

Sociophonetics and Class Differentiation:
A study of working- and middle-class English in
Cape Town's Coloured Community

Tracey Toefy

Supervisor: Professor Rajend Mesthrie

Thesis Presented for the Degree of

Doctor of Philosophy

in the School for African & Gender Studies, Anthropology & Linguistics,

Linguistics Section

University of Cape Town

February 2014

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Acknowledgements

Upon completion of what seemed at times to be an endless endeavour, I humbly acknowledge that without the help and support of many people, far more than are mentioned on the list below, this thesis would not have become a reality.

Firstly, I acknowledge my supervisor, Professor Raj Mesthrie, who has supported me throughout my postgraduate career by creating an enabling environment for me as an aspiring scholar to persevere with one degree after another. I am particularly grateful for the financial assistance provided to me as a product of Prof Mesthrie's own research success as a SARChI chair, and to SARChI and the National Research Foundation for the funding made available which has allowed me to pursue my studies.

As a colleague and friend, Alida Chevalier has been a source of immeasurable help and support to me over the past few years, and to her I am also so thankful. Academics all over the world, some of whom I have never met face-to-face, have provided me with help in fields as varied as computer programming and education, and I express sincere gratitude to the following for their assistance: Keelan Evanini, Thomas Hoffmann, Martin Hilpert, Rochelle Kapp, Ian Bekker and Tyler Kendall.

I am grateful to the 40 speakers who made up the sample for this project for sparing the time to talk to me, allowing me to collect the data needed for this analysis. In that vein, I duly thank Cloe Dennis, Melanie Zeederberg and Chantel Erfort for introducing me to many of the people I interviewed. Thanks are also due to Kabelo Rametse, who patiently and very competently compiled the programs I used to do my analysis.

My friends and family have been a constant source of strength and support throughout this project, and I am thankful to all of them, though they are not mentioned by name here. I am particularly grateful to my father, Stan Dennis, and father-in-law, Hardie Toefy, for driving me to interviews in the more 'dangerous' parts of Cape Town. To my mother, Maureen Dennis, who provided me with editorial help with the thesis itself, I also express my sincere thanks.

Finally, I am most appreciative of the role played by my husband, Shukri Toefy, in the completion of this project. His own sacrifice and hard work enabled me to be a full-time scholar *and* full-time mother, while he provided us with a life that is comfortable, exciting and happy. I am eternally grateful for his unwavering support, encouragement and love.

Abstract

Tracey Toefy | February 2014

Thesis Title: Sociophonetics and Class Differentiation: a study of working- and middle-class English in Cape Town's Coloured Community.

This thesis provides a detailed acoustic description of the phonetic variation and changes evident in the monophthongal vowel system of Coloured South African English in Cape Town. The changes are largely a result of South Africa's post-apartheid socio-educational reform. A detailed acoustic description highlights the most salient changes (compared with earlier reports of the variety), indicating the extent of the change amongst working-class and middle-class speakers.

The fieldwork conducted for this study consists of sociolinguistic interviews, conducted with a total of 40 Coloured speakers (half male, half female) from both working-class and middle-class backgrounds. All speakers were young adults, born between 1983 and 1993, thus raised and schooled in a period of transition from apartheid to democracy. Each of the middle-class speakers had some experience of attending formerly exclusively White schools, giving them significant contact with White peers and teachers, while the educational careers of the working-class speakers exposed them almost solely to Coloured peers and educators.

The acoustic data were processed using methods of Forced Alignment and automatic formant extraction – methods applied for the first time to any variety of South African English. The results of the analysis were found generally to support the findings of scholars who have documented this variety previously, with some notable exceptions amongst middle-class speakers. The changes are attributable to socio-educational change in the post-apartheid setting and the directionality of the changes approximate trends amongst White South African English speakers. The TRAP, GOOSE and FOOT lexical sets show most change: TRAP is lowering, while GOOSE and FOOT are fronting. Although the changes approximate the vowel quality used by White speakers, middle-class Coloured speakers use an intermediate value between White speakers and working-class Coloured speakers i.e. they have not fully adopted White norms for any of the vowel classes. Working-class speakers were found to have maintained the monophthongal vowel system traditionally used by Coloured speakers.

Plagiarism Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is one's own.
2. I have used the Harvard-UCT (author-date) convention for citation and referencing. Each significant contribution to and quotation in, this thesis from the work, or works, of other people has been acknowledged through citation and reference.
3. This thesis is my own work.
4. I have not, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Full Name: Tracey Lynn Toefy

Student number: DNNTRA003

Signature:

Signed by candidate

Date: February 2014

Signature removed

Contents

Acknowledgements	i
Abstract	iii
Plagiarism Declaration	iv
Contents	v
List of Figures	x
List of Tables	xiv
List of Abbreviations	xvi
Chapter One: Introduction	1
1.1 Research Aim	2
1.2 Research Objectives	2
1.3 Research Questions	2
1.4 The Origins of the Coloured Community	3
1.5 Political History of the Coloured community	5
1.6 Apartheid South Africa	10
1.7 Coloured Educational Structures: 1652-1990	16
1.8 Desegregation of Education	20
1.9 ‘Coloured’: A Problematic Label	24
1.10 Coloureds and Language Use	25
1.11 A Note on Methodology	28
1.12 Delimitations	29
1.13 Chapter Outline	29
Chapter Two: Coloured South African English: History and Vowel Quality	31
2.1 History of English in South Africa	32
2.1.1 The Colonial Era	32
2.1.2 Immigrant Influx	33
2.1.3 The Apartheid Era	34
2.1.4 The ‘New’ South Africa	35
2.2 SAE as a ‘Southern’ Variety	35
2.3 Review of SAE literature	36
2.4 Literature on CSAE	39
2.5 Phonetic profile: CSAE	49
2.5.1 Short Vowels	50

2.5.2 Long Monophthongs	52
Chapter Three: Data Collection, Speakers and Social Categories	54
3.1 Criteria for Sample Selection	54
3.1.1 Coloured	55
3.1.2 Cape Town Resident	55
3.1.3 English / English-Afrikaans bilingual	58
3.1.4 The Generation before the ‘Born-frees’	60
3.1.5. Gender	63
3.1.6 Social class	64
3.2 Degree of Integration	76
3.2.1 Schooling	76
3.2.2 Area of Residence	81
3.2.3 Close Contacts	84
3.2.4 Degree of Integration Index	85
3.3 Data Collection and Preparation	88
3.3.1 Interview Procedure and Materials	88
3.3.2 Recording Equipment	90
3.3.3 Transcription	91
Chapter Four: Data Analysis using Automatic Vowel Measurement	92
4.1. Forced Alignment using P2FA	92
4.1.1 Forced Alignment Requirements	94
4.1.2 CMU Pronouncing Dictionary	95
4.1.3 Preparation of Data for P2FA	98
4.1.4 Forced Alignment Accuracy	100
4.2 Automatic Extraction of Formant Measurements using extractFormants	103
4.2.1 Selecting an Automatic Measurement Point	104
4.2.2 Formant Prediction	106
4.3 Exclusions and Isolated Environments	108
4.4 Outliers	109
4.5 Normalisation using NORM	109
4.6 Statistical analysis, Data Manipulation and Presentation in R	110
4.6.1 Graphical representation of F1 and F2	111
4.7 Conclusion	111

Chapter Five: Short Vowels	112
5.1 Introduction	112
5.2 SAE Chain Shift	114
5.3 Acoustic Overview of the Short Vowels of CSAE	117
5.4 KIT	119
5.4.1 WSAE Reports	119
5.4.2 CSAE Reports	121
5.4.3 CSAE Acoustic Data	122
5.4.4 Summary: KIT Findings	133
5.5 DRESS	135
5.5.1 WSAE Reports	135
5.5.2 CSAE Reports	135
5.5.3 CSAE Acoustic Data	136
5.5.4 Summary: DRESS Findings	141
5.6 TRAP	143
5.6.1 WSAE Reports	143
5.6.2 CSAE Reports	143
5.6.3 CSAE Acoustic Data	144
5.6.4 Summary: TRAP Findings	149
5.7 LOT	151
5.7.1 WSAE Reports	151
5.7.2 CSAE Reports	151
5.7.3 CSAE Acoustic Data	152
5.7.4 Summary: LOT Findings	158
5.8 STRUT	159
5.8.1 WSAE Reports	159
5.8.2 CSAE Reports	159
5.8.3 CSAE Acoustic Data	160
5.8.4 Summary: STRUT Findings	165
5.9 FOOT	166
5.9.1 WSAE Reports	166
5.9.2 CSAE Reports	166
5.9.3 CSAE Acoustic Data	167
5.9.4 Summary: FOOT Findings	171

5.10 Conclusion	172
Chapter Six Long Monophthongs	174
6.1 Introduction	174
6.2 Acoustic Overview of the Long Monophthongs of CSAE	175
6.3 FLEECE	176
6.3.1. WSAE Reports	176
6.3.2 CSAE Reports	176
6.3.3 CSAE Acoustic Data	177
6.3.4 Summary: FLEECE Findings	183
6.4 NURSE	184
6.4.1. WSAE Reports	184
6.4.2. CSAE Reports	184
6.4.3 CSAE Acoustic Data	185
6.4.4 Summary: NURSE Findings	189
6.5 GOOSE	190
6.5.1 WSAE Reports	190
6.5.2 CSAE Impressionistic Reports	190
6.5.3 CSAE Acoustic Data	191
6.5.4 Summary: GOOSE Findings	196
6.6 THOUGHT	197
6.6.1 WSAE Reports	197
6.6.2 CSAE Reports	197
6.6.3 CSAE Acoustic Data	198
6.6.4 Summary: THOUGHT Findings	202
6.7 BATH	203
6.7.1 WSAE Reports	203
6.7.2 CSAE Reports	203
6.7.3 CSAE Acoustic Data	204
6.7.4 Summary: BATH Findings	210
6.8 SQUARE: A Rogue Monophthongal Diphthong	211
6.8.1 WSAE Reports	211
6.8.2 CSAE Reports	211
6.8.3 CSAE Acoustic Data	212
6.8.4 Summary: SQUARE Findings	217

6.9 Conclusion	218
6.10 CSAE Vowel System	219
Chapter Seven: Summary and Conclusions	221
7.1 Overview of Research Aim	221
7.2 Research Objectives Recapped	221
7.3 Research Questions Recapped	221
7.4 Research Questions Answered	221
7.5 Significance of the Research	222
7.6 A Methodological Note	222
7.7 Summary of Findings	223
7.8 Relevance of the Findings	226
7.9 Directions for Future Research	227
7.10 Conclusion	228
Appendix A: Demographic Data	229
Appendix B: Word List	230
Appendix C: Reading Passage	231
References	232

List of Figures

1.1 Provinces of South Africa	4
2.1 Lectal Hierarchy of SAE	38
3.1 Districts of Cape Town Metropole	57
3.2 Bar chart of birth year for all 40 speakers, 1983-1992	61
4.1 Force-aligned depiction of the phrase <i>A third cat joined the party</i>	95
4.2 Misalignment of the phrase <i>staff was</i> (speaker M6)	100
4.3 LPC formant tracker: number of formants set to 5 (default value)	107
4.4 LPC formant tracker: number of formants set to 6	107
5.1 Sample boxplot (box-and-whisker plot) for F1 data	114
5.2 Short front vowel raising in WSAE	115
5.3 An acoustic overview of the short vowel system of CSAE, via mean values for middle- and working-class speakers	117
5.4 Range of variation for short vowels of WSAE in WL style	120
5.5 Mean values for IT~SIT in Interview Style (IS), disaggregated by class and gender	122
5.6 Mean values for IT and IT_/l/ in Interview Style (IS), disaggregated by class and gender	123
5.7 Mean values for SIT and SIT_/l/ in Interview Style (IS), disaggregated by class and gender	124
5.8 Boxplot for KIT F1 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by class	126
5.9 Boxplot for KIT F2 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by class	127
5.10 Boxplot for KIT F1 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by gender	128
5.11 Boxplot for KIT F2 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by gender	129
5.12 Boxplot for KIT F1 values in all speech styles, disaggregated by class	130

5.13	Boxplot for KIT F2 values in all speech styles, disaggregated by class	131
5.14	Boxplot for KIT F1 values in all speech styles, disaggregated by gender	132
5.15	Boxplot for KIT F2 values in all speech styles, disaggregated by gender	133
5.16	Mean values for DRESS and DRESS_/l/ in Interview Style, disaggregated by class and gender	136
5.17	Boxplot for DRESS, <i>jealous</i> , <i>jelly</i> and <i>yes</i> F2 values, disaggregated by class .	137
5.18	Boxplot for DRESS F1 values in all speech styles, disaggregated by class	138
5.19	Boxplot for DRESS F2 values in all speech styles, disaggregated by class	139
5.20	Boxplot for DRESS F1 values in all speech styles, disaggregated by gender ..	140
5.21	Boxplot for DRESS F2 values in all speech styles, disaggregated by gender ..	141
5.22	Mean values for TRAP in Interview Style, disaggregated by class and gender	144
5.23	TRAP tokens for speakers F7 and F9, showing the position of tokens before velar /l/	145
5.24	Boxplot for TRAP F1 values in all speech styles, disaggregated by class	146
5.25	Boxplot for TRAP F2 values in all speech styles, disaggregated by class	147
5.26	Boxplot for TRAP F1 values in all speech styles, disaggregated by gender	148
5.27	Boxplot for TRAP F2 values in all speech styles, disaggregated by gender	149
5.28	Mean values for LOT and LOT_/l/ in Interview Style, disaggregated by class and gender	152
5.29	Mean values for LOT and <i>want</i> tokens, disaggregated by class and gender	153
5.30	Mean values for LOT and <i>non-</i> tokens, disaggregated by class and gender	154
5.31	Boxplot for LOT F1 values in all speech styles, disaggregated by class	155
5.32	Boxplot for LOT F2 values in all speech styles, disaggregated by class	156
5.33	Boxplot for LOT F1 values in all speech styles, disaggregated by gender	157
5.34	Boxplot for LOT F2 values in all speech styles, disaggregated by gender	158
5.35	Mean values for STRUT and STRUT_/l/ in Interview Style, disaggregated by class and gender	160
5.36	Mean values for STRUT and ONE in Interview Style, disaggregated by class and gender	161
5.37	Boxplot for STRUT F1 values in all speech styles, disaggregated by class	162
5.38	Boxplot for STRUT F2 values in all speech styles, disaggregated by class	163
5.39	Boxplot for STRUT F1 values in all speech styles, disaggregated by gender ..	164
5.40	Boxplot for STRUT F2 values in all speech styles, disaggregated by gender ..	165

5.41 Mean values for FOOT and FOOT_/l/ in Interview Style, disaggregated by class and gender	167
5.42 Boxplot for FOOT F1 values in all speech styles, disaggregated by class	168
5.43 Boxplot for FOOT F2 values in all speech styles, disaggregated by class	169
5.44 Boxplot for FOOT F1 values in all speech styles, disaggregated by gender	170
5.45 Boxplot for FOOT F2 values in all speech styles, disaggregated by gender	171
6.1 An acoustic overview of the long vowel system of CSAE, showing mean values for middle- and working-class speakers	175
6.2 Mean values for FLEECE in Interview Style (IS), disaggregated by class and gender	177
6.3 Boxplot for FLEECE F1 values in all speech styles, disaggregated by class	180
6.4 Boxplot for FLEECE F2 values in all speech styles, disaggregated by class	181
6.5 Boxplot for FLEECE F1 values in all speech styles, disaggregated by gender ...	182
6.6 Boxplot for FLEECE F2 values in all speech styles, disaggregated by gender ...	183
6.7 Mean values for NURSE in Interview Style (IS), disaggregated by class and gender	185
6.8 Boxplot for NURSE F1 values in all speech styles, disaggregated by class	186
6.9 Boxplot for NURSE F2 values in all speech styles, disaggregated by class	187
6.10 Boxplot for NURSE F1 values in all speech styles, disaggregated by gender ..	188
6.11 Boxplot for NURSE F2 values in all speech styles, disaggregated by gender ..	189
6.12 Mean values for GOOSE in Coronal and Non-Coronal environments in Interview Style (IS), disaggregated by class and gender	191
6.13 Mean values for GOOSE Non-Coronal environments and GOOSE_/l/ in Interview Style (IS), disaggregated by class and gender	192
6.14 Boxplot for GOOSE F1 values in all speech styles, disaggregated by class	193
6.15 Boxplot for GOOSE F2 values in all speech styles, disaggregated by class	194
6.16 Boxplot for GOOSE F1 values in all speech styles, disaggregated by gender ..	195
6.17 Boxplot for GOOSE F2 values in all speech styles, disaggregated by gender ..	196
6.18 Mean values for THOUGHT in Interview Style (IS), disaggregated by class and gender	198
6.19 Boxplot for THOUGHT F1 values in all speech styles, disaggregated by class	199
6.20 Boxplot for THOUGHT F2 values in all speech styles, disaggregated by class.	200
6.21 Boxplot for THOUGHT F1 values in all speech styles, disaggregated by gender	201

6.22 Boxplot for THOUGHT F2 values in all speech styles, disaggregated by gender	202
6.23 Mean values for BATH in Interview Style (IS), disaggregated by class and gender	204
6.24 BATH tokens for speakers F2 and M12, showing the position of tokens before velar /l/	205
6.25 Mean values for BATH and DANCE in Interview Style (IS), disaggregated by class and gender, and single token of <i>chance</i> by speaker M15 ...	206
6.26 Boxplot for BATH F1 values in all speech styles, disaggregated by class	207
6.27 Boxplot for BATH F2 values in all speech styles, disaggregated by class	208
6.28 Boxplot for BATH F1 values in all speech styles, disaggregated by gender	209
6.29 Boxplot for BATH F2 values in all speech styles, disaggregated by gender	210
6.30 Mean values for SQUARE in Interview Style (IS), disaggregated by class and gender	212
6.31 Boxplot for SQUARE F1 values in all speech styles, disaggregated by class ...	214
6.32 Boxplot for SQUARE F2 values in all speech styles, disaggregated by class ...	215
6.33 Boxplot for SQUARE F1 values in all speech styles, disaggregated by gender	216
6.34 Boxplot for SQUARE F2 values in all speech styles, disaggregated by gender	217
6.35 An acoustic overview of the simple vowel system of CSAE, mean values for middle- and working-class speakers	219

List of Tables

1.1 Relative per capita income as a percentage of White level, 1917-1995	14
1.2 Per Capita School Spend, by Race, 1970-1990, in ZAR	19
1.3 Annual fees charged in schools in Western Cape in ZAR, 2001, classified by former department	22
2.1 Diphthong quality in CSAE, according to Wood (1987)	43
2.2 The short and long monophthongs of CSAE and WSAE	50
3.1 Male and Female speakers' year of birth, L1 and L2	62
3.2 Socioeconomic Status Index: Four components	68
3.3 Male Speakers' occupations and parents' occupations	71
3.4 Female speakers' occupations and parents' occupations	72
3.5 Monthly Household income for Coloured households in Athlone, 2011 census data	73
3.6 SES scores for 40 speakers, by component	75
3.7 Socioeconomic Scale	76
3.8 Speakers' Socioeconomic ranking	76
3.9 Speakers' Primary Schools, Fees and Former Department	78
3.10 Speakers' High Schools, Fees and Former Department	79
3.11 Schools attended by the speakers	80
3.12 Area of Residence	82
3.13 Profile of four suburbs, 2011 data	84
3.14 Average monthly household income in four suburbs (2011 data), percentage .	84
3.15 Degree of Integration Index	85
3.16 Speaker scores on Degree of Integration Index	86
3.17 Degree of Integration Scale	87
3.18 Speakers' ranking on Degree of Integration Scale	87
4.1 The 20 most frequent misaligned words	102
5.1 Distribution of the allophones of KIT across the three lects of WSAE (Lass 2002: 114-115)	119
5.2 Six allophones of KIT (Bekker 2009: 267)	120
5.3 Vowel length of two IT and two SIT tokens for four speakers (msec)	125
5.4 Phonological environments for KIT tokens	125

5.5 Summary of acoustic findings for the short vowels of CSAE	172
6.1 Comparative mean values for F1 and F2 in regular and pre-/l/ tokens of FLEECE (Hertz)	178
6.2 Comparative formant values of six tokens of FLEECE for four speakers (Hertz; unnormalised)	179
6.3 Comparative mean values for F1 and F2 in regular and pre-/l/ tokens of NURSE (Hertz)	185
6.4 Comparative mean values for F1 and F2 in regular and pre-/l/ tokens of THOUGHT (Hertz)	199
6.5 Comparative formant values for six tokens of SQUARE for four speakers (Hertz; unnormalised)	213
6.6 Summary of acoustic findings for the long vowels of CSAE	218

List of Abbreviations

AfrE	Afrikaans English
ANAE	Atlas of North American English
ANC	African National Congress
APO	African People’s Organisation (originally African Political Organisation)
AVM	Automatic Vowel Measurement
BC	Black Consciousness
BEE	Black Economic Empowerment
BSAE	Black South African English
BrE	British English
CAC	Coloured Advisory Council
CAD	Coloured Affairs Department
CE	Cape English
CMU	Carnegie Mellon University
CRC	Coloured Representatives Council
CSAE	Coloured South African English
CPUT	Cape Peninsula University of Technology
DE	Dutch English
DET	Department of Education and Training
ETE	Extraterritorial English
IS	Interview Style
GMM	Gaussian Mixture Model
HMM	Hidden Markov Model
HOA	House of Assembly
HOD	House of Delegates
HOR	House of Representatives
L1	First language
L2	Second language
LPC	Linear Predictive Coding
MC	Middle-class
MOI	Medium of Instruction
NAE	North American English
NE	Natal English
NEUM	Non-European Unity Movement
NP	National Party
P2FA	Penn Phonetics Lab Forced Aligner

P2TK	Penn Phonetics Toolkit
PLP	Perceptual Linear Prediction
RP	Reading Passage (Style)
RSA	Republic of South Africa
SA	South Africa
SAE	South African English
SARP	South African Received Pronunciation
SECS	South-East England short vowel chain shift
SES	Socioeconomic Status
SGB	School Governing Body
SRC	Students' Representative Council
TLSA	Teachers' League of South Africa
UCT	University of Cape Town
UK	United Kingdom
US	United States (of America)
UWC	University of the Western Cape
WC	Working-class
WL	Word List (Style)
WSAE	White South African English

Chapter One

Introduction

The research presented in this thesis is framed historically within the social and political context of apartheid South Africa. The consequences of legislated population segregation and subsequent democratisation has impacted significantly on every aspect of the lives of South African citizens. The phonetic systems of the languages spoken by these citizens is one of the areas impacted, and this study focuses on one such system, viz. the phonetic variation within the vowel system of English spoken in the Coloured¹ community of Cape Town. By drawing these aspects together, the original contribution to knowledge presented in this dissertation is an acoustic account of the monophthongal vowels of an under-studied variety of South African English (henceforth SAE), viz. Coloured SAE using modern acoustic methodology including automatic methods of forced alignment and formant extraction.

Speech is imbued with social meaning because of the intrinsic nature of the relationship between language use and social context (Hay and Drager 2007: 90). The present study aims to investigate what social meaning is attached to the speech of Coloured Capetonians in a post-apartheid setting. Naturally, Coloureds are not a homogeneous group, and one of the primary factors which differentiates members of the community from one another is social class. While the boundaries of class are somewhat indistinct, it is possible to categorise members into one of two broad groups: working-class and middle-class. Dividing the community in this way allows for a more accurate account of the changes occurring in the phonetic system.

In order to set the context for the present study, the research aim, objectives and research questions are presented in sections 1.1, 1.2 and 1.3 respectively. The origins and political history of the community of Coloured people is detailed in sections 1.4 and 1.5. Following that, in section 1.6, is a general background into apartheid South Africa and its segregatory laws, through which the social conditions (which provide the basis for this research) were created. This section focuses on the impact of these laws on Coloured people in particular. Section 1.7 is an account of the educational

¹ The term Coloured is not an uncontested label for the community being studied. For the sake of simplicity, the term will be used without scare quotes throughout the thesis, and the complexity and contestations around the term will be discussed in section 1.9 below.

² 'San' is a term which replaced the derogatory ethnonym 'Bushman' (Traill 2002: 45). The Khoe (also 'Khoi' or 'Khoikhoi') were previously referred to as 'Hottentots' (du Pré 1994: 11).

structures catering for Coloureds prior to and during apartheid, significant because the people who make up the sample in this study are largely a product of these structures. This is followed in section 1.8 by a discussion of the process of desegregation in schools following the collapse of the apartheid government and the inception of democracy in South Africa. Section 1.9 problematises the label ‘Coloured’. Language use in the Coloured community is described in section 1.10, both historically and as it is currently used. A brief note on the methodology employed for data analysis follows in section 1.11, and section 1.12 discusses the delimitations of the study. The chapter concludes with an outline of the chapters that comprise this thesis (section 1.13).

1.1 Research Aim

The ultimate aim of the research is to provide a detailed acoustic description of the phonetic variation and changes evident in the vowel system of Coloured SAE (henceforth CSAE) speakers in greater Cape Town. The investigation highlights the most salient changes, with reference to earlier accounts on the variety (which used traditional auditory methods of analysis), and reveal whether or not similar changes are apparent amongst the middle-class and working-class youths.

1.2 Research Objectives

To realise the aim stated above, the research is guided by the following objectives:

- i. To produce a complete acoustic analysis of the monophthongal vowel system used by CSAE speakers, both middle-class and working-class, using modern methods of acoustic analysis.
- ii. To highlight the most salient changes in the vowel system of middle-class speakers.
- iii. To investigate whether changes in the middle-class system are evident in the working-class vowel system, and if so, the extent of the change in the latter system.

1.3 Research Questions

The following research questions were proposed in order to maintain a clear focus on the aim and objectives:

- i. How does an acoustic phonetic analysis of present day CSAE differ from the existing accounts of CSAE phonetic and phonological systems, if at all?
- ii. What are the most salient changes that have arisen in the simple vowel system of middle-class CSAE speakers in a post-apartheid social setting?

- iii. To what extent are such changes evident in the simple vowel system of working-class CSAE speakers?

Bearing in mind that these questions frame the research endeavour, the following sections present essential background information which provide the social, historical and political context for this study.

1.4 The Origins of the Coloured Community

Before proceeding, it is necessary to provide a historical description of who is actually designated by the term 'Coloured'. In compiling this section and the following one on the political history of the Coloured community, I have drawn on a number of accounts which detail the history of the group of people classified as 'Coloured' by the apartheid government. Al J. Venter's *Coloured: A profile of two million South Africans* (1974) is a very detailed account, but very much a product of its time in terms of its outlook and descriptions of the apartheid system. Roy Du Pré writes another detailed volume, *Separate but Unequal: The 'Coloured' people of South Africa – a political history* (1994), which expresses his flagrant disapproval of the apartheid system, and particularly the plight of the Coloured people, through frequent comparisons with Hitler's Nazi Germany. Mohamed Adhikari discusses the development of racial identity of Coloured people in *Not White Enough, Not Black Enough* (2005). Two histories with a political bent are *Between the Wire and the Wall* by Gavin Lewis (1987) and *The Rise and Decline of Apartheid* by R. E. van der Ross (1986). These volumes are very useful in providing information for this general overview of the origins and political history of the Coloured people.

Coloureds make up 8.9 percent of the South African population, numbering just under four and a half million in a country of close to 52 million people (Statistics South Africa 2012: 21). In the Western Cape Province (see figure 1.1 below for a map of the provinces of South Africa), Coloureds are a numerical majority, making up 48.8 percent of the population (Statistics South Africa 2012: 21). Whites make up 15.7 percent of the Western Cape population, 32.8 percent are Black and one percent Indian (Statistics South Africa 2012: 21). Sixty-one percent of all Coloured South Africans live in the Western Cape (Statistics South Africa 2012: 21), with relatively small Coloured communities in other parts of the country. The geographical base of this study is thus the City of Cape Town, capital city of the Western Cape Province.



Figure 1.1: Provinces of South Africa

Source: <http://www.issafrica.org/pubs/monographs/no73/map.html>

The Western Cape is essentially the birthplace of Coloured people, and remains their main area of concentration within the country. The Cape Colony grew through arrivals of Dutch and English settlers and shipments of slaves to add to the local inhabitants at the time (Venter 1974: 13-22). Upon the arrival of the first settler in the Cape, Jan van Riebeeck, in 1652, he and his small team were met by the local inhabitants of the area at the time, the Khoe and San². Khoe and San were distinguished primarily by occupation: the Khoe were herders while the San were hunter-gatherers (Besten 2009: 135), though some scholars are skeptical about such an easy division. Collectively they are referred to as Khoesan.

² 'San' is a term which replaced the derogatory ethnonym 'Bushman' (Traill 2002: 45). The Khoe (also 'Khoi' or 'Khoikhoi') were previously referred to as 'Hottentots' (du Pré 1994: 11).

Due in part to the fact that very small numbers of European women were present at the settlement, miscegenation between the Dutch (and other European) settlers and local Khoesan was common. In 1658, of a total of 360 Dutch inhabitants of the colony, women and children numbered only 20 (Kies 1939: 5). The slave community at the Cape was very diverse in terms of ethnic origin (Worden 1985), having been brought to the colony from parts of Africa, including Mozambique and Madagascar, and South-East Asia, including India (and Bengal), Ceylon (modern day Sri Lanka), and Indonesia (Lewis 1987: 8; McCormick 2002: 15). Despite attempts to prohibit marriage and other union between Europeans and slaves, many 'mixed blood' children were born in the colony during this time. There was also interaction between the slave community and the local Khoesan. It is the descendents of these local indigenes, settlers and slaves who eventually came to be classified 'Coloured' under the apartheid system.

Because of the very diverse origins of the Coloured people, the community remains phenotypically diverse, ranging from very fair in complexion, light eyes, straight and sleek hair to very dark of complexion, dark eyes and *kroes* hair i.e. coarse, curly hair. The former might easily be thought to be White, and the latter Black, and in fact, these ambiguities were the basis for appeals following the racial categorisation in terms of the Population Registration Act of 1950 (details in section 1.5 below). The Coloured community is also religiously diverse, comprised of Christians and Muslims (and some Rastafarians). The two primary religious groupings within the community were introduced to South Africa during the period of Dutch colonial administration, with the Europeans bringing the message of Christianity, and Islam imported along with many East Indian slaves who practised the religion (Lewis 1987: 8).

