MSB 853-4-3/ MSB 853-4-1]
Van Huyssteen, Matthys P. Z.	letter to his sister Katie Van Huyssteen (10/1/1905) [SAL-MSB 853-4-1]
Van Oosthuizen, Lettie	letter to Katie Van Huyssteen (5/3/1907) [SAL-MSB 1-6]
Van Zyl, Susan	letter to Katie Van Huyssteen (24/3/1905) [SAL-MSB 853-1-3]

*Vermaak, Hester S.	letters to Katie Van Huyssteen (Nov. 1905, 27/1/1906, 19/2/1906, 13/6/1907, 22/6/1907) [SAL-MSB 853-4-1/ MSB 853-4-3/ MSB 853-1-6]
Viljoen, Hendrina	letters to her-parents-in-law C. P. Hoogenout and C. M. Hoogenhout (22/1/1911, 19/1/1911, 14/2/1912, 13/5/1919) [SAL-MSC 1-18-6/ MSC 1-27-2/ MSC 1-18-4]
'Hester'	to 'Maria' (7/11/1900) [UCT-B28 F19]
'Gaybba'	to Catharina Maria Hoogenhout (circa 1900) [SAL-MSC 1-26-4]
'Auntie'	letter to her niece (27/4/1904) [SAL-MSB 744-1-43]

Writers included in the final multivariate analysis (Chapter 7) are marked by an asterisk (*).

References

Manuscript sources

CA The State Archives Depot, Cape Town
- A 583 F. S. Malan Collection

SAL South African Library (Cape Town) Manuscript Collection
- MSC 1 Hoogenhout Collection
- MSC 7 J. H. De Villiers
- MSC 33 Rae Steyn Collection
- MSB 162 Du Preez Family Collection
- MSB 229 Heroldt Family Papers
- MSB 437 A. S. Schabort
- MSB 510 Volkscongres, Worcester (1900)
- MSB 744 South Africa Letters (Miscellaneous Collection)
- MSB 852 D. J. Malan Collection
- MSB 853 Katie Malan Collection
- MSB 892 Women's Studies (Miscellaneous Collection)
- MSB 957 Mimie Reyneke Collection

UCT Manuscript Collection of the University of Cape Town
- B28 De Villiers Collection
- BCZA 77 Smuts Collection
- BC94 Louis Leipoldt Papers

NGK Archives of the Nederlands Gereformeerde Kerk, Cape Town
- P Private Papers
- S 5 Sendingskerk

G Archives of the Moravian Church, Genadendal

Official publications

CPP Cape Parliamentary Papers
- G42-1876 Cape Census for 1875
- G6-892 Cape Census for 1891
- G19-1905 Cape Census for 1904

G75-1880 Report of the Commission appointed to inquire into and report upon the working of the Education Acts in force in this Colony (1879–1880).
G12-1883 Preliminary Report on the State of Education in the Colony of the Cape of Good Hope (1882).
G3-1892 Education Commission, First and Second Report (1891).
G 682 EI-1892 Third and Final Report and Proceedings, with Appendices, of a Commission appointed to Enquire into and Report upon Certain Matters concerned with the Educational System of the Colony (1891).

Union of South Africa
UG32-1912 Census of the Union of South Africa (1911).

Voter's list and directories

Cape of Good Hope Voters List 1903. Vol. I–V. Cape Town: The Cape Times Government Printers.
The General Directory of South Africa for 1903. Comprising separate directories for nearly 500 townships, cities and districts of the Cape Colony, Natal, Orange River Colony, The Transvaal, Basutoland, Mashonaland, Rhodesia etc. Cape Town: Dennis Edwards & Co., Proprietors and Publishers.
The General Directory for a United South Africa for 1911. Separate Directories for over 1550 townships, cities and districts of the Cape Colony, the Transvaal, Natal, Orange Free State, Basutoland, Swaziland, Rhodesia, German Southwest Africa, Portuguese Africa etc. Cape Town/Johannesburg: Dennis Edwards & Co., Proprietors and Publishers.
The South African Directory for 1883–1884. Embracing separate alphabetical directories for the OFS, the Transvaal State and Natal. Cape Town: Saul Solomon & Co., Printers and Publishers.

Periodicals

Die Afrikaanse Patriot (Paarl)
Die Burger (Cape Town)
The Cape Monthly Magazine (Cape Town)
The Cape Times (Cape Town)
Ons Land (Cape Town)
De Zuid Afrikaan (Cape Town)
Het Zuid-Afrikaansche Tijdschrift (Cape Town)
De Goede Hoop (Cape Town)
De Ebenezer (Cape Town)
De Bode van Genadendal (Genadendal)

Books and journals

Abraham, W. & Conradie, J. C. (2001). *Präteritumsschwund und Diskursgrammatik. Präteritumsschwund in gesamteuropäischen Bezügen: areale Ausbreitung, heterogene Entstehung, Parsing sowie diskursgrammatische Grundlagen und Zusammenhänge.* Amsterdam: John Benjamins.

Adhikari, M. (1989). *Identity and Assimilation in the Malay Community of Nineteenth Century Cape Town.* History of Cape Town – Work in Progress Workshop. University of Cape Town.

—— (1992). The sons of Ham: Slavery and the making of coloured identity. *South African Historical Journal, 27*, 95–112.

—— (Ed.). (1996). *Straatpraatjes. Language, politics and popular culture in Cape Town, 1909–1922.* Pretoria: Van Schaik.

Aldenderfer, M. S. & Blashfield, R. K. (1984). *Cluster Analysis* [Series no. 07-044]. London: Sage Publications.

Alderson, A. W. (1909). *The Worst Tax of All. The real cause of the South African war.* London: King and Son.

Alexiou, M. (1982). Diglossia in Greece. In W. Haas (Ed.), *Standard Languages. Spoken and written* (pp. 156–191). Manchester: Manchester University Press.

An. (1873). *Benigna van Groenekloof of Mambre. Een verhaal voor de Christen kleuringen van Zuid-Afrika. Door een hunner leeraars.* Genadendal: Genadendal Press.

Anderberg, M. R. (1973). *Cluster Analysis for Applications*. New York: Academic Press.

Anderson, B. (1991). *Imagined Communities. Reflections on the origin and spread of nationalism* (second revised edition). New York: Verso.

Anderson, S. R. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.

Annamalai, E. (1989). The linguistic and social dimensions of purism. In B. J. Jernudd & M. J. Shapiro (Eds.), *The Politics of Language Purism* (pp. 225–231). Berlin: Mouton de Gruyter.

Arabie, P. & Hubert, L. J. (1995). Clustering from the perspective of combinatorial data analysis. In W. J. Kzranowski (Ed.), *Recent Advances in Descriptive Multivariate Analysis* (pp. 1–13). Oxford: Oxford University Press.

Armstrong, J. & Worden, N. (1989). The Slaves 1652–1834. In R. Elphick & H. Giliomee (Eds.), *The Shaping of South African Society 1652–1820* (pp. 109–183) (second edition). Cape Town: Maskew Miller Longman.

Auer, P. (1999). From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism, 3*, 309–332.

Bailey, C.-J. (1973). *Variation and Linguistic Theory.* Washington, DC: Center for Applied Linguistics.

—— (1980). Conceptualizing dialects as implicational constellations rather than as entities bounded by isoglossic bundles. In P. Ivic Göschel & K. Kehr (Eds.), *Dialekt und Dialektologie. Ergebnisse des Internationalen Symposions 'Zur Theorie des Dialekts' Mahrburg/Lahn, 5–10 September 1977* (pp. 234–272). (= Beiheft, N. F., *Zeitschrift für Mundartforschung*.)

Bailey, G. & Ross, G. (1992). The evolution of a vernacular. In M. Rissanen, O. Ihalainen, T. Nevalainen, & I. Taavitsainen (Eds.), *History of Englishes. New methods and interpretations in historical linguistics* (pp. 519–531). Berlin: Mouton de Gruyter.

Baker, P. (1990). Off target? *Journal of Pidgin and Creole Languages, 5*, 107–120.

—— (1993). Assessing the African contribution to French-based Creoles. In S. Mufwene (Ed.), *Africanisms in Afro-American Language Varieties* (pp. 123–155). Athens: University of Georgia Press.

—— (1997). Directionality in Pidginization and Creolization. In A. K. Speers & D. Winford (Eds.), *The Structure and Status of Pidgins and Creoles* [Creole Language Library 19] (pp. 91–109). Amsterdam: John Benjamins.

—— (2000). Theories of creolization and the degree and nature of restructuring. In I. Neumann-Holzschuh & E. Schneider (Eds.), *Degrees of Restructuring in Creole Languages* [Creole Language Library 22] (pp. 41–63). Amsterdam: John Benjamins.

Baker, P. & Corne, C. (1982). *Isle de France Creole. Affinities and origins.* Ann Arbor: Karoma.

Bakhtin, M. M. (1981). *The Dialogic Imagination*. Austin: University of Texas Press.

Barnouw, A. J. (1934). *Language and Race Problems in South Africa.* The Hague: Martinus Nijhoff.

Basso, K. (1989 [1974]). The ethnography of writing. In R. Bauman & J. Scherzer (Eds.), *Explorations in the Ethnography of Speaking* (pp. 425–432). Cambridge: Cambridge University Press.

Batten, C. L. (1978). *Pleasurable Instruction. Form and convention in eighteenth century travel literature.* Berkeley: California University Press.

Baugh, J. (1980). A re-examination of the Black English copula. In W. Labov (Ed.), *Locating Language in Time and Space* (pp. 83–106). New York: Academic Press.

—— (1992). Hypercorrection: Mistakes in production of vernacular African American English as a second dialect. *Language and Communication, 12*, 317–326.

Bax, R. C. (2000). A network strength scale for the study of eighteenth-century English. *European Journal of English Studies, 4*(3), 277–289.

Bechert, J. & Wildgen, W. (1991). *Einführung in die Sprachkontaktforschung*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Belcher, R. (1987). Afrikaans en kommunikasie oor die kleurgrens. In H. Du Plessis & T. Du Plessis (Eds.), *Afrikaans en Taalpolitiek* (pp. 17–35). Pretoria: HAUM.

Bell, A. (1999). Styling the other to define the self: A study in New Zealand identity marking. *Journal of Sociolinguistics, 3*, 523–541.

Berdan, R. (1978). Multidimensional analysis of vowel variation. In D. Sankoff (Ed.), *Linguistic Variation. Models and methods* (pp. 149–160). New York: Academic Press.

Berg-Seligson, S. (1986). Linguistic constraints on intersentential code-switching. *Language in Society, 15*, 313–348.

Bernstein, C. (1993). Measuring social causes of phonological variation in Texas. *American Speech, 68*, 227–240.

Berruto, G. (1987). Varietät. In U. Ammon, N. Dittmar, & K. J. Mattheier (Eds.), *Soziolinguistik/Sociolinguistics. Ein Internationales Handbuch* (pp. 263–267). Berlin: de Gruyter.

Besnier, N. (1988). The linguistic relationship of spoken and written Nukulaelae registers. *Language, 64*, 707–733.

Besselaar, G. (1914). *Zuid-Afrika in de Letterkunde.* Amsterdam: J. H. De Bussy.
Bester, M. (1989). Die Afrikaanse kindertydskrif as kultuurhistoriese inlichtingsbron. *Suid Afrikaanse Tydskrif vir Kultuur en Kultuurgeskiedenis, 3*, 19–30.
Bex, T. & Watts, R. J. (Eds.). (1999). *Standard English. The widening debate.* London: Routledge.
Biber, D. (1988). *Variation across Speech and Writing.* Cambridge: Cambridge University Press.
Bickerton, D. (1973). The nature of a Creole continuum. *Language, 49*, 640–669.
—— (1975). *Dynamics of a Creole System.* Cambridge: Cambridge University Press.
—— (1981). *Roots of Language.* Ann Arbor: Karoma.
—— (1984). The language bioprogram hypothesis. *The Behavioral and Brain Sciences, 7*, 173–188.
—— (1995). The origins of variation in Guyanese. In G. Guy, C. Feagin, D. Schiffrin, & J. Baugh (Eds.), *Towards a Social Science of Language. Papers in honor of William Labov* [Current Issues in Linguistic Theory 127], Vol. I (pp. 311–327). Amsterdam: John Benjamins.
Bickford-Smith, V. (1991). *Labour, Ethnicity and Conflict in Cape Town 1891–1902.* Cape Town History Project, Workshop 11–12 November 1991, University of Cape Town.
—— (1995). *Ethnic Pride and Racial Prejudice in Victorian Cape Town.* Johannesburg: Witwatersrand University Press.
Biewenga, A. (1996). Alfabetisering aan de Kaap de Goede Hoop omtreeks 1700. *Tydskrif vir Nederlands en Afrikaans, 3*, 109–121.
Blom, J.-P. & Gumperz, J. J. (1972). Social meaning in linguistic structures: Code-switching in Norway. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics* (pp. 407–434). New York: Holt, Rinehart and Winston.
Blommaert, J. (1999). The debate is open. In J. Blommaert (Ed.), *Language Ideological Debates* (pp. 1–38). Berlin: Mouton de Gruyter.
Bloomfield, L. (1933). *Language.* New York: Henry Holt.
—— (1944). Secondary and tertiary responses to language. *Language, 20*, 45–55.
Boëseken, A. (1977). *Slaves and Free Blacks at the Cape 1658–1700.* Cape Town: Tafelberg.
Boniface, C. E. (1979 [1832]). *De Nieuwe Ridderorde of De Temperatisten. Kluchtig Blijspel in Vier Bedrijven.* Facsimile reprint. Cape Town: South African Library.
Booij, G. E. & Van Santen, A. (1995). *Morfologie. De Woordstructuur van het Nederlands.* Amsterdam: Amsterdam University Press.
Boonzaier, W. J. C. (1989). *'n Taalkundige Karakterisering van Bolandse Afrikaans.* Ph.D. Thesis, University of the Orange Free State.
Borgatti, E. & Freeman, L. (1996). *UCINET IV.* Natick, MA: Analytic Technologies.
Boissevain, J. (1974). *Friends of Friends. Networks, manipulators, and coalitions.* Oxford: Basil Blackwell.
Bosker, A. (1961). *Het Gebruik van het Imperfectum en het Perfectum in het Nederlands, het Duits, het Frans en het Engels.* Groningen: J. B. Wolters.
Bosman, F. C. L. (1930). *Hollandse Joernalistiek in Suid Afrika Gedurende die 19de Eeu.* Cape Town: Ons Land.
Bouman, A. C. & Pienaar, E. C. (1925). *Afrikaanse Spraakkuns.* Stellenbosch: Pro Ecclesia Drukkery.

Bourdieu, P. (1984). *Distinction. A social critique of the judgement of taste.* London: Routledge.

—— (1985). The social space and the genesis of groups. *Theory and Society, 14*, 723–744.

—— (1991). *Language and Symbolic Power*. Ed. and introduced by J. B. Thompson. Cambridge: Polity Press.

Brenner, P. J. (Ed.). (1989). *Der Reisebericht. Die Entwicklung einer Gattung in der deutschen Literatur.* Frankfurt/M.: Suhrkamp.

Brill, J. (1909 [1875]). *De Landstaal.* Reprinted in Du Toit (1909 [1880]), *Geskiedenis van die Afrikaanse Taalbeweging* (pp. 28–39). Paarl: Paarl Drukpers Maatschappij.

Brink, M. (1905). *Voor Volk en Taal.* Cape Town: Van Sandt de Villiers.

Britto, F. (1986). *Diglossia. A study of the theory with application to Tamil.* Washington, DC: Georgetown University Press.

Bryn, A. & Veenstra, T. (1993). The creolization of Dutch. *Journal of Pidgin and Creole Languages, 8*(1), 29–80.

Buccini, A. F. (1992). *The Colonial Dutch Dialects. New evidence for the origins of Afrikaans and the development of standard Dutch in seventeenth century Holland.* Paper presented to the Germanic Linguistics Roundtable, University of California at Berkeley, April 1992.

—— (1996). New Netherland Dutch, Cape Dutch, Afrikaans. *Taal en Tongval* (Themanummer), *9*, 35–51.

Buchan, J. (1903). *The African Colony*. Edinburgh/London: William Blackwood and Sons.

Bundy, C. (1986). Vagabond Hollanders and runaway Englishmen: White poverty in the Cape before poor whiteism. In W. Beinart, P. Delius, & S. Trapido (Eds.), *Putting a Plough to the Ground. Accumulation and dispossession in rural South Africa, 1850–1930*. Johannesburg: Ravan Press.

Burchell, W. J. (1953 [1822–1824]). *Travels in the Interior of Southern Africa.* London: Batchworth Press.

Burke, P. & Porter, R. (Eds.). (1987). *The Social History of Language*. Cambridge: Cambridge University Press.

Burke, P. & Porter, R. (Eds.). (1993). *Language, Self and Society. A social history of language*. Cambridge: Polity Press.

Burke, P. & Porter, R. (Eds.). (1995). *Languages and Jargons. Contributions to a social history of language*. Cambridge: Polity Press.

Burton, A. R. E. (1903). *Cape Colony for the settler. Its urban and rural industries, their development, and extension.* London and Cape Town: P. S. King & Son and J. C. Juta & Co.

Burridge, K. (1992). *Syntactic Change in Germanic* [Current Issues in Linguistics Theory 89]. Amsterdam: John Benjamins.

Butler, C. (1985). *Statistics in Linguistics*. Oxford: Basil Blackwell.

Bybee, J. & Hopper, P. (2001). Introduction to frequency and the emergence of linguistic structure. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* [Typological Studies in Language 45] (pp. 1–24). Amsterdam: John Benjamins.

Carstens, W. A. M. (1991). *Norme vir Afrikaans. Enkele riglyne by die gebruik van Afrikaans.* Pretoria: Academic.

Catalogus van Boekwerken enz., ook geschickt voor school- en dorpsbibliotheeken in Zuid-Afrika 1897. With a preface by C. F. J. Muller (second edition). Amsterdam/Cape Town: J. Dusseau & Co.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioural Research, 1*, 245–276.

Cedergren, H. J. & Sankoff, D. (1974). Variable rules: Performance as a statistical reflection of competence. *Language, 50*, 333–355.

Chafe, W. L. (1982). Integration and involvement in speaking, writing and oral literature. In D. Tannen (Ed.), *Spoken and Written Language. Exploring orality and literacy* (pp. 35–54). Norwood: Ablex Publishing.

Chambers, J. K. (2002). Dynamics of dialect convergence. *Journal of Sociolinguistics, 6*, 117–130.

Chambers, J. K. & Trudgill, P. (1980). *Dialectology*. Cambridge: Cambridge University Press.

Changuion, A. N. E. (1844). *De Nederduitsche Taal in Zuid-Afrika Hersteld*. Cape Town: JH Collard.

Chaudenson, R. (2001). *Creolization of Language and Culture.* Revised in collaboration with S. S. Mufwene. London: Routledge.