1.5 Political History of the Coloured community

In the early years of the colony, racially based discrimination was not as common as that which was based on religion. The Khoesan and slaves were admonished for being heathens (du Pré 1994: 13), and upon manumission, slaves who had converted to Christianity were allowed to marry Europeans (du Pré 1994: 14). However, colour was becoming increasingly prevalent as a criterion for subjugation in the colony. People of colour, including Khoesan, slaves and offspring from interracial relationships, were subjected to restrictions on their movement and land ownership (du Pré 1994: 38; Soudien et al., forthcoming: 57; 60).

The British colonial administration initially allowed Coloured and White men equal rights as voters, but introduced franchise restrictions, in respect of income and property ownership, towards the end of the 19th century (Adhikari 2005: 3). The franchise qualification applied to everyone in the Colony, but excluded many more Coloured than White men. The British kept the franchise qualification reasonably low because it benefitted them to have a Coloured electorate, as the British were outnumbered by the Dutch in the Cape Colony (du Pré 1994: 42). Women were disenfranchised until the 1930s, and then it was only White women who were admitted to the common roll of voters, effectively doubling the White vote, and greatly diluting the power of the Coloured vote (van der Ross 1986: 97).

At the turn of the 20th century, the political rights of Coloureds were equal to Whites; there were, however, many discriminatory practices which prevented Coloureds from exercising those rights in the same way that Whites were able to. These include lack of education, expertise and job opportunities (van der Ross 1986: 248). The 20th century saw the erosion of the civil rights accorded to Coloureds, as the government began to realise the threat that a potential alliance between Coloureds and Blacks (who make up the overwhelming majority of the population) would pose to White dominance.

The first step in the attrition of Coloureds' rights came at the time when the Union of South Africa was formed in 1910. The British colonies of the Cape and Natal along with the Boer republics of the Transvaal and Orange Free State were to make up the Union. Upon settling the matter of Union voting rights, the Cape 'liberals' fought for the franchise to be extended to people of colour as it had been in the Cape Colony. The representatives from the other colony and republics were opposed to this, however. As a compromise, each of the provinces of the Union retained the franchise laws in place before the establishment of the Union, so Coloureds in the Cape retained their voting rights, but were precluded from candidature for the Union parliament (du Pré 1994: 48-49; Adhikari 2005: 3).

Unhappy with their treatment as second class citizens, many Coloureds banded together in support of an organisation led by Dr Abdullah Abdurahman, the African Political Organisation (APO; later African People's Organisation), established in 1904. The APO was the first substantive Coloured political body, and it dominated Coloured protest politics for 40 years (Adhikari 2005: 4). The APO actively attempted

to have the subordinate position of Coloured people revised, asserting that their primary desire was to assimilate into the dominant society. The ruling government at the time paid lip service to the concerns of the APO because the Coloured vote remained valuable in strengthening its parliamentary power. Coloureds were assured that they would never be subject to political and economic segregation, only to social segregation (van der Ross 1986: 121).

In 1937, a Coloured Affairs Department (CAD) was set up to deal with all matters pertaining to Coloureds (du Pré 1994: 58-59), but the CAD was viewed as a tool of the government to implement further segregation and discrimination against Coloured people. An anti-CAD movement was established to oppose the increasing discrimination against Coloureds, based on the belief that nothing but full democratic rights for Coloured people should be acceptable and that bodies that concern themselves exclusively with Coloured affairs could not achieve this. General Jan Smuts attempted to diffuse the hostility of the anti-CAD movement by establishing the Coloured Advisory Council (CAC) in 1944. The CAC was tasked with advising government on matters pertaining to the welfare of the Coloured community. Shortly after the formation of the CAC, prominent resistance movements including the Teachers' League of South Africa (TLSA) and the Non-European Unity Movement (NEUM) were established to mobilise against the government's policy of segregation of Coloured and White people. The formation of the CAC caused a schism within the resistance movements: one faction believed that improved conditions for Coloureds could be secured through cooperation with the government through the CAC, while the other faction saw it as a forerunner to further segregation and discriminatory practice, as government was not taking active steps to reform the problems about which Coloured people were complaining (du Pré 1994: 60).

Members of the CAC took a decision in 1950 to resign as a body because they felt that the system of apartheid espoused by the NP government was at odds with the desires of the Coloured community, and since the government was unwilling to revise this policy, their role of advising government as to the welfare of Coloureds was effectively pointless (van der Ross 1986: 257). In 1951, the Separate Representation of Voters Bill was proposed to remove Coloured voters from the common voters' roll. When enacted in 1956, 48,000 Coloured people were denied the right to vote (du Pré 1994: 139-140; van der Ross 1986: 250) and for the first time since 1834, Coloureds in South Africa were completely disenfranchised.

A Coloured Representatives Council (CRC) was established in 1964 to oversee and administrate the affairs of Coloureds, such as education and social welfare *inter alia* (Venter 1974: 7-8; van der Ross 1986: 305). As with each turn of events, some Coloured people were optimistic that the new Council would pave the way towards equality and democratic rights for Coloureds, but were again disappointed as the Council was largely impotent in effecting any real change – the persisting barrier was that Coloured people opposed the policy of apartheid and the government was not prepared to review this policy. A *Commission of Inquiry into the Matters relating to the Coloured Population Group* was appointed by government in 1973, and three years later, produced a report with 178 recommendations to government, which essentially requested that government reform its discriminatory practices against Coloureds. Most recommendations were rejected, with then Prime Minister Vorster stating that ‘any recommendations ... that direct representation [in parliament] be granted to [C]oloureds ... is not acceptable to the government’ (cited in du Pré 1994: 170).

Coloured people were angry, and this decision of the government came at a time when the black³ South African masses were beginning to unite in protest against the racism and discrimination that plagued every aspect of their lives. It marked the start of a period in South Africa’s history that is characterised by violence, boycotts and general unrest. On 16 June 1976, thousands of Black high school students took to the streets of the largest township in Johannesburg, Soweto, protesting against the use of Afrikaans as a medium of instruction (MOI) in Black schools. While the Soweto Uprisings did not directly concern the Coloured people in the Cape, it did mark the beginning of a period of solidarity between all oppressed groups of South Africans (du Pré 1994: 171), and a popularisation of Coloured rejectionism (Adhikari 2009: 2). Unity of blacks had since the 1940s been the mantra of the NEUM, which argued that the ruling government was using a ‘divide and rule’ tactic which split the black majority to maintain its dominance (Kies 1945: 15; Adhikari 2009: 3). The NEUM had always encouraged Coloureds to relinquish their Coloured identity and reconceive of themselves and other blacks as a commonly oppressed people (Kies 1945: 15), but

³ I use ‘black’ (with lowercase ‘b’) to denote a collective group of Coloured, Indian and Black African people (for the latter group I reserve the label ‘Black’).

it was only when Black Consciousness⁴ (BC) ideology gained popular support amongst Coloureds towards the end of the 1970s that there large-scale Coloured rejectionism (Adhikari 2009: 3).

In efforts to alleviate the ongoing unrest amongst black South Africans, it was proposed that a tricameral parliament be introduced under the leadership of Prime Minister P.W. Botha. The newly constituted parliament would consist of three houses, viz. the House of Assembly (HOA; for Whites), the House of Representatives (HOR; for Coloureds) and the House of Delegates (HOD; for Indians). Coloureds and Indians were to vote for representatives to sit in their own chambers of parliament. Each chamber was responsible for its 'own affairs': issues pertaining to health, education and community affairs for their respective groups. While this seemingly represented progress for Coloured people, as there was finally Coloured representation in parliament, the tricameral system was regarded by Coloured people to be a new guise for the same apartheid structures to persist: full political rights were not granted to Coloureds and Indians; the HOR and HOD only had jurisdiction over matters pertaining to their own communities; and very significantly, the new system completely excluded Blacks, leaving their affairs to be run by a separate Department of Training and Education (DET). Coloureds, who had united with Blacks as a common oppressed victim of apartheid, revolted against the proposed system by organising a boycott on the day of the Coloured election for the HOR (du Pré 1994: 181). Despite a mere 30 percent voter turnout for the HOR, and 24 percent for the HOD (van der Ross 1986: 356-357), the tricameral system was launched in 1984.

As mentioned previously, it was towards the end of the 1980s that the movement towards democratisation of South Africa began, and the decades that many South Africans spent fighting for freedom began to pay off. Before discussing the transition to democracy, however, an overview of the apartheid regime and its primary effects upon the Coloured population is provided in the following sections.

⁴ Under the leadership of Steve Biko, a union of students formed called the South African Students Organisation (SASO), which was at the forefront of the Black Consciousness Movement. The BC movement stressed that all blacks, including Coloureds and Indians, should unite as a common victim of racial oppression and overcome feelings of racial inferiority (Lewis 1987: 278).

1.6 Apartheid South Africa

At no point in the more than 350 years since Europeans first settled on South African soil has the country been free of discrimination based essentially on colour. During the development of the Cape Colony, life amongst the various groups of inhabitants became progressively segregated along racial lines. This segregation reached its peak in the mid 20th century when the system known as apartheid was formally introduced to legislate racial segregation. In 1948 the National Party (NP) won the national election and assumed leadership of the country under Prime Minister D.F. Malan. At this point the segregation which permeated South African society was promulgated through a series of laws that served to ensure that the different ‘nations’ that resided within the borders of South Africa, developed separately and ostensibly, equally. The discussion that follows details the most divisive of these laws, and focuses particularly on their impact upon the Coloured population.

The first of the laws to affect Coloureds was the Prohibition of Mixed Marriages Act of 1949, which disallowed marriage between Whites and blacks. The Immorality Amendment Act of 1950⁵ criminalised intercourse between White and black people. Mixed race couples were, as result of these acts, subjected to humiliation and abuse by the police who were tasked with ensuring that the law was not contravened (du Pré 1994: 66). A law enacted a few years later, the Reservation of Separate Amenities Act of 1953, prevented Coloureds from using the same public amenities as Whites. These amenities included buses, parks, public toilets, beaches and entrances to buildings and were reserved either for Whites and others for ‘non-Whites’, a common apartheid-speak term for black South Africans. The best of each of these amenities was reserved for Whites; black people’s resources were often inferior and substandard.

Another law which had a significant impact upon the lives of Coloured people was the Population Registration Act of 1950, which assigned each member of the population of South Africa into one of three racial groups: White, Black, Coloured. In a subsequent amendment to the Act, a fourth group, Indian or Asian, was added. Until then, this group had been subsumed within the category ‘Coloured’, which also included the sub-groups Malay, Griqua, other Coloured, Chinese and other Asiatic. The act of classifying South Africans into population groups was a precursor to the solidification of entrenched racism against people of colour. Subsequent to the

⁵ This law was originally enacted in 1923, and a new law with the same name promulgated in 1957.

passing of this Act, there could no longer be any ambiguities as to the race of any South African, as one's classification was recorded in one's national identity document.

The Minister of the Interior at the time declared the test of race to be subject to the 'judgement of society' and that 'the classification of a person should be made according to the views held by the members of that community' (HOA debates 17/3/1967 cited in Posel 2001: 55). Racial categorisation was thought to be a matter of 'common sense', and citizens were assigned to racial categories based on two factors, viz. appearance and general acceptance within a community. Though vague and unscientific, the following definition of race was supplied in the Population Registration Act:

A white person is one who in appearance is, or who is generally accepted as, a white person, but does not include a person who, although in appearance obviously a white person, is generally accepted as a Coloured person.

A native⁶ person is a person who is in fact or is generally accepted as a member of any aboriginal race or tribe of Africa.

A Coloured person is a person who is not a white person nor a native.

(Population Registration Act no 30 of 1950)

The fact that the Coloured 'race' was negatively defined i.e. not White and not Black, was quite significant for the community. Marike de Klerk, wife of Former State President F.W. de Klerk, infamously described Coloured people as 'a negative group', the 'leftovers', and 'people that were left after the nations were sorted out' (Sunday Tribune 5 February 1983 cited in Erasmus 2001: 18).

The first round of mass classification of the population into racial categories occurred in 1951 when a national population census was conducted. The classification was ostensibly the responsibility of the Director of Census, and photographs were attached to each completed census form in order for the classification to be made. In reality, however, it was the census enumerators whose judgements on the race of those with whom they conducted the survey, determined which racial category people were assigned to (Posel 2001: 58). These classifications were thereafter recorded in each

⁶ 'Native' was the term used by the apartheid government to refer to Black Africans.

citizen's national identity document, and essentially determined what opportunities would be available to citizens and their children for the next four decades and beyond.

Once classified, people were able to appeal their classification through the Race Classification Appeal Board, and at such appeals, hardly any aspects of one's social life and physical appearance was excluded as a potential signifier of race (Posel 2001: 59). The process of having one's race reviewed was often humiliating, as the tests administered by officials allowed them to measure and inspect any part of the applicant's anatomy. One of the more infamous tests is the pencil test, where a pencil was pushed into the hair of an applicant, who was then instructed to shake his head. If the pencil fell out, the person would be declared 'Coloured', but if it remained in place, the person would be declared 'Native' (du Pré 1994: 70). Many fair-skinned Coloureds applied to be reclassified 'White' and if they were successful, were forced to leave their neighbourhood and cut ties with schools, churches and their families (du Pré 1994: 71). The lingering emotional damage this caused to the fabric of such families, and to the community more generally, cannot be underestimated (see February 1991).

A law enacted also in 1950, the Group Areas Act, forced members of the four population groups to reside in separate areas in urban centres of South Africa, according to the race to which they had been assigned. Many black South Africans were forced to move away from residential areas in which their families had lived for generations, to areas designated for people of their race group. It became a criminal offence for a member of one racial group to reside in an area designated for another group. Those who owned property in areas zoned for members of other population groups were forced to rent their homes to members of the 'qualifying' group. Upon death of the registered owner, the family was forced to sell the property to someone who could legally reside in the area. As a result of the Group Areas Act, most Coloureds were moved to, and to a large degree, still reside in, the Cape Flats – a large expanse of sandy wasteland away from Cape Town city centre.

Significantly for the current study, another facet of life that was segregated during apartheid, was education. The education of Coloured people during apartheid will be discussed in detail in section 1.7 below.

Apartheid was also effective in dividing citizens in terms of economics. More than simple separate development, black South Africans were greatly disadvantaged by highly inequitable distribution of the government's resources. White South Africans, who make up less than 10 percent of the total population of the country (Statistics South Africa 2003: 26), enjoyed the lion's share of resources, while the black majority suffered as a result of policies which actively induced poverty (Leibbrandt, Woolard & Woolard 2007: 1). The system of job reservation ensured that Whites were favoured over Coloureds when competing for employment opportunities. In cases where Coloured people did the same work as Whites, their salaries were significantly lower e.g. White railway workers earned an annual average of R3,821 in 1970, compared with Coloureds who earned R854 per annum (Venter 1974: 142). In some industries, such as the clothing, footwear, food, catering and motor assembly industries, jobs were reserved exclusively for Coloureds (DOI 1974: 83). They were not professional jobs, however, and in this way Coloureds were largely confined to the domains of semi-skilled labour.

Generally speaking, high skill and high wage jobs were reserved for Whites, and low skill, low wage jobs for blacks. Table 1.1 shows the relative per capita personal income by race group for the period 1917 to 1995. The table shows that Coloured, Indian and Black economically active citizens earned an average of between 24 and 30 percent of what their White counterparts earned. The data in the table show that this trend pre-dates apartheid by more than 30 years. While the figures only reflect the trend until shortly after the advent of democracy in 1994, economic inequality persists in post-apartheid South Africa and the disparities remain closely correlated with race (Leibbrandt, Woolard & Woolard 2007).

Year	White	Coloured	Indian⁷	Black	Average
1917	100	22	24.1	9.1	30.2
1924	100	20	19.4	7.9	29.9
1936	100	15.6	24.1	7.6	27.9
1946	100	16.3	23	8.9	28.8
1956	100	16.9	21.9	8.6	28
1959	100	15.7	17.1	7.7	26.7
1960	100	15.9	17.1	8.1	26.8
1970	100	17.3	20.2	6.8	24.3
1975	100	19.4	25.4	8.6	25.5
1980	100	19.1	25.5	8.5	24.4
1987	100	20.9	30.2	8.5	24.3
1993	100	19.3	42	10.9	24
1995	100	20	48.4	14.5	26

Table 1.1: Relative per capita income as a percentage of White level, 1917-1995.
Source: Leibbrandt, Woolard & Woolard (2007: 4).

While Coloureds earned significantly less than Whites, the population group which had the lowest per capita income was the Black group. South Africa's remaining resources, once Whites had been catered for, were not even equitably distributed between all 'non-Whites': there existed a racial hierarchy, in which Whites held a superior position to Coloureds, and Coloureds in turn, held a superior position to Blacks⁸. This was believed to be acceptable because of Coloureds' ancestral links with Europeans (details on the history of the Coloured community are provided in section 1.4). As Prime Minister Hertzog stated in 1925 'It must not be lost sight of that, in the case of the Cape Coloured, we have to do with a class of our population which, in many respects, are close to the European ... he knows no other civilisation than that of the Europeans ... he has an outlook on life which in fact is that of the European and not of the native; and speaks the language of the European as his mother-tongue' (cited in du Pré 1994: 52). Jan S. Marais also discusses the position of Coloureds in relation to Whites, stating that '[a] Coloured community as distinct from the European does not exist in any realistic interpretation of the term. White and Coloured are, and have been from the beginning, inextricably mixed up together'

⁷ I have replaced the label 'Asian' as used in the source of this table, with 'Indian' for the sake of consistency within this dissertation. Indians are numerically dominant within the group which is often labeled Asian, but includes other South Africans of Asian descent.

⁸ The same is true of Indians. I omit mention of Indians due to its irrelevance in the present study of the Coloured population.

(Marais 1939: 283). As a result of this, Coloureds did not suffer the worst of the discrimination under the apartheid dispensation, even though the relative privilege they were afforded was incommensurate with that enjoyed by Whites (du Pré 1994: 68-69). The hierarchical structure of discrimination held not only in the practicalities of everyday life in South Africa, it was also entrenched in the minds of Coloureds. Sociologist Zimitri Erasmus states that ‘growing up [C]oloured meant knowing that I was *not only* not white, but *less than white*; *not only* not black, but *better than black*’ (Erasmus 2001: 13; original italics). This mental entrenchment of apartheid doctrine was one of the victories of the apartheid proponents.

Towards the end of the 1980s, South Africa was plagued with violent uprisings and protests by black South Africans who were dissatisfied with their continued subjugation in the land of their birth. In addition, there was immense pressure from the international community for South Africa to reform. The United Nations had imposed cultural, academic, sporting and economic sanctions on South Africa. The United States of America, which had also imposed sanctions, agreed to lift them on condition that all apartheid laws be repealed by the end of June, 1991 (du Pré 1994: 206). It was in this context that the incumbent leader of the NP and state president, F.W. de Klerk, began the process of negotiations with the African National Congress (ANC) and other resistance movements to end apartheid.

Political groups that had been banned because of their resistance to the apartheid state, were unbanned, political prisoners were released and the way was paved for political exiles to return to South Africa. The formal transition from apartheid to democracy thus began in 1990 and continued until the first democratic elections were held in 1994, where every South African aged 18 years and older was permitted to cast a vote. Nelson Mandela, leader of the previously banned ANC, was subsequently sworn in as president of the Republic of South Africa.

The democratic government inherited a very unequal South Africa from its predecessors, and correcting the imbalances of the past is an unenviably difficult task – one that is likely to take many generations to complete. Twenty years after democratisation, the relics of oppression and segregation are strongly present in everyday South African life. One area that has struggled to reform, and which I turn to next, is education.

1.7 Coloured Educational Structures: 1652-1990⁹

Until reasonably late in the history of South Africa, education was not compulsory for Coloureds. Under British colonial rule, there were attempts to ensure that all children in the Cape Colony attended school, but many of them (predominantly children of slaves and Khoesan) remained uneducated or had very low levels of education. This trend continued well into the 20th century. A snapshot of the late 1950s illustrates the extent of educational impoverishment amongst Coloureds: of nearly 48,000 pupils who had entered Grade One¹⁰ in 1946, only 735 completed Grade 12 by 1957, a mere 1.5 percent of the initial number (van der Ross cited in Venter 1974: 310) and in 1959, only 750 Coloured people held university degrees, out of a total population of over 1.5 million (Venter 1974: 338). The history of this dismal educational performance can be traced back to the early days of the Cape Colony, and I begin this account in the mid-1600s.

The primary function of the schools established by the Dutch settlers was to teach (non-European) slaves the Dutch language, and to promote Christianity through instruction of the Gospel (Kies 1939: 5). During the Dutch colonial period, one's skin colour was not as significant as one's religious affiliation: all non-Christians were considered to be heathens, and educational institutions were established in order to inform slaves (and to some degree, Khoesan indigenes) about the language and moral code subscribed to by the Dutch colonisers (Kies 1939: 7).

In 1685, colour-based segregation entered the educational system for the first time. Children of Europeans were to attend the ordinary public school, and a second school was established in the colony for slaves (Kies 1939: 7). Even though segregated education had been decreed, there were instances of slave children attending public schools alongside White children (Kies 1939: 9). This remained the state of educational affairs in the colony until the end of Dutch rule in the Cape in 1795.

⁹ The education of Black Africans was administered by the Department of Education and Training (DET). The Bantu Education Act of 1953 was passed, effectively ensuring that the education of Blacks was impoverished and sub-standard. The apartheid government felt that the education of Blacks need not prepare them for a life and job that they would never achieve, as Blacks were to be employed only as unskilled labourers. There is a significant body of literature that deals with Bantu Education and the protests against it, such as Christie (1991) and Kallaway (2002).

¹⁰ At this time, year levels at South African schools were not referred to as Grades, but rather as Sub A (Grade One), Sub B (Grade Two) and Standard One to Ten (Grades Three to 12). For the sake of consistency, however, I will refer to Grades throughout the thesis.

Under the English administration, education of slaves was not prioritised. New slaves were able to learn the dominant languages, Dutch and English, from older slaves, and the English were not as concerned with winning converts to Christianity. The South African Missionary Society took over this concern, however, and requested permission to establish schools for slaves. A number of these schools had been established after 1807 (Kies 1939: 12). The Khoesan were also accommodated in these mission schools (Kies 1939: 17).

So in the 1800s, mission schools played an important role in the education of Coloured children, having been established in response to a need for schooling for slaves and indigenous people (Soudien et al., forthcoming: 71). It was widely acknowledged that these schools provided an inferior quality of education to the schools intended for White children (Marais 1939: 270; Kies 1939: 30).

These schools were independently financed until 1841, when the state began to provide financial assistance in the form of a grant. This grant proved quite advantageous for Coloureds, because the number of schools increased as a result of the funding, and with it, the number of children attending schools. In 1843, the number was 3,322, which increased more than ten-fold to 38,389 by 1883 (Venter 1974: 315-6). At this stage, even though racial segregation was becoming more established, there was still a significant degree of racial mixing in schools in the Cape Colony. Kies (1939: 30) states that large numbers of White children attended mission schools until the late 1890s, primarily because the fees were inexpensive, and payment was essentially optional. Also, public schools were attended by all who agreed to 'conform to the general rules, pay the fees, and [were] decently clad and well-behaved' (Kies 1939: 30), so Coloured children who fulfilled these criteria, with especial focus on the ability to pay the fees, were not precluded from attending public schools. Racial mixing in Cape schools thus characterised the education system until the end of the 19th century.

Racially segregated schooling became a reality before the time of Union in 1910. White children were removed from the mission schools, and the justification for keeping Coloured children in these inferior schools was that the standard of education suitable for White children was in its range and character beyond their [the Coloured children's] needs' (Special Report of the Superintendent-General of Education 1890

cited in Soudien et al., forthcoming: 79). Within this system, Coloured children rarely progressed beyond Grade Six (Kies 1939: 46).

In 1905, education in the Cape was formally segregated under the School Board Act, effectively legislating the status quo. Very significantly, the Act made provision for compulsory public schooling for Whites only (Kies 1939: 49; Adhikari 2005: 3). The lack of compulsory education for Coloureds created a situation in which Coloured children generally joined school later and left earlier than White children, with the result that levels of education in the Coloured community were very low (Kies 1939: 60).

When the Union of South Africa was established in 1910, mission schools in the Cape remained under the administration of the Cape Province, rather than being administered nationally. In 1953, it was recommended by a Coloured Education Commission that compulsory free education be made available for Coloured children until the end of Grade Seven. In order to facilitate this, an additional 10,000 pupils and 300 teachers would need to be accommodated each year, for ten years. It was suggested that the Province start by taking over mission schools. The MOI up until the end of Grade Eight was to be the child's home language (Venter 1974: 318).

It was only in 1974 that schooling became compulsory for Coloured children. Prior to this, a very small percentage of children attained a Junior Certificate (Grade 10), and even fewer a Senior Certificate (Grade 12). Compulsory education was confounded by the reality that the family income could be augmented if children found work rather than attended school, so parents often encouraged children to work from quite a young age (Venter 1974: 329). After making education compulsory, increased numbers of pupils attending schools created a shortage of qualified teachers (Venter 1974: 331).

University admission for Coloureds was very limited at White institutions, which admitted Coloureds in small numbers only if their course of choice was not available at an institution for Coloureds. The University of the Western Cape, which was opened in 1962, was the only university for Coloureds in South Africa.

Despite the fact that educational provision was being made for Coloured South Africans, it was far from equitable compared with the standards of White education.

The per capita spend on education for the different racial groups (see table 1.2 below) exemplifies this inequality. The table shows that in 1970, the value spent on the education of each Coloured child was less than one quarter of that which was being spent on the education of a White child. By 1980, it was 20 percent that of the White total. There was a noticeable improvement by 1990, but the per capita spend on Coloured education was still at just 64 percent of the per capita spend on White education. Teachers' salaries also reflected the inequity of the education system: White teachers in 1972 earned an average of R5,100 per annum, while their Coloured counterparts earned R3,360 (Venter 1974: 323).

	Black	Coloured	Indian	White
1970	25.31	94.41	124.40	461.00
1980	91.29	234.00	389.66	1169.00
1990	930.00	1984.00	2227.01	3084.00

Table 1.2: Per Capita School Spend, by Race, 1970-1990, in ZAR.
Source: MacKenzie (1993: 289)

Pupil-teacher ratios for the different population groups also reflect the inequity of the system: in 1989, the average figures stood at 17:1 for Whites, 23:1 for Coloureds and 50:1 for Blacks (Smit and Hennessy 1995: 10).

In 1980, by way of protest of 'Gutter Education'¹¹ for black South Africans, Coloured students boycotted and effectively assumed control of Coloured schools. The boycotts were employed as a means to strongly express the demand for democratic, non-racial and equal education for all South Africans (Molteno 1987: 3). The control of the school shifted away from the teaching staff to the students, who replaced the existing institutional structures with democratic structures, usually led by a Students' Representative Council (SRC) (Molteno 1987: 8-9).

Intensifying unrest characterised South Africa for much of the 1980s, with regard to education and all other sectors. Towards the end of the decade, the long awaited movement towards constitutional freedom for all South Africans in a democratic South Africa was in its embryonic phase, and the prospect of desegregated society

¹¹ "Gutter education" is a term used to refer to inferior education for Coloured, Indian and Black South Africans, a segregated system designed by the apartheid government to prevent black South Africans from attaining necessary educational achievement to compete with Whites on equal footing.

surfaced as a real possibility. I turn next to a discussion of the process of educational desegregation in South African schools.

1.8 Desegregation of Education

As with other basic services provided by the post-apartheid government, the provision of education ranges from excellent to extremely sub-standard. In well-staffed, highly resourced schools, the failure rate at secondary school level is very low, and many pupils from these schools gain admission to higher education (Naidoo 2005: 21). By contrast, there are millions of adult South Africans who are illiterate, and millions of children whose learning conditions are severely impoverished in respect of the learner-teacher ratios, physical conditions of schools and teacher qualifications (Naidoo 2005: 21).

Prior to the 1994 elections, fifteen different education ministries were in place: four in independent homelands; six in non-independent homelands; one catering for Blacks outside homelands; one each for Whites, Coloureds and Indians in their respective Houses within the tricameral parliament (see section 1.5 for a details about this system); and one responsible for national co-ordination and countrywide norms and standards (DOE 1995a). Despite the fragmented state of the system, three main categories of schools were identified in 1994: state schools, state-aided schools, and independent or private schools. State schools were those that were ex-Black, Coloured and Indian and entirely government funded. State-aided schools were those formerly White schools that were partially state-funded; and independent schools received no state funding (DOE 1995b).

In 1990, at a time when South Africa was just starting the transition into democracy, Education and Culture minister in the House of Assembly (HOA), Piet Clase, declared that racial exclusivity in schools would no longer suffice as a condition for admission to state schools (Soudien 1992: 281). He introduced the Clase models to HOA schools. There were three models – A, B and C – from which White parent communities could choose in order to control admissions policy, as well as future finance and governance policies for the schools with which they were involved. Schools had to hold elections, and obtain at least 72 percent consensus within the parent body in order to adopt one of the models.

Model A was the privatisation option, in which schools would close and reopen as private schools run by a School Governing Body (SGB) which would dictate the admissions policy. The second option, Model B, was to remain fully state-funded with open admission. The model which the majority of schools opted for was Model C, essentially a semi-private/semi-state option (Naidoo 2005: 23; Tikly and Magoboane 1997: 161-2). Under this model, the schools were financially aided by the state, which paid teachers' salaries and between 75 and 85 percent of the running costs of the school. The SGB was responsible for raising the balance of the funds through school fees (as well as donations or other sources), and had considerable power over school policy including admissions procedures (Tikly and Magoboane 1997: 162).

It is argued that the advent of model C schools (as they came to be known) effectively shifted the boundaries of access to well resourced (HOA) schools away from a racially based system: class was the new criterion which would determine the quality of education a child received (Fiske & Ladd 2003: 15). Model C schools began charging fees in order to raise the balance of the school's running costs. After 1994, the post-apartheid government, despite declarations during their election campaign that free basic education would be accessible for all South Africans, decided to encourage public schools to continue charging fees in order to supplement the public funds available for education (Fiske & Ladd 2003: 4).

Table 1.3 provides the average school fees charged by public primary and secondary schools in the Western Cape in 2001, aggregated by the department which had formerly administered them. The House of Assembly (HOA) was responsible for White education, the House of Representatives (HOR) for Coloured education, and the House of Delegates (HOD) for Indian education. As the table shows, fees charged at HOA schools (R2,701 per secondary school pupil) are significantly higher than those charged at HOR schools (R333 per secondary school pupil). These additional financial resources enabled HOA schools to employ more teachers in SGB posts. HOA schools had lower pupil-teacher ratios, teachers had higher average qualifications and a lower proportion of under-qualified teachers (Fiske & Ladd 2003: 15). Though differential access to educational resources strongly suggests corresponding differences in the quality of education, this assumption is corroborated by the educational outcomes achieved by the different schools: higher Matric (Grade 12) pass rates were consistent with better resourced schools, and vice versa (Fiske & Ladd 2003: 17).