Chen, M. Y. (1977). The time dimension contribution toward A theory of sound change. In W. S.-Y. Wang (Ed.), *The Lexicon in Phonological Change* (pp. 197–251). The Hague: Mouton.

Cheshire, J. (1987). Syntactic variation, the linguistic variable and sociolinguistic theory. *Linguistics, 25*, 257–282.

—— (1998 [1995]). English negation from an interactional perspective. In P. Trudgill & J. Cheshire (Eds.), *The Sociolinguistics Reader. Volume I: Multilingualism and variation* (pp. 127–144). London: Arnold.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: The MIT Press.

—— (1980). *Rules and Representations.* Brighton: Harvester Press.

Clyne, M. (1987). Constraints on code-switching: How universal are they? *Linguistics, 25*, 739–764.

—— (2003). *Dynamics of Language Contact*. Cambridge: Cambridge University Press.

Cobley, A. G. (1990). *Class and Consciousness. The black petty bourgeoisie in South Africa, 1924–1950.* New York: Greenwood.

Cochran, E. P. (1997). An instance of triglossia? Codeswitching as evidence for the present state of Greece's 'language question'. *International Journal of the Sociology of Language, 126*, 33–62.

Coetzee, C. (1996). Early Afrikaans theatre and the erasure of the black face: Some notes on *Susanna Reyniers. Tydskrif vir Nederlands en Afrikaans, 3*, 146–157.

Colquhoun, A. R. (1906). *The Afrikaner Land.* London: John Murray.

Combrink, J. G. H. (1978). Afrikaans: Its origin and development. In L. W. Lanham & K. P. Prinsloo (Eds.), *Language and Communication Studies in South Africa* (pp. 69–95). Cape Town: Oxford University Press.

Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.

Conradie, C. J. (1979). *Die Diachronie van die Afrikaanse Voltooide Deelwoord*. Unpublished Ph.D. dissertation, University of Witwatersrand.

—— (1981). Die ontwikkeling van die Afrikaanse voltooide deelwoord I: 1652–1875. *Tydskrif vir Geesteswetenskappe, 21*, 275–284.
—— (1982). Die ontwikkeling van die Afrikaanse voltooide deelwoord II: 1875–1978. *Tydskrif vir Geesteswetenskappe, 22*, 96–109.
—— (1986). *Taalgeskiedenis*. Pretoria: Academica.
—— (1994). Patrone van [t]-apokopee in Kaapse en Europese Nederlands. In C. Van der Merwe, H. Waher, & J. Hambidge (Eds.), *Rondom Roy. Studies opgedra aan Roy H. Pheiffer* (pp. 133–141). Cape Town: Department of Afrikaans and Nederlands.
—— (1998). Tempusgebruik in Afrikaanse narratiewe. *South African Journal of Linguistics, 16*, 37–43.
Cook, N. M. (1992). Form, function and the 'perfective' in German. In I. Rauch, G. F. Carr, & R. L. Kyes (Eds.), *On Germanic Linguistics. Issues and methods* (pp. 45–60). Berlin: Mouton de Gruyter.
Cooper, R. L. (1989). *Language Planning and Social* Change. Cambridge: Cambridge University Press.
Corbett, G. G. (1991). *Gender.* Cambridge: Cambridge University Press.
Coseriu, E. (1974). *Synchronie, Diachronie und Geschichte. Das Problem des Sprachwandels.* München: Wilhelm Fink Verlag.
Coulmas, F. (1992). *Die Wirtschaft mit der Sprache. Eine sprachsoziologische Studie*. Frankfurt/M.: Suhrkamp.
Coupland, N. (1984). Accommodation at work. *International Journal of the Sociology of Language, 4*, 49–70.
Coupland, N. (2001). Dialect stylization in radio talk. *Language in Society, 30*, 345–375.
Croft, W. (2000). *Explaining Language Change. An evolutionary approach*. London: Longman.
Crowley, T. (1989). *The Politics of Discourse. The standard language question and British cultural debates*. London: Macmillan.
—— (2001). Bakhtin and the history of language. In K. Hirschkop & D. Sheperd (Eds.), *Bakhtin and Cultural Theory* (pp. 177–200) (second edition). Manchester: Manchester University Press.
Crystal, D. (1966). Specification and English tenses. *Journal of Linguistics, 2*, 1–34.
Daltas, P. (1980). The concept of diglossia from a variationist point of view with reference to Greek. *Archivum Linguisticum, 11*, 65–88.
Davenport, R. & Saunders, C. (2000). *South Africa. A modern history.* Houndsmills/London: Macmillan.
Davids, A. (1991). *The Afrikaans of the Cape Muslims from 1815 to 1915*. Unpublished MA Thesis, University of Natal (Durban).
Davis, H. (1996). Ordinary people's philosophy: Comparing lay and professional metalinguistic knowledge. *Language Sciences, 19*, 33–46.
Davis, L. M. (1982). American social dialectology: A statistical appraisal. *American Speech, 57*, 83–94.
DeBose, C. E. (1992). Codeswitching: Black English and standard English in the African-American linguistic repertoire. In C. Eastman (Ed.), *Codeswitching* (pp. 157–167). Clevedon: Multilingual Matters.
Debrabandere, F. (1997). De n-apocope in het Kortrijks. *Taal en Tongval, 49*, 139–144.

DeCamp, D. (1971a). Toward the generative analysis of a post-creole speech continuum. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 349–370). Cambridge: Cambridge University Press.

—— (1971b). Implicational scales and sociolinguistic linearity. *Linguistics, 73*, 30–43.

Dekeyser, X. (1980). The diachrony of the gender systems in English and Dutch. In J. Fisiak (Ed.), *Historical Morphology* (pp. 97–111). The Hague: Mouton.

Dekker, G. (1926). Ds. S. J. Du Toit as Leier van die Eerste Afrikaanse Taalbeweging. In *Gedenkboek ter Ere van Die Genootskap van Regte Afrikaners* (pp. 71–80). Published by the Afrikaanse Studentenbond. Potchefstroom: Die Weste Drukkery.

De Kleine, C. (1997). The verb phrase in Afrikaans: Evidence of creolization? In A. K. Spears & D. Winford (Eds.), *The Structure and Status of Pidgins and Creoles* [Creole Language Library 19] (pp. 289–307). Amsterdam: John Benjamins.

De Klerk, V. & Bosch, B. (1998). Afrikaans to English: A case study of language shift. *Journal of South African Linguistics, 16*, 43–51.

Delanty, G. & O'Mahony, P. (2002). *Nationalism and Social Theory. Modernity and the recalcitrance of the nation.* London: Sage.

De Lima, S. (1844). *De Taal der Kapenaren tegen de Schandelijke Aanraking Derzelve van Professor Changuion, Verdedigd door Mr. J. Suasso de Lima.* Cape Town: J. Suasso de Lima.

Den Besten, H. (1978). Cases of possible syntactic interference in the development of Afrikaans. *Amsterdam Creole Studies, 2*, 5–56.

—— (1981). Marking WH-movement in Afrikaans. In P. Muysken (Ed.), *Generative Studies on Creole Languages* (pp. 141–180). Dordrecht: Foris.

—— (1985). Die doppelte Negation im Afrikaans und ihre Herkunft. In N. Boretzky, W. Enninger, & T. Stolz (Eds.), *Akten des 1. Essener Kolloquium über Kreolsprachen und Sprachkontakte* (pp. 9–42). Bochum: Studienverlag Dr. N. Brockmeyer.

—— (1986). Double negation and the genesis of Afrikaans. In P. Muysken & N. Smith (Eds.), *Substrata Versus Universals in Creole Genesis* [Creole Language Library 1] (pp. 185–230). Amsterdam: John Benjamins.

—— (1987a). Die niederländischen Pidgins der alten Kapkolonie. In N. Boretzky (Ed.), *Beiträge zum 3.Essener Kolloquium über Sprachwandel und seine bestimmenden Faktoren, vom 30.9. – 2.10.1987 an der Universität Essen* (pp. 9–40). Bochum: Studienverlag Brockmeyer.

—— (1987b). Review Article (*Einführung in Geschichte und Struktur des Afrikaans*, by Edith H. Raidt. Darmstadt: Wissenschaftliche Buchgesellschaft). *Journal of Pidgin and Creole Languages, 2*, 67–82.

—— (1988). Universal-Grammatik und/oder Zweitsprachenerwerb: Der Fall Afrikaans. In N. Boretzky, W. Enninger, & T. Stolz (Eds.), *Beiträge zum 4. Essener Kolloquium über Sprachkontakt, Sprachwandel, Sprachwechsel, Sprachtod* (pp. 11–44). Bochum: Studienverlag Dr. N. Brockmeyer.

—— (1989). From Khoekhoe foreignertalk via hottentot Dutch to Afrikaans: The creation of a novel grammar. In M. Pütz & R. Dirven (Eds.), *Wheels within Wheels. Papers of the Duisburg Symposium on Pidgin and Creole Languages* (pp. 207–250). Frankfurt/M.: Peter Lang.

—— (1996). Afrikaans relatief 'wat' en de West-Germaanse relativisatiesystemen. *Taal en Tongval* (Themanummer), *9*, 12–34.
—— (2000). The slaves' languages in the Dutch Cape colony and Afrikaans *vir*. *Linguistics, 38*, 949–971.
—— (2002). Khoekhoe syntax and its implications for L2 acquisition of Dutch and Afrikaans. *Journal of Germanic Linguistics, 14*, 3–56.
De Rooij, J. (1980). Ons bruin(e) paard. *Taal en Tongval, 32*, 3–25, 109–129.
De Silva, M. W. S. (1974). Convergence in diglossia: The Sinhalese situation. *International Journal of Dravidian Linguistics, 3*, 60–91.
Deumert, A. (2001). Variation and standardization at the Cape (1888–1922) – A contribution to Afrikaans socio-historical linguistics. *Journal of Germanic Linguistics, 13*, 301–352.
—— (2002). Standardization and social networks. In N. McLelland & A. Linn (Eds.), *Standardization. Studies from the Germanic languages* [Current Issues in Linguistic Theory 235] (pp. 1–26). Amsterdam: John Benjamins.
—— (2003a). Bringing speaker's back in? Epistemological reflections on speaker-oriented explanations of language change. *Language Sciences, 25*, 15–76.
—— (2003b). Describing language standardization. Models and methods. In K. Arneson (Ed.), *Standardisering og variation i vestnordisk* (pp. 9–32). Reykjavik: Iceland University Press.
—— (2003c). Standard languages as civic rituals. Theory and examples. In U. Ammon & K. J. Mattheier (Eds.), *Sprachstandards / Language Standards* [Sociolinguistica 17]. Tübingen: Niemeyer.
—— (2003d). Salience and markedness in second language acquisition and pidginization – Evidence from a non-canonical contact language. *Language Sciences, 25*(6), 531–669.
Deumert, A. & Vandenbussche, W. (Eds.). (2003). *Germanic Standardizations. Past to Present* [IMPACT: Studies in Language and Society 18]. Amsterdam: John Benjamins.
Deutsch, K. W. (1968). The trend of European nationalism. The language aspect. In J. A. Fishman (Ed.), *The Sociology of Language* (pp. 598–606). Rowley, MA: Newbury House.
De Villiers, A. (1934). *Die Hollandse Taalbeweging in Suid Afrika*. Annale van die Universiteit van Stellenbosch XIV, Reeks B, Aflewering 2.
De Villiers, M. (1884–1894). *Fondamentssteenen. Spel-en leesboek voor de Afrikaansche jeugd* (Deel I–II). Paarl: D. F. du Toit.
De Villiers, M. & Ponelis, F. A. (1987). *Afrikaanse Klankleer.* Cape Town: Tafelberg.
Devitt, A. J. (1989). *Standardizing Written English.* Cambridge: Cambridge University Press.
De Vooys, C. G. N. (1913). *Historiese Schets van de Nederlandse Letterkunde* (sixth improved edition). Groningen: J. B. Wolters.
—— (1953). *Nederlandse Spraakkunst* (third edition by M. Schönfeld). Groningen: J. B. Wolters.
De Vos, P. J. G. (1891). *Zullen wij nog voor onze Moedertaal ijveren?* Cape Town: Townshend en Zoon.
De Vries, J. (1974). De slot -t in consonantclusters te Leiden. Een sociolinguïstisch onderzoek. *Forum der Letteren, 15*, 235–350.
De Vries, M. (1853). *De Nederlandsche Taalkunde in Haren Aard en Hare Strekking*. Haarlem.

De Waal, J. H. H. (1910). *Africander Grammar.* Cape Town: Cape Times.
—— (1932). *My Herinnerings van ons Taalstryd.* Cape Town: Nasionale Pers.
De Wet, G. C. (1981). *Die Vryliede en Vryswartes in die Kaapse Nedersetting 1657–1707.* Cape Town: Die Historiese Publikasie Vereniging.
Diaconis, P. (1995). Theories of data analysis. From magical thinking through classical statistics. In W. J. Krzanowski (Ed.), *Recent Advances in Descriptive Multivariate Analysis* (pp. 1–36). Oxford: Clarendon Press.
Di Eerste Afrikaanse Taalkongres. Gehou aan die Paarl, 15. en 16. Jannewari 1896. Paarl: Paarl Drukpers Maatschappij, Beperkt.
Di Tweede Afrikaanse Taalkongres. Gehou aan die Paarl, 27. en 28. Jannewari 1897. Paarl: Paarl Drukpers Maatschappij, Beperkt.
Diehl, E., Christen, H., Leuenberger, S., Pelvat, I., & Studer, T. (2000). *Grammatikunterricht. Alles für die Katz? Untersuchungen zum Zweitspracherwerb Deutsch*. Tübingen: Niemeyer.
Dillard, J. L. (1976). *American Talk.* New York: Random House.
Dines, E. R. (1980). Variation in discourse – 'and stuff like that'. *Language in Society, 9*, 13–31.
Donaldson, B. C. (1981). *Dutch Reference Grammar.* The Hague: Nijhoff.
—— (1991). *The Influence of English on Afrikaans. A case study of linguistic change in a language contact situation* (second edition). Pretoria: Academica.
—— (1993). *A Grammar of Afrikaans.* Berlin: Mouton de Gruyter.
—— (1995). Language contact and linguistic change. The influence of English on Afrikaans. In R. Mesthrie (Ed.), *Language and Social History. Studies in South African sociolinguistics* (pp. 222–229). Cape Town: David Phillips.
Dorian, N. C. (1982). Defining the speech community to include its working margins. In S. Romaine (Ed.), *Sociolinguistic Variation in Speech Communities* (pp. 25–34). London: Arnold.
—— (2002). Diglossia and the simplification of linguistic space. *International Journal of the Sociology of Language, 157*, 63–69.
Dunn-Rakin, P. (1983). *Scaling Methods.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Du Plessis, I. D. (1944). *The Cape Malays*. Cape Town: Maskew Miller.
Du Plessis, L. T. (1986). *Afrikaans in Beweging.* Bloemfontein: Patmos.
Du Toit, A. (1983). No chosen people: The myth of the Calvinist origins of Afrikaner nationalism and racial ideology. *American Historical Review, 88*, 920–952.
—— (1985). Workshop on research in intellectual history: 19th century discourse on nationality, religion and language at the Cape – Or how to get one hand clapping in time. In B. F. Nel, R. Singh, & W. M. Venter (Eds.), *Focus on Quality. Selected proceedings of a conference on qualitative research methodology in the social sciences* (pp. 190–243). Durban: Institute for Social and Economic Research, University of Durban-Westville.
Du Toit, P. S. (1938). *Onderwys aan die Kaap onder die Kampanjie, 1652–1795.* Cape Town: Juta.
Du Toit, S. J. (1980 [1876/1897]). *Eerste Beginsels van die Afrikaanse Taal. Fergelijkende Taalkunde fan Afrikaans en Engels.* Facsimile edition, introduced by J. E. Loubser. Cape Town: Tafelberg.
—— (1891). *Afrikaans ons Volkstaal. 71 Theses, of stellinge.* Paarl: D. F. Du Toit & Co.

—— (1909 [1880]). *Geskiedenis van die Afrikaanse Taalbeweging*. Paarl: Paarl Drukpers Maatschappij, Beperkt.

Eckert, P. (1997). Age as a Sociolinguistic Variable. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 151–167). Oxford: Basil Blackwell.

—— (2000). *Language variation as social practice*. Oxford: Blackwell.

Eckert, P. & McConnell-Ginet (1992). Think practically and look locally. Language and gender as a community based practice. *Annual Review of Anthropology, 21*, 461–490.

—— (2003). *Language and Gender*. Cambridge: Cambridge University Press.

Ehrenberg, A. S. C. (1982). *A Primer in Data Reduction. An introductory statistics textbook.* Chichester: John Wiley & Sons.

Eisenstein, E. L. (1979). *The Printing Press as an Agent of Change. Communications and cultural transformations in early modern Europe*. Cambridge: Cambridge University Press.

Elffers, H. (1893). *Practical Dutch grammar and analysis. A grammar for English and Cape students to learn the Dutch language*. Cape Town: Juta.

—— (1903). *The Englishman's guide to the speedy and easy acquirement of Cape Dutch* (*Grammar, useful information, conversation*). *For the use of travellers, settlers, and military men.* Cape Town: Juta.

Ellis, M. (1994). Literary dialect as linguistic evidence: Subject-verb concord in nineteenth-century southern literature. *American Speech, 69*, 128–144.

Elphick, R. (1972). *The Cape Khoi and the First Phase of South African Race Relations*. Ph.D. Thesis, Yale University.

—— (1977). *Kraal and Castle: Khoikhoi and the founding of white South Africa*. New Haven: Yale University Press.

Elphick, R. & Malherbe, V. C. (1989). The Khoisan to 1828. In R. Elphick & H. Giliomee (Eds.), *The Shaping of South African Society 1652–1820* (pp. 3–65) (second edition). Cape Town: Maskew Miller Longman.

Elspaß, S. (2002). Standard German in the 19th-century? (Counter-)evidence from the private correspondence of 'ordinary people'. In A. Linn & N. McLelland (Eds.), *Standardization. Studies from the Germanic languages* [Current Issues in Linguistic Theory 235] (pp. 43–65). Amsterdam: John Benjamins.

Engelenburg, F. V. (1897). De spelling-kwestie met het oog op Zuid-Afrika. *De Gids, 4*, 357–365.

Erben, J. (1980). *Deutsche Grammatik. Ein Abriss.* München: Hüber.

Erickson, B. H. (1997). Social networks and history. A review essay. *Historical Methods. A Journal of Quantitative and Interdisciplinary History, 30*, 149–157.

Escure, G. (1997). *Creole and Dialect Continua. Standard acquisition processes in Belize and China* (*PRC*) [Creole Language Library 18]. Amsterdam: John Benjamins.

Everett, M. G. & Borgatti, S. P. (2000). Peripheries of cohesive subsets. *Social Networks, 21*, 397–407.

Everitt, B. S. (1986). Numerical approaches to classification. In R. J. Brook & G. C. Arnold (Eds.), *The Fascination of Statistics* (pp. 51–64). New York: Marcel Dekker.