	HOA	HOR	HOD	DET¹²
Primary school	2077	99	327	45
Secondary school	2701	333	283	105

Table 1.3: Annual fees charged in schools in Western Cape in ZAR, 2001, classified by former department.

Source: Fiske & Ladd (2003: 27,28)

High school fees, coupled with the issue of cost and logistics of travelling distance and transport to schools, which were situated in formerly Whites-only residential areas, put these schools out of the financial reach of many black children whose parents could not afford the extra expense. Thus the decision of the government to charge fees in HOA schools had the effect of preventing a high degree of desegregation in these schools, retaining middle-class families within the public school system and avoiding large-scale ‘white flight’ to private schools (Fiske & Ladd 2003: 17).

So the schools which benefited from greatest financial support during apartheid, HOA schools, still offer the highest quality of education and produce the best academic results. The reality is then that HOA schools remain most desirable because good quality education paves the way for access to more opportunities throughout one’s life. Access to these schools is subject to parents’ ability to pay higher fees.

This then begs the question: to what extent have South African schools become desegregated in the past 20 years? And are the opportunities to which only White children were once privy being accessed by black children today? In short, the answer is that post-apartheid restructuring and development of the education system has, above all else, and despite expressed intentions of redress for the poor, benefited the expanding, racially mixed middle-class (Chisholm 2004: 7).

In a quantitative statistical survey, Chisholm and Sujee (2006) report the figures of desegregation in South African schools based on data from the national Education Management Information Systems database. A significant number of respondents (41 percent in HOA schools; 8 percent in HOR schools) in the Western Cape schools chose not to select one of the four racial categories and instead ticked the non-

¹² The DET (Department of Education and Training) was the body responsible for the education of Black South Africans.

racialised box labelled 'other'¹³, thus the full picture of deracialisation is slightly obscured in the data, but it nevertheless provides insight into the desegregation process as it stood in 2001. In former HOA schools in the Western Cape, 38 percent of the student body was White and 17 percent Coloured. The joint Black and Indian population came to four percent, and the balance (41 percent) was unclassified (Chisholm and Sujee 2006: 150). So from the available evidence, it seems that the majority of the pupils in HOA schools remained White, but that there were a number of Coloured pupils who had been admitted since the change in admission policy. These figures show the trend over the whole of the Western Cape Province and due to the fact that they are aggregated, the reality is masked that there still exist some former HOA schools that have remained largely White; and conversely, there are some which now have a predominantly black student population (Soudien 2004: 104). Factors which determine the racial make-up of the student body include proximity to former Coloured residential areas and the rate at which school fees were set (which was at the discretion of the SGB).

In HOR schools in 2001, the majority of pupils (86 percent) were still Coloured. Six percent were Black and eight percent classified themselves 'other'. No pupils placed themselves in the categories White or Indian (Chisholm and Sujee 2006: 151). The data suggest that the process of desegregation is slower in HOR schools in the Western Cape than in HOA schools.

The tendency of post-apartheid movement within the school system is for children of middle-class black parents to enter the former White system, in essence, a realignment of socioeconomically similar groupings (Soudien 2004: 106). It is in this context of post-apartheid egalitarian mixing between Whites and Coloureds in HOA schools in the Western Cape that the present research is set. The inevitable social change which results from far-reaching political changes marks one of the first steps in the long process of building a deracialised South Africa, which is free of the bondage of racism and segregation. The aim of this thesis is to track the changes which occur in the monophthongal vowel system of Coloured SAE as a result of desegregated social networks within former HOA schools and to compare this phonetic system to that of

¹³ Issues of racial identity are fraught with the trouble of non-acceptance of the definition of what races are. The Constitution of the RSA allows people the freedom of not classifying themselves into one of four boxes – a freedom most South Africans did not have during the apartheid era. Resistance to these categorisations was central to the apartheid resistance struggle.

Coloured people whose social networks have remained predominantly Coloured i.e. working-class Coloureds.

Due in part to differences in origin and then enforced segregation (which persists today despite political reform and new fluidities), the distinct SAE phonetic systems still exist robustly, especially in working-class and other largely segregated communities. It is the purpose of this study to look into the phonetic system of the Coloured community of Cape Town, to ascertain how post-apartheid fluidities have impacted upon the vowel system in this community. As discussed above, the community is very diverse, and the label 'Coloured' is laden with connotations which were largely undesirable to those so designated. I turn now to the problematisation of nomenclature in relation to the Coloured community.

1.9 'Coloured': A Problematic Label

Sensitivities still abound concerning the use of apartheid racial classifications and categorisations to label groups of South Africans. While the categories were once the basis for privilege and discrimination, the self-same categories are now, rather ironically, used for issues of redress and to measure advancement away from the disadvantages caused by apartheid policies (Posel 2001: 51). Despite general difficulties of labelling people according to redundant classifications, the label 'Coloured' contains intricacies of its own, which I attempt to unpack in the following paragraphs.

The first documented use of the term 'Coloured' occurs in legal documents dated 1834, and it was used to refer collectively to Khoesan and slave people after the emancipation of slaves (du Pré 1994: 9). It was essentially a catch-all category which included anyone who was not European and not a Nguni speaker, i.e. a Black African. As mentioned above, Coloureds were defined in the Population Registration Act as a 'leftover' category: "A Coloured person is a person who is not a white person nor a native" (Population Registration Act, no. 30 of 1950). Coloured people have been variously labelled throughout South Africa's history, including the appellations *Gekleurdes* ('Coloureds'), *Kleurling* ('Coloured'), *bruinmense* ('brown people'), *the Coloured People* and *people of colour inter alia*. The boundaries of 'Colouredness' were always vague, though, and not until 1950 with the passing of the Population Registration Act, was this *nation* of people 'created' and officially labelled 'Coloured' (du Pré 1994: 9). As a result of the negative definition, the label was always offensive

to those thus labelled and became associated with inferiority (Adhikari 2005: 13; du Pré 1994: 10). The inferiority associated with Colouredness is captured in the title of a novel by Sarah Gertrude Millin (1924), where she refers to Coloured people as *God's stepchildren*. To express their dissatisfaction with the label 'Coloured', it became common for people to make use of the prefix 'so-called' and the term 'so-called Coloured' is still used by many Coloured people today.

Historians and other scholars have not had an easy task grappling with which terminology to use to refer to this group of people who were classified 'Coloured' under apartheid. The two most common terms are firstly, 'Coloured' with uppercase 'C', which denotes an officially defined population group. The second, 'coloured', with lowercase 'c', is a general denotation of people of mixed race. The former label essentialises a racial identity which, it could be argued, does not actually exist, and the latter is problematic because not only those who were formally classified 'Coloured' are 'mixed race'. In fact, it has been shown that any family whose ancestors have been in South Africa for more than 200 years are likely to have some degree of racial infusion in their ancestry (Jeffreys 1959, cited in du Pré 1994: 36). This is partly because during colonial times when miscegenation was common, offspring from marriages and sometimes from illicit unions between White settlers and slave or local women, were often absorbed into White society (Kies 1939: 9; van der Ross 1986: 2; du Pré 1994: 36).

Despite the problems attached to racial terminology, a study of the people classified 'Coloured' during apartheid requires the use of a term to describe this group. In order to spare the reader from confusion with awkward references throughout the thesis, I refer to these people as Coloured, without scare quotes and with an uppercase 'C'. For consistency, other population groups are referred to as White, Indian and Black (see footnote 3 for an explanation of how 'black' – with lowercase 'b' – is used in this thesis).

1.10 Coloureds and Language Use

The predominant languages spoken by Coloured South Africans today are local varieties of Afrikaans and English. In the 2011 population census survey, 75.8 percent of Coloureds reported that Afrikaans was their first language, and 20.8 percent

English¹⁴ (Statistics South Africa 2012: 27). This is rather unsurprising given that the dominant colonial powers in the Cape, where most Coloured people reside, were the Dutch and the British, although a brief history of the language use from colonial days until present shows that a multitude of languages have contributed to the linguistic character of greater Cape Town.

The indigenous Khoe and San people spoke loosely related Khoe and San languages respectively (Traill 2002: 45). The Dutch settlers arrived at the Cape in the mid 1600s, and because they controlled the colony, Dutch was the primary language of the colony at the time (McCormick 2002: 21). Over time its form changed to include features which now characterise Afrikaans¹⁵. The developing language carried the label ‘Cape Dutch’ before it became known as ‘Afrikaans’ (Combrink 1978).

The slaves brought to the colony from various parts of Africa and Asia spoke the languages of their place of origin. The use of most of these languages could not be maintained, however, as the need for a lingua franca amongst the slaves, the settlers and locals increased (McCormick 2002: 15).

With successive occupations of the Cape in 1795 and 1806, the British took control of the colony from the Dutch. An influx of British immigrants in 1820 preceded an intense process of Anglicisation in the Cape under the leadership of Governor Charles Somerset in 1824. The aim of Anglicisation was to weaken the power of the Dutch-speaking people, and its effect was felt throughout the colony. English became the only official language of the colony and dominated in all domains, including government, commerce, education, religion and public life (McCormick 2002: 17). The language of the home amongst most Coloured families remained a local dialect of Afrikaans, but the MOI in most schools was English.

English was also the language of business and employment in the city of Cape Town which had developed and where many people sought employment. Those interested in doing business in the city or interacting with residents in the suburbs south of the city required some command of English. Most residents of the inner city areas learnt it either from neighbours, who may themselves have been L2 (second language)

¹⁴ These figures are for all Coloured people in South Africa, and are not exclusive to the Western Cape population.

¹⁵ For a history of the development of Afrikaans see Den Besten 1989, Ponelis 1993, Combrink 1978, Roberge 1995, Deumert 1999.

speakers, or from contact with immigrants, many of whom spoke non-standard dialects of English (Finn 2004: 966). In the 1950s under the apartheid government, schools were forced to introduce Afrikaans as a MOI as a result of the mother-tongue education policy. Even though Afrikaans was the home language for most Coloured children, parents believed that education in English schools would provide better opportunities for further study and employment, thus did what they could to ensure that their children remained in English medium classes, wherever possible (Finn 2004: 967).

As a result of the change in education policy regarding MOI, there are clear intergenerational differences in English proficiency within the Coloured community. Older speakers who were educated in English before the mother-tongue education policy was effected in the 1950s, are comfortable speaking English. Their children would not be as comfortable having been educated in Afrikaans and having little opportunity outside of school to use English. This generation, in turn, often chose to raise their children as L1 English speakers so that they could be admitted to English schools (Finn 2004: 967). As a result, in many instances Coloured children's L1 was acquired with an L2 variety as the primary input (Malan 1996: 126).

The conditions of L2 acquisition of English account partly for the non-standard variety of English which is now spoken by working-class Coloured people in Cape Town, and the persistence of some of those non-standard features in L1 CSAE today (Anthonissen 2013: 33). Afrikaans remains an important lingua franca amongst members of the Coloured community, however.

The varieties of English spoken by different members of the Coloured population today is dependent on factors such as class, level of schooling and type of school attended (McCormick 2004: 993). I hypothesise that middle-class speakers who attended HOA or private schools will use a variety which is more similar to that spoken by White SAE speakers, while working-class speakers, for whom English is often an L2, use a vowel system that is more traditional within the Coloured community. As outlined previously, the purpose of this study is to provide illumination into precisely this matter, with particular reference to the monophthongal vowel system of the varieties.

1.11 A Note on Methodology

The methodology employed for data analysis is a key point in this thesis. While use of acoustic methodology is not new to SAE, with scholars such as Mesthrie (e.g. 2010) and Bekker (e.g. 2009) using acoustic techniques, this is the first time that modern methods of automatic vowel measurement (AVM; detailed in chapter four) have been employed in an acoustic study of SAE.

Discussing first the choice to use acoustic rather than more traditional aural methods of data analysis, it must be noted that technological advancement allows easy access to acoustic methodological tools, which were once limited to big phonetic laboratories. The advancement creates an imperative for researchers to move into a space where acoustic work is not only possible, but for particular fields of research, for example the one in which the present study is located, it has become something of a norm. As a phonetician, the value of using one's ear to verify results of acoustic processes cannot be underestimated, but the use of acoustic methodology to provide accountability and replicability to the research is invaluable. Foulkes and Docherty (1999: 23) emphasise the importance of being able replicate the results of empirical research, and Boberg (2005: 136) advises that auditory impressionistic analysis has inherent limitations, including potential problems of intertoken and intercoder reliability and objectivity. Acoustic techniques of analysis in essence remove the problem of annotator subjectivity (even more so using AVM).

The programs used in the AVM process (detailed in chapter four), viz. P2FA Forced Aligner and extractFormants, were developed for use on North American English (NAE), and had to be adapted by the present researcher for use on CSAE. Use of AVM allows for a larger scope of analysis to be performed, as its goal is to save time spent manually aligning audio and text files, as well as manually extracting formant measurements.

As an under-researched variety of SAE, Coloured SAE will benefit from application of modern acoustic techniques to provide a detailed description of the monophthongs of the variety.

1.12 Delimitations

In this thesis, only the speech of adolescent Coloured speakers is analysed. The youthful sample was chosen because of the time-frame of the socio-political change, which is a central factor in this study.

A further delimitation is on the phonetic variables that are analysed. In order to provide a detailed thorough analysis, I have concentrated the analysis on short and long monophthongs only. The reason for omitting diphthongs is also practical, and concerns the methodology used to extract formant measurements¹⁶ i.e. only one measurement is extracted for the vocalic nucleus. I have also excluded analysis of consonants, as well as morphological and syntactic features. While investigation into all of these features would indubitably provide a remarkable overview of CSAE, I believe it is the charge of a doctoral thesis to be narrow in its scope, and comprehensive in its address of the chosen issue. The reader can thus expect the analysis in this study to be unwaveringly thorough in the content that it purports to cover i.e. provide an acoustic account of the current state of CSAE monophthongs.

1.13 Chapter Outline

This thesis is made up of seven chapters, each one unique in its scope. Having introduced the topic of the research and provided a background to the study in this chapter, the following chapters are structured as follows: Chapter Two is a review of the literature relevant to the study, in terms of the history of SAE and Coloured SAE in particular, as well as the vowel quality recorded by scholars of both SAE generally, and CSAE. Following this, in Chapter Three, is a detailed account of how the data were collected, and the composition of the sample in terms of their socio-economic status, schooling, area of residence *inter alia*. Chapter Four provides details of the methodology employed to analyse the data collected during the interview process. This includes discussion of the automated process of alignment and formant extraction, statistical techniques employed and all software required to perform these various functions.

In the two chapters that follow, Chapters Five and Six, the data are presented. Chapter Five covers the short vowels, and Chapter Six covers the long monophthongs. The

¹⁶ Different methods of automatic formant extraction are available which would allow analysis of the nucleus and glide of diphthongal vowels, e.g. extracting numerous measurements throughout the vowel's duration (see e.g. Risdal and Kohn 2013), but such methodology was not employed in the present study.

presentation of each lexical set is very detailed, including a recap of the literature on WSAE, the literature on CSAE and then a presentation of the acoustic results from the present study. The final chapter, Chapter Seven, provides a conclusion detailing the relevance of the study and its pertinent findings. Appendices and references are located after the concluding chapter.

Chapter Two

Coloured South African English:

History and Vowel Quality

The variety of English spoken by Coloured people in the Western Cape of South Africa has not been very well documented since its emergence and during its development into the distinct dialect it is today. Lanham (1996: 22) suggests that Coloureds, who had ‘a tradition of multilingualism, had probably developed some precursor to what is easily recognised today as [C]oloured English¹⁷’ during the colonial period, but little or no evidence is available to suggest what form this earlier variety might have taken. Accounts of the historical development of SAE generally focus on those who originally brought the language to South African shores, i.e. British settlers and their descendants i.e. White South Africans. In this chapter, all pertinent literature is reviewed in an attempt to map the origins and development of CSAE.

In the first section, 2.1, a brief history of South African English is presented, followed in section 2.2 by an outline of the characteristics that make SAE a ‘southern’ variety. SAE literature is reviewed in section 2.3, and in section 2.4, a detailed review of the literature on CSAE is presented. The final section of this chapter is a vowel-by-vowel phonetic profile of CSAE, telling the phonetic ‘story’ of CSAE as reported by earlier scholars who have worked on the variety. This provides a basis of comparison for the results of the acoustic analysis presented in chapters five and six.

In this chapter, as in the remainder of the thesis, vowel categories are identified in terms of the ‘standard lexical sets’ posited by Wells (1982: xviii). In this framework, vowel categories are labelled by an item which exemplifies the class e.g. STRUT is the label for the lexical set of words whose vowel quality is the same as the word *strut*, such as *cut*, *son*, *up*. The class is symbolised using SMALL CAPS.

¹⁷ What Lanham calls ‘Coloured English’ is referred to throughout this thesis as ‘Coloured SAE’ or ‘CSAE’. Other scholars have also used the term ‘Cape Flats English’ (e.g. Malan 1996). For the sake of consistency, I have replaced all other terms with CSAE, although the subtleties regarding the use of particular labels for different varieties of a language are noted.

2.1 History of English in South Africa

English was introduced to the Cape colony in the early 1800s. In the 200 years since then, the language has taken on different forms and functions, and its status has varied, as has its social distribution amongst residents of South Africa (Lanham 1996: 19). Lanham (1996) divides the history of SAE into four distinct eras, with social and political change acting as the catalyst for the changing nature of the use of English in the region. The first era was defined by colonialism; the second, by an influx of British and other immigrants towards the end of the nineteenth century. The third era is what Lanham refers to as the Post-war era (with reference to World War II), but since the war had little relevance to the history of English in South Africa, this period might more accurately be labelled the Apartheid era, as it was the internal political landscape which affected the use and distribution of English in the country. The final era is the era of a ‘New’ South Africa (1990 onwards), in which the present study is located and seeks to document. Each of the eras will be described briefly below, drawing primarily on historical accounts by Mesthrie (1993), Lanham (1996) and Lass (2002).

2.1.1 The Colonial Era

With successive occupations in 1795, 1803 and 1806, the British eventually seized power of the Cape Colony from the Dutch on their third attempt. It was then that English was introduced to the southern tip of Africa, in what is modern day South Africa. There were relatively few English speakers in the colony at this time, so Dutch remained a prominent language despite efforts to anglicise the colony. The first local variety of English in South Africa was thus an L2 variety – Dutch English (DE) – as Dutch speakers were required to learn the language of the new imperial power (Lanham 1996: 20). The people who came to be known as ‘Coloured’ of course predated the arrival of the British at the Cape and were thought to be ‘part of the Dutch speech community’ (Lanham 1996: 22). Cape Dutch, later evolving into and called Afrikaans, was the primary language spoken by Coloured people at the time.

A group of permanent settlers arrived from Britain in 1820, numbering approximately 5,000. This was the first sizable settlement since the British assumed control of the colony, and shortly after their arrival, in 1822, English was proclaimed the only official language of the colony. It was amongst the children of the 1820 settlers that the first mother tongue variety of SAE emerged. The majority of these settlers spoke a southern variety of British English (BrE), hence the ‘southern’ features that still

characterise SAE today (Lass 2002: 105; Mesthrie 2012a: 2097). (Details of these features are provided in section 2.2 below.)

This early variety of SAE was influenced by Dutch (and DE), because of the close contact between speakers of the different languages. Certain ‘Afrikaans’ features entered SAE at this time, and are still featured in present-day SAE (Lanham and Macdonald 1979: 73) e.g. the prevalence of schwa in positions where other varieties favour [ɪ]. In the mid-nineteenth century, a large group of settlers arrived in Natal (now KwaZulu Natal) – another British colony on the east coast of the region. Two distinct varieties were thus present in South Africa at the close of the colonial period: Cape English (CE) and Natal English (NE). A small number of Coloureds were educated in English in church schools during this time, so English was introduced to the community through education, and as Lanham (1996: 22) suggests, Coloured people may have developed a precursor to modern CSAE by the end of the colonial period.

2.1.2 Immigrant Influx

The second era in the history of SAE begins with the discovery of diamonds and gold in Kimberley and Johannesburg respectively. Hundreds of thousands of immigrants from Britain and continental Europe flocked to South Africa from the 1870s onwards with the hopes of cashing in on the economic opportunities to build wealth in the mining industrial society that was developing (Lanham 1996: 22). The result of the mineral discovery on SAE was to catalyse the development of class dialects within the variety (Mesthrie 1993: 28): differing levels of success led to the social stratification of the mining society. Standard Southern BrE was upheld as the most prestigious variety spoken in the region at the time, but many local varieties began to emerge due to the movement of people within modern South African borders. A local standard, which differed in minor ways to the British Standard (Lanham 1996: 23), was beginning to emerge as the distinct varieties (CE and NE, influenced by BrE and DE) merged into a single system of SAE. Natalian colonials were ‘more obviously English’ (Lanham 1996: 23), so NE variables were regarded to be more prestigious than CE variables (Mesthrie 1993: 28). The average Cape settler descendant was not highly educated and lacked artisan skills, so found themselves, and by extension their variety of English, filling the lower ranks of the mining society (Lanham 1996: 23).

In 1910, when the two Boer Republics, the Cape Colony and Natal became unionised, Dutch and English were declared to be the official languages of the Union (Lanham 1996: 25). Towards the middle of the 20th century, Afrikaner nationalists became actively hostile in their defence of Afrikaner purity and against English dominance in the Union. It was these Afrikaners who assumed governmental power as the ruling party in 1948. This event drew to a close the second era, giving rise to the Apartheid Era. With reference to the development of CSAE, Lanham (1996: 25) notes that no evidence exists to suggest that the variety changed greatly during this period.

2.1.3 The Apartheid Era

Despite its co-official status, Afrikaans played a subordinate role in the Union while English dominated most areas of public life in the mid-1900s. Unhappy with this, the Afrikaner rulers actively attempted to replace English with Afrikaans in public domains (Lanham 1996: 26). Their attempts were largely successful, and as a direct result, English proficiency declined significantly, particularly in the Afrikaner community. The effect upon English speaking White South Africans was that they tended to become bilingual. The Bantu Education Act of 1953 was enacted in the interest of promoting Afrikaans over English. It was a means of ensuring that Black South Africans were denied the opportunity to be educated in English or their mother tongue: the policy stipulated that Blacks were to be educated in Afrikaans, while all other population groups were to be educated in their mother tongue. As a result of this policy, many schools for Coloured children were required to change to Afrikaans as a MOI because the children's parents were L1 Afrikaans speakers. Parents were resentful of this policy because they had chosen to raise the children as L1 English speakers, hence their choice of English as MOI, believing that this would increase the opportunities for future success (Finn 2004: 967). Popular dissent over the Act formed the basis for the Soweto Uprisings in 1976 – a violent protest that began a spate of subsequent protests against the apartheid state throughout the 1980s.

At this time, a local standard SAE (General SAE on Lanham's (1978) lectal hierarchy, depicted in figure 2.1, section 2.3 below) was regarded as the 'educated standard' and was only marginally different to BrE. Speakers of Broad SAE, which was largely indistinguishable from AfrE, held lower social status, and were distinguished from Conservative speakers primarily by occupation. During this period, Lanham (1996: 30) points out that socio-economic advancement in the Coloured community led the language shift from Afrikaans to English. This supports

reports that language use amongst Coloured speakers correlates with class i.e. middle-class speakers tended to use English while the working-class retained Afrikaans as L1 (Finn 2004: 968).

2.1.4 The ‘New’ South Africa

As discussed in chapter one, significant political and social change characterised South Africa in the early 1990s. The post-apartheid South African government made a commitment to promoting all languages in South Africa, and went as far as declaring 11 of them official languages: nine African languages, English and Afrikaans. Despite a guise of linguistic equality for all of these languages, there is a *de facto* preponderance of English in South African media, politics and education, where it dominates as MOI in secondary and tertiary educational institutions (Lanham 1996: 30; Bowerman 2004: 934).

Much effort has been made during this new era in the history of South Africa to eliminate the relics of racial segregation and discrimination amongst South Africans. Segregatory laws were repealed, one consequence of which was to open all schools to children regardless of race. Schools thus became a significant site for racial integration for a new generation of South Africans. Most significantly for the present study, speakers of different varieties of English (and Afrikaans) came into contact with one another as peers. Lanham (1996:32), looking forward, suggested that post-apartheid linguistic change in the Coloured community might have included a ‘resurgence of Afrikaans as a matter of identity and solidarity’. This has in fact not played out as Lanham suggested. Rather, English is increasingly being used as a language of middle-class interaction in South Africa (Mesthrie 2008), and my own experience in the Coloured community confirms that this is true for young Coloured people too, as young parents consistently report raising their children to be L1 English speakers, even when they themselves are L1 Afrikaans speakers.

2.2 SAE as a ‘Southern’ Variety

All varieties of SAE are ‘southern’ in that the British settlers who introduced English to the colony spoke a southern variety, and specific features of their dialects persist in descendant varieties today. The most important southern features are (Lass 2002: 105; Mesthrie 2012a: 2097):

1. [æ] or a higher vowel in TRAP.
2. STRUT/FOOT split: Most words before the split started in the vicinity [u], but shifted to the STRUT set, with a large range of realisations from lower mid back [ʌ] to centralised variants e.g. central [ɶ] (Lass 2002: 106). This split is the reason that there are very few words in SAE that form part of the FOOT set, hence very few tokens of FOOT in the interview data.
3. Lengthening I (TRAP/BATH split): /æ/ lengthened before /f, θ, s/ and also /nt, ns/ to that TRAP has a short vowel and BATH a long one (which differs in quality) (Lass 2002: 105).
4. Lengthening II: /æ/ was lengthened before voiced stops and nasals (except /ŋ/) so that TRAP is typically realised with [æ], and [æ:] in words like *bad, bag, man* (Lass 2002: 105).
5. Rhoticity: SAE is non-rhotic.

It is due to these features that a collapsing of some of Wells' lexical sets is in order for analysis of SAE. Lass (1990: 274) provides an overview of the contrastive classes in SAE:

Short monophthongs: KIT, DRESS, TRAP, LOT, STRUT, FOOT

Long monophthongs: FLEECE, NURSE, GOOSE, THOUGHT, BATH, SQUARE

Diphthongs: FACE, PRICE, CHOICE, NEAR, CURE, GOAT, MOUTH

So in terms of the vowels investigated in this study, the classes that make up Wells' (1982: xviii) original list are collapsed as follows: BATH becomes a 'superset' (Mesthrie et al. 2013) redistributed over BATH proper (e.g. in *bath, fast, class*), [ɑ:] in the START set (this is not a pre-/r/ environment because SAE is non-rhotic), [ɑ:] in the PALM set, as well as in open syllables like *ah, ma* and *pa*. BATH in this study thus refers to the superset. Similarly, NORTH and FORCE are collapsed into the THOUGHT set, and CLOTH into the LOT set, as there is no general distinction in vowel quality for these sets in South African varieties of English.

2.3 Review of SAE literature

SAE is described by Lass (1990: 272) as 'an enormously complex and grossly under-described dialect cluster, comprising both mother-tongue and L2 varieties.' The country's divided socio-political history resulted in each of the population groups

developing distinct ethnolects. Because English diffused from the White community into other population groups, having been brought into the country by the ancestors of WSAE speakers, it is mere pragmatism that results in the use of WSAE as a reference point for other varieties of SAE (Lass 2002: 104). It is these ethnolects, viz. Coloured, Black and Indian SAE, that are ‘grossly under-described’, more so than WSAE, which has a number of dedicated accounts dating back to 1928. Although varieties spoken by black South Africans have not received much scholarly attention historically, this has begun to change in recent years as research is undertaken to provide descriptions of black SAEs, and to document the linguistic changes occurring as a result of the recent socio-political change in the country. Scholars such as Wissing (2002), van Rooy (2004), da Silva (2007) and Morreira (2012) have produced reports on Black SAE; Mesthrie (e.g. 2004; 2006) and Chevalier (2011) on Indian SAE; and Dennis (2008) and Brown (2012) on CSAE. Mesthrie has undertaken a large-scale project documenting changes in SAE country-wide, and has produced significant work on the various ethnolects in recent years (see e.g. Mesthrie 2008; 2010).

Because early works providing a linguistic description of SAE pronunciation are historical records of WSAE, only a brief summary of the literature will be presented here, as the work does not bear directly on the present study. More detailed summaries of the studies of CSAE are presented in the following section, 2.4.

A continuum of SAE lects was first proposed by Lanham (1967; 1978) to mirror the work of Mitchell and Delbridge (1965) on Australian English. As is common to the Southern Hemisphere Englishes, three major lects comprise the continuum: Conservative, General and Broad SAE¹⁸. A revised version of Lanham’s (1978) original descriptions of the lects is provided in Lass (2002: 111), who updates the lectal hierarchy with more modern descriptions of the speakers said to use them. The continuum is depicted in figure 2.1 below:

¹⁸ Mitchell and Delbridge (1965) used the term ‘Cultivated’ rather than ‘Conservative’, and Lanham (1967) originally used the terms ‘Conservative’, ‘Respectable’ and ‘Extreme’, with the latter two lects providing a parallel for ‘General’ and ‘Broad’. Throughout this thesis, the trichotomy of lects is referred to using the terminology ‘Conservative’, ‘General’ and ‘Broad’, as these terms seem least laden with nuanced connotations.

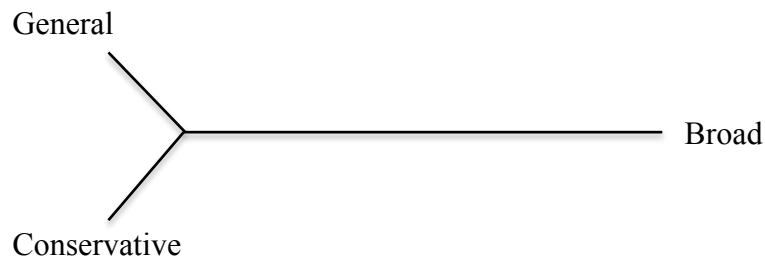


Figure 2.1: Lectal Hierarchy of SAE

The upper end of the continuum is split into two standard lects: Conservative SAE, and General SAE. The conservative variety is least distinguishable from Southern BrE, and represents a prestigious, ‘transplanted’ norm based on Received Pronunciation. Traditional speakers of Conservative SAE include upper middle-class White South Africans. The General lect is a local, high status standard variety, which is used by other middle-class White English speaking South Africans. On the lower end of the continuum, Broad SAE is associated with low socio-economic status, low levels of education and non-professional occupations. The most Broad variety is almost indistinguishable from L2 AfrE. This trichotomy of lects describes WSAE, although there is cross-over with speakers from other ethnic varieties of SAE, particularly Coloured and Indian SAE (Lass 2002: 111).

The earliest account of SAE is Hopwood (1928), who describes SAE pronunciation, but fails at times to distinguish between L1 and L2 varieties. He conflates features occurring in AfrE with general SAE features. His account of SAE provides values for the variety that are consistent with Broad SAE. A later study by Lanham and Traill (1962) describes two lects within what they call South African Received Pronunciation (SARP): SARP A and SARP B correlate with Lanham’s (1978) later distinction between Conservative and General SAE (discussed above). Lanham and MacDonald (1979) provide an analysis of formal SAE in an attempt to correlate variation in SAE with socio-historical factors in the country’s history.

Lass and Wright (1985; 1986) discuss the uniquely South African split of the KIT lexical set in SAE, and Lass (1990) provides a description of the standard vowel system of SAE spoken by Upper Middle Class White Capetonians. Lass (1995) and Lass (2002) are overviews of White SAE in terms of its vowel, consonant and

morphosyntactic systems. Wells (1982) is a general overview of SAE, as is Mesthrie (1993) and Branford (1994). Bowerman's (2004) article on WSAE in *the Handbook of Varieties of English* provides a systematic overview of the vowel qualities in the variety.

Bekker and Eley (2007) performed an acoustic analysis on a small sample of 10 White female speakers, half from Johannesburg and the other half from East London. Looking at monophthongs of these speakers (Word List style only), they suggest that lowered and retracted TRAP might be emerging as a new prestigious value in the Northern suburbs of Johannesburg.

Ian Bekker (2009) provides a detailed description of Word List style, White SAE vowel system, using acoustic methods of data analysis. His sample was comprised of 27 White female speakers, aged 18-19, from various urban centres in South Africa. Despite the limited style reported on in this thesis, the description of the variety is situated in explicit detail within historical accounts of the variety, as well as reports of emerging trends in SAE. This work, along with Lass (2002), provides a point of comparison to ascertain whether any departures amongst CSAE speakers from traditional Coloured vowel qualities are moving in the direction of the standard qualities for WSAE speakers.

2.4 Literature on CSAE

Details regarding the development of the distinct variety of SAE that is labelled Coloured SAE are rather scant in the SAE literature. Of course, its features are 'southern', having shared a historical origin with the other varieties of SAE (see section 2.2 for a description of southern features), but with regard to more specific features of the variety, there are only a handful of sources on whom we must rely, none of them dating back more than 35 years. The literature reviewed in this section reveals the factors that are unique to the development of CSAE, and provides details of the characteristics that make it distinctive.

English in the Coloured community must be contextualised in terms of its intimate relationship with Afrikaans (and earlier, Cape Dutch), and also its contact with other languages that were spoken in the Cape during the colonial period. As described in Chapter One, the population of the colony was incredibly diverse, both ethnically and linguistically: the native inhabitants of the Cape, slaves from many parts of the world

and the European colonists came together and interacted in socially complex ways. A Cape Dutch vernacular evolved during the 17th and 18th centuries, and was used as a lingua franca amongst the Dutch, slave and indigenous populations (Malan 1996: 127). After the British took control of the colony in 1806 and abolished slavery in 1834, Coloured people migrated en masse to the areas surrounding Cape Town city centre, many of them to an area known as District Six. This suburb was inhabited not only by Coloured people, but also by many immigrants from Eastern and Western Europe and migrants from other parts of Southern Africa. Trade and commercial enterprise in this vibrant and diverse community became dependant on the use of L2 varieties of English as lingua franca, because of the multilingual nature of the environment (Malan 1996: 130).

Code-switching and code-mixing became a notable feature of Cape Coloured speech. Because of the multiplicity that characterised their ancestry, and the lack of acceptance into any of the contributing ancestral communities, Coloured people had no vested interest in maintaining the purity of either Afrikaans or English (McCormick 1989a: 206): they themselves were not regarded to be ‘pure’ ancestors of either Dutch or English speakers. McCormick (1989a: 207) describes the ‘Coloured’ vernacular to be comprised of three distinct codes: non-standard Afrikaans, non-standard English and code-switching between the two. The non-standard variety of Afrikaans contains numerous English loanwords and has a distinctive phonological system compared with standard Afrikaans, although the syntactic structure is similar. The non-standard English variety, i.e. CSAE, displays more morpho-syntactic differences compared with the standard, and is influenced strongly by Afrikaans (Malan 1996: 132). Coloured parents have, throughout the decades, been consistent in their belief that English-medium education would offer their children better economic and social opportunities, so encouraged children to speak English by speaking English to them. As a result, children of Afrikaans speaking parents acquired English as L1, with an L2 variety as their primary input (Malan 1996:135). Historically, then, CSAE developed with a constant and very close relationship with Afrikaans – so much so that code-switching and code-mixing are pervasive, unmarked features of language use in the community.

Student Research Papers (1979-1984)

Four senior undergraduate research papers by final year speech therapy students (all cited in Wood 1987) at the University of Cape Town provide what seems to be among

the earliest reports on specifically CSAE features. The papers are rather restricted in scope due to the nature and level of the projects, but provide a useful starting point nonetheless. The first of these studies, conducted by Hastings (1979 cited in Wood 1987: 109-111), looked at the vowel system of a small sample of six Coloured children, concluding that the influence of Afrikaans is prevalent in the Coloured community, and is responsible for CSAE developing into a distinct variety of SAE. Wood (1987: 111) carefully dissects her findings, suggesting that the small sample size as well as methodological issues might have compromised the validity of some of her findings, although some of the results are consistent with his own.

Saffery (1986 cited in Wood 1987: 111-112) undertook a study similar to Hastings (1979), but focused her analysis on the consonantal system of CSAE. She reports similar findings to those reported by Hastings, save for a few key features such as the presence of linking /r/ in intervocalic positions.

Steenkamp (1980 cited in Wood 1987: 112-115) investigated the pronunciation of /r/ as a function of socio-economic status, using 16 speakers from middle and low socio-economic suburbs in Cape Town. Steenkamp found that class, area of residence and the preponderance of Afrikaans as L1 in the lower socio-economic suburb, were the major determinants of the variation.

Douglas (1984 cited in Wood 1987: 116-117) looked into the intonation patterns of three Coloured children, comparing their intonation patterns to that of RP speakers. She found a definite contrast in the intonation patterns of the CSAE child speakers compared with RP speakers.

Wood (1987)

The most important reference work for the present study is Tahir Wood's (1987) Masters thesis entitled 'Perceptions of, and Attitudes towards, Varieties of English in the Cape Peninsula, with particular reference to the 'Coloured community''. Wood presents a description of English spoken in the Coloured community with the objective of identifying the particular characteristics of a specifically Coloured variety of English at the time of his study.

Wood (1987: 100) proposes a continuum of CSAE lects, which parallels the continuum for White SAE proposed by Lanham (1978) (described in 2.3 above). At

the most non-standard pole of the continuum lies Broad CSAE, and the most standard speakers of CSAE fall at the opposite pole, labelled General CSAE. On the scale, the interference of Afrikaans is directly proportional to the standardness of the lect, so that General CSAE is least influenced by Afrikaans, and the Broad lect most influenced. Wood (1987: 99-100) further suggests a close correlation between linguistic variables with social variables, most notably social class, level of education, amount of English spoken in the home, area of residence and political affiliation. Broad speakers of CSAE tend to be at the lower end of the socio-economic scale, have lower levels of education, live in areas characterised by sub-economic housing and crowded conditions and where vernacular Afrikaans predominates residential communication i.e. L2 English speakers. General speakers, conversely, occupy a higher socio-economic status, have higher levels of education, reside in more affluent areas and tend to be L1 English speakers. Due to the social segregation between Coloureds and Whites during apartheid, the varieties remained distinct, but there is overlap between the Broad and General varieties on each continuum.

Wood collected data in two stages, aiming to elicit a relatively formal style from each speaker, keeping the variable of style constant. The first set of data was collected at a Coloured secondary school, and was comprised of approximately six hours of speech of Grade 11 and 12 pupils, male and female, both L1 and L2 speakers. Each pupil was recorded presenting an oral on a topic of his or her choice. The second body of data was used primarily for the language attitude component of Wood's analysis, and consisted of both White and Coloured people, although he states that 'these recordings served to provide support for the data drawn from the school recordings' (1987: 108). He does not expressly state that he excludes from the analysis the speech of the White speakers he interviewed, although I imagine that he must have done so. The second group of speakers was made up of members of Wood's own social networks viz. friends and acquaintances, as well as some hitchhikers to whom he offered a lift in exchange for their co-operation in his study. Some of these speakers were asked to read a short passage, others simply asked to describe a simple task such as making a cup of tea or changing the tyre on a car. Although Wood claims to have achieved consistency of style, he also states that he asked the visitors in his vehicle to 'just talk naturally' (1987: 109), so it is unlikely that speakers performing these rather dissimilar elicitation tasks would truly have used the same, formal style. Despite this, his report on the various features of CSAE is very detailed, making this a very useful reference work.

Among the features reported by Wood (1987) include vowel raising, vowel lowering and fronting, pertaining largely to the monophthongal vowels in the variety. These three features will be detailed in section 2.5 under the relevant vowel subsections. The phonological features pertaining to diphthongs and consonants, which do not bear directly on the focus of the present study, are outlined here.

Table 2.1 below summarises Wood's (1987: 123-125) findings concerning diphthongal vowels. Wood (1987: 123) reports that the first variant of PRICE (strong glide) is characteristic of CSAE, even though both realisations occur. Glide-weakened/glideless MOUTH is one of four features that Lanham and Macdonald (1979: 37-40) mention as being definitive of old CE, the others being obstruent /r/, back raised and glide-weakened /aɪ/ and backed, raised /a:/. Wood (1987: 125) finds it unsurprising that three of these features (all besides glide-weakened PRICE) persist in CSAE.

Vowel	Realisation
FACE	Onset lowered to [ɛ]; occasionally glide-weakened and backed
PRICE	Two realisations: <ol style="list-style-type: none"> 1. Raised onset and strong glide 2. Low, glide-weakened [a]
CHOICE	High glide, culminating in the region of [i:] as a result of the Afrikaans influence
MOUTH	Raised, glide-weakened or glideless
GOAT	Onset backed and lowered
NEAR	Four realisations: <ol style="list-style-type: none"> 1. [ɪjə] in <i>serious</i> 2. [ɛ:] in <i>gear</i> 3. [jə:] in <i>near</i> 4. [i:ɜ] in <i>ears</i>
CURE	Two realisations: <ol style="list-style-type: none"> 1. [jɔ:] 2. [uə]

Table 2.1: Diphthong quality in CSAE, according to Wood (1987).

Wood (1987: 126-133) reports the following to be stigmatised features of CSAE, attributable largely to the influence of Afrikaans upon the variety, and used predominantly by Broad speakers:

- De-aspiration of unvoiced stops /p, t, k/ (excluding phrase-final or word-final /p, t, k/).
- Use of /h/ to replace a glide e.g. *piano* [pihænəʊ].
- Use of /j/ to replace word-initial /h/ e.g. *hell* [jæɪ].
- Word-final devoicing, mainly with /z/ and /d/ e.g. *seconds* [sekəndʒ].
- Use of alveolar fricatives instead of palato-alveolar fricatives e.g. *change* [tʃeɪns]; *finish* [finis].
- Use of the dental stop /t/ instead of dental fricative /θ/ e.g. *thirty* [tɜːti].
- Elision, comprised of reduction of consonant clusters e.g. *parents* [peːrəns]; omission of word final nasals e.g. *plan* [plæ̃]; and syllable reduction e.g. *reasonable* [riːzəˈbəl].
- Alternative stress placement rules, most commonly shifting stress in polysyllabic words to the final syllable e.g. *participate* [pɑːtɪsɪˈpeɪt].
- Peculiar pronunciation of certain words e.g. *skip* [skəp], *asthma* [æfɪmə], *non-* (e.g. *non-smokers*) [nʌn], *aren't* [ɑːrənt].
- Realisation of /f/ as a bilabial fricative by speakers with no front teeth¹⁹.

Some features are also reported by Wood (1987: 130-133) to be prestigious, and used predominantly by General speakers:

- Use of more palatal variants instead of palato-alveolar fricatives e.g. *shy* [[ʃ^h]. Wood (1987: 130) states that this variant occurs in many ‘idiolects of Received Pronunciation’, so suggests that it carries prestige in CSAE. It is reported to be more prevalent amongst women, particularly Muslim women.
- Lengthening and stressing of continuants occurring before word-final voiced alveolar consonants can occur e.g. *friends* [frennz]. Occurs primarily in women’s speech; could be a form of hypercorrection contrasting to the stigmatised devoicing mentioned above.
- Use of resonant /r/ (rather than the tapped or trilled variant).

¹⁹ Removal of the top four incisor teeth is very widespread amongst working-class Coloureds, both male and female (Allen et al. 1990: 335). The resultant gap is often referred to as a ‘passion gap’ (often realized with raised TRAP as [peʃəŋgep]). The removal is thought, by dentists, to be as a result of decay and subsequent extraction, although it seems to hold some prestige, and indeed indexes beauty, amongst certain parts of the Coloured community. Two of my male speakers, M8 and M15, had had their front teeth removed.

The following features are mentioned as characteristic of CSAE, but not recorded to be either stigmatised or prestigious (Wood 1987: 126-133):

- Vowels in unstressed syllables are not reduced to schwa, rather retained ‘as spelt’.
- Strongly rounded GOOSE vowel (although less rounded, more centralised variants are said to occur, determined by the linguistic context).
- Variable realisation of obstruent /r/, including as a tap, trill, resonant /r/, as well as the fricative and uvular variants.
- Articulation of voiceless interdental fricative /f/ with the lower lip slightly in front of the top teeth rather than below them.

Wood (1987: 126) argues that the use of unreduced vowels in unstressed syllables is among the most salient characteristics of CSAE. Unfortunately this feature cannot be tested in the present analysis as unstressed vowels were not included in the analysis, the focus of the thesis being monophthongal vowels with primary stress. It does present an interesting opportunity for future research, however.

Wood (1987: 138-140) continues with a short account of non-standard syntactic features of CSAE. A single criticism of Wood’s thorough phonological account of CSAE is that it is rather unsystematic at times: while he seems to provide all possible variants for vowel, consonant and other features of the variety, he fails to discuss their distribution e.g. for the FACE set, he gives three examples *late* [lɛɪt], *take* [tɛk], *came* [kəɪm], each with a different onset, without discussing distribution of the allophones. Nevertheless, his work remains very useful as a reference work for this study, as it is a very detailed account of all features of CSAE.

Malan (1981; 1996)

Malan (1981) studied the non-standard morpho-syntactic properties of speech produced by 20 Coloured children, aged 12 to 13 years (1996) from two working-class suburbs in Cape Town. She found seven constructions to be frequently occurring:

- (a) absence of the auxiliary *are*
- (b) absence of copula *are*
- (c) absence of third person singular present tense marker *-s*
- (d) absence of regular past tense marker *-ed*

- (e) simple present replacing future conditional tense
- (f) simple future replacing future conditional tense

In a later paper, Malan (1996) draws on the work of many of the scholars reviewed here (Malan 1981; Shirk 1985; Wood 1987; McCormick 1989) to describe the linguistic features of CSAE. The discussion covers morpho-syntactic, lexical, phonological and discourse features of the variety. The phonological features she lists are drawn from Wood's (1987) analysis, details of which are provided in section 2.5 below.

Shirk (1985)

Another study of non-standard syntax was conducted by Shirk (1985), who compared the presence or absence of grammatical features in the speech of ten White and ten Coloured children from lower middle-class backgrounds in Cape Town. She found that her Coloured subjects used more non-standard constructions than their White counterparts.

McCormick (1989b; 1995; 2002; 2004)

McCormick (1989b; 1995; 2002) revealed interesting patterns of English Afrikaans code-switching and code-mixing in a remnant community of District Six²⁰. As mentioned previously, she found the use of three distinct codes comprising the linguistic repertoire of the community, viz. non-standard Afrikaans, non-standard English, and third code that is characterised by code-switching between the first two.

The *Handbook of Varieties of English* contains a chapter on CSAE in each of its volumes: Volume I contains the phonological account (Finn 2004), which is discussed in detail below (section 2.5). Volume II contains a review of the morphology and syntax of the variety (McCormick 2004). McCormick (2004) reports that while the standard forms for all features do occur in CSAE, the non-standard constructions seem to occur more frequently, particularly in informal speech. Amongst others, these are a few of the morphosyntactic features that characterise the variety:

²⁰ District Six was a vibrant suburb on the outskirts of Cape Town city centre, home to people of many races, including many Coloured families. The residents of District Six were forcibly removed from their homes, which were razed to the ground, upon implementation of the Group Areas Act. Residents were relocated to far flung areas around the Cape Peninsula, most notably, the Cape Flats. A very small section of this community was preserved, and it is here where McCormick conducted her fieldwork.

- (a) non-standard use of auxiliaries
- (b) contraction and deletion of auxiliaries
- (c) deletion of adverbial suffix *-ly*
- (d) omission of complementiser *that*
- (e) non-standard use of concord
- (f) double negation

Mesthrie (1999; 2007; 2010; 2012b)

Mesthrie (1999) studied the history of unstressed ‘do’, which is pervasive in CSAE e.g. *I did eat the apple* (where ‘*did*’ is unemphatic). He suggests an alternative to the commonly held view that the feature is a result of a transfer effect from Afrikaans *het* (e.g. *Ek het die appel geëet* – lit: ‘I did the apple eat’), arguing instead that it is most likely a relic of a centuries old standard English norm in both BrE and SAE.

Mesthrie has also produced some phonetic work on CSAE. Using acoustic methods, he confirmed impressionistic beliefs that Cape Town Coloured speakers are raising the vowel in BATH (Mesthrie 2007). In a study of the GOOSE lexical set, Mesthrie (2010) found that Coloured speakers used a consistently backer vowel than White speakers, while Black speakers show greatest fronting (among Black, Indian and Coloured subjects), accommodating to the White norm for this vowel. The back GOOSE is typical of older Coloured speakers, so the use of the less fronted variant by young, middle-class Coloured speakers is attributed to the assertion of a positive Coloured identity by these speakers. Mesthrie (2012b) compared the varieties of English spoken by Coloured and Indian communities in five South African cities, focusing on the variable /t/. He found that in Cape Town, both groups show greater fronting of /t/ than in all other cities viz. Port Elizabeth, Johannesburg, Kimberley and Durban, ordered as such on a sliding scale.

Finn (2004)

Finn (2004) wrote a chapter on the phonological aspects of CSAE in the *Handbook of Varieties of English*. The source of his data is not revealed, though it is likely to be based on his PhD at Leeds University. The findings are presented in section 2.5.

Dennis (2008)

My own Masters thesis (Dennis 2008) surveyed 20 middle-class Coloured speakers who were educated in private or model C (former Whites-only) schools for some or

all of their schooling careers i.e. had mixed social networks. The acoustic phonetic analysis was done in Praat by manually extracting formant measurements for each token of the GOOSE, BATH and PRICE lexical sets. Phonetically, these speakers showed connections to both the Coloured and White communities. The GOOSE set ranged from very back /u:/ to a fronted value nearing /y:/. The latter variant is a prominent feature amongst young White, predominantly female, speakers, while the former is historically the typical value used by Coloured speakers. The data were not isolated into different segmental environments, however, so further work is required to elucidate exactly what phonological factors play a role in the fronting of this vowel. The analysis in the present study will interrogate this further.

The speakers' use of the BATH set seemed to place them firmly in the Coloured community, quite consistently using a back, raised and rounded variant which is typical of Coloured speakers. The diphthong PRICE was found not to be glide-weakened, which is a prominent feature of WSAE speakers.

In addition to the phonetic analysis, conclusions were drawn about how these speakers construct the speakers' identities as Coloured youth in deracialising spaces: they tended to assume a very proud Coloured identity, strongly asserting their membership to the Coloured community. This suggests a revitalisation of the formerly negative identity associated with being Coloured in South Africa. Although there was some evidence of WSAE features entering the vowel system of CSAE, the speech of the young speakers tended to support these sentiments phonetically.

Brown (2012)

Brown (2012) conducted a sociophonetic study of three lexical sets: GOOSE, BATH and KIT. His sample was comprised of 12 females and 8 males, all but two of whom were over the age of 30. Most (18 of the 20) were monolingual English speakers, and all of the speakers were middle-class, split into upper and lower middle-class for the purpose of analysis. Brown (2012: 54 – 70) reports an extensive distribution of the GOOSE tokens, ranging from very back variant to centralised, and even fronted variants by many of his speakers. He found that lower middle-class speakers more readily broke with the traditional norm for Coloured speakers of using a back vowel, replacing it with a more fronted variant. For the BATH set, most of Brown's speakers used a backed, rounded and raised variant, although a lower variant does occur in some of his speakers. The findings for the KIT set showed a division between lower

middle-class speakers, and upper middle-class speakers: the former group use a more centralised variant, while the latter group tend to use a more raised and fronted variant. His findings for all three lexical sets are consistent with earlier reports on CSAE. Brown (2012: 82) further reports a pre-occupation amongst his speakers with notions of ‘correctness’, with non-standard features of CSAE regarded as highly stigmatised.

Having reviewed all the literature on CSAE of which I am aware, the following section proceeds with details of the phonetic distribution of the vowel classes that are being investigated in this study.

2.5 Phonetic profile: CSAE

In table 2.2 below, values for the long and short monophthongs of CSAE are presented, as reported by Finn (2004) and Wood (1987). WSAE variants are also included in the table (from Bowerman 2004). Together, these sets of variants provide the basis for comparison with the data in the present study. Where speakers in the sample display a departure from traditional Coloured vowel qualities, it is expected that the shift will be in the direction of the WSAE norms. This is due to the increased exposure of middle-class speakers to White peers in model C or private school environments or as a result of moving into former White residential areas.

Vowel retraction is very common in SAE, and is thought to have a prestige value in CSAE (because retraction before /l/ is not a feature of Broad SAE or AfrE (Lanham 1978)). Wood (1987: 127) claims that vowels of CSAE, particularly the DRESS set, are retracted before /l/. Central vowels are reported to be replaced by one of the back vowels e.g. *uncle* [ʌŋkəl], *children* [tʃɒlɾən], *girls* [go:lz], *people* [pi:pəl]. In order to accurately report on this phenomenon, following /l/ was isolated in the present dataset as one potentially significant segmental environment. The effect of following /l/ in this dataset is reported in chapters five and six.

Vowel	Exemplary word	CSAE Variants	WSAE variants
KIT	<i>kit</i>	ɪ ~ i	ɪ ~ ī, ə, i
	<i>bit</i>	ī	ī ~ ə
DRESS		e > ε ~ æ	e
TRAP		ε > æ > æ ^e	æ > ε
LOT / CLOTH		ɔ ~ ɒ > ɒ(:) ^e	ɒ̄ > ɔ, ʌ
STRUT	<i>strut</i>	a ~ ɐ > ʌ > ɒ	ä > ɐ
	<i>one</i>	a ~ ɐ > ɒ	
FOOT		u > ʊ > ʁ	ʊ > ʁ
BATH / PALM / START	<i>bath</i>	a ~ ɑ > a:(ə) ~	ä:, ɑ:
	<i>dance</i>	ɑ:(ə) > ɒ:(:)ə	
		æ:(ə) > a:ə ~	
		ɑ:(ə)	
NURSE		ɜ: > ø(ə) > ɐ ~	ɜ:, œ:
		ɒ:ə ~ ə > o:	
FLEECE		i: ~ i > i:(:)ə ~ iʊ	i:
GOOSE		u: > ʊ > u > ʊ	u:, ʊ: > y:
THOUGHT / NORTH /		ɔ ~ o: > o:ə > ɒ	ɔ:, œ:
FORCE			
SQUARE		e: > e: ^ə > ε:	ε:, e:

Table 2.2: The short and long monophthongs of CSAE and WSAE.

Key: Tilde (~) indicates co-variance; a greater-than sign (>) indicates that the variant to the left occurs more frequently than the one to the right of the sign.

Sources: Finn (2004: 968-969); Bowerman (2004: 936)

The descriptions below shed some light on the distribution of the CSAE allophones listed in table 2.2.

2.5.1 Short Vowels

KIT

All varieties of SAE, including CSAE, display the KIT-split (Lass 1995: 97; Wood 1987:122-123). The split involves two distinct variants for this set: (a) the *IT* subset as [ɪ] ~ [i] word-initially, after /h/, following velar consonants and before /f/ and (b) the *SIT* subset as centralised [ī] elsewhere. In Broad varieties of SAE, centralised [ī] is realised as low schwa [ə] (Lass 1995: 97); for CSAE speakers, [ə] occurs even amongst more General speakers, not exclusively in Broad lects (Wood 1987: 111). Before /l/, KIT is typically retracted to [ɪ̠]. Wood (1987: 122) notes further that KIT can

be raised to [i], but this is a stigmatised feature, characteristic of the lower end of his respectability scale i.e. Broad CSAE.

DRESS

Wood (1987: 122) reports that in CSAE this vowel is ‘raised’ to [e], implying that a lower variant is the norm. This is strange since [e] is the general quality of SAE for DRESS (Lass 2002: 115; Bowerman 2004: 936), and Wood does use WSAE as a point of reference to describe the distinct features of CSAE. DRESS in this variety is actually more prone to lowering, with realisations as [ɛ], and even [æ] (Finn 2004: 968), although these are reported to occur mainly before /l/. Wood (1987: 122) reports realisation of the word *yes* as [jə(:)s].

TRAP

Wood (1987: 122) reports a marked tendency of raising this set towards [ɛ], although [æ] and even [æ̞] are reported to occur. TRAP retains this quality pre-/l/ in CSAE, but there is some evidence of retraction of TRAP before /l/ in General SAE.

LOT

L2 speakers typically use [ɔ] ~ [ɒ], whereas L1 speakers use [ɒ] more consistently (Wood 1987: 122). [ɒ(:)̞] is also reported to occur. Retraction before /l/ is not reported to affect this set.

STRUT

L2 speakers use one of two variants: [a] ~ [ɐ], with [ʌ] occurring sporadically (Finn 2004: 970; Wood 1987: 122), so this set is typically lowered by CSAE speakers. The low realisation is not a stigmatised feature. Finn (2004: 970) reports that the *ONE*-subset (comprised of *(-)one, once*) contains realisations which vary between [a] for L2 speakers and [ɐ, ɒ] for L1 speakers.

FOOT

Realised in the region of [u] – very back and rounded, though less so for L1 speakers, for whom it may also be realised as [ɯ] and [ɤ] (Finn 2004: 970). This set is not affected by following /l/.

2.5.2 Long Monophthongs

BATH

This set is typically realised as [a] ~ [ɑ], both variants often lengthened and optionally followed by schwa i.e. [a:(ə)] ~ [ɑ:(ə)] (Finn 2004: 970). The following is also reported to occur: [ɒ:(ə)]. There is a subset which is fronted for Broad CSAE speakers: *dance* and *chance* are realised with [æ:] (Wood 1987: 123).

NURSE

Wood (1987: 127) does not record use of [ɜ] except as an equivalent of schwa in *teacher*. Finn (2004: 971) suggests that L1 speakers use [ø:(ə)] with instances of [ɛ:], [ɒ:ə] and [ə]. Before /l/, this set tends to be raised and retracted to [o:].

FLEECE

Typically realised as [i:] when stressed, and [i] in unstressed positions (Finn 2004: 971). This set is reported to be largely unaffected by following /l/ (Wood 1987: 128), and no diphthongisation is reported for the set.

GOOSE

Wood (1987: 128) reports that GOOSE is almost always backed and rounded – realised in the region of [u:]. The centralised variant of other SAE varieties was not reported to occur in Wood's data. Lass (1995: 98-99) confirms that frontier values are avoided by CSAE speakers. Finn (2004: 972), however, found that some L1 speakers do use a more centralised [ʊ:], as occurs in WSAE (as well as 'new' BSAE – see Mesthrie 2010). Again, this set has not been reported to be affected by following /l/.

THOUGHT

According to Wood (1987: 122), Broad speakers typically use [ɔ], while Finn (2004: 971) reports that a higher [o:] is more common. THOUGHT is also not reported to be affected by following /l/.

SQUARE²¹

[e:] occurs most commonly in Finn's (2004: 973) data, with instances of [e:ə] also occurring. Wood (1987: 126) reports [ɛ:] for his speakers.

²¹ While not traditionally a monophthong, SQUARE is included in this analysis because of its long history of monophthongisation in SAE. See section 6.8 for details.

The descriptions above will be used as a basis for comparison for the acoustic results presented for the short vowels in chapter five and the long monophthongs in chapter six. Further to this, where speakers show departure from the traditional values reported here, Bekker's (2009) report on WSAE will be referenced to compare any divergence with WSAE norms.

Before the analysis is presented, however, methodologies of data collection and data analysis are discussed in chapters three and four respectively.

Chapter Three

Data Collection, Speakers and Social Categories

The procedures employed to collect the data for this thesis will be described in this chapter. In order to achieve the aims and objectives outlined in Chapter One, it was necessary to gather suitable data from a representative sample of the population under investigation. The chapter begins with a discussion of the criteria for sample selection, along with explanations for the delimitations on the sample, in section 3.1. This includes an analysis of the sample according to the speakers' scores on a socio-economic index developed for this purpose. In section 3.2, the sample is described in terms of the schools the speakers attended, the areas in which they live and their friends and closest contacts. This information is used to measure the degree to which the Coloured speakers have integrated into White society, using a Degree of Integration index. The chapter concludes with a discussion of the data collection techniques (section 3.3).

3.1 Criteria for Sample Selection

The study is focused on a small part of the Coloured population, viz. adolescents / young adults, so the required sample had to be representative of this subsection of the community. The general approach followed in soliciting appropriate speakers for this study was to ensure that they fit within the parameters delimited below. I selected 40 speakers to fit into pre-defined categories, using personal judgement as an insider to the community to determine potential interviewee's appropriateness for inclusion in the various categories.