Fairman, T. (2002). Riting these fu lines: English overseers' correspondence, 1800–1835. In R. Willemyns (Ed.), *De taal in Vlaanderen in de 19de eeuw. Historisch-sociolinguïstische* onderzoekingen [Verslagen en Mededelingen KANTL 112, 3] (pp. 557–573). Gent: KANTL.

February, V. A. (1991). *The Afrikaners of South Africa*. London: Routledge.

Feinauer, L. (1994). Die plasing van nie-1 en nie-2 in die Afrikaanse spreektaal. *South African Journal of Linguistics*, *19*, 92–119.

Ferguson, C. A. (1959). Diglossia. *Word, 15*, 325–340.

—— (1963). Problems in teaching languages with diglossia. In A. S. Dil (Ed.), *Language Structure and Language Use. Essays by Charles A. Ferguson* (pp. 71–86). Stanford: Stanford University Press.

—— (1983). Language planning and social change. In J. Cobarrubias & J. A. Fishman (Eds.), *Progress in Language Planning* (pp. 29–40). Berlin: Mouton De Gruyter.

—— (1988). Standardization as a form of language spread. In P. Lowenberg (Ed.), *Language Spread and Language Policy. Issues, implications, and case studies. Georgetown University Round Table on Languages and Linguistics 1987* (pp. 119–132). Washington, DC: Georgetown University Press.

—— (1991). Diglossia revisited. *Southwest Journal of Linguistics, 10*, 214–234.

—— (1995). Variation and Drift. Loss of Agreement in Germanic. In G. Guy (Ed.), *Towards a Social Science of Language. Papers in honor of William Labov* [Current Issues in Linguistic Theory 127], Vol. I (pp. 173–197). Amsterdam: John Benjamins.

First Report and Proceedings with Appendices of a Commission to Enquire into and Report upon Certain Matters concerned with the Educational System of the Colony. Cape Town: The Cape Times Government Printers, 1891.

Fishman, J. A. (1967). Bilingualism with and without diglossia. Diglossia with and without language contact. *Journal of Social Issues, 23*, 29–38.

—— (1980). Bilingualism and biculturalism as individual and societal phenomena. *Journal of Multilingual and Multicultural Development, 1*, 3–15.

—— (1987). *Ideology, Society and Language. The odyssey of Nathan Birnbaum.* Ann Arbor: Karoma Publisher.

—— (1989). *Language and Ethnicity in a Minority Sociolinguistic Perspective.* Clevedon: Multilingual Matters.

Fitch, W. M. (1979). Numerical taxonomy. A special project (II). *Systematic Zoology, 28*, 254–255.

Fitzmaurice, S. (2000a). Coalitions and the investigation of social influence in linguistic history. *European Journal of English Studies, 4*, 265–276.

—— (2000b). The Spectator, the politics of social networks, and language standardisation in eighteenth-century England. In L. Wright (Ed.), *The Development of Standard English, 1300–1800* (pp. 195–218). Cambridge: Cambridge University Press.

Foucault, M. (1984). *Language and Politics.* Oxford: Basil Blackwell.

Frangoudaki, A. (1992). Diglossia and the present language situation in Greece. A sociological approach to the interpretation of diglossia and some hypothesis on today's linguistic reality. *Language in Society, 21*, 365–381.

Francken, A. (1900). *An Easy Guide to the South African Dutch Language.* Bloemfontein: Borckenhagen.

—— (1908). *Susanna Reyniers.* Amsterdam/Pretoria: Du Bussy.
Franken, J. L. M. (Ed.). (1938). *Duminy-dagboeke.* Cape Town: Van Riebeek Vereniging.
—— (1953). *Taalhistoriese Bydraes.* Amsterdam: A. A. Balkema.
Frings, T. (1957). *Grundlegung einer Geschichte der deutschen Sprache*. Halle: Niemeyer.
Gal, S. (1989). Language and political economy. *Annual Review of Anthropology, 18*, 345–367.
—— (1995). Cultural bases of language-use among German speakers in Hungary. *International Journal of the Sociology of Language, 111*, 93–102.
Gardner-Chloros, P. (1995). Code-switching in community, regional and national repertoires. The myth of the discreteness of linguistic systems. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages* (pp. 68–89). Cambridge: Cambridge University Press.
Garvin, Paul (1964). The standard language problem. Concepts and methods. In D. Hymes (Ed.), *Language in Culture and Society* (pp. 521–526). Harper: New York.
—— (1991). A conceptual framework for the study of language standardization. *Indian Journal of Applied Linguistics, 17*, 1–20.
Gee, J. P. (1994). Orality and literacy. From the savage mind to ways with words. In J. Maybin (Ed.), *Language and Literacy in Social Practice* (pp. 168–192). Clevedon: Multilingual Matters.
Geerts, G., Haeseryn, W., De Rooij, J., & Van den Toorn, M. C. (1984). *Algemene Nederlandse Spraakkunst*. Groningen: Wolters Noordhoff.
Gellner, E. (1983). *Nations and Nationalism*. Oxford: Basil Blackwell.
Giacalone Ramat, A. (1995). Code-switching in the context of dialect-standard relations. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages* (pp. 45–67). Cambridge: Cambridge University Press.
Giesbers, H. (1989). *Code-switching tussen Dialect en Standaardtaal.* Amsterdam: P. J. Meertens-Instituut.
Giles, Howard & Powesland, Peter (1975). A social psychological model of speech diversity. In H. Giles & P. Powesland (Eds.), *Speech Style and Social Evaluation* (pp. 154–170). New York: Harbourt Brace.
Giliomee, H. (1989). Aspects of the rise of Afrikaner capital and Afrikaner nationalism in the Western Cape 1870–1915. In W. G. James & M. Simons (Eds.), *The Angry Divide* (pp. 63–79). Cape Town: David Phillips.
Girard, D. & Larmouth, D. (1993). Some applications of mathematical and statistical models in dialect geography. In D. Preston (Ed.), *American Dialect Research* (pp. 107–131). Amsterdam: John Benjamins.
Glorie, I. (1996). 'n Klein alleenstaande koppie op 'n kaal vlakte? Melt Brink en zijn venster op de wereld. *Tydskrif vir Nederlands en Afrikaans, 3*, 122–145.
Gloy, K. (1975). *Sprachnormen I. Linguistische und soziologische Analysen*. Stuttgart/Bad Cannstatt: Fromann-Holzboog.
—— (1997). Sprachnormen als 'Institutionen im Reich der Gedanken' und die Rolle des Individuums in Sprachnormierungsprozessen. In K. J. Mattheier (Ed.), *Norm und Variation* (pp. 27–26). Frankfurt/M.: Peter Lang.
Godée-Molsbergen, E. C. (Ed.). (1916). *Reizen in Zuid-Afrika in de Hollandse Tijd.* Two volumes. The Hague: Martinus Nijhoff.

Goeman, A. C. M. (1983 [1976]). Aspecten van de vervoeging van her presens. In J. Stroop (Ed.), *Nederlands Dialectonderzoek. Artikelen uit de periode 1927–1982* (pp. 185–210). Amsterdam: Huis aan de Drie Grachten.

—— (1992). De presensvervoeging in de dialecten van Nederland. Ja-nee versus gradueel. *Taal en Tongval, 44*, 97–139.

Goode, E. & Nachman, B. Y. (1994). *Moral Panics. The social construction of deviance*. Oxford: Basil Blackwell.

Gracía-Bermejo Giner, M. F. & Montgomery, M. (1997). British regional English in the nineteenth century. The evidence from emigrant letters. In A. R. Thomas (Ed.), *Issues and Methods in Dialectology* (pp. 167–183). Bangor: University of Wales Bangor, Department of Linguistics.

Graff, H. J. (1979). *The Literacy Myth. Literacy and social structure in the nineteenth century city.* New York: Academic Press.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*, 1360–1380.

—— (1982). The strength of weak ties. A network theory revisited. In P. Marsden & N. Lin (Eds.), *Social Structure and Network Analysis* (pp. 105–130). Beverly Hills, CA: Sage.

Grayson, J. D. (1962). *The Remains of Inflection in Afrikaans*. Unpublished Ph.D. dissertation, University of New York.

Grebe, H. P. (2002). Standaardafrikaans. Daar's 'n vlieg in die salf. *South African Linguistics and Applied Linguistics, 20*, 37–46.

Griessel, J. S. (1981). *Die Voorbepaling in Vroeë Afrikaans*. Unpublished Masters Thesis, University of South Africa (UNISA).

Grootaers, W. A. (1959). Origin and nature of the subjective boundaries of dialects. *Orbis, 8*, 355–384.

Grosjean, F. (1989). Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3–15.

—— (1995). A psycholinguistic approach to code-switching. The recognition of guest-words by bilinguals. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages* (pp. 259–275). Cambridge: Cambridge University Press.

—— (2001). The bilingual's language modes. In J. Nicol (Ed.), *One Mind, Two Languages: Bilingual language processing* (pp. 1–22). Oxford: Blackwell.

Guelke, L. (1989). Freehold farmers and frontier settlers, 1657–1780. In R. Elphick & H. Giliomee (Eds.), *The Shaping of South African Society* (pp. 66–108). Cape Town: London Maskew Miller.

Gumperz, J. J. (1962). Types of linguistic communities. *Anthropological Linguistics, 4*, 28–40.

—— (1968). The speech community. In D. L. Sills (Ed.), *International Encyclopaedia of Social Sciences* (pp. 381–386). New York: MacMillan.

Gumperz, J. J. & Levinson, S. C. (1996). Introduction. Linguistic relativity re-examined. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 1–18). Cambridge: Cambridge University Press.

Guttman, L. (1944). A basis for scaling quantitative data. *American Sociological Review, 9*, 139–150.

Guy, G. R. (1980). Variation in the group and individual. The case of final stop deletion. In W. Labov (Ed.), *Locating Language in Time and Space* (pp. 1–36). New York: Academic Press.

Haarmann, H. (1990). Language planning in the light of a general theory of language. A methodological framework. *International Journal of the Sociology of Language, 86*, 103–126.

Häcki-Buhofer, Annelies (2000). Psycholinguistische Aspekte der Variation. Das Sprachbewusstsein in der Variationstheorie. In A. Häcki-Buhofer (Ed.), *Vom Umgang mit sprachlicher Variation. Soziolinguistik, Dialektologie, Methoden und Wissenschaftsgeschichte* (pp. 173–185). Tübingen: Francke.

Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & van den Toorn, M. C. (1997). *Algemene Nederlandse Spraakkunst* (second edition). Groningen/Deurne: Martinus Nijhoff/Wolters Plantyn. Two volumes. (http://oase.uci.kun.nl/~ans/; last accessed 5 July 2003)

Hakkarainen, H. J. (1983). Linguistic changes as diffusion of innovations. *Neuphilologische Mitteilungen, 84*, 25–35.

Haugen, E. (1956). *Bilingualism in the Americas. A bibliography and research guide.* Alabama: University of Alabama Press.

—— (1966). Linguistics and language planning. In J. A. Fishman (Ed.), *Readings in the Sociology of Language* (pp. 673–687). The Hague: Mouton.

—— (1972). *The Ecology of Language*. Stanford: Stanford University Press.

Havránek, B. (1964). The functional differentiation of the standard language. In P. Garvin (Ed.), *A Prague School Reader on Aesthetics, Literary Structure, and Style* (pp. 3–16). Washington: Georgetown University Press.

Hawkins, P. (1983). Diglossia Revisited. *Language and Communication, 3*, 219–232.

Hechter, M. & Opp, K.-D. (2001). Introduction. In M. Hechter & K.-D. Opp (Eds.), *Social Norms* (pp. xi–xx). New York: Sage.

Heese, H. F. (1984). *Groep sonder Grense. Die rol en status van die gemengde bevolking aan die Kaap 1652–1795.* Bellville: Institute of Historical Research.

Henn-Memmesheimer, B. (1986). *Nonstandardmuster. Ihre Beschreibung in der Syntax und das Problem ihrer Arealität.* Tübingen: Max Niemeyer.

Hesseling, D. C. (1905). *Het Negerhollands der Deense Antillen. Bijdrage tot de geschiedenis der Nederlandse taal in Amerika*. Leiden: Sijthoff.

—— (1923 [(1899]). *Het Afrikaansch. Bijdrage tot de geschiedenis der Nederlandsche taal in Zuid-Afrika* (second edition). Leiden: E. J. Brill.

Hill, J. H. (1993). Structure and practice in language shift. In K. Hyltenstam & Å. Viberg (Eds.), *Progression and Regression in Language. Sociocultural and linguistic dimensions* (pp. 68–93). Cambridge: Cambridge University Press.

Hobsbawm, E. (1983). Introduction. Inventing traditions. In E. Hobsbawm (Ed.), *The Invention of Tradition* (pp. 1–14). Cambridge: Cambridge University Press.

—— (1990). *Nations and Nationalism since 1780.* Cambridge: Cambridge University Press.

Hoekstra, J. & Tiersma, P. M. (1994). Frisian. In E. König & J. Van der Auwera (Eds.), *The Germanic Languages* (pp. 505–531). London: Routledge.

Hoeksema, J. (1995). Negation and negative concord in Middle Dutch. In D. Forget, P. Hirschbühler, F. Martineau, & M.-L. Rivero (Eds.), *Negation and Polarity. Syntax and semantics* [Current Issues in Linguistic Theory 155] (pp. 139–156). Amsterdam: John Benjamins.

Hoenigwald, H. M. (1966). A proposal for the study of folk-linguistics. In W. Bright (Ed.), *Sociolinguistics* (pp. 16–26). The Hague: Mouton.

Hofmeyr, I. (1987). Building a nation from words. Afrikaans language, literature and ethnic identity, 1902–1924. In S. Marks & S. Trapido (Eds.), *The Politics of Race, Class and Nationalism in Twentieth-Century South Africa* (pp. 95–123). London: Longman.

Hol, A. R. (1955). Het meervoud van het praesens in onze oostelijke dialecten. *Taal en Tongval, 7*, 86–96, 160–175.

Holliday, L. (1993). The first language congress for Afrikaans. In J. A. Fishman (Ed.), *The Earliest Stage of Language Planning: The 'first congress' phenomenon* (pp. 11–30). Berlin: de Gruyter.

Holm, J. A. (1988/89). *Pidgins and Creoles.* Two volumes. Cambridge: Cambridge University Press.

—— (2001). The semicreole identity of Afrikaans lects. Parallel cases of partial restructuring. *Journal of Germanic Linguistics, 13*, 353–379.

Holm, J. A. (2002). The social and linguistic origins of Brazilian vernacular Portuguese. Parallel cases of partial restructuring. *Revista Portuguesa De Filologia, 14*, 123–145.

Holmes, P. & Hinchliffe, I. (1997). *Swedish. An essential grammar*. London: Routledge.

Hoogenhout, C. J. (1876). *Geskiedenis van Josef voor Afirkaanse Kinders en Huisouwens, in Hulle Eige Taal Geskywe deur een Vrind*. Cape Town: N. H. Marais.

—— (1880). *Eerste Afrikaanse Printjies Boekie ver Soet Kinders, deur Jan wat Versies Maak.* Paarl: D. F. Du Toit & Co.

Hoogenhout, N. M. (1904). *Praktisches Lehrbuch der Kapholländischen Sprache (Burensprache). Sprachelehre, Gespräche, Lesestücke und Wörterbuch.* Wien: A. Hartleben.

Horn, L. R. (1989). *A Natural History of Negation.* Chicago: Chicago University Press.

Horne, C. (2001). Sociological perspectives on the emergence of social norms. In M. Hechter & K.-D. Opp (Eds.), *Social Norms* (pp. 3–34). New York: Sage.

Horvath, B. (1985). *Variation in Australian English. The sociolects of Sydney.* Cambridge: Cambridge University Press.

Hudson, A. (2002). Outline of a theory of diglossia. *International Journal of the Sociology of Language, 157*, 1–48.

Hudson, R. H. (1996). *Sociolinguistics*. Cambridge: Cambridge University Press.

Hymes, D. (1974). *Foundations in Sociolinguistics. An ethnographic approach*. London: Travistock.

Irvine, J. T. (2001). "Style" as distinctiveness. The culture and ideology of linguistic differentiation. In P. Eckert & J. R. Rickford (Eds.), *Style and Sociolinguistic Variation* (pp. 21–43). Cambridge: Cambridge University Press.

Jackson, J. E. (1991). *A User's Guide to Principal Components Analysis.* New York: John Wiley & Sons.

Jacobs, N. G., Prince, E. F., & Van der Auwera, J. (1994). Yiddish. In E. König & J. Van der Auwera (Eds.), *The Germanic Languages* (pp. 388–419). London: Routledge.

Jahr, E. H. (1989). Language planning and language change. In L. E. Breivik & E. H. Jahr (Eds.), *Language Change: Contributions to the study of its causes* (pp. 99–113). Berlin: Mouton de Gruyter.

Janda, R. D. & Auger, J. (1992). Quantitative evidence, qualitative hypercorrection, sociolinguistic variables – and French speakers. *Language and Communication, 12*, 195–236.

Janson, T. (1991). Language change and metalinguistic change: Latin to Romance and other cases. In R. Wright (Ed.), *Latin and Romance in the early Middle Ages* (pp. 19–28). London: Routledge.

—— (2002). *Speak. A Short history of languages*. Oxford: Oxford University Press.

Jasso, G. (2001). Rule finding about rule making. Comparison processes and the making of rules. In M. Hechter & K.-D. Opp (Eds.), *Social Norms* (pp. 348–393). New York: Sage.

Jernudd, B. H. (1989). The texture of language purism. An introduction. In B. J. Jernudd & M. J. Shapiro (Eds.), *The Politics of Language Purism* (pp. 1–19). Berlin: Mouton de Gruyter.

Jespersen, O. (1925). *Mankind, Nation and Individual from a Linguistic Point of View*. Oslo: H. Aschehoug & Co.

—— (1945). *Growth and Structure of the English Language*. Oxford: Blackwell.

Jobson, J. D. (1992). *Applied Multivariate Data Analysis*. New York: Springer.

Jolliffe, I. T. (1986). *Principal Components Analysis.* New York: Springer.

Joseph, J. E. (1984). The elaboration of an emerging standard. *Language Sciences, 6*, 39–52.

—— (1987). *Eloquence and Power. The rise of language standards and standard languages.* Oxford: Basil Blackwell.

—— (1989). Popular and scientific beliefs about language status. An historical sketch. In U. Ammon (Ed.), *Status and Function of Languages* (pp. 243–255). Berlin: Mouton de Gruyter.

Joyce, P. (1991). The people's English. Language and class in England, c. 1840–1920. In P. Burke & R. Porter (Eds.), *Language, Self, and Society. A social history of language* (pp. 154–190). Cambridge: Polity Press.

Kähler, H. (1971). *Studien über die Kultur, die Sprache, und die Arabisch-Afrikaanse Literatur der Kap-Maleien*. Berlin: Reimer.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151.

Kannemeyer, J. C. (1984). *Geskiedenis van die Afrikaanse Literatuur* (second edition). Volume one. Pretoria: Academica.

Kastovsky, D. (1994). Typological differences between English and German morphology and their causes. In T. Swan, E. Mørck, & O. J. Westvik (Eds.), *Language Change and Language Structure. Older Germanic languages in a comparative perspective* (pp. 135–157). Berlin: Mouton de Gruyter.

Katzen, M. F. (1969). White settlers and the origin of a new society 1652–1778. In M. Wilson & L. Thompson (Eds.), *The Oxford History of South Africa* (pp. 187–232). (Volume One: South Africa to 1870). Oxford: Claredon Press.

Kaufman, L. & Rousseeuw, P. J. (1990). *Finding Groups in Data. An introduction to cluster analysis*. New York: John Wiley & Sons, Inc.

Kazazis, K. (1992). Sunday Greek Revisited. *Journal of Modern Greek Studies, 10*, 57–69.

Keller, R. E. (1982). Diglossia in German-speaking Switzerland. In W. Haas (Ed.), *Standard Languages: Spoken and written* (pp. 70–93). Manchester: Manchester University Press.

Kempton, W. (1978). Category grading and taxonomic relations. A mug is a sort of cup. *American Anthropologist, 5*, 44–65.

Kloeke, G. G. (1950). *Herkomst en groei van het Afrikaans.* Leiden: Universitaire Pers.

—— (1956). *Een oud sjibboleth. De gewestelijke uitspraak van 'heeft'*. Verhandelingen der Koninglijke Nederlandse Akademie van Wetenschappen, Afdeling Letterkunde, NR 63, 1.

Kloss, H. (1969). *Research Possibilities on Group Bilingualism. A report.* Quebec: International Center for Research on Bilingualism.

—— (1978). *Die Entwicklung neuer germanischer Kultursprachen seit 1800*. Düsseldorf: Pädagogischer Verlag Schwann.

Kolbe, M. P. (1719). *Caput Bonae Spei Hordiernum. Das ist die vollständige Beschreibung des Africanischen Vorgebürges zur Guten Hoffnung.* Two volumes. Nürnberg: Monath.

Kollewijn, R. A. (1903). *Opstellen over Spelling en Verbuiging* (second edition). Amsterdam.

Kooiman, H. (1954). Ik heb geweest, ik ben geweest. *De Nieuwe Taalgids, 13*, 183–188.

Kotzé, E. (1983). *Variasiepatrone in Maleier-Afrikaans*. Ph.D. Thesis, University of Witwatersrand.

—— (1989). How creoloid can you be? Aspects of Malay Afrikaans. In M. Pütz & R. Dirven (Eds.), *Wheels Within Wheels. Papers of the Duisburg symposium on Pidgin and Creole languages* (pp. 251–264). Frankfurt/M.: Peter Lang.

—— (1994). Afsonderlik ontwikkel deur verbastering? Afrikaans en Middelnederduits in die kreoolse konteks. In G. Olivier & A. Coetzee (Eds.), *Nuwe Perspektiewe op die Geskiedenis van Afrikaans* (pp. 130–138). Johannesburg: Southern Books.

Kouwenberg, S. (1993). Berbice Dutch. In P. Kahrel & R. Van den Berg (Eds.), *Typological Studies in Negation* [Typological Studies in Language 29] (pp. 236–266). Amsterdam: John Benjamins.

—— (1994). Berbice Dutch. In J. Arends, P. Muysken, & N. Smith (Eds.), *Pidgins and Creoles. An introduction* [Creole Language Library 15] (pp. 233–243). Amsterdam: John Benjamins.

Krackhardt, D., Lundberg, M., & O'Rouke, L. (1993). Krackplot. A picture's worth a thousand words. *Connections, 16*, 37–47.

Kretzschmar, W. A. & Schneider, E. W. (1996). *Introduction to Quantiative Analysis of Linguistic Survey Data. An atlas by the numbers.* London: Sage.

Krumbacher, K. (1902). *Das problem der modernen griechischen schriftsprache.* München: K. B. Akademie.

Kruskal, J. B. & Wish, M. (1978). *Multidimensional Scaling* [Sage University Paper Series on Quantitative Applications in the Social Sciences 07-011]. Beverly Hills: Sage.

Kuryłowicz, J. (1949). La nature des procès dits analogiques. *Acta Linguistica, 5*, 15–37.

Kytö, M. (1991). *Variation and Diachrony, with Early American English in Focus. Studies on* CAN/MAY *and* SHALL/WILL. Frankfurt/M.: Peter Lang.

—— (1993). Third-person present tense singular verb inflection in early British and American English. *Language Variation and Change, 5*, 113–139.

Kytö, M. & Rissanen, M. (1983). The syntactic study of early American English. The variationist at the mercy of his corpus? *Neuphilologische Mitteilungen, 89*, 470–490.

Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
—— (1969). Contraction, deletion, and inherent variability of the English copula. *Language, 45*, 715–762.
—— (1971). The notion of 'system' in creole languages. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 447–472). Cambridge: Cambridge University Press.
—— (1991 [1972]). *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.
—— (1972). Some principles of linguistic methodology. *Language in Society, 1*, 97–120.
—— (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (Eds.), *Language in Use* (pp. 28–53). Englewood Cliffs, NJ: Prentice Hall.
Labov, W. & Harris, W. A. (1986). De facto segregation of black and white vernaculars. In D. Sankoff (Ed.), *Diversity and Diachrony* [Current Issues in Linguistic Theory 53] (pp. 1–24). Amsterdam: John Benjamins.
Lätti, M. C. (1978). *Die Taalgebruik in 'Ons Klyntji' 1896–1906*. Unpublished MA thesis, University of Witwatersrand.
Langenhoven, C. J. (1922 [1914]). Afrikaans as voertaal. In *Ons Weg deur die Wêreld* (pp. 9–68). Cape Town: Nasionale Pers.
Langenhoven, C. J. (1935). *My Aandeel aan die Taalstryd.* Cape Town: Nasionale Pers.
Lass, R. (1976). Variation studies and historical linguistics. *Language in Society, 5*, 219–256.
—— (1987). *The Shape of English: Structure and history.* London: J. M. Dent.
—— (1990). How to do things with junk: Exaptation in language evolution. *Journal of Linguistics, 26*, 79–102.
—— (1992). What are language histories histories of? In H.-H. Lieb (Ed.), *Prospects for a New Structuralism* [Current Issues in Linguistic Theory 96] (pp. 243–272). Amsterdam: John Benjamins.
—— (1997). *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
—— (2000). *Language Periodization and the Concept 'Middle'*. In I. Taavitsainen, T. Nevalainen, & M. Rissanen (Eds.), *Placing Middle English in Context* (pp. 7–42). Berlin: Mouton de Gruyter.
—— (forthcoming). Texts as linguistics objects. In C. Kay & J. J. Smith (Eds.), *Proceedings of the Glasgow Conference on Linguistic Categorisation*. Amsterdam: John Benjamins.
Lass, R. & Wright, S. (1986). Endogeny vs. contact. 'Afrikaans influence' on South African English. *English World Wide, 7*, 201–223.
Lavandera, B. R. (1978). Where does the social variable stop? *Language in Society, 7*, 171–182.
Lebbe, D. (1997). Schoon Vlaams. Eeen onderzoek naar syntactische en morfosyntactische aspecten van tussentaal door Ieperlingen gesproken. *Taal en Tongval, 49*, 158–175.
Le Cordeur, B. A. (1981). *The Politics of Eastern Cape Separatism 1820–1854*. Cape Town: Oxford University Press.
Lee, J. & Kretzschmar, W. A. (1993). Spatial analysis of linguistic data with GIS function. *International Journal of Geographical Information Science, 7*, 541–560.
Legassick, M. (1989). The emergence of the Griqua people. In R. Elphick & H. Giliomee (Eds.), *The Shaping of South African Society 1652–1820* (pp. 358–420). Cape Town: Longman.

Leith, D. (1995). Tense variation as a performance feature in a Scottish folktale. *Language in Society, 24*, 53–77.
Lenker, U. (2000). The monasteries of the Benedictine Reform and the 'Winchester School'. Model cases of social networks in Anglo-Saxon England. *European Journal of English Studies, 4*, 225–238.
LePage, R. B. (1988). Some premises concerning the standardization of languages, with special reference to Caribbean Creole English. *International Journal of the Sociology of Language, 71*, 25–36.
LePage, R. B. & Tabouret-Keller, A. (1985). *Acts of Identity. Creole-based approaches to language and ethnicity.* Cambridge: Cambridge University Press.
LeRoux, J. J. (1923). *Oor die Afrikaanse Sintaksis.* Amsterdam: Swets and Zeitlinger.
—— (1944). *Praatjies oor ons Taal*. Cape Town: Nasionale Pers.
LeRoux, T. H. (1910). *Beschrijvende Klankleer van het Afrikaans*. Leiden: A. W. Sijthoff.
LeRoux, T. H. & LeRoux, J. J. (1973). *Middelnederlandse Grammatika.* Pretoria: Van Schaik.
Lévi-Strauss, C. (1972). *The Savage Mind*. London: Weidenfield and Nicolson.
Lichtenstein, H. (1967 [1811–1812]). *Reisen im südlichen Afrika in den Jahren 1803, 1804, 1805 and 1806.* Facsimile edition; introduced by W. Drascher. Stuttgart: Brockhaus.
Lieb, H.-H. (1993). *Linguistic Variables. Towards a unified theory of linguistic variation* [Current Issues in Linguistic Theory 108]. Amsterdam: John Benjamins.
Lightfoot, D. (1979). *Principles of Diachronic Syntax.* Cambridge: Cambridge University Press.
—— (1991). *How to set Parameters. Arguments from language change.* Cambridge, MA: The MIT Press.
Links, T. (1989). *So praat ons Namakwalanders.* Cape Town: Tafelberg.
Lippi-Green, R. (1994). *Language Ideology and Language Change in Early Modern German. A sociolinguistic study of the consonantal system of Nuremberg* [Current Issues in Linguistic Theory 119]. Amsterdam: John Benjamins.
Lodge, R. A. (1993). *French. From dialect to standard*. London: Routledge.
Logeman, W. S. (1902). *How to Speak Dutch. A simple method of learning to speak, read, and understand the language* (fourth edition). With translations into Cape Dutch of all sentences and conversations by J. F. Van Oordt. Cape Town/Amsterdam: Hollandsch Afrikaansche Uitgevers Maatschappij.
Lombard, R. T. J. (1984). *Handbook for Genealogical Research in South Africa.* Pretoria: Human Sciences Research Council/Institute for Historical Research.
Loubser, J. E. (1954). *Aspekte van die Afrikaanse Spreektaal Gebruik in en om Loriesfontein: 'n Streekproef in die taal van die Noordweste*. MA Thesis, University of Stellenbosch.
Louden, M. L. (2000). Umlaut, ablaut, and phonetic symbolism in German. *General Linguistics, 37*, 1–22.
Lowenstamm, J. (1977). Relative clauses in Yiddish. A case for movement. *Linguistic Analysis, 3*, 197–216.
Lub, J. (1908). *Eenvoudige Mense. Hollandse-Afrikaanse schetsen.* Potchefstroom: Het Westen.
Lubbe, H. J. (1983). Rule generalisation. Further supporting evidence from the history of Afrikaans. *Tijdschrift voor Nederlands en Afrikaans, 2*, 15–22.

Lüdtke, H. (1989). Invisible-hand processes and the universal laws of language change. In L. E. Breivik & E. H. Jahr (Eds.), *Language Change. Contributions to the study of change and its causes* (pp. 131–136). Berlin: Mouton de Gruyter.

Luijks, C. (1998). *The Pronominal Paradigm in Oorlam-Nama Dutch.* Paper presented to the Linguistic Society of South Africa, Potchefstroom University, 29 June – 3 July 1999.

—— (2000). *The Realisation of Syntactic Principles in Non-standard Afrikaans. The correspondence of Jan Jonker Afrikaner* (*1820–1899*). Ph.D. dissertation, University of Cape Town.

Lyons, J. (Ed.). (1970). *New Horizons in Linguistics*. Harmondsworth: Penguin.

Ma, R. & Herasimchuk, E. (1972). Speech styles in Puerto Rican bilingual speakers. In J. A. Fishman (Ed.), *Advances in the Sociology of Language*, Vol. 2 (pp. 268–297). The Hague: Mouton.

Maas, U. (1988). On the Interpretative Analysis of Historical Records. Linguistic Relations in Seventeenth Century Osnabrück. In P. Auer & A. Di Luzio (Eds.), *Variation and Convergence. Studies in Social Dialectology* (pp. 11–43). Berlin: Mouton de Gruyter.

Mabin, A. (1983). *Strikes in the Cape Colony, 1854–1899*. African Studies Seminar Paper, 9 May 1983, University of the Witwatersrand.

Macaulay, R. K. S. (1977). *Language, Social Class, and Education*: *A Glasgow study.* Edinburgh: Edinburgh University Press.

—— (1988). The rise and fall of the vernacular. In C. Duncan-Rose & T. Venneman (Eds.), *On Language: Rhetorica, phonologica, syntactica. Festschrift for Robert P. Stockwell* (pp. 106–115). London: Routledge.

MacMillan, W. M. (1930). *Complex South Africa. An economic footnote to history.* London: Faber and Faber.

Maeder, G. A. (1919). Die invloed van die bybel op ons volk. In *Populair Wetenskaplike Leesboek* II (pp. 115–130). Cape Town: De Nasionale Pers.

Mahajan, V. & Peterson, R. A. (1985). *Models for Innovation Diffusion* [Sage University Paper series on Quantitative Applications in the Social Sciences, 07-048]. London: Sage.

Maher, J. (1984). Sociolinguistic factors in the direction of language change. *Language Problems and Language Planning, 8*, 281–293.

Makhudu, D. P. (1984). *Is Afrikaans a Creole Language?* MA thesis, Southern Illinois University. Ann Arbor: University Microfilms International.

Malherbe, D. F. (1917). *Afrikaanse Taalboek. Praktiese wegwijser bij die vernaamste moeilikhede in verband met die Afrikaanse grammatica.* Bloemfontein: Nasionale Pers.

Malherbe, E. G. (1925). *Education in South Africa (1652–1922)*. Cape Town: Juta.

Mansvelt, N. (1884). *Proeve van een Kaapsch-Hollandsch Idioticon*. Cape Town: Van De Sandt de Villiers.

Markey, T. L. (1982). Afrikaans: Creole or non-creole? *Zeitschrift für Dialektologie und Linguistik, 49*, 169–207.

Marsh, C. (1988). *Exploring Data. An introduction to data analysis for social scientists*. Oxford: Basil Blackwell.

Martinet, A. (1961). *Éléments de linguistique générale.* Paris: Colin.

Mattheier, K. J. (1985). Dialektologie van de dialektsprekers. Overwegingen omtrent een interpretatiewe aanpak in de dialektologie. In J. Taeldeman & H. Dewulf (Eds.), *Dialekt, Standaardtaal en Maatschappij* (pp. 23–44). Leuvenfoort: Acco.

—— (1991). Standardsprache als Sozialsymbol. Über kommunikative Folgen gesellschaftlichen Wandels. In R. Wimmer (Ed.), *Das 19. Jahrhundert. Sprachgeschichtliche Wurzeln des heutigen Deutsch* (pp. 41–73). Berlin: De Gruyter.

—— (2003). German. In A. Deumert & W. Vandenbussche (Eds.), *Germanic Standardizations. Past to Present* [IMPACT: Studies in Language and Society 18] (pp. 213–248). Amsterdam: John Benjamins.

McClean, R. J. (1947). *Swedish. A Grammar of the modern language.* London: English University Press Ltd.

McCormick, K. (1995). Code-switching, code-mixing and convergence in Cape Town. In R. Mesthrie (Ed.), *Language and Social History. Studies in South African sociolinguistics* (pp. 193–208). Cape Town: David Philip.

—— (2002). *Languages in Cape Town's District Six.* Oxford: Oxford University Press.

McEntegart, D. & Le Page, R. B. (1982). An appraisal of the statistical techniques in the Sociolinguistic Survey of Multilingual Communities. In S. Romaine (Ed.), *Sociolinguistic Variation in Speech Communities* (pp. 105–124). London: Edward Arnold.

McNeill, D. & Freiberger, P. (1993). *Fuzzy Logic*. New York: Simon and Schuster.

McWorther, J. (1998). Identifying the creole prototype. Vindicating a typological class. *Language, 74*, 788–818.

Meillet, A. (1967). *The Comparative Method on Historical Linguistics*. Paris: Librairie Honoré Champion.

—— (1970). *General Characteristics of the Germanic Languages.* Trans. by W. P. Dismukes. Coral Gables, FL: University of Miami Press.

Mentzel, O. F. (1921/25, 1944 [1785–1787]). *A Geographical and Topographical Description of the Cape of Good Hope.* Two volumes (in three parts). Trans. by H. J. Mandelbrote. Cape Town: Van Riebeeck Society.

Mesthrie, R. (1992). *English in Language Shift. The history, structure and sociolinguistics of South African Indian English.* Cambridge: Cambridge University Press.

—— (1993). Nineteenth-century attestations of English-Afrikaans code-mixing in the Cape. *Language Matters, 24*, 47–61.

—— (Ed.). (2002). *Language in South Africa*. Cambridge: Cambridge University Press.

Meurant, L. H. (1861). *Zamenspraak tusschen Klaas Waarzegger en Jan Twyfelaar over het Onderwerp van Afscheiding tusschen de oostelyke en westelyke Provincie*. Cradock: J. S. Bold & Co.

—— (1885). *Sixty Years Ago; or, reminiscences of the struggle for the freedom of the press in South Africa and the establishment of the first newspapers in the eastern province.* Cape Town: Paul Solomon.

Meyer-Benfey, H. (1904). Die Burensprache und ihre Literatur. *Preußische Jahrbücher*, Nov. 1904, 209–234.

Mihm, A. (1998). Arbeitersprache und gesprochene Sprache im 19. Jahrhundert. In D. Cherubim, K. J. Mattheier, & S. Grosse (Eds.), *Sprache und bürgerliche Nation* (pp. 282–316). Berlin: Mouton de Gruyter.

Milroy, J. (1979). Review article (Language, social class and education. A Glasgow study. By R. K. S. Macaulay. Edinburgh: Edinburgh University Press). *Language in Society, 8*, 88–96.