Forty speakers made up the sample to represent the young Coloured working-class and middle-class. The speakers who were included in the sample were sourced primarily through a network of personal contacts. Initially, I interviewed speakers who were acquaintances of mine. Some of these speakers put me in contact with friends and in some cases, cousins or other family members, who fulfilled the criteria. Several interviews resulted from these leads. The rest of the speakers were found through personal contacts who put me in touch with appropriate interviewees. In this way I managed to complete the interview quota.

The 40 speakers in the sample were required to be:

- 3.1.1. Coloured
- 3.1.2. Resident of greater Cape Town, Western Cape Province
- 3.1.3. English speaking or English-Afrikaans bilinguals
- 3.1.4. Born between 1983 and 1992

In addition to the criteria above, the sample was subject to a further two restrictions in terms of gender and social class:

- 3.1.5. 20 male, 20 female
- 3.1.6. Working-class or middle-class

Each of the delimitations will be discussed in turn, below.

3.1.1 Coloured

Because the focus of the study is the Coloured community of Cape Town, the respondents that make up the sample were all Coloured. Despite the intricacies of post-apartheid racial categorisation detailed in chapter one, identifying members of this community was reasonably simple. To corroborate my assumption as to the race of the participants, each person was asked to identify their ‘race’ as part of the demographic data collected at the beginning of each interview (see Appendix A). Each speaker thus self-identified as Coloured during the interview process, and there did not seem to be any discomfort with the use of the term.

3.1.2 Cape Town Resident

The geographical delimitation on the study is a result of the concentration of Coloured people within the Western Cape Province of South Africa (reasons for this are outlined in section 1.4). The Western Cape is home to 61.6 percent of the Coloured population of South Africa (Statistics South Africa 2012: 21), and the population of the province is concentrated in the capital city, Cape Town, and its immediate surrounds. This project was thus focused within the greater Cape Town area, as it is currently comprised politically under the management of the City of Cape Town Metropolitan Municipality. There are eight districts that make up greater Cape Town (their positioning is depicted in figure 3.1 below):

- a. Table Bay District
- b. Southern District
- c. Blaauwberg District

- d. Northern District
- e. Tygerberg District
- f. Cape Flats District
- g. Khayelitsha / Mitchells Plain District
- h. Helderberg District

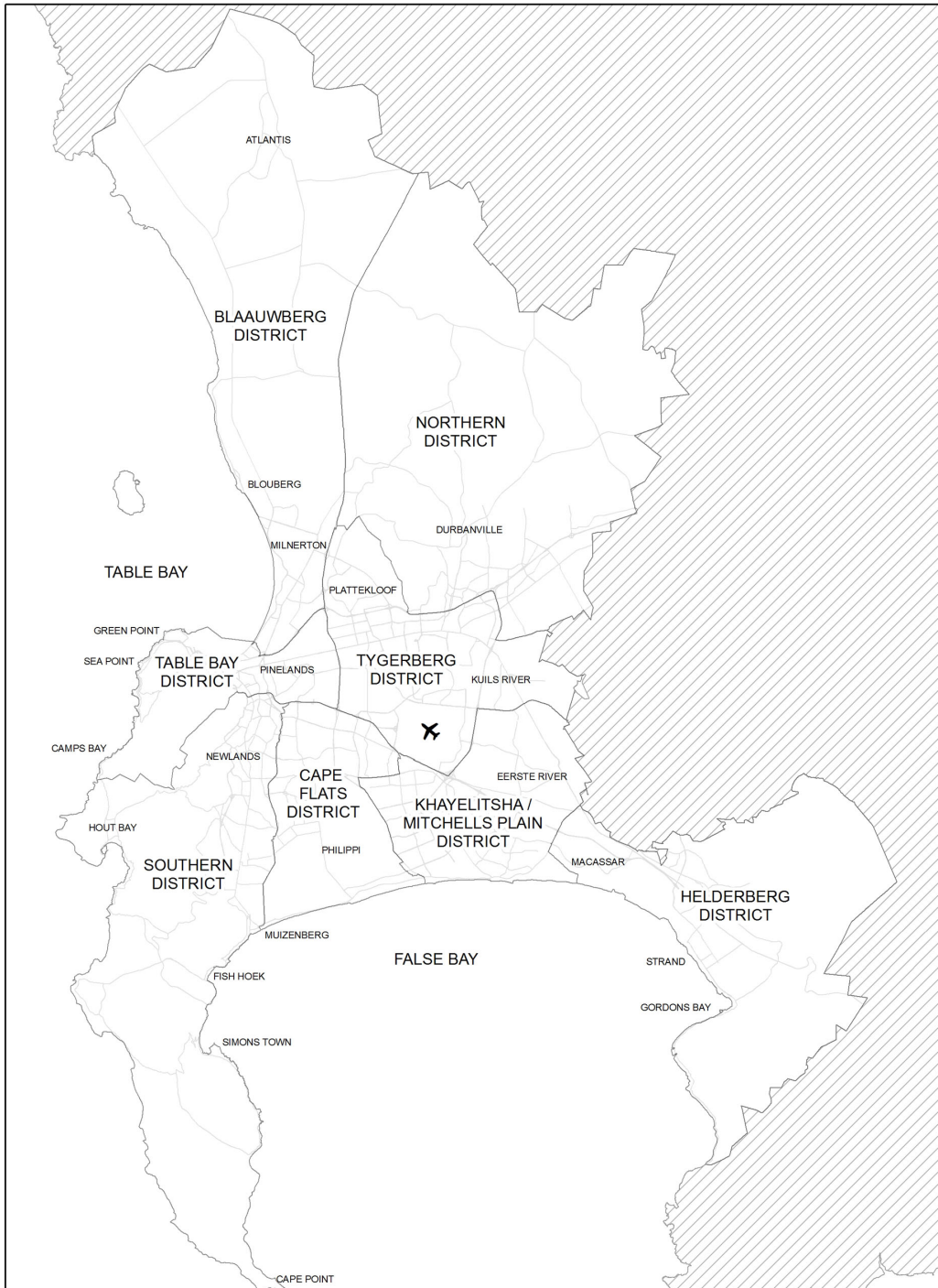


Figure 3.1: Districts of Cape Town Metropole
Source: <http://www.capetown.gov.za/en/sdf/Pages/Maps.aspx>

Each of the above-mentioned districts is represented in the sample. All of the informants lived and were schooled in one (or more) of these areas throughout their schooling careers and until the time of the interview, with two reasonably short and minor exceptions. Two informants left Cape Town with their families during primary school: one to England for one year and the other to Port Elizabeth and East London for two years. In both cases the speakers were very young and upon their return to Cape Town, completed a significant portion of their primary school career in their respective local schools. One informant was excluded from the sample after it was discovered that he had lived in the United Kingdom for five years during his primary school career. Because the focus of the project is on Cape Town residents and the features exhibited in their use of English, I felt that too much sustained exposure to other varieties of English, as in the case of this speaker, might skew the results. This speaker is not included in the total of 40 speakers.

3.1.3 English / English-Afrikaans bilingual

The research for this thesis forms part of a larger research project which focuses on English Social and Regional Dialectology in South Africa²². As a result, the sample was restricted to English users. For most Coloureds in South Africa, Afrikaans is the dominant language: in 2011, 75.8 percent of Coloureds had Afrikaans as L1, compared with 20.8 percent whose L1 was English²³ (Statistics South Africa 2012: 27). As documented by various scholars, English-Afrikaans bilingualism is on the increase, and in some instances English has become the dominant language within the Coloured community (Finn 2004: 968; McCormick 2004: 993; Stone 2002: 382; Anthonissen 2013: 28). This speaks to a noted trend amongst South Africans for English to replace other languages as a mother tongue (Mesthrie 2008), especially in the emerging middle-class black communities where children attend former HOA or private schools.

The changing nature of code choice is not a new phenomenon in South Africa. Stone (2002: 382) notes that since the 1960s, functional differentiation between English and Afrikaans has become increasingly common in Coloured homes: parents converse with one another in the local variety of Afrikaans, with their children in a variety of

²² This project, funded by the National Research Foundation (FA2005031800008), is currently running in the Linguistics Section of the School of African and Gender Studies, Anthropology and Linguistics at the University of Cape Town under the leadership of Professor R. Mesthrie.

²³ These figures are for all Coloured people in South Africa, not only those in Cape Town or the Western Cape Province.

English and prefer English mass media. Stone (2002: 382) further states that use of English amongst working-class Coloureds is increasing. McCormick (2004: 993) confirms that it is common for Coloured parents to speak to one another in Afrikaans, but speak only English to their children. This is due to the ubiquitous belief that being proficient in English would provide children with better opportunities for further study and employment (Finn 2004: 967). A number of speakers in the sample reported that they were raised and schooled in English, even though one or both parents was an L1 Afrikaans speaker. Such parents reportedly spoke Afrikaans to each other, but spoke English when addressing their children. This is consistent with the findings of Malan (1996: 126), who states that many Coloured children acquire English as an L1 from Afrikaans dominant parents and recently, Anthonissen (2013: 33) has found that even ‘new generation’ L1 CSAE speakers i.e. those whose parents were not L1 speakers, use grammatical structures that are typical of L2 speakers.

As a consequence of the L2 input, a non-standard variety of English emerged and is maintained in the Coloured community, and labelled CSAE (Malan 1996; Anthonissen 2013). McCormick (2004) identifies and illustrates the unique morphological and syntactic constructions that characterise this variety of English. Many examples of these non-standard constructions were used by some of the speakers in the sample during their interviews. It is beyond the scope of this study to provide a detailed discussion of morphosyntactic features used by these speakers, so just two examples are provided here by way of illustration.

(1) That’s just how it is in that communities. (F6)

(2) He don’t eat healthy. (F15)

Both (1) and (2) are typical CSAE morphological constructions, according to McCormick’s (2004: 997) findings. Speaker F6 uses the singular form of the demonstrative adjective *that* with a plural noun *communities* in (1). In (2), speaker F15 uses the plural form *don’t* in a third person negative construction, where the Standard English form would be *doesn’t*. Another non-standard feature is the use of the adjective *healthy* instead of the adverb *healthily* in (2).

Broad CSAE carries rather negative connotations as a result of the low socio-economic status (SES) of its typical users, and is regarded by many, including Coloured people of higher SES, to be improper. As such, its use is highly stigmatised. This is demonstrated in the following quotes from middle-class Coloureds from the sample in (3) and (4) below:

- (3) They [stereotypical Coloureds] have a strange accent ... not like mine. Yeah, it sounds very funny and inappropriate. It's just not proper English, it's very ugly. (F5)
- (4) I think [the way I talk] is more kind of proper ... than other Coloured people. (F16)

In the sample, 27 of the speakers declared English to be their L1, and Afrikaans their L2. For the remaining 12 speakers, the reverse was true: their L1 was Afrikaans and their L2, English (see table 3.1). Finn (2004: 968) claims that a linguistic division between Coloureds along class lines is in evidence: L2 CSAE speakers tend to be working-class, while L1 speakers are usually middle-class.

The L2 English speakers were comfortable speaking English during the interviews, and all of them except one managed to read the Word List and Reading Passage (discussed in section 3.3.1 below) without much difficulty. Speaker F17 struggled to read the list of words and the reading passage, guessing a lot of words and in some cases, waiting for prompts from the interviewer. The L2 Afrikaans speakers have differing levels of proficiency in Afrikaans: some reported that they could not speak Afrikaans fluently, but had a comprehensive understanding of it, while others feel completely comfortable with both speaking and comprehension.

3.1.4 The Generation before the 'Born-frees'

One of the expressed aims of this research project was to investigate the nature of post-apartheid phonetic changes in the vowel system of the Coloured community, with the primary site for racial integration being the education system. For this reason it was necessary for the speakers in the sample to have been legally able to attend any school in South Africa, for the majority of their schooling career. As detailed in chapter one, legislated racial segregation in schools made it impossible for Coloured children to attend public schools other than those designated for Coloureds (HOR schools) until 1990. Private schools set admission policies at their own discretion, but it was the practice of most of these schools to preserve exclusively White pupil populations until at least the mid-1980s, when the Sacred Heart College in Johannesburg led the way in admitting children of colour as an act of defiance against the apartheid government (Soudien and Sayed 2003: 29).

The speakers in the sample who are at the upper limit of the age parameter (those born in 1983) would have started Grade One in 1989, at the age of six. They would thus have been able to attend public HOA or any other schools from Grade Two onwards. The lower age limit (speakers born not later than 1992) was set so that speakers in the sample will have had the opportunity to complete their formal schooling at the time of the interview. This enabled them to reflect holistically on their schooling experience. Those born in 1992 would have started their formal schooling in 1999 and, barring failure and dropout, completed Matric in 2010. The speakers in the sample were thus all born between 1983 and 1992, as illustrated in figure 3.2, which is a bar chart of the number of speakers per birth year. They are thus the generation before the ‘born-frees’, which is a term commonly used to refer to South African children who were born into democracy i.e. post-1994. Table 3.1 overleaf provides details of each speakers’ year of birth, their L1 and L2.

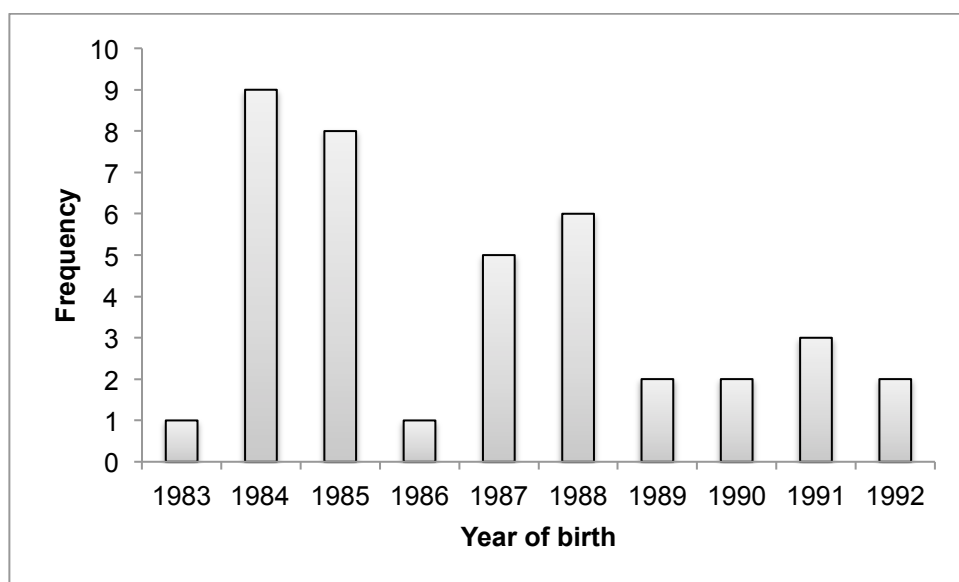


Figure 3.2: Bar chart of birth year for all 40 speakers, 1983-1992.

Code	Year of birth	L1	L2
M1	1987	English	Afrikaans
M2	1985	English	Afrikaans
M3	1990	English	Afrikaans
M4	1983	Afrikaans	English
M5	1991	Afrikaans	English
M6	1985	English	Afrikaans
M7	1984	English	Afrikaans
M8	1984	Afrikaans	English
M9	1986	English	Afrikaans
M10	1984	English	Afrikaans
M11	1985	English	Afrikaans
M12	1984	English	Afrikaans
M13	1988	Afrikaans	English
M14	1989	English	Afrikaans
M15	1990	Afrikaans	English
M16	1984	English	Afrikaans
M17	1984	English	Afrikaans
M18	1991	Afrikaans	English
M19	1987	English	English
M20	1982	English	Afrikaans
F1	1985	English	Afrikaans
F2	1985	English	Afrikaans
F3	1988	Afrikaans	English
F4	1988	English	Afrikaans
F5	1988	English	Afrikaans
F6	1987	Afrikaans	English
F7	1985	English	Afrikaans
F8	1987	Afrikaans	English
F9	1985	English	Afrikaans
F10	1988	English	Afrikaans
F11	1989	English	Afrikaans
F12	1987	English	Afrikaans
F13	1992	Afrikaans	English
F14	1992	Afrikaans	English
F15	1991	English	Afrikaans
F16	1984	English	Afrikaans
F17	1984	Afrikaans	English
F18	1985	English	Afrikaans
F19	1988	English	Afrikaans
F20	1984	English	Afrikaans

Table 3.1: Male and Female speakers' year of birth, L1 and L2.

3.1.5. Gender

Both sex and gender are relevant to the present study, as separate factors which should not be confounded. Sex is biological and determined prenatally, while gender is a sociocultural entity that is acquired postnatally (Chambers 2009: 116). In a phonetic study such as the present one, one sex difference that must be accounted for is that, on average, the length of women's vocal tracts is shorter than that of men, and as a result, female speakers typically display higher formant frequencies than their male counterparts (Flynn 2011: 2). In their raw forms, formant values for different speakers are incomparable, so the data must be normalised in order to eliminate differences that result from physiological factors, and preserving only socially salient differences (details about vowel formant normalisation follow in section 4.5).

Many sociolinguistic studies have shown clear gender differentiation in the use of linguistic variables, for example Labov (1972a) and Wolfram (1969). The same gendered pattern emerges consistently in these studies, as Chambers (2009: 114) summarises: 'women use fewer stigmatised and non-standard variants than do men of the same social group in the same circumstances'. Such clear patterning suggests that the possibility of gender distinction with regard to linguistic variables should be taken into account in sociolinguistic studies, lest the researcher fail to account accurately for patterns of linguistic variation and change.

More recently, South African based studies have shown gender to be a significant social variable in sociophonetic research. Results from my Masters thesis show that young Coloured males are leading the shift for the GOOSE lexical set, using a fronter variant than the females in the study (Dennis 2008). This finding seemingly contradicts the statement by Chambers (2009) above, unless it is argued that GOOSE fronting is in fact not a prestigious feature. (This is unlikely to be the case, however, and the results of the acoustic analysis of this sample do not find gender distinction to be significant for the GOOSE set.) Studying the same vowel category across all four South African population groups, Mesthrie (2010) found gendered patterns of variation: Black speakers most closely approximate the fronted GOOSE variant used by their White counterparts, females more so than males, while Indian and Coloured females show some resistance to the trend of fronting this vowel.

Results from the present study will show whether similar patterns are emerging for Coloured speakers with respect to all monophthongal vowels in the variety of English

under investigation. The sample was divided equally by gender (20 males, 20 females) in order to establish whether there is alignment of this social variable with significant phonetic differences.

3.1.6 Social class

Social class refers to the ordering of people in relation to others, according to factors such as income, education, occupation, residence or lifestyle (Milroy 1980: 13). Inequalities in the distribution of wealth, privilege and opportunity is ubiquitous in all societies (Chambers 2009: 39), thus every society can be stratified into classes, depending on the larger society's evaluation of each of the factors.

Where race was once the primary fault line upon which social division in South Africa was based, there has been a definite shift towards class as the defining factor. That said, Harold Wolpe (1988) cautions against the use of either class or race as a reductionist category through which to explain social difference in South Africa, arguing that neither race nor class can, in isolation, account for the inequitable distribution of privilege and power in South African society. Acknowledging the delicate dialectic that exists between class and race (and a host of other social factors too) in modern day, post-apartheid South Africa, I proceed nonetheless to categorise the speakers in the sample into class categories in order to facilitate analysis of this social variable with their linguistic correlates.

One aim of this research project is to compare the rate of change in middle-class and working-class English spoken in the Coloured community of Cape Town, therefore my sample contains speakers who fall into each of these categories. The final delimitation on the sample thus concerns the social class, or socio-economic status (SES)²⁴, of the speakers.

Under apartheid the South African class system was rigidly defined along colour lines. Apartheid labour and education policies engineered the class system in a way that preserved the highest ranking for Whites only; Coloured, Indian and Black South Africans occupied the lowest rungs of the socio-economic scale. That said, and even though there were limits to the extent of upward mobility for blacks, some social class

²⁴ The terms 'social class' and 'socio-economic status' (SES) are, strictly speaking, not synonyms. Boundaries of social class are largely untraversable, whereas one's SES can be improved or worsened through education and life choices (Rubin et al. 2014). For the purposes of this thesis, however, the terms are used interchangeably.

continua were inevitable e.g. doctors, lawyers, lecturers, accountants, etc. occupied a higher position on the social class continuum within each race group compared with those who were employed in traditionally working-class occupations.

In the post-apartheid setting, South African society is still largely divided into the haves and the have-nots. When apartheid proper, in its legislated form, ended in 1994, it gave way to 'economic apartheid' under which the inequality of the past is perpetuated largely along the same racial lines (Schneider 2003: 54) i.e. the South African black majority remains economically repressed as a result of the imbalances of the apartheid system. There is, however, an emerging black middle-class who have managed to improve their SES post-apartheid, due to a combination of factors including access to high-quality education, removal of occupational restrictions and the introduction of affirmative action policies (Seekings and Natrass 2002: 12). Those who make up this growing middle-class have benefited greatly from economic policies aimed at redress (Southall 2004: 539) and those black South Africans who are able to access such opportunities are encouraged to aspire to higher SES than their ancestors would have been able to achieve.

The implementation of employment equity policies, such as Black Economic Empowerment (BEE), were developed in an attempt to balance the representation of South Africans in all occupational grades to more accurately reflect the country's racial demographics. Despite the intention behind policy development, however, the masses of impoverished black South Africans have not experienced improvement in their standard of living as a result of these economic policies. The primary beneficiaries of these policies have been a handful of Black business magnates who have become incredibly wealthy in the post-apartheid economic setting (Ponte et al. 2007: 947).

The nature of inequality in South Africa is changing: inter-racial inequality is declining slowly, but intra-racial inequality has in fact increased because the gap between the richest and poorest black people has widened significantly (Seekings and Natrass 2002: 11, 26; Southall 2004: 531). As a result of this, the proportion of black South Africans who are upwardly mobile is relatively small.

The process of middle-class formation amongst black South Africans is a dynamic and nuanced area, making social class quite a complex attribute to judge. Some

speakers in the sample are quite clearly working-class, as they live in distinctly working-class neighbourhoods and their occupations (or lack of employment in some cases) are those typically associated with the working-class. Accurate classification of other speakers is more tricky because speakers who come from working-class backgrounds, and whose parents are working-class, have begun to elevate their SES through access to high quality education. Area of residence is no longer a reliable means of judging the SES of a speaker because some people who may have been financially able to move their families to more affluent neighbourhoods chose not to do so for reasons of community cohesion and familiarity.

Class inequality is closely linked with education: for those who are able to access it, tertiary education paves the way towards middle-class occupations, while those who fail to complete secondary school tend to struggle either with low-paying employment or intermittent unemployment throughout their lives (Seekings and Natrass 2002: 20; Borat 2004: 47).

For the purpose of categorising speakers into social classes, the intricacies of the class continuum need to be mapped in terms of the theoretical accounts of what constitutes each of the

categories. In this thesis thus far, I have made frequent reference to the middle-class and working-class. These two classes are generally accepted as categories of social division. An upper class does exist, but almost more as a theoretical construct than a social reality, especially in South Africa. It is constituted by people who have inherited wealth and privilege – an aristocracy in essence – who need not work in order to maintain high standards of living. The basic distinction between the middle-class and working-class is that the working-class earn their living by working with their hands, while the middle-class earn theirs through administrative work and services (Chambers 2009: 41). Sociologist Roger Southall defines South Africa's middle-class as those who draw their 'primary income (directly or indirectly) from non-manual employment, as 'white-collar employees', managers, self-employed business persons, or professionals' (Southall 2004: 522). Further subdivision of the broad categories allows for finer distinction between members of the middle-class and working-class, although it must be noted that social class is best thought of as a continuum rather than a set of rigid categories, as boundaries between the subgroups are inherently vague and fuzzy.

Judging the social class of others, such as the speakers in the sample, is a rather subjective task, but using an objective measure such as occupation as the primary determining factor reduces the subjectivity of the classification. Occupation could be regarded as the most significant of the factors determining social class, as it is a function of the other factors – one’s level of education largely determines the job one will have, and this, in turn determines the value of one’s remuneration. Categorising people into classes is not a simple task, however, and it is best practice to make use of a multi-factor scale rather than determining class solely on the basis of occupation²⁵.

A common approach used by sociolinguists to determine the SES of speakers in a sample is to develop a multi-index scale on which speakers are scored (e.g. Labov 1966; Trudgill 1974). This becomes necessary when judgement about an individual’s class category cannot reliably be based on the intuition of the researcher. Because of South Africa’s dynamic situation of middle-class formation, class is particularly difficult to judge. For this reason, I developed a SES index with which to rank the speakers into class categories.

The three factors typically considered in such indices are occupation, income and education. It is generally agreed that together they are a better determinant of SES than either component on its own (Deonandan et al. 2000: 1; Bradley and Corwyn 2002: 373). The SES index was thus made up of these three factors, as well as a fourth component, viz. parent’s occupation. Because the speakers in the sample are relatively young, and many were financially dependent upon their parents at the time of the interview, the occupation of one parent was taken into account as an additional factor indexing their social class.

The SES index was thus comprised of following four components:

1. Occupation
2. Parent’s occupation
3. Level of education
4. Income

²⁵ There are, however, instances in which occupation can be used as a sole determinant of class e.g. Macaulay (1976, cited in Chambers 2009: 51). In his sociolinguistic study of Glasgow, Macaulay found clear correlations between occupational groups and linguistic class markers.

The information for the first three components, speaker’s occupation, parent’s occupation and level of education, was gathered during the interview (discussed in section 3.3.1 below). Because income is a sensitive topic of discussion, I did not ask about personal or household income during the interview to prevent the interviewees becoming uncomfortable. Asking directly about income also runs the risk of false reporting, either through a desire to project a higher than actual SES, or through lack of accurate knowledge regarding family income, or both. As is practiced internationally in research studies where either level of education or income is unknown (Deonandan et al. 2000: 1), I made use of 2011 census survey data for each of the suburbs in which the speakers lived to determine the average household income for families in those areas, and placed the speakers on the income scale of the SES index according to the area in which they resided at the time of the interview.

Details for each speaker about each of the four components that comprise the SES index are provided below. The SES index is presented in table 3.2 below. Each component has six possible scores: the lowest is zero (0) and the highest is five (5). The minimum number of points a speaker can score on the SES index is zero (0) and the maximum number of points is twenty (20). Details of how each speaker scored is provided in table 3.6, after a discussion of each of the components of the index.

Occupation	Points	Parent’s Occupation	Points
University student / Professional	5	Professional worker	5
Skilled	4	Employer / manager	4
Semi-skilled	3	Skilled (non-manual)	3
Unskilled	2	Semi-skilled	2
Intermittent worker	1	Unskilled	1
Unemployed	0	Unemployed	0

Level of Education	Points	Income (projected for 2011)	Points
Three + years tertiary education	5	R145 000 – R170 000	5
Some tertiary education	4	R120 001 – R145 000	4
Grade 12	3	R95 001 – R120 000	3
Grade 11	2	R70 001 – R95 000	2
Grade 9	1	R40 001 – R70 000	1
Grade 7 or less	0	20 000 – 40 000	0

Table 3.2: Socio-economic Status Index: Four components.

Tables 3.3 and 3.4 show the occupations of male and female speakers respectively, as well as the occupation of one parent, as reported by each speaker during the interview. Each occupation is also ranked according to the scales in table 3.2 above. I consulted

the International Socio-economic Index of Occupational Status (Ganzeboom et al. 1992: 33-50) in order to rank the occupations of both the speakers and their parents. The parent whose occupation ranked higher was included in the index, or in the case of single parent families, the mother's occupation was included. There were no instances of single parent families in which the father was the sole parent. In the sample of 40 speakers, 11 mothers were housewives, and therefore not involved in activities that provide income for the household. Speaker F6 had recently lost her father, who was employed at the time of his death, so her mother's occupation is listed as housewife. The only speaker for whom both parents were not working was F15: her father had been medically boarded and her mother is a housewife.

Nineteen of the speakers were registered university students at the time of the interview, so accommodation had to be made for this in the scale. University enrolment is prestigious and available to only a handful of South African school leavers, so on the occupation scale, 'student' attained the maximum number of points, alongside professional. This analysis is supported by the causal link between achieving a university degree and becoming a professional worker. The tertiary institution at which the speaker was enrolled is provided in parentheses after their occupation listed as 'student' in tables 3.3 and 3.4. The University of Cape Town (UCT) is a former White university while the University of the Western Cape (UWC) is a former Coloured university – the only Coloured university in South Africa during apartheid. The Cape Peninsula University of Technology (CPUT) is a technical university that historically was in fact two separate institutions, one catering for Coloured students, the other for White students²⁶. M9 was the only speaker who had completed his degree at the time of the interview, so was classified as a 'professional' on the occupation scale, even though his incumbent occupation as a business analyst might have been classified as 'skilled' rather than 'professional'. Five of the speakers were employed as 'Service Delivery Officers', which is the title assigned to the operators of machines which dispense *Allpay* – the South African government's social grant payout system.

The parental occupation scale differs from the speakers' occupation scale in that none of the parents were students, so the top ranking on the scale was 'professional'.

²⁶ Cape Technikon and Peninsula Technikon merged in 2003 to form the Cape Peninsula University of Technology, as part of a restructuring of the Higher Education system in South Africa in the early 2000s (Jansen 2003).

Occupations which were ranked professional include doctor and lawyer. Business owners who employ staff, and company managers were ranked below professionals. Below this ranking were skilled, semi-skilled, and unskilled labourers.

On the third scale in the index, speakers were rated according to their level of education. Those who had completed three or more years of tertiary education (whether or not they had completed the degree) were highest ranked. Following that were those who had begun their tertiary studies, but not yet completed three years. The third ranking was a matriculation pass, and following that, various levels of school completion. On the education scale, none of the speakers scored zero as all of them had successfully completed at least one year of high school.

Code	Occupation	Ranking	Parent's occupation*	Ranking
M1	Service Delivery Officer	Semi-skilled	Shop assistant	Semi-skilled
M2	Student (UCT)	University student	Business owner: Imports / exports	Employer
M3	Student (UWC)	University student	Municipal officer	Skilled (non-manual)
M4	Student (UCT)	University student	Primary school teacher	Skilled (non-manual)
M5	Unemployed	Unemployed	Cleaner	Unskilled
M6	Student (UCT)	University student	Attorney	Professional
M7	Student (UCT)	University student	Local government manager	Manager
M8	Unemployed	Unemployed	Cleaner	Unskilled
M9	Business analyst	Professional	Machine repairman	Semi-skilled
M10	Disability payment administrator	Skilled	Business owner: Civil engineer	Employer
M11	Student (UCT)	University student	Disaster management supervisor	Manager
M12	Hotel Front Office Manager	Semi-skilled	Lecturer (CPUT)	Professional
M13	Maintenance worker	Semi-skilled	Pastor	Skilled (non-manual)
M14	Stock Replenisher	Semi-skilled	Garment cutter	Semi-skilled
M15	Unemployed	Unemployed	Factory worker	Semi-skilled
M16	Student (UCT)	University student	Business owner: travel and tourism	Employer
M17	Student (CPUT)	University student	Nursery school teacher	Skilled (non-manual)
M18	Construction worker	Intermittent worker	Factory worker	Semi-skilled
M19	Student (UCT)	University student	Doctor	Professional
M20	Tiler	Semi-skilled	Quantity surveyer	Skilled (non-manual)

Table 3.3: Male Speakers' occupations and Parents' occupations.

* In single parent families, the mother's occupation is listed. In cases where speakers have both parents, the occupation that is rated higher (according to the International Socio-economic Index of Occupational Status (Ganzeboom et al. 1992: 33-50) is included in this scale.

Code	Occupation	Ranking	Parent's Occupation	Ranking
F1	Student (UCT)	University student	Anglican Diocese administrator	Skilled (non-manual)
F2	Service Delivery Officer	Semi-skilled	Cleaner	Unskilled
F3	Service Delivery Officer	Semi-skilled	Painter	Semi-skilled
F4	Admin clerk	Semi-skilled	Local government clerk	Skilled (non-manual)
F5	Student (UCT)	University student	Publisher	Manager
F6	Supervisor	Skilled	Housewife	Unemployed
F7	Student (UCT)	University student	Company Manager	Manager
F8	Admin Clerk	Semi-skilled	Nurse	Semi-skilled
F9	Student (UCT)	University student	Local government manager	Manager
F10	Service Delivery Officer	Semi-skilled	Painter	Semi-skilled
F11	Student (UCT)	University student	Bank Area manager	Manager
F12	Student (UCT)	University student	Training and development manager	Manager
F13	Unemployed	Unemployed	Factory worker	Semi-skilled
F14	Unemployed	Unemployed	Factory worker	Semi-skilled
F15	Admin clerk	Semi-skilled	Unemployed (medically boarded)	Unemployed
F16	Student (UCT)	University student	Company Director	Manager
F17	Service Delivery Officer	Semi-skilled	Cleaner	Unskilled
F18	Student (UWC)	University student	Magistrate	Professional
F19	Student (UCT)	University student	Doctor	Professional
F20	Student (UCT)	University student	Business owner: Electrical	Employer

Table 3.4: Female speakers' occupations and Parents' occupations.

* In single parent families, the mother's occupation is listed. In cases where speakers have both parents, the occupation that is rated higher (according to the International Socio-economic Index of Occupational Status (Ganzeboom et al. 1992: 33-50) is included in this scale.

The final component of the index was a rating according to the annual household income for the neighbourhood in which the speakers resided at the time of the interview. Because I did not ask about household or personal income during the interview (for reasons outlined above), I made use of 2011 census data for each suburb in order to arrive at the figures in the scale. Table 3.5 below is an example of how the census data is presented for each suburb. The figures in the table are for Athlone.

Income range	Number	Percentage
No income	726	9.9
R1 – R1 600	834	11.4
R1 601 – R3 200	729	10
R3 201 – R6 400	837	11.9
R6 401 – R12 800	1176	16.1
R12 801 – R25 600	1482	20.2
R25 601 – R51 200	1059	14.5
R51 201 – R102 400	345	4.7
R102 401 or more	102	1.4
Total	7326	100

Table 3.5: Monthly Household income for Coloured households in Athlone, 2011 census data. Source: <http://www.capetown.gov.za/en/stats/Pages/2011-Census-Suburb-Profiles-land.aspx>

In order to arrive at the income scale for the SES index, I multiplied the number of Coloured households in each income range by the median in that range. For the highest income range, I multiplied the number of households by the minimum value in that range (R102,401). These figures were added together to get an estimated value of the total monthly income for Coloured households in the suburb, and then divided by the total number of (Coloured) households in the suburb. Using this calculation, the average monthly household income for Athlone was R16,920.75. The suburb with the highest average annual household income was Pinelands, with a value of R40,404.46 and the suburb with the lowest, was Delft, with a value of R4,276.39. The income component of the SES index was thus calibrated at R7,000 increments beginning at R0 and ending at R42,000.

The scores for each speaker on the four components are provided in table 3.6 below. Based on their total score, each speaker was placed into one of four class categories as indicated in table 3.7: lower working-class, upper working-class, lower middle-class and upper middle-class. Table 3.8 shows which speakers fit into each of the four class categories. The upper middle-class is over-represented in the sample, with a total of 15 speakers falling into this category (seven male, eight female). Nine speakers fall into the lower middle-class category (five male, four female), 10 into the upper

working-class (four male, six female) and five into the lower working-class (three male, two female). Removing the subdivisions, the total number of working-class speakers in the sample is thus 15 (seven male, eight female) and the total number of middle-class speakers is 24 (12 male, 12 female). The uneven split of the sample in terms of class was due to erroneous class judgement on the part of the researcher during the speaker selection process (due to the dynamic state of the South African class system, discussed above). Many speakers who were selected to fill pre-defined working-class categories in fact turned out to be middle-class based on the index developed for class assignment.

Code	Occupation	Parent's Occupation	Level of education	of Income range	Total
M1	3	2	3	1	9
M2	5	4	5	2	16
M3	5	3	4	1	13
M4	5	3	5	1	14
M5	0	1	2	0	3
M6	5	5	5	2	17
M7	5	4	5	2	16
M8	0	1	1	0	2
M9	5	2	5	1	13
M10	4	4	4	2	14
M11	5	4	5	3	17
M12	3	5	4	5	17
M13	2	4	3	0	9
M14	3	2	3	0	8
M15	0	2	2	0	4
M16	5	4	5	4	18
M17	5	3	4	1	13
M18	1	2	3	0	6
M19	5	5	4	5	19
M20	3	3	2	1	9
F1	5	3	5	2	15
F2	3	1	3	1	8
F3	3	2	3	1	9
F4	3	3	3	3	12
F5	5	4	4	2	15
F6	4	0	5	0	9
F7	5	4	5	2	16
F8	3	2	5	1	11
F9	5	4	4	5	18
F10	3	2	3	1	10
F11	5	4	4	4	17
F12	5	4	4	2	15
F13	0	2	3	0	5
F14	0	2	3	0	5
F15	3	0	3	1	7
F16	5	4	5	5	19
F17	3	1	2	0	6
F18	5	5	5	4	19
F19	5	5	4	2	16
F20	5	4	5	2	16

Table 3.6: SES scores for 40 speakers, by component.

Class category	Points
Lower Working Class	0-5
Upper Working Class	6-10
Lower Middle Class	11-15
Upper Middle Class	16-20

Table 3.7: Socio-economic Scale.

Category	Male Speakers	Female Speakers
Lower WC	M5, M8, M15	F13, F14
Upper WC	M1, M13, M14, M18, M20	F2, F3, F6, F10, F15, F17
Lower MC	M3, M4, M9, M10, M17	F1, F4, F5, F8
Upper MC	M2, M6, M7, M11, M12, M16, M19	F7, F9, F11, F12, F16, F18, F19, F20

Table 3.8: Speakers' Socio-economic ranking.

3.2 Degree of Integration

Social class, dubbed ‘the primary social variable in sociolinguistics’ by Chambers (2009: 74) often shows clear correlations with linguistic patterns of variation. It is possible, however, that speakers in the present study might group more naturally based on their degree of integration into White society. For reasons outlined in chapter one and in section 3.1.6 above, social segregation is still a reality for most South Africans, and that segregation remains largely racially based. The primary site of desegregation is former HOA educational institutions, and access to such institutions is available only to middle-class black South Africans, for economic reasons. For the majority of Coloured people, their social networks remain predominantly Coloured. What follows is a discussion of the composition of the sample in terms of the diverse array of schools the speakers attended (section 3.2.1) and the areas in which they reside (section 3.2.2). In Section 3.2.3 I provide an overview of the friendship circles and close contacts of the speakers, based on the information they offered on the topic during their interviews. Section 3.2.4 brings this information together into a Degree of Integration index, in order to group speakers based on the extent to which they have integrated into White society.

3.2.1 Schooling

The 40 speakers attended a total of 43 primary schools and 27 high schools. Of the primary schools, 25 of them are former HOR schools, 11 are former HOA schools and seven are independent (private) schools (see table 3.9). Of the high schools, 14 are

former HOR schools, 10 are former HOA schools and three are independent schools (see table 3.10).

The schools attended by the speakers in the sample represent the very best and among the worst of South Africa's educational offering. In tables 3.9 and 3.10, the primary and high schools are listed in order of ascending annual school fees (based on 2012 fees), so the school that charges the lowest school fees is listed first on each respective table, and the school with the highest fees is listed last. The tables clearly show that HOR schools charge the lowest school fees, and private schools charge the highest. HOA school fees are in between those charged by schools in the other two categories. The only school that does not fit neatly into this patterning is Madrassa Tu Tarbiyyah Islamic School (table 3.9 number 26), which is a private school, but is less expensive than all the HOA schools listed in the table. This is probably because the school caters for the Muslim population in and around Grassy Park, a relatively low socio-economic area on the Cape Flats. Table 3.11 lists the schools attended by each speaker.

As mentioned in chapter one (section 1.6), there is a close correlation between positive educational outcomes and financial resources available to the school administrative bodies (Fiske and Ladd 2003: 17). The better resourced schools are those that charge higher fees, because they are able to employ more teachers, have smaller class sizes and lower pupil to teacher ratios (Fiske and Ladd 2003: 17). Despite post-apartheid redistribution of governmental spending on education, with more resources being assigned to poorer schools, little has changed in terms of matriculation output in the South African schools system. Schools perform much as they did under apartheid: black schools have generally maintained their low matriculation outputs while former HOA schools perform as well as they did in the past (van der Berg and Burger 2003: 497). The correlation with fees is also strong: schools that charge higher fees produce the best matriculation results, and schools at the bottom of the socio-economic ladder produce much poorer results (van der Berg and Burger 2003: 504-505).

	School	Fees*	Former Department
1	Sonderend Primary	R100	HOR
2	Silverstream Primary	R150	HOR
3	Delft Primary	R150	HOR
4	Voerspoed Primary	R150	HOR
5	Bergsig Primary	R250	HOR
6	Hyde Park Primary	R280	HOR
7	Cascade Primary	R300	HOR
8	Arcadia Primary	R300	HOR
9	Eastville Primary	R350	HOR
10	Welcome Primary	R350	HOR
11	Vanguard Primary	R400	HOR
12	Jamaicaweg Primary	R400	HOR
13	Jan Bosman Primary	R420	HOR
14	Mitchells Plain Primary	R450	HOR
15	Goeiehoop Primary	R500	HOR
16	Plantation Primary	R600	HOR
17	Silverlea Primary	R650	HOR
18	Norma Road Primary	R660	HOR
19	St Augustine's Primary	R980	HOR
20	Zonnebloem Girls Primary	R1 000	HOR
21	St John's Primary	R1 050	HOR
22	Muhammadiyah Primary	R1 200	HOR
23	Habibia Primary	R1 300	HOR
24	York Road Primary	R2 000	HOR
25	Turfhall Primary	R2 350	HOR
26	Madrasa Tu-Tarbiyyah Islamic School	R3 000	Private
27	Koos Sadie Primary	R4 200	HOA
28	Mountain Road Primary	R4 400	HOA
29	Labiance Primary	R4 510	HOA
30	Golden Grove Primary	R5 171	HOA
31	Blouberg Ridge Primary	R8 880	HOA
32	Bergvliet Primary	R9 075	HOA
33	Kenridge Primary	R10 980	HOA
34	Wynberg Girls High	R13 172	HOA
35	Rustenburg Girls Junior	R17 800	HOA
36	Rondebosch Boys Preparatory	R20 790	HOA
37	El Shaddai Christian	R23 460	Private
38	SACS	R27 000	HOA
39	St Joseph's Marist College	R30 609	Private
40	Springfield Convent	R34 355	Private
41	Constantia Waldorf	R34 653	Private
42	St George's Grammar	R42 388	Private
43	Bishops Diocesan College	R57 303	Private

Table 3.9: Speakers' Primary Schools, Fees and Former Department.

Source: <http://wcedemis.pgwc.gov.za/wced/findaschool.html>

*2012 rates

	School	Fees	Former Department
1	Leiden Secondary	R300	HOR
2	Beacon Hill Secondary	R400	HOR
3	Phoenix High	R400	HOR
4	Crystal Secondary	R480	HOR
5	Arcadia High	R500	HOR
6	Athlone Secondary	R950	HOR
7	Garlandale Secondary	R1 250	HOR
8	Belgravia High	R1 500	HOR
9	Pelican Park High	R1 500	HOR
10	Belgravia Secondary	R1 500	HOR
11	Wittebome High	R1 640	HOR
12	Fairmount Secondary	R1 800	HOR
13	Maitland Secondary	R2 000	HOA
14	Livingstone High	R4 400	HOR
15	Cape Town High	R5 000	HOA
16	Bellville High	R8 600	HOA
17	Fairburn College	R11 050	HOA
18	Fairmont High	R16 000	HOA
19	Pinelands High	R17 800	HOA
20	Wynberg Girls High	R20 500	HOA
21	Westerford High	R21 165	HOA
22	Rustenburg Girls High	R24 000	HOA
23	Rondebosch Boys High	R25 600	HOA
24	SACS	R27 500	HOA
25	Springfield Convent	R34 355	Private
26	Reddam House	R52 725	Private
27	Bishops Diocesan College	R82 560	Private

Table 3.10: Speakers' High Schools, Fees (2012 rates) and Former Department.
Source: <http://wcedemis.pgwc.gov.za/wced/findaschool.html>

Code	Primary School*	High school**
M1	23, 26	8
M2	38	24
M3	10	15
M4	5, 15	16
M5§	6	12
M6	31	27
M7	13, 30, 33	18
M8§	6	12
M9	18	7, 13
M10	11, 30	21
M11	14, 24	27
M12	30	23
M13	8	6
M14	16	9
M15§	6	12
M16	41, 30	21
M17	19, 43	27
M18	4	4
M19	32, 42	19
M20	22	10,11
F1	19, 33	18
F2	11	14
F3	2	3
F4	21, 27	17
F5	34	26
F6	1	3
F7	35	22
F8	9	7
F9	20, 28	14
F10	12	2
F11	34	20
F12	40	25
F13	3	1
F14	3	1
F15	17	8
F16	37, 33	18
F17§	7	5
F18	39, 35	22
F19	25	21
F20	25, 35	22

Table 3.11: Schools attended by the speakers.

* See table 3.9 for school names and details.

** See table 3.10 for school names and details.

§ Did not complete high school.

3.2.2 Area of Residence

Implementation of the Group Areas Act of 1950 saw South Africa's population segregated residentially, and even though the Group Areas Act was repealed in 1991, the population is still largely segregated as designed by the apartheid government. As is the case with the HOA schools, the residential areas that were reserved for Whites are still most desirable, and have the highest property values. Economic constraint is one of the primary factors preventing movement from former Coloured areas into former White areas. Where such movement has occurred, it is amongst upwardly mobile families who could afford to buy or rent property in formerly Whites-only areas. Again, this is a function of class and it highlights the class distinction within the Coloured community, which is discussed above (section 3.1.6).

Most of the Coloured community of Cape Town lives in a vast area known as the Cape Flats. The area is named for its flat terrain, and is the area to which many Coloured families were relocated when the Group Areas Act was implemented after its enactment in 1950. The Cape Flats, notorious for gangsterism and drug abuse (Kinnes 2000), is comprised of a number of Coloured and Black townships. Many former Coloured suburbs listed in table 3.12 form part of the Cape Flats and are labelled with an asterisk (*). Involvement in gangs is one of the reasons why many young people from the Cape Flats fail to complete their formal schooling. In 2011, only 35.7 percent of Coloured adults (aged 20+) in Cape Town had completed Grade 12; and just 8.7 percent were educated at tertiary level (Statistics South Africa 2012).

In the sample, the majority of speakers ($n = 27$; 67.5 percent) still lived in areas zoned 'Coloured' under the Group Areas Act at the time of the interview (see table 3.12). The remaining 13 speakers had moved (along with their families) into former White areas at various stages during their schooling careers.

Of the suburbs listed in table 3.12 below, both former White and former Coloured, some are more affluent than others. Rondebosch East, for example, is situated adjacent to a very affluent former White suburb (Rondebosch) and catered for a higher socio-economic subsection of the Coloured community, while Mitchells Plain is home to a lower socio-economic component of the Coloured population, and is one of the suburbs renowned for gangsterism and its affiliate social ills.

Code	Former Coloured Suburb	Former White Suburb
M1	Grassy Park*	
M2	Walmer Estate	University Estate
M3	Heideveld*	
M4	Glenhaven	
M5	Parkwood*	
M6	Atlantis	Durbanville, Rondebosch
M7	Kuilsriver	Durbanville
M8	Parkwood*	
M9	Athlone*	Brooklyn
M10	Surrey Estate*, Athlone*	
M11	Mitchells Plain*, Lansdowne	
M12	Fairways, Rondebosch East	Pinelands
M13	Bonteheuwel*	
M14	Skaapkraal*	
M15	Parkwood*	
M16	Rondebosch East	Plumstead
M17	Strandfontein*	Somerset West
M18	Hanover Park*	
M19	Retreat*	Pinelands
M20	Wynberg, Belhar*	
F1	Belhar*	Durbanville
F2	Manenberg*, Mitchells Plain*	
F3	Heideveld*	
F4	Kensington	Goodwood
F5	Wetton, Wynberg	
F6	Manenberg*	
F7	Rondebosch East	
F8	Mitchells Plain*	
F9	Northpine, Woodstock	Sybrand Park
F10	Mitchells Plain*	
F11	Wynberg, Zeekovlei, Grassy Park	Plumstead
F12	Retreat*, Heathfield	
F13	Delft*	
F14	Delft*	
F15	Mitchells Plain*	
F16	Glenhaven	Durbanville, Pinelands
F17	Bishop Lavis*	
F18	Northpine	Goodwood, Plumstead
F19	Crawford	
F20	Rondebosch East	

Table 3.12: Area of Residence.

Key: *Cape Flats.

The speakers whose families moved to White areas generally moved into reasonably affluent White areas (such as Durbanville, Pinelands, Plumstead). One exception is speaker M9, whose family moved to Brooklyn, a suburb home to the lowest socio-economic subsection of Whites.

In tables 3.13 and 3.14, a profile of two former Coloured and two former White suburbs, based on 2011 census data, is presented to illustrate the diversity of the various suburbs. Durbanville is still a predominantly White area (81 percent White) with a small Coloured population (11.4 percent) and even smaller Black population (5.5 percent). The level of education is high: 84.1 percent of residents have Grade 12 or higher, and the distribution of residents in terms of income is concentrated in the four highest income ranges (see table 3.13), where 65.3 percent are reported to be placed. Brooklyn is a lower socio-economic area, with only 22.1 percent of the population falling into the four highest income ranges – the rest (77.9 percent) earn below R12,800 per month. More than half the residents (55.6 percent) of Brooklyn had achieved Grade 12 or above. The suburb is no longer predominantly White, in fact the White population is smaller than both the Black (35.6 percent) and Coloured (31.5 percent) populations, with just 30.3 percent in Brooklyn now represented in this category.

The two former Coloured suburbs, Athlone and Delft, contrast quite starkly with one another, and with the former White suburbs. In Delft, 86.8 percent of the population earn less than R6,400 per month, while the income statistics for Athlone are higher than for the former White suburb, Brooklyn: 40.8 percent of the population fall into the four highest income categories. Almost 60 percent of the residents of Athlone have an educational qualification equal to or higher than Grade 12, while in Delft, this figure stands at just 26.7 percent. Both Athlone and Delft remain predominantly Coloured (62.3 and 51.5 percent respectively), with small White populations (2 and 0.1 percent respectively). Delft has a sizable Black population, however (46.2 percent), while the Black population of Athlone makes up 8.5 percent of the suburb's population. Athlone is the only one of the four suburbs listed here which has a sizable Indian population (22 percent).

Suburb	White		Coloured	
	Durbanville	Brooklyn	Athlone	Delft
% Coloured residents	11.4	31.5	62.3	51.5
% White residents	81	30.3	2	0.1
% Black residents	5.5	35.6	8.5	46.2
% Indian/Asian	1	1.2	22	0.3
% residents with Gr 12+	84.1	55.6	59.1	26.7

Table 3.13: Profile of four suburbs, 2011 data.

Compiled from 2011 Population Census data (Statistics South Africa 2012), available at <http://www.capetown.gov.za/en/stats/Pages/2011-Census-Suburb-Profiles-land.aspx>.

Income range	White		Coloured	
	Durbanville	Brooklyn	Athlone	Delft
No income	7.2	9.2	10.7	16.6
R1 – R1 600	2.8	15.9	10.6	28.5
R1 601 – R3 200	3.9	13.9	10.1	23.9
R3 201 – R6 400	6.3	17.8	11.9	17.8
R6 401 – R12 800	14.3	21	16.1	9.4
R12 801 – R25 600	22.1	15.1	19.9	2.9
R25 601 – R51 200	23.9	5.9	14	0.7
R51 201 – R102 400	14.6	0.6	5.2	0.1
R102 401 or more	4.8	0.5	1.7	0.1
Total	100	100	100	100

Table 3.14: Average monthly household income in four suburbs (2011 data), percentage.

Compiled from 2011 Population Census data (Statistics South Africa 2012), available at <http://www.capetown.gov.za/en/stats/Pages/2011-Census-Suburb-Profiles-land.aspx>.

3.2.3 Close Contacts

In this section I extrapolate information about the speakers' circles of close contacts based on what they reported during their interviews. A useful sociolinguistic tool for examining contacts of individuals in a community is the concept of the social network, which was introduced to the field of sociolinguistics by Lesley Milroy (1980). Applying this theory requires the researcher to establish, in detail, the nature of people's relationships with members of their community and outsiders to their community. Such detailed information is most easily attainable through the use ethnographic fieldwork methods. As such methodology was not employed in the data collection process, social network theory was not applied in the present study. Instead, a rather more basic presentation of the speakers' closest ties is presented. Close contacts are considered to be friends, colleagues, sports teammates and the like.

A number of the speakers (n = 16; 8 male, 8 female) were educated exclusively in former HOR schools and have lived in former Coloured residential areas from birth until adulthood. For these speakers, their circle of friends and close contacts was

almost exclusively Coloured. This is because HOR schools have retained predominantly Coloured student populations, with almost no White or Indian pupils and only a small percentage of Black pupils (Chisholm and Sujee 2006: 151). These speakers worked in environments in which most of their colleagues were either Coloured or Black. The speakers in this category who were unemployed spent most of their time with Coloured peers. None of the speakers in the sample reported that they had close Black friends or colleagues.

For the balance of the speakers (n = 24), their closest friends and contacts were a mixture of predominantly Coloured and White people. The different speakers had varying degrees of contact with White people. Speaker F1, for example, reported that all of her school friends were White as she did not feel comfortable in the company of the other Coloured pupils who attended her school. From the age of seven, her family lived in a formerly White residential area and attended an HOR school for only two years. Speaker F20, by contrast, lived in a former Coloured area all her life, had predominantly Coloured friends, and attended an HOR school for three years before moving to a former HOA school.

3.2.4 Degree of Integration Index

Descriptions of the speakers' schooling experiences, areas of residence and the makeup of their closest contacts were used to rank them according to their degree of integration into White society. The index for ranking speakers is provided in table 3.15. On each of the three scales of the index, speakers scored between zero (0) and four (4) points. The maximum number of points any speaker could score was thus 12, and the minimum was zero. The speakers' scores on the Degree of Integration index are presented in table 3.16.

Schooling	Points	Residence	Points	Close contacts	Points
HOR all along	4	Coloured all along	4	Only Coloured	4
HOR up to Matric	3	Coloured up to Matric	3	Mainly Coloured	3
HOR up to Grade 7	2	Coloured up to Grade 7	2	Coloured and White equally	2
HOR up to Grade 3	1	Coloured up to Grade 3	1	Mainly White	1
HOA all along	0	White all along	0	Only White	0

Table 3.15: Degree of Integration Index.

Speaker Code	Schooling	Residence	Friends	Total
M1	4*	4	4	12
M2	0	2	2	4
M3	2	4	3	9
M4	2	4	3	9
M5	4	4	4	12
M6	0	2	2	4
M7	0	2	2	4
M8	4	4	4	12
M9	4	4	3	11
M10	1	4	3	8
M11	2	4	3	9
M12	0	2	3	5
M13	4	4	4	12
M14	4	4	4	12
M15	4	4	4	12
M16	2	3	2	7
M17	0	2	3	5
M18	4	4	4	12
M19	0	1	3	4
M20	4	4	4	12
F1	1	1	1	3
F2	4	4	4	12
F3	4	4	4	12
F4	2	2	3	7
F5	0	4	1	5
F6	4	4	4	12
F7	0	4	2	6
F8	4	4	4	12
F9	3	2	2	7
F10	4	4	4	12
F11	0	3	3	6
F12	0	4	3	7
F13	4	4	4	12
F14	4	4	4	12
F15	4	4	4	12
F16	0	1	2	3
F17	4	4	4	12
F18	0	1	2	3
F19	2	4	3	9
F20	1	4	3	8

Table 3.16: Speaker scores on Degree of Integration Index.

*Speaker M1 attended Madrassa Tu Tarbiyyah Islamic School, which is private, but in this instance classified along with HOR schools because its pupil population is exclusively Coloured (and/or Indian).

Category	Points
Completely integrated	0-2
Largely integrated	3-5
Slightly integrated	6-9
Unintegrated	10-12

Table 3.17: Degree of Integration Scale.

Category	Male Speakers	Female Speakers
Completely integrated	-	-
Largely integrated	M2, M6, M7, M12, M17, M19	F1, F5, F16, F18
Slightly integrated	M3, M4, M10, M11, M16	F4, F7, F9, F11, F12, F19, F20
Unintegrated	M1, M5, M8, M9, M13, M14, M15, M18, M20	F2, F3, F6, F8, F10, F13, F14, F15, F17

Table 3.18: Speakers' ranking on Degree of Integration Scale.

The scale for ranking speakers is provided in table 3.17. The speakers' rankings on the Degree of Integration scale (table 3.18) suggests that none of the speakers are completely integrated into White society i.e. they all retain some ties with the Coloured community. Ten of the speakers (six male, four female) scored between three and five points, attaining the classification 'largely integrated'. Twelve of the speakers (five male, seven female) were classified as 'slightly integrated', and the balance (18 speakers: 9 male, 9 female) were classified as 'unintegrated'.

There is considerable overlap between the SES-index scale and Degree of Integration scale: all speakers classified as 'unintegrated' on the latter scale were either upper or lower working-class speakers, save for M9 and F8, who were lower middle-class speakers. The upper and lower middle-class speakers (as indicated on the SES-index scale in table 3.8) were all classified as either 'slightly integrated' or 'largely integrated'.

Due to this overlap in the two scales, I have maintained the division of speakers in their class categories for the purposes of analysis in chapters five and six. The 'upper' and 'lower' categories are collapsed so that just two classes are represented, viz. middle-class and working-class.

3.3 Data Collection and Preparation

Using the most common method of data collection in sociolinguistics, I conducted a sociolinguistic interview with each speaker in the sample in order to elicit the data necessary for phonetic analysis. What follows is a discussion of the interview procedure and the materials used in the interviews (section 3.3.1). Section 3.3.2 is an overview of the equipment used and the shortcomings thereof. Details about how the interviews were transcribed are provided in section 3.3.4.

3.3.1 Interview Procedure and Materials

The sociolinguistic interview is the traditional method for collecting data in sociolinguistic studies (Feagin 2002: 26; Milroy and Gordon 2003: 57). Developed by Labov (1966) and employed by countless sociolinguists the world over, this methodology was suitable for the purposes of this research project because of its semi-structured and relatively informal nature, and the fact that it elicits a sample of speech from each speaker that can be analysed using acoustic technology.

All interviews were conducted face-to-face in a location most convenient for the interviewee. Wherever possible, the interviews were conducted in the speakers' homes in order for them to feel most comfortable and relaxed, which is important when trying to elicit speech that is as close as possible to the speakers' vernacular (Labov 1972a: 208). Interviews that were not conducted in the speakers' homes were conducted either at their workplace or university campus, and a few were conducted at the home of the researcher. The interviews typically lasted about 45 minutes, although they ranged in length from 30 to 90 minutes. In order to keep track of the different speakers I was interviewing, I asked each interviewee to complete a form of demographic information (appendix A). Following established ethical practices, prior to the commencement of each interview, I obtained the informed consent of the interviewees to the following:

- Recording the interview
- Using their demographic data in aggregate form
- Using the recordings and the information contained therein for the purposes of academic research

The interview was comprised of three sections: a Word List, a Reading Passage and a casual discussion. Respectively, these sections elicit increasingly informal speech styles within the setting of a controlled interview. The Word List was a slightly

amended version of the lexical set proposed by Wells (1982). It consists of a set of words which represent each vowel in the English vowel system. The Word List is appended as appendix B. Speakers were asked to read each word carefully and slowly, as it was revealed to them. In this way, a very careful speech style was elicited from the interviewees.

The second component of the interview, the Reading Passage (appendix C), was a short passage which interviewees were asked to read aloud. The style of speech typically used for this genre tends to be less careful than for Word List style, as speakers have the added concern for the fluency with which they read the passage. The style remains more careful than the casual style that speakers use when they are speaking rather than reading, however. After discovering that beginning the interviews with these formal components tended to make the interviewees more nervous about being ‘analysed’, in subsequent interviews I requested the interviewees to read the Word List and Reading Passage only after we had concluded our casual conversation.

The function of the final component of the interview, the casual discussion, was two-fold: firstly, to gather personal information about such things as their parents’ occupations and other details about their nuclear and extended families, their areas of residence and their schooling experiences. Questions which elicit such data work well as ‘ice breakers’ and help both interview parties to feel more comfortable at the beginning of the interview. The second function of this component is to elicit a more casual style of speech, in the hope of gathering speech data as near as possible to the vernacular. Recognising that the desire of the researcher is to observe the very speech style that speakers use while not under observation, a phenomenon known as the ‘Observer’s Paradox’ (Milroy and Gordon 2003: 49), I do not presume to have gathered this style from all, or perhaps any, of the speakers in the sample, but I did attempt to get as close to it as possible. A method of circumventing this constraint, developed by Labov (1972b), is to get the speaker to relate a narrative. Speakers can become so involved in telling a story that they essentially forget their immediate context and revert to their most natural speech style (Schiffrin 1996: 41). To this end, I asked the speakers questions about preparations for their Matric Ball²⁷, games they played as children and other questions that the interview context suggested might

²⁷ A Matric Ball (or Matric Dance) is a celebration for pupils in their final year of school, similar to the American *prom*.

yield interesting narrative. While Labov's (1996) 'danger of death' question may well have provided interesting content from some of the working-class (especially male) speakers, for most of the speakers in the sample who lived a reasonably sheltered existence, I didn't expect that they would have had much in answer to this question, hence its exclusion from the interview schedule.