—— Language ideologies and the consequences of standardization. *Journal of Sociolinguistics, 5*, 530–555.
Milroy, J. & Milroy, L. (1985a). *Authority in Language. Investigating language prescription and standardisation.* London: Routledge.
—— (1985b). Linguistic change, social network and speaker innovation. *Journal of Linguistics, 21*, 339–384.
—— (1992). Social network and social class. Towards an integrated sociolinguistic model. *Language in Society, 21*, 1–26.
Milroy, L. (1980). *Language and Social Networks.* Oxford: Basil Blackwell.
—— (1988). Language varieties and situations (Review of B. Horvath (1985), Variation in Australian English: The sociolects of Sydney). *Language in Society, 17*, 577–581.
Milroy, L. & Gordon, M. (2003). *Sociolinguistics. Method and interpretation*. Oxford: Blackwell.
Milroy, L. & Li Wei (1995). A social network perspective on code-switching and language choice. The example of the Tyneside Chinese community. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages. Cross-disciplinary perspectives on code-switching* (pp. 136–157). Cambridge: Cambridge University Press.
Molewijk, G. C. (1992). *Spellingverandering van Zin naar Onzin (1200-heden)*. 's-Gravenhage: SDU Uidgeverij.
Molnárfi, L. (1997). Preposisie of kasusmarkeerder? Oor die status van *vir* in Afrikaans. *Tydskrif vir Geesteswetenskappe, 37*, 89–106.
—— (1999). Zur Diskurskomponente in der Sprechsprache. *Vir* als Rhemamarkierer im Afrikaans. *Folia Linguistica* (= Special issue: Characteristic Properties of Spoken Vernaculars, ed. by W. Abraham), *33*, 75–100.
Mougeon, R. & Beniak, É. (1995). Social class and language variation in bilingual speech communities. In G. Guy, C. Faegin, D. Schiffrin, & J. Baugh (Eds.), *Towards a Social Science of Language. Papers in honor of William Labov* [Current Issues in Linguistic Theory 127], Vol. I (pp. 69–99). Amsterdam: John Benjamins.
Mühlhäusler, P. (1997). *Pidgin and Creole Linguistics* (second edition). London: University of Westminster Press.
—— (2003). Pacific pidgins/creoles. In A. Deumert & W. Vandenbussche (Eds.), *Germanic Standardizations. Past to Present* [IMPACT: Studies in Language and Society 18] (pp. 359–386). Amsterdam: John Benjamins.
Mufwene, S. S. (1988). Starting on the wrong foot. *Journal of Pidgin and Creole Languages, 3*, 109–117.
—— (1996). The founder principle in creole genesis. *Diachronica, 13*, 83–134.
—— (2001). *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
Muller, P. J. (1962). Die Maleier-Afrikaanse Taalbeweging. *Tydskrif vir Volkskunde en Volkstaal, 8*, 1–6.
Muysken, P. (1995). Studying variation in older texts. Negerhollands. *Journal of Pidgin and Creole Languages, 10*, 335–347.
—— (1997). Code-switching processes. Alternation, insertion, congruent lexicalization. In M. Pütz (Ed.), *Language Choices. Conditions, Constraints and Consequences* [IMPACT: Studies in Language and Society 1] (pp. 361–380). Amsterdam: John Benjamins.

Myers-Scotton, C. (1986). Diglossia and code switching. In J. A. Fishman (Ed.), *The Fergusonian Impact. In honor of Charles A. Ferguson on the occasion of his 65th birthday*, Vol. 2 (pp. 403–415). Berlin: Mouton de Gruyter.

—— (1992). Comparing codeswitching and borrowing. In C. Eastman (Ed.), *Codeswitching* (pp. 19–39). Clevedon: Multilingual Matters.

—— (1993a). Common and uncommon ground. Social and structural factors in codeswitching. *Language in Society, 22*, 475–503.

—— (1993b). *Duelling Languages. Grammatical structure in codeswitching*. Oxford: Claredon Press.

Nerbonne, J. & Heeringa, W. (1998). Computationele vergelijking en classificatie van dialecten. *Taal en Tongval, 50*, 164–193.

Nettle, D. (1999). *Linguistic Diversity.* Oxford: OUP.

Neuburg, V. E. (1973). The literature of the streets. In H. J. Dyos & M. Wolff (Eds.), *The Victorian City. Images and realities* (pp. 191–209). London: Routledge.

Nevalainen, T. (2000). Mobility, social networks and language change in early modern England. *European Journal of English Studies, 4*, 253–264.

Nevalainen, T. & Raumolin-Brunberg, H. (1989). A corpus of early modern standard English. *Neuphilologische Mitteilungen, 90*, 67–110.

—— (Eds.) (1996). *Sociolinguistics and Language History. Studies based on the corpus of early English correspondence*. Amsterdam: Rodopi.

—— (1998). Reconstructing the social dimension of diachronic change. In R. Hogg & L. Van Bergen (Eds.), *Historical Linguistics 1995* [Current Issues in Linguistic Theory 162] (pp. 189–209). Amsterdam: John Benjamins.

—— (2003). *Historical Sociolinguistics. Language change in Tudor and Stuart England*. London: Routledge.

Niedzielski, N. & Preston, D. R. (2000). *Folk Linguistics*. Berlin: Mouton de Gruyter.

Nienaber, G. S. (1939). *Van Roem tot Selfmoord.* Johannesburg: Voortrekkerpers.

—— (1942). *Die Afrikaanse geskrifte van Samuel Zwaartman*. Johannesburg: Voortrekkerpers.

—— (1950). *Taalkundige Belangstelling in Afrikaans tot 1900*. Johannesburg: Afrikaanse Pers-Boekhandel.

—— (1960). *Die opkoms van Afrikaans as kultuurtaal*. Pretoria: Van Schaik.

—— (1963). *Hottentots.* Pretoria: Van Schaik.

—— (1965). Iets naders oor die ontkenning in Afrikaans. In P. J. Nienaber (Ed.), *Taalkundige Opstelle* (pp. 22–38). Cape Town: A. A. Balkema.

—— (1966). *Register van Afrikaans voor 1900. Deel 1- Kaapland*. Pretoria: Suid-Afrikaanse Akademie vir Wetenskap en Kuns.

—— (1967). *Register van Afrikaans voor 1900. Deel 2- Transvaal, Oranje-Vrystaat en Natal.* Pretoria: Die Suid-Afrikaanse Akademie vir Wetenskap en Kuns.

—— (1968). *Louis Henri Meurant: 'n vroeë Afrikaanse journalis.* Cape Town: Nasionale Pers.

—— (1971). *Afrikaans in die Vroeër Jare* (second edition). Johannesburg: Voortrekkerspers.

—— (1994a). Die ontstaan van Khoekhoe Afrikaans. In G. Olivier & A. Coetzee (Ed.), *Nuwe Perspektiewe op die Geskiedenis van Afrikaans* (pp. 139–152). Johannesburg: Southern Books.

—— (1994b). Na 60 jaar weer by nie-2. *South African Journal of Linguistics/Suid-Afrikaanse Tydskrif vir Taalkunde, 21*, 68–149.
—— (1994c). Pa-Hulle is kreools. *South African Journal of Linguistics/Suid-Afrikaanse Tydskrif vir Taalkunde, 21*, 14–66.
Nienaber, G. S. & Nienaber, P. J. (1941). *Geskiedenis van die Afrikaanse Beweging.* Pretoria: J. L. van Schaik.
Nienaber, P. J. (1936). *D'Arbez as Skrywer.* Cape Town/Pretoria: H. A. U. M./de Bussy.
—— (1945–1967). *Bibliographie van Afrikaanse Boeke*. Six volumes. Johannesburg: The Author.
—— (Ed.). (1974). *Notulen van die Genootskap van Regte Afrikaners 1875–1878.* Cape Town: Tafelberg.
Noordegraaf, J. (1985). *Norm, Geest en Geschiedenis. Nederlandse taalkunde in de negentiende eeuw.* Dordrecht: Foris.
Norde, Muriel (1998). Grammaticalization versus reanalysis. The case of possessive constructions in Germanic. In R. M. Hogg & L. van Bergen (Eds.), *Historical Linguistics 1995* [Current Issues in Linguistic Theory 162], Vol. II (pp. 211–222). Amsterdam: John Benjamins.
Oakes, M. P. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
Odijk, J. (1992). Uninflected adjectives in Dutch. In R. Bok-Bennema & R. Van Hout (Eds.), *Linguistics in the Netherlands 1992* (pp. 197–208). Amsterdam: John Benjamins.
Oosthuizen, J. (1998). The final *nie* in Afrikaans negative sentences. *Stellenbosch Papers in Linguistics, 31*, 61–92.
Oosthuizen, J. & Waher, H. (1994). On the syntax of the *se*-construction in Afrikaans. *Stellenbosch Papers in Linguistics, 28*, 21–43.
Panchen, A. L. (1992). *Classification, Evolution and the Nature of Biology.* Cambridge: Cambridge University Press.
Pannevis, A. (1880). Afskrif van 'n Lys van Afrikaanse Woorde en Uitdrukkings, versamel deur Arnoldus Pannevis. In H. J. J. M. van der Merwe (Ed.), *Vroeë Afrikaanse woordelyste* (pp. 61–128). Pretoria: Van Schaik.
Paolillo, J. C. (1997). Sinhala diglossia. Discrete or continuous variation? *Language in Society, 26*, 269–296.
Pappageotes, G. C. & Macris, J. (1964). The language question in modern Greece. *Word, 20* (special publication 5, *Papers in Memory of G. C. Pappageotes,* ed. by R. Austerlitz), 53–59.
Paradis, M. (1985). On the representation of two languages in one brain. *Language Sciences, 7*, 1–39.
Patrick, P. L. (1995). The urbanization of creole phonology. Variation and change in Jamaican (KYA). In G. Guy, C. Faegin, D. Schiffrin, & J. Baugh (Eds.), *Towards a Social Science of Language. Papers in honor of William Labov* [Current Issues in Linguistic Theory 127], Vol. I (pp. 329–355). Amsterdam: John Benjamins.
—— (1999). *Urban Jamaican Creole. Variation in the mesolect* [Varieties of English around the World 17]. Amsterdam: John Benjamins.
Paul, H. (1920). *Prinzipien der Sprachgeschichte* (fifth edition). Halle: Max Niemeyer.
Pauwels, J. L. (1958). *Het dialect van Aarschoten en omstreken.* Tongeren: Drukkerij Michiels.

Pellowe, J. (1976). The Tyneside linguistic survey: Aspects of developing a methodology. In W. Viereck (Ed.), *Sprachliches Handeln – Soziales Verhalten. Ein Reader zur Pragmalinguistik und Soziolinguistik* (pp. 203–217). München: Wilhelm FinkVerlag.

Pells, E. G. (1938). *European, Coloured and Native Education in South Africa* (*1652–1938*). Cape Town: Juta.

Penn, N. G. (1995). *The Northern Cape Frontier Zone, 1700-c. 1815*. Ph.D. dissertation, University of Cape Town.

Perkin, H. (1989). *The Rise of Professional Society. England since 1880*. London: Routledge.

Pheiffer, R. (1980). *Die Gebroke Nederlands van Franssprekendes aan die Kaap in die Eerste Helfte van die Agtiende Eeu.* Pretoria: Academica.

Pienaar, E. C. (1931). *Taal en Poësie van die Tweede Afrikaanse Taalbeweging.* Cape Town: Nasionale Pers.

Pfaff, C. (1979). Constraints on language mixing. Intra-sentential code-switching and borrowing in Spanish/English. *Language, 55*, 291–318.

Platt, J. (1975). The Singapore English speech continuum and its basilect 'Singlish' as a 'Creoloid'. *Anthropological Linguistics, 17*, 363–374.

—— (1977). A model for polyglossia and multilingualism (with special reference to Singapore and Malaysia). *Language in Society, 6*, 361–379.

—— (1991). Social and linguistic constraints on variation in the use of two grammatical variables in Singapore English. In J. Cheshire (Ed.), *English around the World: Sociolinguistic perspectives* (pp. 376–387). Cambridge: Cambridge University Press.

Ponelis, F. A. (1979). *Afrikaanse Sintaksis*. Pretoria: Van Schaik.

—— (1981). Review of Van Selms 1979. *Tydskrif vir Geesteswetenskappe, 21*, 71–79.

—— (1984). Aspecte van die taalgebruik van Johanna Maria van Riebeeck. *Tydskrif vir Geesteswetenskappe, 24*, 237–255.

—— (1987). Uit die geskiedenis van die Afrikaanse betreklike konstruksie. *South African Journal of Linguistics*, Occasional Papers 5.

—— (1988). Afrikaans en Taalversteuring. *Tydskrif vir Geesteswetenskappe, 28*, 119–129.

—— (1993). *The Development of Afrikaans*. Frankfurt/M.: Peter Lang.

—— (1995). Die Ontstaan van Afrikaans. *Tydskrif vir Geesteswetenskappe, 34*, 218–230.

—— (1996). Codes in Contradiction. The sociolinguistics of 'Straatpraatjes'. In M. Adhikari (Ed.), *Straatpraatjes. Language, Politics and Popular Culture in Cape Town, 1909–1922* (pp. 129–140). Pretoria: Van Schaik.

—— (2002). Caveats and Comments. *Journal of Germanic Linguistics, 14*, 95–102.

Poplack, S. (1993). Variation theory and language contact. Concepts, methods and data. In D. Preston (Ed.), *American Dialect Research* (pp. 251–286). Amsterdam: John Benjamins.

Preston, D. R. (1989). *Perceptual Dialectology. Nonlinguists' views of areal linguistics.* Dordrecht: Foris Publications.

—— (1992). Talking black and talking white. A study in variety imitation. In J. H. Hall, N. Doane, & D. Ringler (Eds.), *Old English and New. Studies in the honor of Fredric G. Cassidy* (pp. 237–355). New York: Garland Publishing.

Prijslist bij de Paarl Drukpers Maatschappij, Beperkt. (Oktober 1908).

Puddu, M. (1996). *Pietistiese Dagboek van Susanna Catherina Smith 1799–1863*. Three volumes. Masters Thesis, University of the Witwatersrand.

Rademeyer, J. H. (1938). *Kleuring Afrikaans. Die taal van die Griekwas en Rehoboth-Basters.* Amsterdam: Swets and Zeitlinger.

Raidt, E. (1968). *Geskiedenis van die Byvoeglike Verbuiging in Nederlands en Afrikaans.* Cape Town: Nasau Beperk.

—— (1983). *Einführung in Geschichte und Struktur des Afrikaans.* Darmstadt: Wissenschaftliche Buchgesellschaft.

—— (1991). *Afrikaans en sy Europese Verlede* (third edition). Cape Town: Nasau.

—— (1993). Linguistic variants and language change. Deictic variants in some German and Dutch dialects vis-à-vis Afrikaans. In J. Van Marle (Ed.), *Historical Linguistics 1991. Papers from the tenth international conference on historical linguistics, Amsterdam 12.-16. August 1991* [Current Issues in Linguistic Theory 107] (pp. 281–293). Amsterdam: John Benjamins.

—— (1994). *Historiese Taalkunde. Studies oor die geskiedenis van Afrikaans.* Johannesburg: Witwatersrand University Press.

Rampton, B. (1995). *Crossing. Language and ethnicity among adolescents*. London: Longman.

Rampton, B. (Ed.). (1999). Styling the Other. Introduction. *Journal of Sociolinguistics, 3*, 421–427.

Ravenhart, R. (1967). *Before van Riebeeck. Callers at South Africa.* Cape Town: Struik Ltd.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. London: Macmillan.

Reinecke, J. E. (1937). *Marginal Languages*. Ph.D. dissertation, Yale University. Ann Arbor: University Microfilms International.

Richter, M. (1985). Towards a Methodology for Historical Sociolinguistics. *Folia Linguistica Historica, 6*, 41–61.

Rickford, J. R. (1974). The insights of the mesolect. In D. DeCamp & I. F. Hancock (Eds.), *Pidgins and Creoles. Current trends and prospects* (pp. 92–117). Washington: Georgetown University Press.

—— (1983). What happens in decreolization. In R. Andersen (Ed.), *Pidginization and Decreolization as Language Acquisition* (pp. 298–319). Rowley, MA: Newbury House.

—— (1987a). *Dimensions of a Creole Continuum. History, texts and linguistic analysis of Guyanese creole.* Stanford: Stanford University Press.

—— (1987b). The haves and have nots. Sociolinguistic surveys and the assessment of speaker competence. *Language in Society, 16*, 149–178.

Rietveld, T. & Van Hout, R. (1992). *Statistical Techniques for the Study of Language and Language Behaviour.* Berlin: Mouton de Gruyter.

Rissanen, M. (1986). Variation and the Study of English Historical Syntax. In D. Sankoff (Ed.), *Diversity and Diachrony* [Current Issues in Linguistic Theory 53] (pp. 97–109). Amsterdam: John Benjamins.

—— (1991). The diachronic corpus as a window to the history of English. In Jan Svartvik (Ed.), *Directions in Corpus Linguistics. Proceedings of Nobel Symposium, Stockholm 4–8 August 1991* (pp. 185–209). Berlin: Mouton de Gruyter.

Robbers, K. (1997). *Non-finite Verbal Complements in Afrikaans. A comparative approach* [HIL dissertations 33]. The Hague: Holland Academic Graphics.

Roberge, P. T. (1986). What's happening in Afrikaans historical linguistics? *Zeitschrift für Dialektologie und Linguistik, 53*, 194–202.

—— (1990). The ideological profile of Afrikaans historical linguistics. In J. E. Joseph & T. J. Taylor (Eds.), *Ideologies of Language* (pp. 131–149). London: Routledge.
—— (1992). Afrikaans and the ontogenetic myth. *Language and Communication, 12*, 31–52.
—— (1994a). On detecting a prior linguistic continuum in Cape Dutch. In G. Olivier & A. Coetzee (Eds.), *Nuwe Perspektiewe op die Geskiedenis van Afrikaans* (pp. 153–165). Johannesburg: Southern Books.
—— (1994b). The formation of Afrikaans. *Stellenbosch Papers in Linguistics PLUS, 23*.
—— (1994c). On the origins of the Afrikaans verbal hendiadys. *Stellenbosch Papers in Linguistics, 28*, 45–81.
—— (1995). The formation of Afrikaans. In R. Mesthrie (Ed.), *Language and Social History. Studies in South African sociolinguistics* (pp. 68–88). Cape Town: David Phillips.
—— (1996). Multilevel syncretism and the evolution of Afrikaans periphrastic possessives with 'se'. *Stellenbosch Papers in Linguistics, 30*, 123–153.
—— (2000). Etymological opacity, hybridization, and the Afrikaans brace negation. *American Journal of Germanic Linguistics and Literatures, 12*, 101–176.
—— (2002a). The modal particles *mos* and *maskie* in Cape Dutch. *Language Sciences, 24*, 397–408.
—— (2002b). Convergence and the formation of Afrikaans. *Journal of Germanic Linguistics, 14*, 57–94.
—— (2002c). Afrikaans: Considering origins. In R. Mesthrie (Ed.), *Language in South Africa* (pp. 79–103). Cambridge: Cambridge University Press.
—— (2003). Afrikaans. In A. Deumert & W. Vandenbussche (Eds.), *Germanic Standardizations. Past to Present* [IMPACT: Studies in Language and Society 18] (pp. 15–40). Amsterdam: John Benjamins.
—— (forthcoming). Reconstructing the Cape Dutch pidgin. In P. Baker, H. Den Besten, & M. Parkvall (Eds.), *Pidgins: Their nature and significance*. London: Battlebridge.
Roberts, C. & Street, B. (1997). Spoken and written language. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 168–186). Oxford: Basil Blackwell.
Rogers, E. M. (1995). *Diffusion of Innovations* (fourth edition). New York: Free Press.
Romaine, S. (1980). A critical overview of the methodology of urban British sociolinguistics. *English World Wide, 1*, 163–198.
—— (1982). *Socio-historical Linguistics. Its status and methodology.* Cambridge: Cambridge University Press.
—— (1995). *Bilingualism* (second edition). Oxford: Basil Blackwell.
Romesburg, H. C. (1984). *Cluster analysis for researchers.* Belmont, CA: Wadsworth.
Roorda, T. (1856). Over het onderscheid tusschen spreektaal en schrijftaal, inzonderheid in onze moedertaal. *Verslagen en Medelingen van de Koninglijke Akademie, I*, 93–118.
Rosch, E. & Mervis, C. B. (1975). Family Resemblances. Studies in the internal structures of categories. *Cognitive Psychology, 7*, 573–605.
Ross, A. S. C. (1954). Linguistic class-indicators in present-day English. *Neuphilologische Mitteilungen, 55*, 20–56.
Ross, R. (1981). The rise of the Cape gentry. *The Societies of Southern Africa in the 19th and 20th Centuries, 12*, 1–14. London: Institute of Commonwealth Studies.
—— (1982). *Cape of Torments*. London: Routledge.

—— (1999a). *Status and Respectability in the Cape Colony, 1750–1870. A tragedy of manners.* Cambridge: Cambridge University Press.

—— (1999b). *A concise history of South Africa.* Cambridge: Cambridge University Press.

Rossouw, F. (1987). *South African Printers and Publishers 1795–1925*. Cape Town: South African Library.

Rubin, J. (1977). New insights into the nature of language change offered by language planning. In B. G. Blount & M. Sabches (Eds.), *Sociocultural Dimensions of Language Change* (pp. 253–270). New York: Academic Press.

Sankoff, D. & Cedergren, H. (1976). The dimensionality of variation. *Language, 52*, 163–178.

Sankoff, D. & Laberge, S. (1978). The linguistic market and the statistical explanation of variability. In D. Sankoff (Ed.), *Linguistic Variation. Models and methods* (pp. 239–250). New York: Academic Press.

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change, 2*, 71–101.

Sankoff, G. (1980). *The Social Life of Language.* Philadelphia: University of Pennsylvania Press.

Santa Ana, O. & Parodi, C. (1998). Modelling the speech community. Configuration and variable types in the Mexican Spanish setting. *Language in Society, 27*, 23–51.

Sapir, E. (1921). *Language.* New York: Harcourt-Brace.

Sassen, A. (1983). Het taaltje wat ik spreek. *Tijdschrift voor Nederlands en Afrikaans, 1*, 38–40.

Sayre, R. T. (1972). The present tense inflection of the Dutch dialects. A study in structural dialectology. *Zeitschrift für Dialektologie und Linguistik, 39*, 16–36.

Schelling, T. C. (1978). *Micromotives and Macrobehaviour*. New York: W. W. Norton & Co.

Schendl, H. (1996b). Text types and code-switching in medieval and Early Modern English. *Vienna English Working Papers, 5*, 50–62.

Schiffman, H. F. (1993). The balance of power in multiglossic languages. Implications for language shift. *International Journal of the Sociology of Language, 103*, 115–148.

—— (1996). *Linguistic Culture and Language Policy.* London: Routledge.

—— (1997). Diglossia as a sociolinguistic situation. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 205–216). Oxford: Basil Blackwell.

—— (1998). Standardization or restandardization. The case for 'Standard' Spoken Tamil. *Language in Society, 27*, 259–285.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to Multidimensional Scaling. Theory, methods, and applications*. London: Academic Press.

Schneider, E. W. & Viereck, W. (1984). The use of computer in American, Canadian and British English dialectology and sociolinguistics. In H. Goebl (Ed.), *Dialectology* (pp. 15–60). Bochum: Studienverlag Dr. N. Brockmeyer.

Schneider, E. W. & Montgomery, M. B. (2001). On the trail of early nonstandard grammar. An electronic corpus of Southern U.S. antebelluum overseers' letters. *American Speech, 76*, 388–410.

Schoeman, K. (1995). *Die wêreld van Susanna Smit 1799–1863.* Cape Town: Human and Rousseau.

—— (1998). *Merksteen. 'n Dubbelbiografie*. Cape Town: Human and Rousseau.

Scholfield, P. (1991). Statistics in Linguistics. *Annual Review of Anthropology, 20*, 377–393.

Scholtz, J. du P. (1963). *Taalhistoriese Opstelle. Voorstudies tot 'n geskiedenis van Afrikaans.* Pretoria: Van Schaik.
—— (1964). *Die Afrikaner en sy Taal 1806–1875.* Cape Town: Nasau.
—— (1965). *Afrikaans uit die Vroeë Tyd. Studies oor die Afrikaanse Taal en Literêre Volkskultuur van voor 1875.* Cape Town: Nasau.
—— (1972). *Afrikaans-Hollands in die Agtiende Eeu.* Cape Town: Nasau.
—— (1980). *Wording en Ontwikkeling van Afrikaans.* Cape Town: Tafelberg.
Schonken, F. F. (1914). *De Oorsprong der Kaapsch-Hollandsche Volksoverleveringen*. Amsterdam: Swets en Zeitlinger.
Schreiner, O. (1991 [1897/1923]). *Thoughts on South Africa.* Introduced and annotated by M. Lenta. Pretoria: A. D. Donker.
Schuchardt, H. (1883). *Kreolische Studien IV.* Vienna.
—— (1885). Review of Mansvelt 1884. *Literaturblatt für germanische und romanische Philologie, 6*, 464–470.
—— (1891). *Kreolische Studien IX.* Vienna.
Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*, 209–230.
Settersten, R. A. & Mayer, K. U. (1997). The measurement of age, age structuring, and the life course. *Annual Review of Sociology, 23*, 233–261.
Seuren, P. A. M. (1997). *Semantic Syntax.* Oxford: Basil Blackwell.
Shaffer, D. (1978). Afrikaans as a case study in vernacular elevation and standardization. *Linguistics, 16*, 51–64.
Shell, R. (1994). *Children of Bondage. A social history of the slave society at the Cape of Good Hope 1652–1838.* Johannesburg: Witwatersrand University Press.
Shetter, W. Z. (1994). *Dutch. An essential grammar.* London: Routledge.
Siegel, J. (1985). Koines and Koineization. *Language in Society, 14*, 357–378.
—— (1998). Substrate reinforcement and dialectal differences in Melanesian pidgin. *Journal of Sociolinguistics, 2*, 347–373.
Siegenbeek, M. (1805). *Verhandelingen over de Nederduitsche spelling, ter bevordering van eenparigheid in deselves.* Amsterdam: Allart.
Simpson, G. G. (1961). *Principles of Animal Taxonomy.* New York: Columbia University Press.
Singler, J. V. (1995). The demographics of Creole genesis in the Caribbean. A comparison of Martinique and Haiti. In J. Arends (Ed.), *The Early Stages of Creolization* [Creole Language Library 13] (pp. 203–232). Amsterdam: John Benjamins.
Smith, A. B. & Pheiffer, R. (1993). *De Khoikhoi at the Cape of Good Hope. Seventeenth Century Drawings in the South African Library.* Cape Town: South African Library.
Smith, A. D. (1993). The nation: Invented, imagined, reconstructed? In M. Ringrose & A. Lerner (Eds.), *Reimagining the Nation* (pp. 9–28). Buckingham: Open University Press.
Smith, J. (1996). *An Historical Study of English. Function, form and change.* London: Routledge.
Smith, J. C. (2002). Middle French. When? What? Why? *Language Sciences, 24*, 423–446.
Smith, N. (1989). *The Twitter Machine. Reflections on language.* Oxford: Basil Blackwell.
Smits, C. (1993). Resistance to erosion in American Dutch inflection. In G. E. Booij & J. Van Marle (Eds.), *Yearbook of Morphology* (pp. 155–184). Dortrecht: Kluwer.

—— (1996). *Disintegration of Inflection. The case of Iowa Dutch* [HIL dissertations 22]. The Hague: Holland Academic Graphics

Smuts, J. (1968). *Vroeë Afrikaans.* Cape Town: Nasau.

Snyman, F. C. (1979). Apocope van slot-[t] in die beskaafde Nederlandse omgangstaal van die 17de en 18de eeu. *Tydskrif vir Geesteswetenskappe, 19*, 3–17.

Sokal, R. (1986). Phenetic taxonomy. Theory and methods. *Annual Review of Ecology and Systematics, 17*, 423–442.

Sokal, R. & Sneath, P. (1963). *Principles of Numerical Taxonomy.* San Francisco: W. H. Freeman.

Sparrman, A. (1785). *A Voyage to the Cape of Good Hope, 1772–1776.* London: Robinson.

STATISTICA (1995). *Manual*, Vol. I–III. Tulsa: Stat.Soft, Inc.

Stanford, M. (1994). *A Companion to the Study of History*. Oxford: Basil Blackwell.

Stedman Jones, G. (1976). *Outcast London*. Harmondsworth: Penguin.

Stein, D. (1994). Sorting out the variants. Standardization and social factors in the English Language 1600–1800. In D. Stein & I. Tieken-Boon van Ostade (Eds.), *Towards a Standard English 1600–1800* (pp. 1–17). Berlin: Mouton de Gruyter.

—— (1997). Syntax and varieties. In J. Chesire & D. Stein (Eds.), *Taming the Vernacular. From dialect to written standard language* (pp. 35–50). Harlow: Longman.

Stein, P. (1989). When Creole speakers write the standard language. An analysis of some of the earliest slave letters from St. Thomas. In M. Pütz & R. Dirven (Eds.), *Wheels within Wheels* (pp. 153–178). Frankfurt/M.: Peter Lang.

—— (1995). Early creole writing and its effects on the discovery of the creole language structure: The case of eighteenth-century Negerhollands. In J. Arends (Ed.), *The Early Stages of Creolization* [Creole Language Library 13] (pp. 43–62). Amsterdam: John Benjamins.

Steinberg, J. (1987). The historian and the *questione della lingua*. In P. Burke & R. Porter (Eds.), *The Social History of Language* (pp. 198–209). Cambridge: Cambridge University Press.

Stemberger, J. P. & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition, 14*, 17–26.

Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Steyn, J. C. (1980). *Tuiste in eie Taal.* Cape Town: Tafelberg.

—— (1989). 'n Paar aantekeninge oor Afrikaanse koerante se grammatiese probleme in die 'owergangstijdperk' tot omstreeks 1917. *South African Journal of Linguistics, 7*, 20–29.

Stoops, Y. (1980). Vraagtekens by die apokopee van [t] in 16de- en 17de-euse Nederlands. *Tydskrif vir Geesteswetenskappe, 20*, 147–156.

Strandskogen, Å., & Strandskogen, R. (1986). *Norwegian. An essential grammar.* London: Basil Blackwell.

Sturgis, J. (1982). Anglicisation at the Cape of Good Hope in the early nineteenth century. *Journal of Imperial and Commonwealth History, 7*, 5–32.

Swart, B. H. & Prinsloo, P. P. J. (1936). *Anglisismens, Stylverbeteringe en Tegniese Woorde.* Cape Town: Maskew Miller.

Swaving, J. G. (1830). *Swavings Zonderlinge Ontmoetingen en Wonderbaarlijke Lotsverwisselingen na zijne Vlucht uit Delft, benevens schetsen van London en deszelfs bewoners; van de Kaap de Goede Hoop en hare omstreken, enz., enz.* Dordrecht: Blusse en Van Brom.

Sweet, H. (1900). *The History of Languages*. London: J. M. Dent.

Taeldeman, J. (1994). Jan Z'N + nomen. Over een bezitsconstructie in de Vlaamse dialecten. *Taal en Tongval, 46*, 220–228.

Talen, J. G., Kollewijn, R. A., & Buitenrust Hettema, F. (1908). *Nederlandse Taal. Proeve van een Nederlandse spraakleer.* Zwolle: W. E. J. Tjeenk Willink.

Tannen, D. (1982). The oral/literate continuum in discourse. In D. Tannen (Ed.), *Spoken and Written Language. Exploring orality and literacy* (pp. 1–16). Norwood: Ablex Publishing.

Tarrow, S. (1992). Mentalities, political cultures and collective action frames. In A. D. Morris & C. M. Mueller (Ed.), *Frontiers in Social Movement Theory* (pp. 174–202). New Haven: Yale University Press.

Ten Rhyne, W. (1686). *Schediasma de Promotorio Bonae Spei.* Schaffhausen: Osvaldus.

Terblanche, H. J. (1968). Die invloed van Engels op Afrikaans. In T. J. R. Botha (Ed.), *Taal en Teken. Bundel opgedra aan Prof. G. S. Nienaber op sy sestigste verjaardag 12.11.1963* (pp. 169–172) (third edition). Cape Town: Academica.

Te Winkel, J. (1896). Het Nederlandsch in Noord-Amerika en Zuid-Afrika. *Vragen van den Dag, 11*, 337–362, 418–442, 483–505.

—— (1897). *De Nederlandsche Taal in Zuid Afrika*. (Rede den 25en Augustus 1897 op het XXIV Ned. Taal- en Letterkundige Congres te Dordrecht gehouden). Algemeen Nederlandsch Verbond.

—— (1904). *Inleiding tot de Geschiedenis der Nederlandsche Taal.* Culemburg: Blom and Olivierse.

Thomason, S. G. (1980). Morphological instability, with and without language contact. In J. Fisiak (Ed.), *Historical Morphology* (pp. 359–372). Berlin: Mouton de Gruyter.

—— (1993). Coping with partial information in historical linguistics. In H. Aertsen & R. J. Jeffers (Eds.), *Historical Linguistics 1989. Papers from the 9th International Conference of Historical Linguistics, Rutgers University, 14–18.08.1989* [Current Issues in Linguistic Theory 106] (pp. 485–496). Amsterdam: John Benjamins.

—— (1997). A typology of contact languages. In A. K. Spears & D. Winford (Eds.), *The Structure and Status of Pidgins and Creoles* [Creole Language Library 19] (pp. 71–88). Amsterdam: John Benjamins.

—— (2001). *Language Contact. An introduction.* Edinburgh: Edinburgh University Press.

Thomason, S. G. & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics.* Berkeley: University of California Press.

Tiflin, E. C. (1984). *Gesproke Afrikaans onder die Kleuringsgemeenskap in Durban. 'n Sosiolinguistiese studie*. MA Thesis, University of Stellenbosch.

Timm, L. A. (1981). Diglossia old and new – a critique. *Anthropological Linguistics, 13*, 356–367.

Togerson, W. S. (1958). *Theory and Methods of Scaling*. New York: Wiley and Sons.

Tottie, G. (1991). *Negation in English Speech and Writing. A study in variation*. London: Academic Press.

Traill, A. (1995). The Khoesan languages of South Africa. In R. Mesthrie (Ed.), *Language and Social History. Studies in South African sociolinguistics* (pp. 1–18). Cape Town: David Phillips.

—— (2002). The Khoesan languages. In R. Mesthrie (Ed.), *Language in South Africa* (pp. 27–49). Cambridge: Cambridge University Press.

Trudgill, P. (1974). *The Social Differentiation of English in Norwich.* Cambridge: Cambridge University Press.

—— (1986). *Dialects in Contact.* Oxford: Basil Blackwell.

—— (1989). Contact and isolation in linguistic change. In L. E. Breivik & E. H. Jahr (Ed.), *Language Change. Contributions to the study of its causes* (pp. 227–237). Berlin: Mouton de Gruyter.

—— (1990). *The Dialects of England.* Oxford: Basil Blackwell.

—— (1996). Language contact and inherent variability. The absence of hypercorrection in East Anglian present-tense forms. In J. Klemola, M. Kytö, & M. Rissanen (Eds.), *Speech Past and Present. Studies in English dialectology in memory of Ossi Ihalainen* (pp. 412–425). Frankfurt/M.: Peter Lang.

Truter, E. (1997). *Tibbie. Rachel Isabella Steyn 1865–1955.* Cape Town: Human and Rousseau.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Reading, MA: Addison-Wesley.

Ueckerman, F. J. (1987). *Die Afrikaanse Geskrifte van Samuel Zwaartman*. Unpublished MA thesis, University of the Witwatersrand, Johannesburg.

Uys, M. D. (1983). *Die Vernederlandsing van Afrikaans*. Unpublished Ph.D. dissertation, University of Pretoria.

Vacheck, J. (1989). *Written Language Revisited.* Ed. by P. A. Luelsdorff. Amsterdam: John Benjamins.

Valkhoff, M. F. (1966). *Studies in Portuguese and Creole. With special reference to South Africa.* Johannesburg: Witwatersrand University Press.

—— (1972). *New Light on Afrikaans and 'Malayo-Portuguese'*. Louvain: Éditions Peeters Imprimerie Orientaliste.

Vandenbussche, W. (2002). Dutch orthography in lower, middle and upper class documents in 19th century Flanders. In A. Linn & N. McLelland (Eds.), *Standardization. Studies from the Germanic languages* [Current Issues in Linguistic Theory 235] (pp. 27–42). Amsterdam: John Benjamins.

Van den Berg, B. (1949). De conjugatie van het praesens in de Noordnederlandse dialecten. *Taal en Tongval, 1*, 6–13.

Van den Berghe, P. L. (1968). Language and nationalism in South Africa. In J. A. Fishman, C. A. Ferguson, & J. Das Gupta (Eds.), *Language Problems in Developing Nations* (pp. 215–223). New York: Wiley and Sons.

Van den Craen, P. & Willemyns, R. (1988). The standardization of Dutch in Flanders. *International Journal of the Sociology of Language, 73*, 45–64.

Van den Heever, R. (Ed.). (1987). *Tree na Vryheid. 'n Studie in alternatiewe Afrikaans*. Kasselsvlei: CPTA.

Van den Toorn, M. C. (1982). *Nederlandse Grammatica* (eighth, revised edition). Groningen: Wolters-Noordhoff.