3.3.2 Recording Equipment

Early interviews were recorded using a Marantz MP3 recorder. Later interviews were recorded on an Olympus, which was more practical as it was smaller than the Marantz and did not need to be connected to an electrical source. Recording the interview allowed me to make use of the digital format in software applications like Praat, P2FA and extractFormants (details of these programs are discussed in chapter four). Of course, the recordings also provide a perfect record of the interview, which I consulted on numerous occasions for information about schooling, parents' occupation and area of residence *inter alia*.

In all interviews, I placed the recorder on a table or other surface near to the interviewee. Due to this positioning, the recorder picked up not only the interviewee and interviewer's speech, but also any background noises that were present in the environment at the time. Having begun this research project with the intention of manually measuring each vowel token, this problem, while not ideal, would not have been as much of an issue as it eventually posed because I, in fact, used a method of Automatic Vowel Measurement (AVM) instead (which is detailed in chapter four).

The first step of AVM is to segment the audio signal into phonemes and align it with a phoneme level transcription of the recording (Evanini 2009: 24). This process, known as forced alignment, performs optimally when the speech signal is clean i.e. free of any noise besides the interviewee's speech. A clean recording could be achieved by conducting interviews in sound studios or similar facilities, which eliminate outside noises from the environment, but the sterility of such an environment would compromise the casual and relaxed atmosphere that is achieved by conducting interviews in people's homes or other equally comfortable environments. A second, more practical method of achieving a clean recording would be to attach a lapel mic to the speaker. Because this was not done in any of the interviews I conducted, I manually edited each interview recording to remove loud background noises (such as ringing phones, family members shouting in another part

of the house, etc.) using a program called *Wavepad*. I also edited out my own speech and all instances of overlapping speech, as well as laughter, coughing, sneezing and the like.

3.3.3 Transcription

Interview transcription is part of the data preparation procedure for AVM, as an orthographic transcription is needed for the forced alignment process. All interviews were transcribed using a playback program called *ExpressScribe* which makes it easy for transcribers to use functions such as pause, rewind and play while working in a *Microsoft Word* document. Both the interviewee and interviewer's speech was transcribed²⁸, indicating details such as overlapping speech.

Pause fillers, such as *um* and *ah*, are always transcribed and details such as coughing, laughter and background noise were transcribed using symbols that the forced alignment system recognises, namely {CG} for cough, {LG} for laughter, {BR} for breath and {NS} for background noise.

The transcription files were then amended in *TextEdit* to remove the interviewer's speech and all overlapping speech to accurately reflect the contents of the edited audio recordings (see section 3.3.2 above for details of this editing), in preparation for forced alignment and automatic vowel extraction.

The full process of AVM, as well as the preparation of the data for AVM, is discussed in the following chapter.

²⁸ Although the present analysis only required transcription of the edited version of the audio files (i.e. only the interviewee's speech), the full interviews were transcribed (i.e. both interviewer and interviewee speech) because these interviews form part of a larger English Social Dialectology Research Project (detailed in footnote 22 on page 58) and complete transcripts are useful in allowing the material to be repurposed).

Chapter Four

Data Analysis using Automatic Vowel

Measurement

This chapter serves to describe and explain the methods used to analyse the data collected²⁹ during the interview process. In Section 4.1, Forced Alignment is discussed, including details of the procedures employed to prepare the interview transcripts and audio files for Forced Alignment. A description of the accuracy with which the process was completed on the data is also provided in this section. The following section (4.2) deals with the process of automatic formant extraction using a program developed for the purpose, `extractFormants`. Details of how an automatic measurement point was selected are provided here, as well as a discussion of the use of the Mahalanobis Distance algorithm as a formant prediction method. Section 4.3 reports on which tokens were excluded from the analysis, and the segmental environments that were isolated in the data prior to analysis. An explanation of how outliers were identified is provided in section 4.4 and normalisation procedures are explained in section 4.5. The final section (4.6) provides details on how the data was manipulated using the statistical package R in order to show the patterns in the data that are reported in the chapters five and six.

4.1. Forced Alignment using P2FA³⁰

Forced Alignment is the first in a two-step process of AVM, and it provides a means of preparing large datasets for analysis relatively quickly by eliminating the cumbersome and time-consuming task of doing alignment manually. Essentially, forced alignment converts orthographic transcriptions into phonemes, before automatically time-aligning words and phonemes to the speech signal in an audio recording. Forced alignment has been used as a tool to investigate various phenomena in recent studies. These include analysis of vowel formants (Konopka and

²⁹ Raw or normalised formant values are available from the author via email, or from rajendmesthrie.com (as part of the NRF Chair Project on English dialectology).

³⁰ P2FA and `extractFormants` are now available as part of a joint program suite called FAVE (Forced Alignment and Vowel Extraction) from <http://fave.ling.upenn.edu> (Rosenfelder et al. 2011). The current repository of FAVE is accessible from <https://github.com/JoFrhwld/FAVE>. Use of FAVE reduces much of the editorial work that I performed on my audio files and orthographic transcriptions. Unfortunately, I only discovered FAVE well after I had completed this work. The current versions of P2FA and `extractFormants` are called FAVE-align and FAVE-extract, respectively.

Pierrehumbert 2008; Yuan and Liberman 2008a; Evanini 2009; Labov et al. 2013), intonation (Anufryk 2008), vowel duration (Tauberer and Evanini 2009), pitch accents (Rosenberg and Hirschberg 2009), and g-dropping and /l/ variation in American English (Yuan and Liberman 2011; Yuan and Liberman 2012). It has also been used in various studies of non-native speech (e.g. Chen et al. 2010; Makino and Aoki 2012; Wu and Shih 2012), and to investigate intonation and tone duration in Mandarin Chinese (Yuan 2012).

The alignment of the data in the present study was performed using the Penn Phonetics Lab Forced Aligner (P2FA), which was developed by Jiahong Yuan at the University of Pennsylvania³¹. The P2FA toolkit contains a set of acoustic models and was trained using 25.5 hours of speech from the SCOTUS corpus (oral arguments from the Supreme Court of the United States) (Yuan and Liberman 2008b: 1). On the said corpus, P2FA produced force-aligned word and phoneme boundaries that were generally as accurate as those produced by human annotators – nearly all disagreements between P2FA word boundaries and human word boundaries differed by less than 50 msec. The authors attribute the exceptional performance of P2FA compared with other acoustic aligners primarily to the fact that it was trained on a large, clean dataset and because of the particular acoustic models used (Yuan and Liberman 2008b: 4). It employs monophone Gaussian Mixture Model based (GMM-based) HMMs (Hidden Markov Models), which were trained using 39 PLP (perceptual linear prediction) coefficients³².

P2FA is one of a handful of automatic alignment tools available. Others include PLA (Prosodylab-Aligner) (Gorman et al. 2011), EasyAlign (Goldman 2011), SPPAS (Speech Phoneticization Alignment and Syllabification) (Bigi and Hirst 2012) and Train&Align (Brognaux et al. 2012). P2FA works on English; EasyAlign is trained to work on French, Spanish, Portuguese and Taiwan Min; and SPPAS is trained to cover French, English, Italian and Chinese (Brognaux et al. 2012). PLA is trained on NAE lab speech, but training data can be input in order to create a new acoustic model (Gorman et al. 2011: 192). Train&Align trains an acoustic model based on the dataset that is input, allowing it to work on any language and style.

³¹ P2FA is available for download as FAVE-align from <https://github.com/JoFrhwld/FAVE>.

³² For details, the reader is referred to Yuan and Liberman (2008b).

Brognaux et al. (2012) compared Train&Align, P2FA and SPPAS, finding that P2FA outperformed the other two aligners as it produced the most accurate alignments. MacKenzie and Turton (2013) compared P2FA to PLA and SPPAS, and also found that P2FA's performance was superior to the others in terms of alignment accuracy on British English data. The decision to use P2FA to perform the forced alignment on this dataset is thus justified. The primary drawback for the choice of P2FA, however, is that it does not have any acoustic models for phonemes that are not part of the General NAE phoneme set, and does not support the training of new acoustic models for different dialects of English (or different languages). Despite this, MacKenzie and Turton (2013) achieved impressively accurate results on British English dialect data. For details about the software requirements and installation for P2FA, see Evanini (2010)³³.

4.1.1 Forced Alignment Requirements

The input requirements for forced alignment are an audio recording and a transcription thereof. The output is a TextGrid file (a Praat format) with two levels of alignment between the audio file and the transcription: viz. word level alignment, and phoneme level alignment. Figure 4.1 is a depiction of the automatically aligned phrase *A third cat joined the party*, displayed using the software package Praat (Boersma and Weenink 2008). The four levels, from top to bottom, are: the oscillogram, the spectrogram, the phoneme level transcription and the word level transcription. The phrase is enclosed on each end by a short pause (sp), which is inserted automatically during the alignment process. The figure clearly shows the alignment of the phonemes with the corresponding speech signal in the spectrogram. The phonemes are provided by a pronouncing dictionary, the details of which I turn to in the following section.

³³ Most of the programs required to run P2FA are released as code. The technicalities of compiling each program's code into a workable program required the technical expertise of a programmer who was familiar with Python programming language. Thanks are thus due to Kabelo Rametse who took on this project as a student programmer, and created an easy-to-use interface from which I could run P2FA and extractFormants.

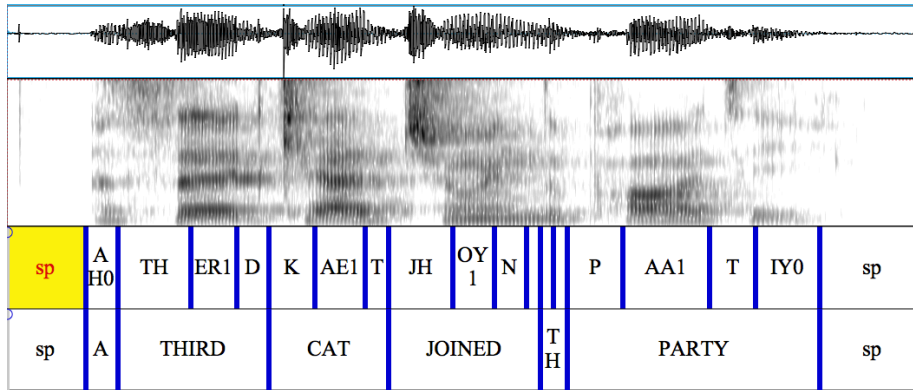


Figure 4.1: Force-aligned depiction of the phrase *A third cat joined the party*

4.1.2 CMU Pronouncing Dictionary

Phonemic transcriptions are taken from the Carnegie Mellon University (CMU) pronouncing dictionary – a pronunciation dictionary for North American English (NAE) that contains over 125,000 words, along with their phonemic transcriptions. The dictionary has 39 phonemes in the current set, and vowel phonemes carry either primary stress, secondary stress, or are unstressed, indicated in the dictionary by 1, 2 or 0 following the phoneme label e.g. UW1 (CMU Pronouncing Dictionary 2013). Below is an excerpt of a few entries in the pronouncing dictionary:

```
LIV L IH1 V
LIVABLE L IH1 V AH0 B AH0 L
LIVE L AY1 V
LIVE L IH1 V
LIVED L AY1 V D
LIVED L IH1 V D
LIVELIER L AY1 V L IY0 ER0
LIVELIEST L AY1 V L IY2 AH0 S T
LIVELIHOOD L AY1 V L IY0 HH UH2 D
LIVELIHOODS L AY1 V L IY0 HH UH2 D Z
LIVELINESS L AY1 V L IY0 N AH0 S
LIVELY L AY1 V L IY0
LIVEN L AY1 V AH0 N
```

In instances in which words have variant pronunciations, all variants are included in the dictionary, as is the case with the word *live* above, whose pronunciation is dependant on its usage as either a verb or an adjective. In such cases, P2FA examines all of the variants and selects one that produces the highest probability match to the acoustic signal (Young et al. 2009).

For P2FA to perform optimally, all words in the transcription must have a corresponding entry in the CMU pronouncing dictionary (Evanini 2010: 7). If there is no dictionary entry for a word present in a transcription, P2FA skips the word. This is problematic because it creates a situation whereby the aligner attempts to match the speech signal with phone labels in the transcript, but the phone labels for that segment

of the audio file are unavailable. Another option would be for P2FA to insert a short pause (sp) or a noise model {NS} in place of the missing word. The aligner would not easily insert a pause when there is a high intensity level in the speech signal (i.e. when there is speech rather than a pause), so the noise model would be a better acoustic match (Evanini 2010: 8).

P2FA provides a warning message at the beginning of the alignment process indicating which words contained in the transcript are not contained in the dictionary. The warning message is displayed as below:

```
SKIPPING WORD COLOURED
SKIPPING WORD BONTEHEUWEL
SKIPPING WORD STURVY
...
```

There were numerous instances of items in the interview transcriptions that did not have entries in the pronouncing dictionary, particularly South Africanisms such as *braai* and *lekker*³⁴, as well as words that were spelt according to UK conventions rather than US conventions, such as *theatre* and *neighbour*. In such cases, the words were added to the dictionary, along with appropriate phonemic transcriptions using the available phoneme set. A further option, in the case of the different spelling conventions (where words had the same pronunciation), was to amend the relevant transcript with the US English spellings. Upon obtaining the warning message described above, I would stop the alignment, make the necessary additions or amendments to the dictionary, and then resubmit the files to P2FA.

Because the CMU pronouncing dictionary is based on NAE, certain amendments had to be made to the dictionary to accommodate features of South African English. The following changes were made to the dictionary before alignment:

³⁴ A 'braai' is a common South African term for a barbeque, and 'lekker' is an adjective meaning 'nice'. Both are Afrikaans loanwords into SAE.

- Post-vocalic /r/ was removed from all relevant environments, as SAE is non-rhotic
- Where SAE has a different vowel quality to NAE, the relevant phoneme was replaced e.g. the original entry for *grass* contained the TRAP vowel, whereas this word, in SAE, typically forms part of the BATH lexical set. The entry was thus amended from <G R AE1 S> to <G R AA1 S>. Such amendments were done while preparing the interview transcripts for alignment (details of these preparations are furnished in section 4.1.3 below). When I encountered a word whose pronunciation deviated from the NAE norm, the word was amended in the CMU pronouncing dictionary.

An additional problem was encountered in that the 39 phonemes available in the dictionary were not sufficient to capture some of the phonemic differences in SAE. This affected two of the phonemes, viz. AA and AH. In the dictionary, AA was used to transcribe words such as *part* (P AA1 R T) and *pot* (P AA1 T). Because SAE is non-rhotic, the /r/ was removed from the transcription of *part*, making it identical to the transcription of *pot*. While this does not present a problem for the forced aligner, the problem arises because SAE is not subject to the low-back vowel merger, as is common in many varieties of US English, so the vowel quality in the words *part* and *pot* is not the same. In SAE, *part* forms part of the BATH lexical set and is a long vowel, while *pot* is a short vowel that forms part of the LOT lexical set. In a study of alignment errors in a corpus of British English (Spoken British National Corpus), researchers added a phone OH to the existing 39 CMU Pronouncing Dictionary phoneme set to accommodate the distinction between BATH and LOT (Baghai-Ravary et al. 2011). In order to include this distinction in my data, I manually amended the output files (from `extractFormants`) for all relevant tokens from AA to OH; thus accommodating the separate categories, albeit after the fact.

In the case of the phoneme AH, which represents the STRUT set, unstressed schwa was transcribed as AH0, so AH1 and AH0 represent separate phonemes, rather than simply indicating a difference in stress. This practice was in place for all occurrences of schwa in the existing dictionary, and I simply continued the pattern when adding new words to the dictionary. Because unstressed vowels are excluded from the analysis however, this does not affect the results reported in this dissertation (see section 4.3 for a discussion of the exclusions.)

There were cases where the pronouncing dictionary contained more than one entry for particular words, such as *get* i.e. <G EH1 T> and <G IH1 T>. The latter pronunciation is not common in SAE, yet when selecting a pronunciation to match the speech signal, the aligner opted for <G IH1 T> over the first option <G EH1 T> when aligning most instances of this word. In such cases, the vowel class was amended in the output files for all speakers.

In the following section, details regarding the preparation of the data to produce optimally accurate results are provided. Following this, the accuracy of the alignment process will be discussed.

4.1.3 Preparation of Data for P2FA

In order to ensure optimal accuracy of the Forced Alignment process, the data had to be prepared before it was aligned. Forced Alignment requires two input files, viz. an audio recording in WAV format (.wav) and a transcription of the recording in plain text format (.txt).

Errors in alignment can be reduced if the audio file exhibits the following qualities (Evanini 2010: 4; Das et al. 2010: 1):

- Contains minimal background noise
- Does not contain overlapping speech
- Recording is of a high quality and the speech signal of the speaker's voice is at an appropriate level throughout the recording
- Contains minimal disfluencies, such as false starts to words e.g. sch- for school
- Contains minimal non-speech sounds such as laughter, coughing, etc.
- Does not contain any other disturbance that detracts from the purity of the speech signal

Besides these factors, which relate to the audio file, various issues with the transcript can also cause alignment errors, such as missing words, wrong words and other discrepancies (Das et al. 2010).

In order to fulfil the requirements listed above as closely as possible, each of the interview recordings was edited in *Wavepad*. The interviewer's speech was deleted, as

well as all background noises, overlapping speech, disfluencies and non-speech sounds. In some cases, this drastically reduced the length of the interview (e.g. speaker M19's interview recording was reduced from 57 minutes to 23 minutes), while others retained the bulk of their length (e.g. speaker F1's interview was reduced from 63 minutes to 48 minutes). The final length of the interview was dependent on the nature of each individual interview: if the interviewee was very talkative and the environment was free of background noise, most of the recording could be retained for analysis. If, on the other hand, the interviewee generally gave shorter answers to questions, the interviewer was required to talk more during the interview, and subsequent deletion of interviewer speech drastically reduced the length of the recording. In some cases there was a lot of background noise, which could not be controlled for, as some interviews were recorded in speakers' homes where, for example, the phone would ring or cars would drive past on a busy road on which the interviewee lived. Such noise was edited out of the affected interviews. The use of the high quality Marantz and Olympus recorders ensured that the recording quality was suitable for use by P2FA.

Once the recordings had been edited, each of the transcriptions was edited so as to reflect the contents of the audio file. Alignment errors can result when the speech signal in the audio file does not correspond with the available phone labels provided by the CMU pronouncing dictionary, so precise transcription is essential for accurate alignment.

Once the audio and transcription files had been edited, they were processed by P2FA. Full interviews, particularly the longer ones which were up to 50 minutes in length, took up to 12 hours to be processed. In order to reduce processing time, the audio files were chunked into four to five minute segments, and the transcripts amended to correspond with the contents of the audio file. The shorter files were then submitted to P2FA for alignment. The processing time was drastically reduced using the shorter segments, with alignments taking between four and six minutes to complete. The Reading Passage and Word Lists were aligned separately from the rest of the interview, and these alignments were completed in a matter of seconds, as the files are very short.

4.1.4 Forced Alignment Accuracy

The benefit of automatic alignment is the relative speed with which large volumes of data can be aligned. Manual alignment approaches unfeasibility on large corpora because of how time-consuming it is: studies report it taking between 130 and 800 times real time (Kawai and Toda 2004; Schiel and Draxler 2003). That said, some degree of inaccuracy is inevitable when employing automatic methods, because acoustic models, no matter how advanced, will struggle to account for all possible variation in a speech signal. The general performance of P2FA is very good, however, with one study by Yuan and Liberman (2008b) reporting that the significant majority of automatically produced word onset boundaries differed from the manual boundaries on the same corpus by less than 50 msec. Evanini (2009: 53-54) manually aligned a small subset of data (two word list recordings) in order to compare the results with that of P2FA. He found that two thirds of the automatically aligned boundaries fell within 20 msec of the manual ones, and all but one fell within 50 msec. The alignment performance of P2FA thus has an excellent success rate on NAE.

To check whether the same could be said of P2FA's performance on SAE, I manually examined the resultant Textgrid files of all 40 interviews in Praat, checking for misalignments. Unfortunately, the results were not as accurate as those reported by Yuan and Liberman (2008b) and Evanini (2009). Where the alignment boundaries were completely inaccurate, i.e. the phoneme boundaries did not overlap at all with the appropriate speech signal, the tokens were considered to be misaligned. One such misalignment is depicted in figure 4.2 below.

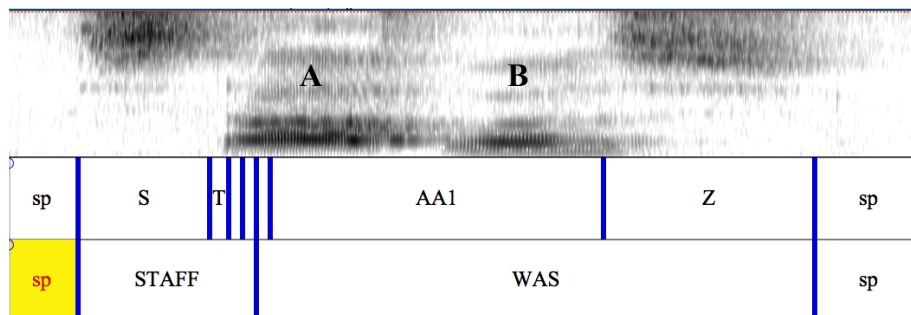


Figure 4.2: Misalignment of the phrase *staff was* (speaker M6)

The figure shows a misalignment between the speech signal and the phoneme boundaries for the phrase *staff was*. In this case, the vowels in *staff* and *was* are aligned as a single vowel, even though it is clear that the speech signal on the

spectrogram that there are two separate vowel sounds (labelled A and B in the figure). This example is typical of the type of misalignments encountered in the TextGrid files.

An average of 5.3 percent of tokens were misaligned in this way, with the highest proportion in a single interview being 20.7 percent (speaker F18) and the lowest being 0.7 percent (speaker M17). The high rate of misalignments for speaker F18 is rather an anomaly, as the next highest percentage was 14.5 percent (speaker F17), and only five of the 40 speakers had a misalignment rate of higher than 10 percent. The aligner worked best on male speakers, with the average misalignment rate of 4.2 percent, while for females the rate was 7.3 percent. This result is not unexpected, as P2FA was trained on the speech of six males and only two females (SCOTUS corpus) (Yuan, personal communication 2013, May 13).

Initially, misaligned phoneme boundaries were manually shifted in order to match the speech signal, but I encountered problems when inputting the edited TextGrid files into `extractFormants`: the automatic extraction would not proceed if the TextGrid had been edited in Praat, only the original output from P2FA would allow successful extraction of formant measurements³⁵. The decision was subsequently taken to remove misaligned tokens from the dataset.

The total number of misaligned tokens was 3,716. Table 4.1 shows the twenty most frequently occurring misaligned words. All of them, except *really*, *everything*, and *just* were excluded from the analysis anyway because they are commonly occurring words, and as such are usually subject to significant reduction (see section 4.3 for details and reasons for exclusions).

³⁵ This problem could no doubt have been resolved with help from `extractFormants` developer, Keelan Evanini, who helped me to resolve many technical problems I experienced with P2FA and `extractFormants`. However, as the affected number of tokens was a small subset of the total number of tokens (5.3%), I elected to remove them from the dataset instead.

	Word	Number		Word	Number
1	and	279	11	we	74
2	the	278	12	in	68
3	was	179	13	at	60
4	to	125	14	it's	58
5	it	119	15	there	58
6	that	116	16	he	57
7	then	111	17	really	56
8	very	98	18	ja ³⁶	54
9	but	95	19	everything	50
10	you	77	20	just	48

Table 4.1: The 20 most frequent misaligned words.

The less than perfect alignment rate of P2FA on this data is conjectured to be the result of two factors: (a) The quality of the speech signal in the recording and (b) The fact that P2FA was trained on North American English. With reference to the first factor, even though the quality of the recording was high because of the good quality recorders employed, the actual placement of the recorder relative to the interviewee's mouth could have affected the quality of the speaker's speech signal in the recording. The recorder was placed on a table or similar item in front of the interviewee. This positioning enabled the recorder to capture both the interviewee and interviewer's speech, as well as any background noises. A more preferable arrangement might have been to attach a lapel mic to the interviewee's clothing, as near as possible to his or her mouth, possibly yielding higher accuracy of alignment. It must be noted that this method was not employed during the interview process because the decision to use AVM over manual measurement, for which the above mentioned method of recording would have been perfectly suitable, was only taken after the interview process had been completed. Nevertheless, an average 94.7% accuracy rate in alignment is reasonably good, and considering that measurements for misaligned tokens were removed from the dataset, the misalignment has not affected the analysis in any way.

Upon investigation into why the rate of misalignment was so high, I discovered that the audio input for P2FA was designed to be mono rather than stereo (Yuan, personal communication, May 2013). Having missed this information in any of the literature which I reviewed on the topic, I proceeded with the alignments using stereo audio input. In order to check whether the stereo signal was the cause of the misalignments, two audio files were re-aligned, inputting only mono signal. The resultant TextGrid files for both interview recordings were identical to when stereo audio was input (as

³⁶ 'Ja' (/jɑ:/) is an Afrikaans loanword meaning 'yes'

were the values for `extractFormants`, which were also checked). It is therefore safe to conclude that the use of stereo signal was not the cause of the misalignments.

4.2 Automatic Extraction of Formant Measurements using `extractFormants`

The second step of AVM, following forced alignment, is to automatically extract formant measurements from the vowels in the TextGrid files. To facilitate this extraction, Keelan Evanini (2009) developed the tool `extractFormants`³⁷, which performs this automatic extraction on phoneme-aligned TextGrid files. As with manual alignment, manually extracting formant measurements is a time-consuming practice, and subject to the natural subjectivity of human annotators performing the analysis. In the analysis of large datasets, this has the potential to be problematic if more than one person is extracting formant measurements. Choosing a stable measurement point for all vowels using automatic extraction provides a solution for both the issue of time constraint, as well as the subjectivity of annotators.

In this section, an explanation is provided for the choice of settings used for automatic extraction of formant measurements on the present dataset. A total of 98,824 tokens of monophthongal vowels were extracted for the 40 speakers, with each measurement consisting of values for F1 and F2, as well as time stamp data, information about vowel duration *inter alia*. This number includes monophthongs with all three levels of lexical stress: primary, secondary and unstressed and also includes the tokens that were subsequently excluded for the various reasons listed in section 4.3. The pre-exclusion per speaker average is 2,471 tokens, ranging from 1,334 to 3,944.

`ExtractFormants` allows the user to make adjustments to various parameters in accordance with the user's desired output. For the most part, the default settings were suitable for the present analysis. Only one parameter was changed, viz. the Formant Prediction Method, which is discussed in section 4.2.2 below. The following section (4.2.1) explains how a point for automatic extraction was selected for the present dataset.

³⁷ The current version of `extractFormants` is the FAVE-extract package available from <https://github.com/JoFrhwld/FAVE>.

4.2.1 Selecting an Automatic Measurement Point

Upon undertaking an acoustic phonetic study, a decision must be taken on how and where to record formant measurements to represent vowel tokens. Although vowels, including monophthongs, are inherently dynamic and not stable throughout their duration, it is accepted as standard practice to represent vowels by a set of formant values extracted at a single measurement point for each token of that vowel. There have been a handful of researchers who have attempted to document the trajectory of vowels by taking numerous measurements for each token (e.g. Watson and Harrington 1999; Nycz and Decker 2005; Baker 2005). Evanini (2009: 59) reports, however, that none of the approaches used by these particular researchers have effectively been used to compare individual vowel tokens from a single speaker, or mean values across speakers. It is more common to select a single point in a vowel's duration to represent the entire vowel. Selecting a single point enables the analyst to disregard the onset and offset of a given vowel, as these may be influenced by neighbouring consonants (although segmental environments in fact affect the entire vowel, not just the onsets and offsets, hence the segmental isolations described in section 4.3). For the purpose of analysis in this study, a single measurement point is selected for each token.

A measurement point must be selected based on one of three criteria: (a) the 'central tendency' of the vocalic nucleus (Labov et al. 2006: 36), (b) steady state detection or (c) a temporal dimension. When measurement is based on central tendency, measurements are taken at a point in the vowel where F1 – or for certain vowels, F2³⁸ – is measured at the point where it changes direction i.e. its maximum value. For example, for most short vowels and long, upgliding vowels, the central tendency is, simply put, a fall and rise of the tongue, which is reflected acoustically as a rise and fall in F1. The highest F1 value thus represents the lowest point that the tongue reaches in the production of the vowel. For these vowels, the formant measurements for F1 and F2 are extracted where F1 is at its maximum (Labov et al. 2006: 38). This approach attempts to record formant values that are most representative of each vowel class.

Secondly, there are methods which select a measurement point based on where formants are in a steady state within the vowel. One such method was devised by

³⁸ The reader is referred to the ANAE (Labov et al. 2006: 36ff) for details of the central tendency for different vowel classes.

Lennig (1978), who calculated a Coefficient of Change at each formant measurement³⁹ and measured the vowel where the difference between formants (in relation to the one preceding and the one following) was smallest. This method is most suitable for monophthongal vowels, as diphthongs are inherently characterised by dynamic formant contours.

A more commonly used method, and one more compatible with AVM, is to select a measurement point based on a temporal dimension, i.e. a point occurring at a percentage of the duration of the vowel. Such methods often try to approximate the steady state methods, with the assumption that a steady state occurs at a specific point within the duration of most vowels (diphthongs excluded). The majority of studies employing this method choose the midpoint for all vowel categories (e.g. Pierrehumbert et al. 2004 and Chen et al. 2009).