Van der Merwe, C. N. (1994). *Breaking Barriers. Stereotypes and the changing of values in Afrikaans writing 1875–1990*. Amsterdam: Rodopi.
Van der Merwe, H. J. J. M. (1971). *Vroeë Afrikaanse woordelyste*. Pretoria: Van Schaik.
Van der Rijst, J. W. (1878). *Spel-en Leesboek vir Afrikaanse Kinders deur Oom Willem*. Paarl: D. F. Du Toit.
—— (1897). *Samesprake in Afrikaans en Engels*, Vol. I–III. Paarl: D. F. Du Toit & Co.
Van der Wal, M. J. (1992). Dialect and standard language in the past. The rise of the Dutch standard language in the sixteenth and seventeenth centuries. In J. A. Van Leuvensteijn & J. B. Berns (Eds.), *Dialect and Standard Language* (pp. 119–129). Amsterdam: North Holland.
Van der Zwan, D. (1986). Die Taalgebruik in die Dagboek van Hendrik Witbooi. MA Thesis, University of South Africa (UNISA).
Van Ginneken, J. J. A. (1928). *Handboek der Nederlandsche Taal. Deel I. De sociologische structuur der Nederlandsche taal* (second edition). 's-Hertogenbosch: L. C. G. Malmberg.
Van Loey, A.(1959). *Historische Grammatica van het Nederlands. Klankleer, vormleer, woordvorming*. Zuthpen: Thieme.
Van Loon, L. G. (1939). *Crumbs from an Old Dutch Closet. The Dutch dialect of New York*. The Hague: Nijhoff.
Van Malssen, P. J. (1900). *Het Leven der Taal, Inzonderheid dat van het Nederlandsch*. 's-Gravenhage: Martinus Nijhoff.
Van Marle, J. (1994). De derivationele morfologie. Een vergeten hoofdstuk uit de geschiedenis van het Afrikaans. In G. Olivier & A. Coetzee (Eds.), *Nuwe Perspektiewe op die Geskiedenis van Afrikaans* (pp. 90–101). Johannesburg: Southern Books.
—— (1995). On the fate of adjectival declension in overseas Dutch (with some notes on the history of Dutch). In H. Andersen (Ed.), *Historical Linguistics 1993. Selected Papers from the 11th International Conference on Historical Linguistics, 16–10 Aug. 1993* [Current Issues in Linguistic Theory 124] (pp. 283–294). Amsterdam: John Benjamins.
Van Marle, J. & Smits, C. (1993). The inflectional system of overseas Dutch. In H. Aertsen & R. J. Jeffers (Eds.), *Historical Linguistics 1989* [Current Issues in Linguistic Theory 107] (pp. 313–328). Amsterdam: John Benjamins.
—— (1995). On the impact of language contact on inflectional systems. The reduction of verb inflection in American Dutch and American Frisian. In J. Fisiak (Ed.), *Linguistic Change under Contact Conditions* (pp. 179–206). Berlin: Mouton de Gruyter.
—— (1997). Dialect versus standard language. Nature versus culture. In J. Cheshire & D. Stein (Eds.), *Taming the Vernacular. From dialect to written standard language* (pp. 13–34). London: Longman.
Van Oordt, J. (1913). *Manual of Cape Dutch for the Use of English Students.* Cape Town: Du Bussy.
Van Oordt, J. F. (1947–1948). *Die Kaapse Taalargief* 1–4. Rondebosch: L. C. Van Oordt.
—— (1949). Die Kaapse Taalargief 5–10. *Tydskrif vir Wetenskap en Kuns* 10–16 (10. 65–88. 11. 55–75. 12. 73–101. 13. 45–73. 14: 33–56. 16: 97–124).
Van Rensburg, M. C. J. (1983). Nie-standaardvorme, variasiepatrone en Afrikaans uit die vorige eeu. In G. N. Claasen & M. C. J. Van Rensburg (Eds.), *Taalverskeidenheid* (pp. 134–161). Johannesburg: Academica.

—— (Ed.). (1984). *Die Afrikaans van die Griekwas van die Tagtigerjare.* Two volumes. Bloemfontein: University of the Orange Free State.

—— (1989). Orange River Afrikaans. A stage in the pidgin/creole cycle. In M. Pütz & R. Dirven (Eds.), *Wheels within Wheels. Papers of the Duisburg Symposium on Pidgin and Creole Languages* (pp. 135–151). Frankfurt/M.: Peter Lang.

—— (1994). Die ontstaan van Afrikaans in 'n intertaal konteks. In G. Olivier & A. Coetzee (Eds.), *Nuwe Perspektiewe op die Geskiedenis van Afrikaans* (pp. 166–170). Johannesburg: Southern Books.

Van Rensburg, M. C. J. & Combrink, J. H. (1984). Transvaalse Afrikaans. In T. J. R. Botha (Ed.), *Inleiding tot die Afrikaanse Taalkunde* (pp. 107–133). Pretoria: Academica.

Van Rensburg, M. C. J., Du Plessis, L. T., & Klopper, A. H. (1987). *'n Verkenning van Afrikaans. Gister.* Bloemfontein: Patmos.

Van Rensburg, M. C. J. & Jordaan, A. (1995). The Growth of Afrikaans in South Africa. In V. N. Webb (Ed.), *Language in South Africa. An input into language planning for a post-apartheid South Africa. The LiCCA (SA) report* (pp. 107–132). Pretoria: University of Pretoria.

Van Rijn, C. J. (1914). *Het zeer nauwe verband tussen het Afrikaans en het Nederlands. Vooral voor onderwijzers.* Cape Town: C. J. Van Rijn.

Van Selms, A. (Ed.). (1979). *Abu Bakr se 'Uiteensetting van die Godsdiens'. 'n Arabies-Afrikaanse teks uit 1869.* Amsterdam: North Holland.

Van Wyk, E. B. (1978). Language contact and bilingualism. Language and communication studies in South Africa. In L. W. Lanham & K. P. Prinsloo (Eds.), *Current Issues and Directions in Research and Inquiry* (pp. 29–52). Oxford: Oxford University Press.

Van Zyl, D. (2000). Afstand en vereenselwiging. Perspektiewe op die veranderende betekenisse van boer and Boer in die Afrikaanse poësie. *Literator, 21*, 1–21.

Varley, D. H. (1952). The 19th century and the 'common reader'. *South African Libraries, 19*, 17–24.

Verdam, J. (1890). *De Geschiedenis der Nederlandsche Taal.* Leeuwarden: Hugo Suringar.

Verslag van het Eerste Congres ter Vereenvoudiging der Nederlandsche Taal. Gehouden te Kaapstad op 4 en 5 Januari 1897. Cape Town: Jacques Dusseau & Co.

Viljoen, W. J. (1891). *De Taalbeweging in de Kaapkolonie.* Amsterdam.

—— (1896). *Beiträge zur Geschichte der Kapholländischen Sprache.* Strassburg: Karl J. Trübner.

Visser, H. E. E. (1992). *Die Verdwyning van die Preteritum in Afrikaans*. Unpublished MA thesis, Johannesburg, RAU.

Von Polenz, P. (1991). *Deutsche Sprachgeschichte. Vom Spätmittelalter bis zur Gegenwart*. Volume One. Berlin: De Gruyter.

—— (1994). *Deutsche Sprachgeschichte. Vom Spätmittelalter bis zur Gegenwart*. Volume two. Berlin: De Gruyter.

Von Wielligh, G. R. (1925a). *Ons Geselstaal.* Pretoria: Van Schaik.

—— (1925b). *Baanbrekerswerk of Die Letterkundige, Kulturele en Politieke Betekenis van die Genootskap van Regte Afrikaners.* Pretoria: J. L. van Schaik.

Von Wright, G. H. (1963). *Norm and Action. A logical enquiry*. London: Routledge.

Vogel, R. (1997). *Aspects of Tense* [HIL dissertations 25]. Den Haag: Holland Academic Graphics.

Waher, H. (1978). *Die Probleem van die Bereik van die Ontkenning met Spesiale Verwysing na Afrikaans*. Ph.D. Thesis, University of Cape Town.

—— (1988). Eenders of anders? 'n Vergelyking tussen die ontkenning in Afrikaans en Afro-Portuguees. *South African Journal of Linguistics, 6*, 109–123.

—— (1994a). Die dialektiese ontkenning in Afrikaans. In G. Olivier & A. Coetzee (Eds.), *Nuwe Perspektiewe op die Geskiedenis van Afrikaans* (pp. 102–109). Johannesburg: Southern Books.

—— (1994b). Vernederlandsing en die Afrikaanse afleidingsaffikse. Getuienis uit Kaapse Moeslim-Afrikaans. In C. Van der Merve, H. Waher, & J. Hambidge (Eds.), *Rondom Roy. Studies opgedra aan Roy H. Pheiffer* (pp. 112–122). Cape Town: Departement van Afrikaans en Nederlands.

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*, 236–244.

Ward, K. (1991). *Catering for Taste in Late-Victorian Cape Town. An investigation into the emergence of factory production in specific consumer industries.* Cape Town History Project, Workshop 11–12 Nov. 1991. Manuscript.

Wardhaugh, R. (1992). *An Introduction to Sociolinguistics* (second edition). Oxford: Basil Blackwell.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis. Methods and applications*. Cambridge: Cambridge University Press.

Wattenmaker, W. D., Nakamura, G., & Medin, D. L. (1988). Relationships between similarity-based and explanation-based categorization. In D. J. Hilton (Ed.), *Contemporary science and natural explanation* (pp. 204–240). New York: New York University Press.

Watts, H. L. (1976). A social and demographic portrait of English-speaking South Africans. In A. De Villiers (Ed.), *English-speaking South Africa* (pp. 41–90). Oxford: Oxford University Press.

Watts, Richard J. (1999). The social construction of Standard English. Grammar writers as a "discourse community". In T. Bex & R. J. Watts (Eds.), *Standard English. The Widening Debate* (pp. 40–68). London: Routledge.

Webb, V. N. (1993). Die herkoms van Oranjerivierafrikaans. In M. Paul & S. von Staden (Eds.), *Linguistica. Festschrift E. B. Van Wyk* (pp. 161–171). Pretoria: Van Schaik.

—— (2002). *Language in South Africa. The role of language in national transformation, reconstruction and development* [IMPACT: Studies in Language and Society 14]. Amsterdam: John Benjamins.

Weinreich, U. (1954). Is a structural dialectology possible. *Word, 10*, 388–400.

—— (1968). *Languages in Contact*. The Hague: Mouton.

—— (1992). *College Yiddish. An introduction to the Yiddish language and to Jewish life and culture* (fifth revised edition). New York: Yivo Institute for Jewish Research.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Makiel (Eds.), *Directions in Historical Linguistics* (pp. 95–195). Austin: University of Texas Press.

Weinstein, B. (1979). Language strategists. Redefining political frontiers on the basis of linguistic choices. *World Politics, 31*, 345–364.

—— (1982). Noah Webster and the diffusion of linguistic innovations of political purposes. *International Journal of the Sociology of Language, 38*, 85–108.

Werlen, I. (1988). Swiss German dialects and Swiss standard High German. In P. Auer & A. di Luzio (Eds.), *Variation and Convergence. Studies in social dialectology* (pp. 93–123). Berlin: Mouton De Gruyter.

Wetherell, C. (1998). Historical Social Network Analysis. *International Review of Social History, 43* (supplement 6), 125–144.

Wexler, P. (1971). Diglossia, language standardization, and purism. Parameters for a typology of literary language. *Lingua, 27*, 330–354.

Whinnom, K. (1971). Linguistic hybridization and the 'special case' of pidgins and creoles. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 91–116). Cambridge: Cambridge University Press.

Willemyns, R. (1989). Language planning as an initiator of linguistic change. In T. J. Walsh (Ed.), *Synchronic and Diachronic Approaches to Linguistic Variation and Change. Georgetown University Round Table on Languages and Linguistics 1988* (pp. 349–357). Washington: Georgetown University Press.

Winer, L. (1984). Early Trinidadian English creole: The Spectator texts. *English World Wide, 5*, 181–210.

—— (1995). Penny Cuts. Differentiation of creole varieties in Trinidad, 1904–1906. *Journal of Pidgin and Creole Languages, 10*, 127–155.

Winford, D. (1985). The concept 'diglossia' in Caribbean creole situations. *Language in Society, 14*, 345–356.

Wodak, R. & Benke, G. (1997). Gender as a sociolinguistic variable. New perspectives on variation dtudies. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 127–149). Oxford: Basil Blackwell.

Wolfram, W. (1993). Identifying and interpreting variables. In D. Preston (Ed.), *American Dialect Research* (pp. 193–221). Amsterdam: John Benjamins.

Wolfson, N. (1979). The conversational historic present alternation. *Language, 55*, 168–182.

Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in Language Studies*. Cambridge: Cambridge University Press.

Woolard, K. A. (1985). Language variation and cultural hegemony. Towards an integration of socio-linguistic and social theory. *American Ethnologist, 12*, 738–748.

—— (1998). Introduction. Language ideology as a field of inquiry. In B. B. Schieffelin, K. A. Woolard, & P. V. Kroskrity (Eds.), *Language Ideologies, Practice and Theory* (pp. 3–47). Oxford: Oxford University Press.

Woolard, K. A. & Schieffelin, B. B. (1994). Language ideology. *Annual Review of Anthropology, 23*, 55–82.

Worden, N. (1985). *Slavery in Dutch South Africa.* Cambridge: Cambridge University Press.

Worden, N., Van Heyningen, E., & Bickford-Smith, V. (1998). *Cape Town. The making of a city.* Cape Town: David Phillips.

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wright, L. (1998). Mixed-language business writing. Five hundred years of code-switching. In E. H. Jahr (Ed.), *Language Change. Advances in historical sociolinguistics* (pp. 99–117). Berlin: Mouton de Gruyter.

Wright, R. (1982). *Late Latin and Early Romance in Spain and Carolingian France*. Liverpool: Cairns.

—— (1991). The conceptual distinction between Latin and Romance. Invention or evolution? In R. Wright (Ed.), *Latin and the Romance Languages in the Early Middle Ages* (pp. 102–113). London: Routledge.

Wright, S. (1994). The Critic and the Grammarians. Joseph Addison and the prescriptivists. In D. Stein & I. Tieken-Boon Van Ostade (Eds.), *Towards a Standard English 1600–1800* (pp. 243–284). Berlin: Mouton de Gruyter.

Zadeh, L. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.

Zerubavel, E. (1991). *The Fine Line. Making distinctions in everyday life.* Toronto: The Free Press.

Zietsman, P. H. (1992). *Die Taal is Gans die Volk.* Pretoria: University of South Africa.

Zwart, C. J.-W. (1997). *The Morpho-Syntax of Verb Movement. A minimalist approach to the syntax of Dutch.* Dordrecht: Kluwer Academic Publishers.

Index

A

abbreviations xvii
Abdurahman, A. 94, 240–243, 292
Abu Bakr Effendi 52, 74 (n. 3), 158, 281
accommodation 3, 5, 37, 38, 40, 131, 295, 299, 300
acquisition planning 35
Acts of Identity 3, 9, 36
African Political Organization 69, 93
Afrikaans language congresses 57, 73, 275, 276
Afrikaanse Taalgenootskap 58
Afrikaanse Taalvereniging 58
Afrikaner Bond 58, 69, 73, 290
Afrikaner, Jan Jonker 17, 213
agreement
  gender 165ff.
  subject-verb 137ff.
akkerboere 27
Alternatiewe Afrikaans 277 (n. 2)
alternation (code-switching) 252–254, 256, 268
amalgamation graph 118, 119, 223, 230
Anderson, B. 54, 55, 74 (n. 5), 281
anglicization 261–263
anglicisms 270–272, 276
analogy 152, 155, 175, 199
Anglo-Boer War 58, 94, 97, 162, 269, 277, 291, 292
apartheid 9, 276
Arabic 52
Arabic Afrikaans *see* Malay Afrikaans
Arabic script 17, 52, 69
Austronesian 25
avant-garde of acculturation 300, 301

B

Bain, A. G. 47, 263
Baker, P. 29, 34
Bakhtin, M. M. 48, 81
Bantu languages 25, 80, 82
basilectalization 30, 35, 43 (n. 11)
Bayân-ud-în *see* Abu Bakr Effendi
Benigna van Groenekloof of Mambre 52, 53, 240–243, 249
Bible 19, 74 (n. 6), 84, 85, 165, 187, 192, 210
Bickerton, D. 29, 30, 43 (n. 10), 230–232
bi-dialectialism 41, 100, 245, 246
bilingualism 100, 41
  Khoe-Dutch 27, 82
  English-Dutch 41, 82, 262, 263, 270, 276
Bloomfield, L. 59, 105
Boniface, C. E. 46–48, 74 (n. 2), 148, 188, 201, 240–243, 281
Berbice Dutch 30, 190
borrowing 22, 38, 39, 233, 258 (n. 4), 264–266, 270ff.
Bourdieu, P. 7, 71, 98, 301
Brink, M. 68, 240–243, 263
bywoners 96, 99

C

Changuion, A. N. E. 60, 145, 180, 209, 210, 272
Cape Census 76 (n. 18), 92ff.
Cape Dutch Pidgin 17, 21, 31–37, 42, 153, 161, 164, *see also* Cape Dutch (Vernacular) – basilect/basilectal
Cape Dutch speech community 59ff., 77ff., 145, 246, 249, 275, 281, 298

Cape Dutch (Vernacular)
acrolect/acrolectal 33, 34, 36, 37, 39–41, 60, 61, 66–68, 100, 101, 159, 176, 181, 182, 185, 194, 200, 203, 205, 214, 227, 228, 236, 238, 240, 243, 245, 257, 282, 298
basilect/basilectal 36, 37, 39, 40, 60, 66, 67, 101, 140, 153, 158, 161, 165, 181, 193, 194, 199, 201, 205, 213–214, 236, 240, 298
mesolect/mesolectal 40, 41, 61, 66, 100, 101, 159, 165, 176, 181, 193, 194, 200, 203, 205, 236, 238, 240, 243, 257, 258, 281, 282, 286, 287, 295, 298, 301, 304 (n. 1)
Cape Town 26, 27, 36, 49, 69, 81, 82, 93, 95, 98, 200, 242, 261, 262, 292
cluster analysis 110ff., 222, 223, 230, 231, 234, 235, 302
code-mixing/switching
Dutch-Afrikaans 249–257
English – Dutch/Afrikaans 263–270
codification 6–9, 54, 56, 58, 62, 70, 74 (n. 4), 203
dictionary 56, 276
grammar 1, 8, 55–57, 61, 145, 175, 176, 188, 202, 203
orthography 56, 57, 70, 88, 293
colonists
British 76 (n. 18), 94, 261, 262
Dutch 23, 24, 262
German 23
Huguenots 23
community of practice 288, 289
competence (linguistic) 105, 106, 136, 232, 245
complexification 39
Congress of Vienna 261
congruent lexicalization (code-switching) 252, 255, 256
convergence 31, 36, 39, 42, 81, 100, 131, 212, 249, 255, *see also* levelling
convergence model 35, 38
copula 182, 204, 205, 210
Corpus of Cape Dutch Correspondence 53, 87ff., 298
correlations 124, 126, 130, 133 (n. 13), 224, 228, 229, 236f.
Creole Portuguese 10, 18, 22, 27, 28, 34, 140, 192, 209, 212–214

D
dialect levelling *see* levelling
decreolization 37, 38, 43, 44 (n. 16)
DeCamp, D. 37, 59, 230–232
De Lima, J. S. 46–48, 60, 148, 240–243, 291
Den Besten, H. 16, 31–35, 38, 43 (n. 12), 140, 153, 161, 163, 192, 193, 197, 199, 208, 209, 211, 218 (n. 2)
dendrogram 113, 117, 118, 222
dialect literature 45ff., 240–243
diffusion *see* language spread, S-curve
diglossia 63ff., 145, 243–249, 298–300
District Six 95, 276
Dorian, N. 78
drift 139, 180, *see also* evolutive language change
Duminy diary 16, 41, 142, 148, 158, 175,185, 200
Dummy High 66
Dutch dialects 23, 24, 92, 139, 140, 146, 158, 180, 197, 218 (n. 10), 219 (n. 14)
Dutch Reformed Church 95, 25, 262
Du Toit, S. J. 56–58, 61, 73, 74 (n. 6), 159, 175, 186, 202, 209, 210

E
early adopters 286, 287
early majority 286, 287
Early Modern Dutch 18, 39, 136, 158, 197, 208, 214
East Indian Dutch 169, 174
Eastern Cape Afrikaans 27, 43 (n. 43), 80
eastern Cape separatism 47
education 19, 20, 83ff., 145, *see also* medium of instruction

Eenvoudig Kaaps(ch) Hollands(ch) 69, 70, 72, 243
eigenvalue 109, 124, 125, 127, 133 (n. 14), 228
eigenvector 124, 125
élite 8, 58, 64, 72, 73, 96, 218 (n. 9), 244, 301
English *see also* bilingualism, borrowing
  African American 49, 115ff.
  attitudes 57, 58, 73
  demonstratives 194
  historical present 163
  Indian English 38
  loss of verbal inflection 139
  loss of gender 169, 191
  medium of instruction 41, 52, 66
  official status 1, 69, 261
  relativizer 190
  standardization 202
Euclidean distance 110, 111, 116ff., 222
exaptation 172
evolutive language change 10, 139, 158, 180, 298, *see also* drift

F

Federasie van Afrikaanse Kultuurvereniginge 275
Ferguson, C. 63, 64, 75 (n. 14), 81, 139, 172, 202, 230, 245, 246, 248, 249, 279
fixity 130–132
focusing 2–6, 37, 51, 130–132, 243, 288, 300, 302, 303
folk linguistics 49, 59ff., 75 (n. 9), 94, 243, 247, 258 (n. 5)
founder principle 33
Frisian 172, 190, 204
fuzzy sets and boundaries 61, 226, 227, 230, 243, 248

G

German 23, 24, 26, 139, 163, 168, 172, 194, 197, 198, 204, 208
  Middle German 219 (n. 13)
  Pennsylvania German 191
  Pidgin German (Namibia) 153, 157, 213
  Upper (Southern) German 158, 177 (n. 7), 180, 190
  Virginia German 140
gemixte taal 69
Genootskap van Regte Afrikaners 1, 55ff., 176, 188, 203, 240–243, 281, 282, 287, 290
Ghommaliedjies 52
Gri (Xri) *see* Khoesaan
Griqua Afrikaans 17, 27, 35, 39, 60, 80, 82, 146, 158, 160, 164, 186, 194, 198, 199, 205, 212

H

Haugen, E. 2, 7, 11 (n. 2), 248
Hesseling, D. C. 21, 84, 192, 209
historical sociolinguistics 8, 9, 11 (n. 1), 83, 87, 101, 102, 297, 303, 304
Hofmeyr, J. H. 58, 68, 69, 94, 289
Holm, J. 22, 43 (n. 12), 163
homestead society 28–30, 39, 42
homophonous diamorphs 252
Hoogenhout, C. P. 56, 73, 89, 240–243, 273, 274, 281, 296 (n. 3)
Hoogenhout, N. M. 88, 145, 188, 195
Hottentots *see* Khoesaan
Huguenots *see* colonists
hypercorrection 144, 146
hyperdemoticism 203, 244

I

implicational scaling 79, 80, 230–234
infinitive clause 135, 204ff., 216, 226, 227
insertion (code-switching) 252, 256
interlanguage 30
Iowa Dutch 142
Islam 26, 93

J

Joseph, J. E. 5, 6, 8, 72, 300, 301

K
Kaapse Taalargief 16, 41, 101, 148, 149, 158
Klaas Waarzegger 47, 50, 241
Kaiser's rule 125
Kaatje Kekkelbek *see* Bain, A. G.
Kaufman, T. 22, 30, 38, 78
Khoe languages *see* Khoesaan
Khoesaan (languages) 17, 22, 24, 25, 27, 28, 31–34, 36, 43 (n. 12), 46, 80, 82, 93, 140, 153, 161, 181, 183, 184, 186, 199, 212, 213, 218 (n. 3)
koinézation *see* levelling
koiné 23, 24, 35
Kloss, H. 2
Kloeke, G. G. 16, 23, 24, 31, 139
k-means clustering 114, 119, 133 (n. 8), 222
Korana (!Ora) *see* Khoesaan

L
Labov, W. 43 (n. 14), 77, 78, 83, 88, 101, 115ff., 135, 225, 248, 282
language and identity 3, 4, 20, 34, 37, 38, 54, 71, 72, 76 (n. 19), 78, 203, 204, 264, 269, 276, 277, 292, 301
language continuum 9, 18, 36–38, 64, 66, 67, 146, 148, 159, 177, 223, 230–234, 247, 257, 258 (n. 6), 298, 299
language demarcation/naming 15, 64, 65
language drift *see* drift
language learning 8, 34, 35, 38, 39, 80, 82, 140, 199, 207, 262
language loyalty 61, 75 (n. 12), 76 (n. 19)
language spread 279ff.
language shift 24, 38–40, 80, 82, 262, 195, 304 (n. 1)
language standard(s) (definition) 2–6, *see also* standard language(s)
language universals 40, 298
Lass, R. 16, 20, 49, 87, 139, 158, 169, 172, 173, 180, 297
late majority 286, 287
Le Page, R. B. 3, 4, 9, 15, 18, 36, 108, 130–132, 226, 228, 300
levelling 2, 3, 6, 18, 24, 131, 139, 140, 153, 180, *see also* koinézation
lingua franca 1, 27, 28, 34
linguistic market(place) 71–73, 303
linguistic variable(s) 135, 136, 221, 222
linkage rules (cluster analysis) 112, 113, 116, 222
literacy levels 19, 20, 42 (n. 5), 75 (n. 15), 83ff., 95, 100, 101
loans 31, 34, 271, 276, 277, *see also* borrowing

M
Malay 18, 22, 27, 28, 34, 43 (n. 9), 82, 140, 192, 209, 212, 213, 218 (n. 2)
Malay Afrikaans 17, 52, 80, 93, 218 (n. 8), 275, *see also* Western Cape Afrikaans
markers (sociolinguistic) 79, 80, 136, *see also* language and identity
markedness, marked features 38, 180, 201, 203, 264, 267, 291
marginalization (of non-standard speakers) 7, 9, 276
Masbieker Afrikaans 80
matrix language 251–253, 264
McCormick, K. 11, 80, 82, 91, 94, 249, 256, 257, 262, 276
McWorther, J. 22
medium of instruction 89, 261, 293, *see also* English – medium of instruction
Mentzel, O. F. 34, 36, 65, 183
Meurant, L. M. 47, 48, 50, 51, 148, 188, 240–243, 281
Middle Dutch 190, 197
Milroy, J. & L. 2, 7, 41, 61, 131, 279, 280, 282, 287, 303
misuse of inflection 144, 146
modernization 6, 279ff., 302
monocentric selection 2
Moravian missionaries 52, 53, 74 (n. 4)

Mufwene, S. S. 24, 28–30, 33, 37, 40, 43 (n. 11, 14 & 16)
multidimensional scaling 120ff., 222–224, 230, 231, 234, 235
Muysken, P. 20, 249, 251, 252, 255–257, 264, 270, 273

N
Nama *see* Khoesaan
Namibia 1, 43 (n. 12), 153, 213
nationalism 4, 9, 54ff., 66, 72–74, 91, 243, 244, 269, 270, 281, 286, 287, 288–293, 295, 300, 303
native-speaker ratio 25, 29
negation 33, 41, 45, 46, 61, 135, 196ff., 216, 217, 225–227, 245
  Khoesaan substrate 199
  preverbal 33, 218 (n. 8)
  Dutch dialects 197
  Middle Dutch 197
Negerhollands 74 (n. 4), 190, 191, 218 (n. 4)
New Jersey Dutch 140
New Netherland Dutch 140, 141, 169, 171, 172
newspapers, periodicals
  Cape Times 69, 70
  De Bode van Genadendal 53, 94
  De Ebenzer 94
  De Verzamelaar 46, 241
  De Zuid-Afrikaan 46, 50, 51, 60–62, 65, 68, 74 (n. 2), 241
  Die Afrikaanse Patriot 55ff., 75 (n. 7), 188, 241, 273
  Die Brandwag 293, 294
  Die Huisgenoot 293
  Friend of the Free State 69
  Ons Land 68, 292
  Ons Moedertaal 293
  Zuid-Afrikaansche Tijdschrift 70, 274
Nienaber, G. S. 23, 48, 49, 51, 59, 186, 199, 242
norms(linguistic) 2–7, 10, 40, 60–67, 77–80, 106, 130–132, 141, 142, 159, 166, 176, 177, 194, 195, 215–217, 227, 228, 242, 244, 245, 257, 279–281, 286, 288, 299, 300
norm awareness 244, 245, 255
North Germanic 139, 172, 191, 194
noun phrase
  adjective inflection 169ff.
  agreement hierarchy 168
  associative plural marker (N-*hulle*) 22, 186, 187
  case marking 177 (n. 8)
  demonstratives 193ff.
  diminutive 167
  gender (grammatical) 165ff., 221ff., 230ff.
  genitive 177 (n. 8), 211, 213
  hybrid nouns 168, 169
  nominativization 180, 185
  objective *vir* 207ff., 221ff.
  possession 211ff.
  pronouns, personal 19, 179ff., 221ff., 230ff.
  pronouns, possessive 187ff., 221ff., 230ff.
  relativizer 189ff.

O
Oosgrens-Afrikaans *see* Eastern Cape Afrikaans
Orange River Afrikaans *see* Griqua Afrikaans

P
Pannevis, A. 74 (n. 6), 76 (n. 18), 282
Paul, H. 67, 68
performance 4, 79, 78, 105, 106
phonology 16, 38, 53
pidgin *see* Cape Dutch Pidgin
preposition stranding 190
prestige planning 57
pro-drop 137
pronominal clitics
  Khoe 43 (n. 13)
  Dutch 179, 180, 182

plantation society 28–30, 34
Ponelis, F. 9, 23, 25, 28, 30, 37, 42 (n. 1), 43 (n. 9), 69, 76 (n. 18), 139, 143, 146, 148, 158, 165, 166, 171, 179, 180, 182–184, 187, 188, 191, 192, 194, 199, 200, 205, 208, 210, 214, 258 (n. 4), 304 (n. 1)
poor whites 99
Portuguese *see* Creole Portuguese
(post-)creole continuum *see* decreolization
principal components analysis 123ff., 229, 230
printing press 19
prototypes 21, 51, 243
Pseudo-H 242
purism (linguistic) 11, 61, 62, 75 (n. 12), 244, 247, 275, 276, 299

Q
questione della lingua 73, 74

R
Rademeyer, J. 27, 146, 160, 164, 198, 212
Raidt, E. H. 9, 10, 15ff., 42 (n. 2 & 6), 45, 49, 53, 54, 56, 76 (n. 18), 89, 92, 139, 145, 148, 158, 166, 171, 176, 180, 181, 183, 184, 192–195, 197, 201, 202, 208–210, 237
reading pronunciation 137
reanalysis 38, 39, 140, 186, 213
reduplication 18, 22
Rehoboth Basters 17, 27, 35, 82
relexified Khoe 31
Roberge, P. T. 9, 11 (n. 4), 15, 17–21, 31, 33, 34, 36–40, 46, 92, 160, 161, 164, 181, 194, 197, 199–201, 205, 207, 213, 214, 218 (n. 9)

S
salience 38, 40, 136, 201
sample size 88, 89, 126, 127, 136, 137
San *see* Khoesaan
schizoglossia 248
Scholtz, J. du P. 9, 10, 15ff., 23, 42 (n. 2 & 6), 48, 49, 145, 147, 152, 158, 176, 183, 185, 197, 210, 211, 213, 262
Schuchardt, H. 15, 35, 36, 209, 218 (n. 13)
scree plot 127, 128, 228, 229
second dialect acquisition 91, *see also* bi-dialectalism
second language acquisition 29, 30, 33, 34–36, 42, 43 (n. 10 & 12), 82, 140, 298
S-curve 142, 216, 217, 286
Sendingskerk 95
semi-creole 22, 39, 40
semi-speaker 78
settlers *see* colonists
Shepard diagram 120, 121,
slaves 17, 25–29, 82, 93, 99, 209
  contact with Khoe 33, 34, 36
  languages *see* Creole Portuguese, Malay
  literacy 42 (n. 5)
  Masbieker 80
  free slaves (*vrije zwarten*) 37
smallpox epidemic 24
social networks 78, 79, 179ff.
social variables
  age 90, 91, 236–240
  gender 91, 92, 236–240
  ethnicity 92–95
  class 95–100, 236–240
societal reduction 78
Somerset, Lord C. 261
South African Academy 58, 59, 293
South African Philological School 15, 41, 42 (n. 1 & 6)
Standard Language Ideology (SLI) 2, 61, 303
standard language(s) (definition) 6, 7, *see also* language standards
standard view (of the history of Afrikaans) 9, 11 (n. 3)
standardization cycle 81
Statenbijbel *see* Bible
stereotypes 49, 60, 79, 80, 136, 194

Straatpraatjes 240ff., 292
strong ties 285, 286
style-shifting 67, 255, 263
superstratist hypothesis 18, 139
Surinamese Dutch 169, 174
stabilization 3, 26, 29, 34, 35, 39, 131
statistics (descriptive and inferential) 105–107, 236
stylization 48, 49, 53, 54, 72, 292

T
target of acquisition 35
Teenstra, M. D. 17, 45, 46, 48, 65, 200, 218 (n. 8)
tertiary hybridization 34
Thomason, S. 22, 30, 38, 39, 78
trade jargon 22
trekboere 27
Trigardt diary 16, 17, 20, 41, 148, 158, 192, 200

U
UCINET 284
urbanization 4, 41, 81, 82, 91, 99, 100, 281, 301

V
Valkhoff, M. F. 35, 82, 140, 184, 213, 270
Van Ginneken, J. 23, 27, 80, 140, 165, 195
Van Oordt, L. C. *see Kaapse Taalargief*
Van Riebeeck, J. 25
variable rules 132 (n. 2), 147
variationist model of language change 20, 21, 43 (n. 12)
verb phrase
  ablaut 146, 152–156
  apocope of *-e(n)* 152ff.
  apocope of *-t* 146ff.
  auxiliaries 39, 138, 239, 142, 158, 177 (n. 1), 198, 242
  historical present 163
  inflectional loss 137ff., 221ff., 230ff.
  infinitive 23, 33, 61, 137ff., 152f., 233, 242
  main verbs 137ff., 221ff., 230ff.
  modal auxiliaries 138, 139, 142, 143
  tense and aspect adverbials 160–165
  past participles 146ff., 152ff., 221ff., 230ff.
  past tense 156–160, 221ff.
  present tense 137ff., 221ff., 230ff.
verbal hendiadys 39, 40
Verenigde Oostindische Compagnie (VOC) 16, 23–27, 31, 34, 148, 262
Vernederlandsing 59, 275
volkseenheid 194

W
weak ties 280, 285, 286, 288, 289
Weinreich, U. 6, 43 (n. 14), 75 (n. 12), 158, 191, 218 (n. 4)
Western Cape Afrikaans 27, 80, 174, 212, 218 (n. 12), *see also* Malay Afrikaans
Witbooi, H. 17, 158
word order 23, 197
written language 2, 19, 20, 42 (n. 6), 63, 65, 75 (n. 15), 83, 84, 92, 137, 181, 217, 298, *see also* education, literacy

Y
Yiddish 11 (n. 3), 21, 158, 190, 191, 194, 262

Z
Zuidafrikaanse Akademie voor Taal, Letteren en Kunst *see* South African Academy
Zuid-Afrikaansche Taalbond 69, 70, 72, 73

In the series IMPACT: STUDIES IN LANGUAGE AND SOCIETY the following titles have been published thus far or are scheduled for publication:

1 **PÜTZ, Martin (ed.):** Language Choices. Conditions, constraints, and consequences. 1997. xxi, 430 pp.

2 **KIBBEE, Douglas A. (ed.):** Language Legislation and Linguistic Rights. Selected Proceedings of the Language Legislation and Linguistic Rights Conference, the University of Illinois at Urbana-Champaign, March, 1996. 1998. xvi, 415 pp.

3 **LINELL, Per:** Approaching Dialogue. Talk, interaction and contexts in dialogical perspectives. .. 1998. xvii, 322 pp.

4 **OWENS, Jonathan:** Neighborhood and Ancestry. Variation in the spoken Arabic of Maiduguri, Nigeria. 1998. xiv, 390 pp.

5 **ANDREWS, David R.:** Sociocultural Perspectives on Language Change in Diaspora. Soviet immigrants in the United States. 1999. xviii, 182 pp.

6 **RICENTO, Thomas (ed.):** Ideology, Politics and Language Policies. Focus on English. 2000. x, 197 pp.

7 **McCAFFERTY, Kevin:** Ethnicity and Language Change. English in (London)Derry, Northern Ireland. . 2001. xx, 244 pp.

8 **ARMSTRONG, Nigel R.:** Social and Stylistic Variation in Spoken French. A comparative approach. 2001. x, 278 pp.

9 **HELLINGER, Marlis and Hadumod BUSSMANN (eds.):** Gender Across Languages. The linguistic representation of women and men. Volume 1. 2001. xiv, 329 pp.

10 **HELLINGER, Marlis and Hadumod BUSSMANN (eds.):** Gender Across Languages. The linguistic representation of women and men. Volume 2. 2002. xiv, 349 pp.

11 **HELLINGER, Marlis and Hadumod BUSSMANN (eds.):** Gender Across Languages. The linguistic representation of women and men. Volume 3. 2003. xiv, 391 pp.

12 **OKITA, Toshie:** Invisible Work. Bilingualism, language choice and childrearing in intermarried families. 2002. x, 275 pp.

13 **OAKES, Leigh:** Language and National Identity. Comparing France and Sweden. 2001. x, 305 pp.

14 **WEBB, Victor:** Language in South Africa. The role of language in national transformation, reconstruction and development. 2002. xxviii, 357 pp.

15 **BOXER, Diana:** Applying Sociolinguistics. Domains and face-to-face interaction. 2002. xii, 245 pp.

16 **BRITAIN, David and Jenny CHESHIRE (eds.):** Social Dialectology. In honour of Peter Trudgill. 2003. x, 344 pp.

17 **TRINCH, Shonna L.:** Latinas' Narratives of Domestic Abuse. Discrepant versions of violence. 2003. x, 315 pp.

18 **DEUMERT, Ana and Wim VANDENBUSSCHE (eds.):** Germanic Standardizations. Past to Present. 2003. vi, 480 pp.

19 **DEUMERT, Ana:** Language Standardization and Language Change. The dynamics of Cape Dutch. 2004. xx, 354 pp. + index. Expected Spring 2004