Evanini (2009: 60-66) compared five different methods of selecting an automatic measurement point in order to ascertain which method provided measurements closest to manual measurements taken on a given corpus, in this case, the Atlas of North American English (henceforth ANAE). Three of the methods were time-based i.e. measurement at one third, one quarter and one half of the vowel's duration. The other two were based on the formant values of vowels: Lennig's (1978) method and the 'central tendency' method used in the ANAE (Labov et al. 2006). Evanini's automation of the Lennig method measures the vowel at the point where the coefficient of change is smallest (see Evanini (2009: 62) for the equation and details about Lennig's method). The automated version of the ANAE method extracted measurements for tokens of the TRAP set where F2 was at its maximum and the NORTH/FORCE set where F2 was at its minimum. For all other vowels, measurements were taken at the F1 maximum.

Evanini found that the method of measurement which produced formant values closest to the manual measurements was the one that takes the measurement one third of the way through the duration of the vowel. The mean difference between the manual and automatic measurements for F1 was 10.4 percent, and 12.5 percent for F2 (Evanini 2009: 65).

³⁹ The reader is referred to Lennig (1978) for details of the equation used to calculate the Coefficient of Change.

Another consideration is that the measurement point should take into account consonantal transitions i.e. not measure the onset or the offset of the vowel. This is potentially problematic when taking measurements automatically at a designated percentage of the vowel's duration, particularly if the vowel is very short. For this reason, `extractFormants` is configured to extract vowels only if their duration is 50 msec or more (based on the output of forced alignment).

Measurement at one third of the duration of the vowel was the method employed for the data in this study, based on Evanini's (2009: 65) findings that it the best method across all vowel categories.

4.2.2 Formant Prediction

Manual formant analysis relies on the judgement of the annotator to ensure that the formant tracker in the acoustic software package, such as Praat, has correctly tracked the formants for a given vowel token. Praat makes use of LPC (Linear Predictive Coding) analysis, which became the standard tool for estimating formant frequencies after two influential papers advocating LPC were published in the 1970s (Atal and Hanauer 1971; Markel and Gray 1976).

When extracting formant measurements, the annotator can view the formants predicted by LPC (the red dots on the spectrogram in figures 4.3 and 4.4 below), as well as the actual formants (represented by dark bands on the spectrogram). In addition, the annotator can listen to the sound file of the particular segment to perform auditory confirmation of the predicted formants. Figure 4.3 below depicts the phrase '*high school was*', uttered by speaker M6. It shows that the formant track produced by Praat's LPC analysis is correct for the words *high* and *was*, but incorrect for the long GOOSE vowel in *school*. A human annotator would modify the settings of the formant tracker before taking a measurement for this vowel. In this case, it requires increasing the number of formants that the tracker searches for from five (5) to six (6). Figure 4.4 shows the formant tracks after this modification, where the LPC analysis has been corrected for GOOSE. The annotator makes such adjustments on a per token basis, based on visual and auditory analysis of each token. These adjustments are frequently necessary, especially for vowels whose F1 and F2 values are close together, such as non-high back vowels (Evanini et al. 2009: 1655). In the ANAE, at least 10 percent of the 125,000+ vowel tokens in the corpus required modification of the number of

formants in the LPC analysis in order to get accurate F1 and F2 measurements (Evanini 2009: 80).

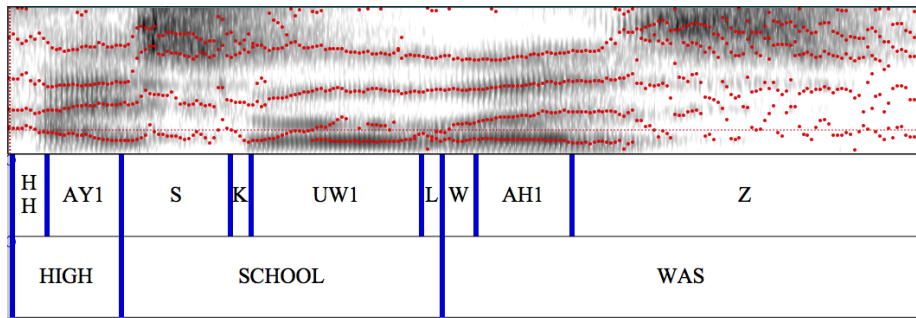


Figure 4.3: LPC formant tracker: number of formants set to 5 (default value).

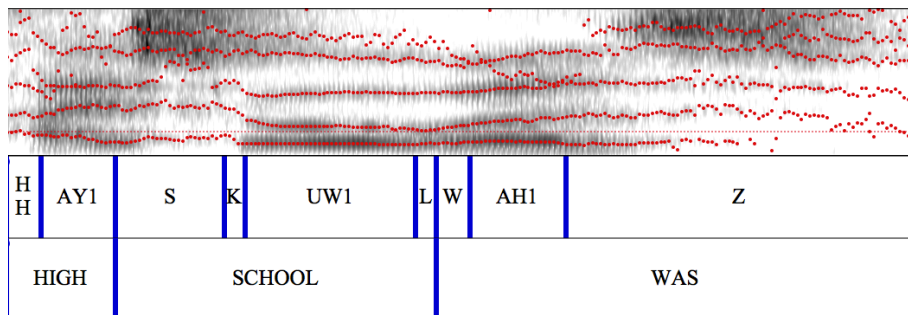


Figure 4.4: LPC formant tracker: number of formants set to 6.

Human annotators thus use their prior knowledge of the formant distribution of a given vowel to determine the accuracy of the LPC analysis. In order to model this in the automatic extraction process, Evanini (2009:80-84) trained a model of formant and bandwidth combinations for each vowel to simulate the procedure in the automatic extraction of formant measurements. The procedure was trained using manual F1 and F2 measurements from the ANAE corpus, using the Mahalanobis Distance equation⁴⁰ to determine which formant measurements are closest to the actual formant peaks in the speech signal.

The default configuration of `extractFormants` measures formants using a standard LPC analysis. The Formant Prediction Method parameter was thus adjusted so that the vowel formants would be predicted using the Mahalanobis Distance algorithm.

⁴⁰ The reader is referred to Evanini (2009) for a detailed explanation of the algorithm used to determine which formant measurements are closest to the actual formant peaks.

4.3 Exclusions and Isolated Environments

Of the total of 98,824 token measurements that were taken by extractFormants, the following tokens were excluded from the analysis:

- Vowels with secondary stress and unstressed vowels (indicated in the phonemic transcription in the CMU pronouncing dictionary).
- Vowels in high frequency function words, most commonly conjunctions and prepositions, which are often subject to phonetic reduction in connected speech. The list includes, but is not limited to, the following words: *but, for, he, she, is, it, its, it's, the, a, an, um, was, you, me, if, in, do, from, of, that, them, then, this, uh, at, as, am, did.*

The reason for the first set of exclusions is to mirror the formant extraction practices of human annotators, who would only extract measurements in syllables with primary stress, as vowels in unstressed syllables and those with secondary stress are usually reduced. Vowels in the frequently occurring function words are also often reduced, so these words were excluded from the dataset.

In addition to the above exclusions, tokens occurring in the following phonological environments were also excluded from the analysis because of known co-articulatory effects of these environments (Labov et al. 2006: 77):

- Vowels before /l/ and /r/
- Vowels after /w/ and /j/
- Vowels after obstruent-liquid onset clusters
- Vowels before nasals
- Vowels before /k/ and /g/
- Vowels in word-initial position

The following phonological environments were isolated in the GOOSE set (following Labov et al. (2006) and Mesthrie (2010)):

- After coronal consonants /t, d, n, l, r, s, z, j, ʃ, ʒ, tʃ, dʒ/
- After non-coronal consonants /k, g, h, m, p, b, f, v/

The KIT set was isolated as follows (following Lanham and Macdonald (1975); Lass (2002)):

- Before and after /k/
- Before and after /g/
- After /h/
- Before /ng/
- Word initially

These particular segmental environments have been found to produce significant effects on the vowels. An elaboration on the effects of the environments is provided in chapters five and six when the actual data are presented.

4.4 Outliers

Despite methods employed to ensure the accuracy of the automatic vowel measurement process, errors in measurement are inevitable. Outliers were thus removed from the dataset by excluding five percent of tokens that were furthest from the mean value for the vowel class, using the Euclidean distance algorithm (Thomas 2011: 159). The formula for Euclidean distance is as follows:

$$\sqrt{(F1_a - F1_b)^2 + (F2_a - F2_b)^2}$$

where $F1_a$ is F1 of point a, $F1_b$ is F1 of point b, $F2_a$ is F2 of point a, and $F2_b$, is F2 of point b. Point a is the mean value and point b the measurement of the individual token (Thomas 2011: 313).

4.5 Normalisation using NORM

Vowel formant normalisation is necessary to enable accurate cross-speaker comparisons. Individual speakers' vocal tracts differ in size and dimension, with the result that formants of phonologically identical vowels occur at different frequencies compared with other speakers. Flynn (2011: 2) presents a list of all the goals of normalisation in the literature on the topic:

1. to minimise or eliminate inter-speaker variation due to inherent physiological or anatomical differences;
2. to preserve in-speaker variation due to social category differences, including age, gender and dialect, or due to sound change;

3. to maintain vowel category and phonemic differences;
4. to model the cognitive processes that allow human listeners to normalise vowels uttered by different speakers

While no method of normalisation fulfils all the above criteria perfectly, Adank et al. (2004: 3105-3106) found that Nearey's (1977) individual logmean formulation (along with Lobanov's (1971) method) performs best in terms of preserving phonemic variation, effectively reducing physiological variation, and preserving nearly all of the sociolinguistic variation in the acoustic measurements. Both Nearey and Lobanov's methods are vowel extrinsic and formant intrinsic.

The formant values were thus normalised according to the individual logmean algorithm in Nearey (1977). Normalisation was performed using *NORM, the normalization suite* (Thomas and Kendall 2007) – a website that makes use of the statistical software package R to normalise data according to one of several normalisation methods available. The values were scaled to produce Hertz-like values using a function made available in NORM⁴¹. The following formulae are used to scale F1 and F2 values respectively:

$$F1 = 250 + 500(F_1^N - F_{1MIN}^N) / (F_{1MAX}^N - F_{1MIN}^N)$$

$$F2 = 850 + 1400(F_2^N - F_{2MIN}^N) / (F_{2MAX}^N - F_{2MIN}^N)$$

Where F_i^N is a normalised value for formant i and F_{1MIN}^N and F_{1MAX}^N are the minimum and maximum normalised values for formant i (Thomas and Kendall 2007).

4.6 Statistical analysis, Data Manipulation and Presentation in R

R is a free, collaboratively designed software package (R core team 2012), based on the programming language S. R's broad range of functions allows for data to be analysed statistically, sorted and ordered, and presented graphically *inter alia*. All statistical analysis for this dissertation was performed in R, and all graphs generated in R. Extensive use was made of Baayen's (2008) book, *Analyzing Linguistic Data: A*

⁴¹ NORM authors Erik Thomas and Tyler Kendall warn that scaling using NORM is not recommended unless all speakers are submitted to NORM at the same time. If speakers are submitted individually, they suggest that researchers use the formulae provided to scale the results after all speakers' vowels have been normalised (Thomas and Kendall 2007). Because all speakers' formant values were submitted to NORM at the same time, I made use of the scaling function.

Practical Introduction to Statistics using R, which is a detailed and methodical manual for using various functions offered in R.

A linear mixed-effects model (using the lme4 package in R (Bates 2005)) was used to produce the p-values which indicate the significance of the differences between speaker groups and speech styles in chapters five and six.

4.6.1 Graphical representation of F1 and F2

Extracting a pair of formant measurements at a single point in the duration of a vowel token to represent the entire vowel is an accepted simplification in standard sociophonetic practice (see section 4.2.1). Plotting these measurements on a two-dimensional vowel plot provides a representation of the vowel space in the mouth: vowel height and advancement (frontness / backness) are in direct relationship with the first and second formants, respectively. Height is inversely proportional to the value of F1, so high vowels have low F1 values. Advancement is directly proportional to F2, so fronter vowels have higher F2 values (Thomas 2011: 145). F1 is presented on the y-axis, and F2 on the x-axis. All scatterplots presented in the following chapters have been generated along these principles.

4.7 Conclusion

In this chapter the methods of data analysis using AVM were explained, beginning with Forced alignment using P2FA, followed by a discussion about the automatic extraction of formant measurements from phoneme-aligned TextGrid files. Following this, details were provided about which data were excluded from the analysis (including outliers), and how the data were normalised using NORM and processed in R. In Chapters five and six, the results of the analysis are presented for the short vowels and long monophthongs respectively.

Chapter Five

Short Vowels

5.1 Introduction

In this chapter, I present the findings of the acoustic analysis with regards to the six short vowels under investigation. The vowels, labelled according to Wells' (1982) lexical sets, are KIT, DRESS, TRAP, LOT, STRUT and FOOT. The acoustic findings are presented alongside the pertinent literature which discusses the trends of each of the vowels for CSAE, as well as trends in WSAE, a variety with which some of the speakers in the sample – viz. middle-class speakers – have had significant contact due to their mixed social networks in model C and private schools, and for some, in the former White neighbourhoods into which they moved with their families. There is also an important historical reason for using WSAE as a variety against which to measure the changes occurring within CSAE: it was the ancestors of WSAE speakers who introduced English to southern Africa, so these speakers can in some ways be regarded as the 'original' SAE speakers.

The chapter begins with a discussion of the SAE chain shift in section 5.2, which affects the front vowels KIT, DRESS and TRAP. Section 5.3 provides a brief acoustic overview of all the short vowels covered in this chapter, discussing their position relative to one another. In the sections that follow, 5.4 to 5.9, an analysis of each vowel is presented. Each section is comprised of an outline of the trends for the relevant lexical set in WSAE based on impressionistic reports by Lass (2002) and an acoustic account of citation-style WSAE by Bekker (2009). This is followed by a summary of the impressionistic literature on CSAE for that vowel (based on Finn (2004) and Wood (1987)). Graphical representations of the acoustic data from the present study are then presented in order to elucidate trends in the data with reference to class and gender, as well as speech style. The three speech styles used in the interviews are: Reading Passage style (RP), Word List style (WL) and Interview style (IS). Phonological conditioning is also indicated, where relevant.

Bekker's (2009) study, which is used as an acoustic point of comparison for CSAE with respect to trends in WSAE, is based on 27 White, female speakers who at the time of being recorded were between 18 and 19 years old. The speakers were recorded reading various Word Lists, so all his data are in citation-form. While the

data are provided to show how CSAE may be changing in relation to WSAE, it must be noted that Bekker's (2009) data and my own are not strictly comparable because of the stylistic difference between the datasets. Citation style is commonly acknowledged to be least like vernacular speech (Labov 1972a), so it would be more accurate to compare Bekker's data to the WL style data in this study, but since the present study is focussed on the analysis of more naturalistic data (i.e. the speech style labelled IS), I do not provide this comparison.

A general note on the graphical representation of data in this chapter and the following one: Data are either presented in a scatterplot, or a boxplot (also known as a box-and-whisker plot). The scatterplots are disaggregated by class and gender, and generally show a mean value for the tokens for each group on the F1-F2 plane (except where otherwise indicated). The boxplots are disaggregated by formant, class and gender. The KIT data are also broken down into different environments to show how the split in this set is phonologically conditioned. Boxplots representing information about the first formant (F1) are presented vertically, with an inverted axis so that they can be read in the same way as the scatterplots i.e. a higher box indicates a higher range of values for the particular vowel. Similarly, boxplots representing the second formant (F2) are presented horizontally, also with an inverted axis, so that vowel advancement can be judged in the same manner as one would when reading a scatterplot i.e. boxes that are further to the right indicate a backer range of values, and vice versa.

Boxplots provide a lot of information about the spread of data. A labelled sample boxplot for F1 is provided (figure 5.1) in order to explain how the data are represented. Point A represents the minimum value in the dataset. The 'whisker' from point A to B represents the lower quartile i.e. 25 percent of the data are less than the value at point B. Point C represents the median, so 50 percent of the tokens are greater than this value, while the mean is represented with an asterisk (point D). The upper quartile is at point E, where 25 percent of the data are greater than this value and point F represents the maximum value (excluding outliers). Outliers, which are tokens outside 1.5 times the range above the upper and below the lower quartile, are indicated by circles beyond the 'whiskers' (e.g. point G). These outliers are those points which remain after the exclusion of extreme outliers using the process described in section 4.4.

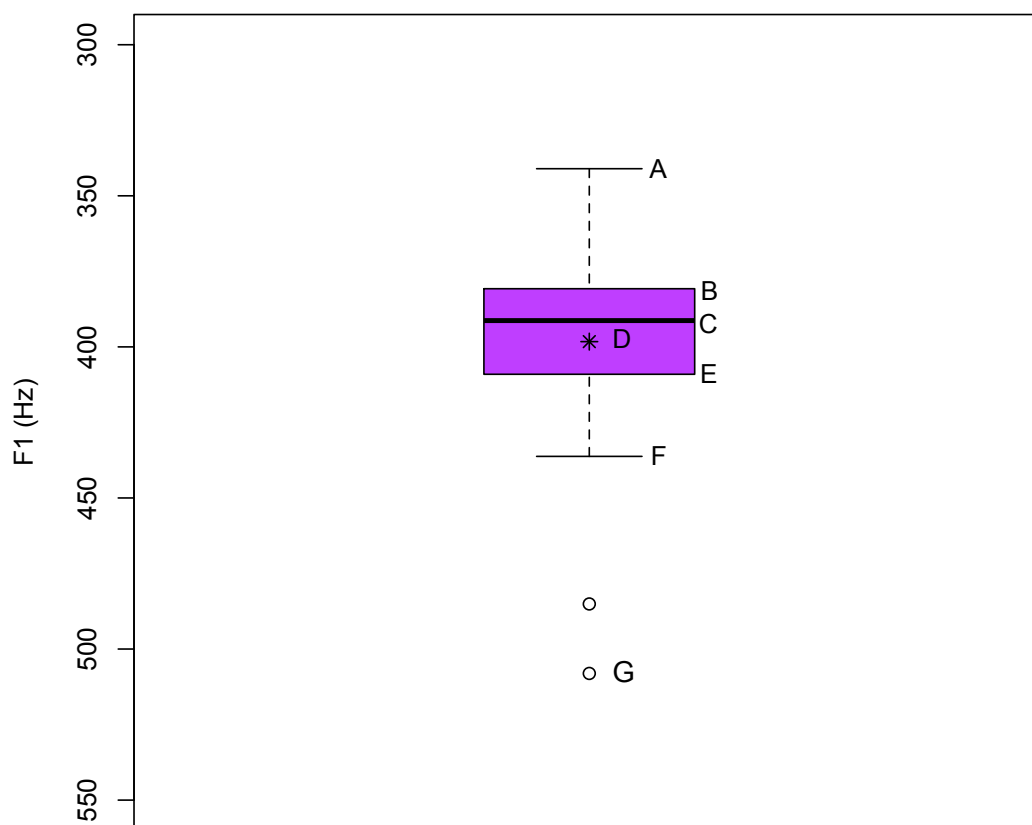


Figure 5.1: Sample boxplot (box-and-whisker plot) for F1 data.

Statistical tests were run on all the data (and subsets of data, as is appropriate for each figure) using a linear mixed-effects model (Bates 2005) and the results reported in conjunction with a visual analysis of the graphs. All data presented in the figures, including Bekker’s (2009) formant values, have been normalised according to Nearey’s (1977) logmean algorithm (see section 4.5 for more details), and scaled to Hertz-like values. The axes are labelled as though the unit of measurement is in fact Hertz (Hz), as in figure 5.1, but the reader is assured that all data presented have been normalised.

5.2 SAE Chain Shift

The short front vowels of SAE, viz. KIT, DRESS and TRAP, which are discussed in subsections 5.4 to 5.6 of this chapter, have been subject to a chain shift, reported to have begun in the 1800s (Lass 2002: 113). Southern extraterritorial Englishes (ETEs) (of which SAE is one – a description is provided in Chapter 2.2) tend to have higher vowels in TRAP and DRESS than other varieties of English e.g. Received Pronunciation (Lass 2002: 113). These raised variants, along with centralised KIT, seem to provide evidence of a SAE chain shift which began in the nineteenth century, after the arrival of the British settlers in the region. Subsequent endogenous developments resulted in

a unique variety – SAE – amidst similarities with Australian and New Zealand Englishes.

Generally, the pattern displayed in the SAE chain shift involves the raising of the short front vowels so that TRAP, raising from [æ] towards [ɛ], impinges on DRESS, which was realised in the region of [ɛ]. DRESS in turn raised to [e] so impinged upon the vowel space of KIT. The majority of the KIT set centralised towards [i] as a result. Some words in the KIT set are not centralised, however, hence the phonologically conditioned split in this set, which is discussed in section 5.4.

Mesthrie (1993: 30) provides a schematisation of the shift (figure 5.2), where words in lowercase represent the original space occupied by each set on a traditional vowel chart, and words in UPPERCASE represent the space to which the set has shifted.



Figure 5.2: Short front vowel raising in WSAE
Source: Mesthrie (1993: 30)

The shift seems to have arisen out of a complex situation in which the categories in question had variable realisations in the input varieties for SAE. While the history is complex and controversial (see Lass and Wright (1985) for a discussion), the developments are likely to have played out as described in Lass (2002: 114), based on evidence from the chronicle of one of the original 1820 Settlers, Jeremiah Goldswain, who was from Buckinghamshire. His non-standard spelling provides insight into the quality of the vowels used by the Settlers. Some of the characteristics of the 1820 input, based on Goldswain's memoirs, include raised TRAP e.g. *contractor* 'contractor', lowered DRESS e.g. *hadge* 'hedge', raised DRESS e.g. *kittle* 'kettle', lowered KIT e.g. *presner* 'prisoner' and retracted KIT e.g. *buld* 'build'. So there seems to have been settler variability, resulting in some overlap between raised TRAP and

lowered DRESS, and raised DRESS and lowered KIT. It seems that over time, the overlap was eliminated by each of the four categories spacing themselves out through raising – in effect, a push chain – and the resultant centralisation of KIT.

Recently, it has been suggested by Bekker (2009) that a reverse shift may be in evidence, based on the acoustic data he provides for a lowered TRAP amongst the young, female WSAE speakers in his study. Mesthrie (2012c) confirms that TRAP is indeed lowering in SAE, and suggests that the same trend is in place for the DRESS and KIT sets. A comparable chain shift, known as the South-East England short vowel chain shift (henceforth SECS), is in place in South-East England (Torgersen and Kerswill 2004). In this chapter, acoustic data are provided to show the extent to which this trend is evident within CSAE.

5.3 Acoustic Overview of the Short Vowels of CSAE

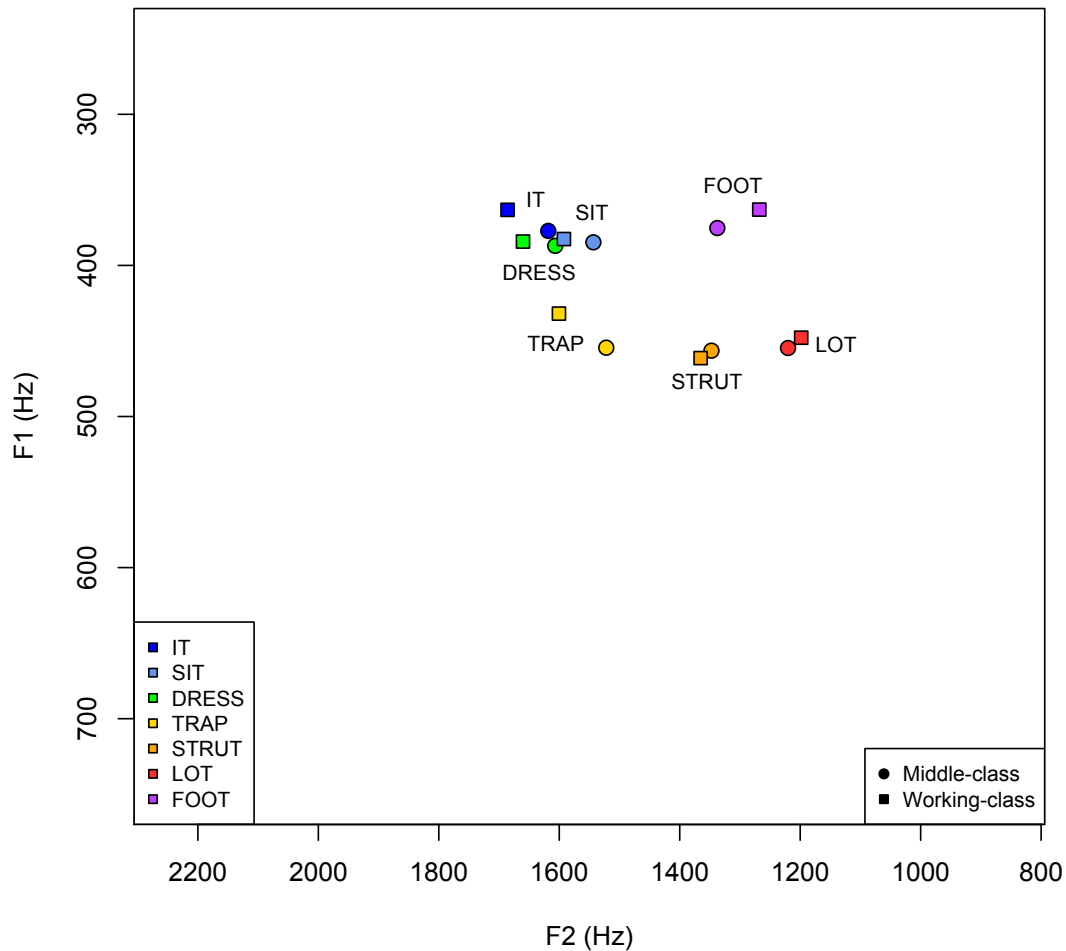


Figure 5.3: An acoustic overview of the short vowel system of CSAE, via mean values for middle- and working-class speakers.

Figure 5.3 provides an overview of the acoustic data for CSAE speakers in this study. It shows two mean values for each vowel class: one for middle-class speakers, the other for working-class speakers. The figure is provided primarily as a means to situate the different vowels in relation to one another in the vowel space. While the range of normalised tokens in the dataset extends from 250Hz to 750Hz on the F1 plane, and 850Hz to 2250Hz on the F2 plane (hence the range of the respective axes), the mean values are situated between 300Hz and 500Hz on the F1 plane, and 1100Hz and 1800Hz on the F2 plane. On the graph, the KIT set is divided into two subsets: IT and SIT, representing the split in the set, which is discussed in section 5.4 below.

Figure 5.3 shows that the mean values for the DRESS and KIT sets are very close together, particularly for middle-class speakers, suggesting that there is some overlap between these lexical sets. This provides evidence of the reverse vowel shift occurring in CSAE. With respect to the other short vowels in the system, it seems, based on the

spread of the mean values, that each occupies a unique position within the vowel space.

Some class variation is also evident: the TRAP and FOOT vowels show particularly clear distinction between middle- and working-class speakers, and to a lesser extent the DRESS vowel and both subsets of the KIT vowel. Class variation will be discussed in detail in the subsections for each vowel, as will gender and stylistic variation.

The sections that follow present, in detail, the acoustic data for each lexical set. Beginning the analysis for each vowel is an overview of the literature for WSAE, followed by a summary of the traditional aural accounts of CSAE. These reports provide a reference point for interpreting the acoustic data that follow. The acoustic account of CSAE discusses, for each vowel, the position of the set on a scatterplot in relation to Bekker's (2009) WSAE female speakers. It also shows the placement of tokens (represented by a mean value) before tautosyllabic velar /l/, which causes vowel retraction in many of the lexical sets. Unique words, which are reported to diverge from the general trend for the rest of the set, are analysed separately to test whether there is acoustic evidence to corroborate the reports. Following this, the data are disaggregated by class, gender and speech style using boxplots in order to show patterns in the distribution of the data.

5.4 KIT

5.4.1 WSAE Reports

KIT is a defining characteristic of SAE, as it is the only variety of English worldwide documented to have split in this vowel class (Lass 2002: 114). The split is phonologically conditioned: a high front variant in the region of [ɪ] occurs in the IT-subset – word-initially, after [h], preceding or following a velar consonant i.e. [k, g, ŋ] and preceding palatal consonants; and a centralised variant i.e. [i̠] occurs in all other environments, labelled the SIT-subset⁴². KIT is used to refer jointly to both subsets.

Across the (WSAE) lectal hierarchy, this vowel is not consistently distributed. Table 5.1 shows the distribution. Conservative WSAE speakers show little or no allophonic variation in this set, consistently using higher and fronter [ɪ]. General WSAE speakers use a centralised variant in the SIT-subset, while Broad speakers use a variant which is further raised and fronted (in the region of high, front [i]) for the IT-subset.

WSAE lect	IT-subset	SIT-subset
Conservative SAE	[ɪ]	[ɪ]
General SAE	[ɪ]	[i̠]
Broad SAE	[i]	[i̠]

Table 5.1: Distribution of the allophones of KIT across the three lects of WSAE (Lass 2002: 114-115).

In some lects of WSAE, there is retraction before /l/ and after /w/ (particularly for Broad speakers), which draws some of the KIT set into the FOOT vowel space. This produces a near merger so that pairs of words like *will/wool*, *bill/bull* become homophonic (Lass 2002: 115). Broad speakers can also display this retraction before and after /f/ so that *fit/foot* become nearly homophonic (Lass 2002: 115).

With respect to this set, Bekker (forthcoming) argues there is a lack of polarisation between the high, front and generally unmarked allophones of General SAE (citation-form) KIT and that the set is better deconstructed into six allophones which he isolates in his data (Bekker 2009: 266ff). He labels the variants KIT1 – KIT6, and their environments are listed and exemplified in table 5.2 below:

⁴² ‘IT’ has been chosen as a label for this category of KIT words to represent the higher, fronted variant within the KIT-split, while SIT refers to the unmarked tokens in the set which tend to be centralised. This follows the labeling used by Finn (2004: 968) for the subsets for this set.

Variant	Environment	Example
KIT1	disyllabic words	<i>city, silly</i>
KIT2	/ʔ_ ; h_ ; velar_ ; _velar/	<i>sing, hid, kit, it</i>
KIT3	/_palato-alveolar/	<i>dish, ditch</i>
KIT4	unmarked	<i>chin, sit</i>
KIT5	/l_ ; r_ /, near bilabials	<i>rid, lit, bit, limp</i>
KIT6	/w_ ; _l/	<i>till, with, fill, pill</i>

Table 5.2: Six allophones of KIT (Bekker 2009: 267).

His analysis shows that KIT4 and KIT5 overlap quite significantly. KIT2 and KIT3 also overlap, although KIT2 partially overlaps with DRESS, so is fronter than KIT3, which is more centralised. KIT6 is the backest variant and partially overlaps with FOOT (see figure 5.4).

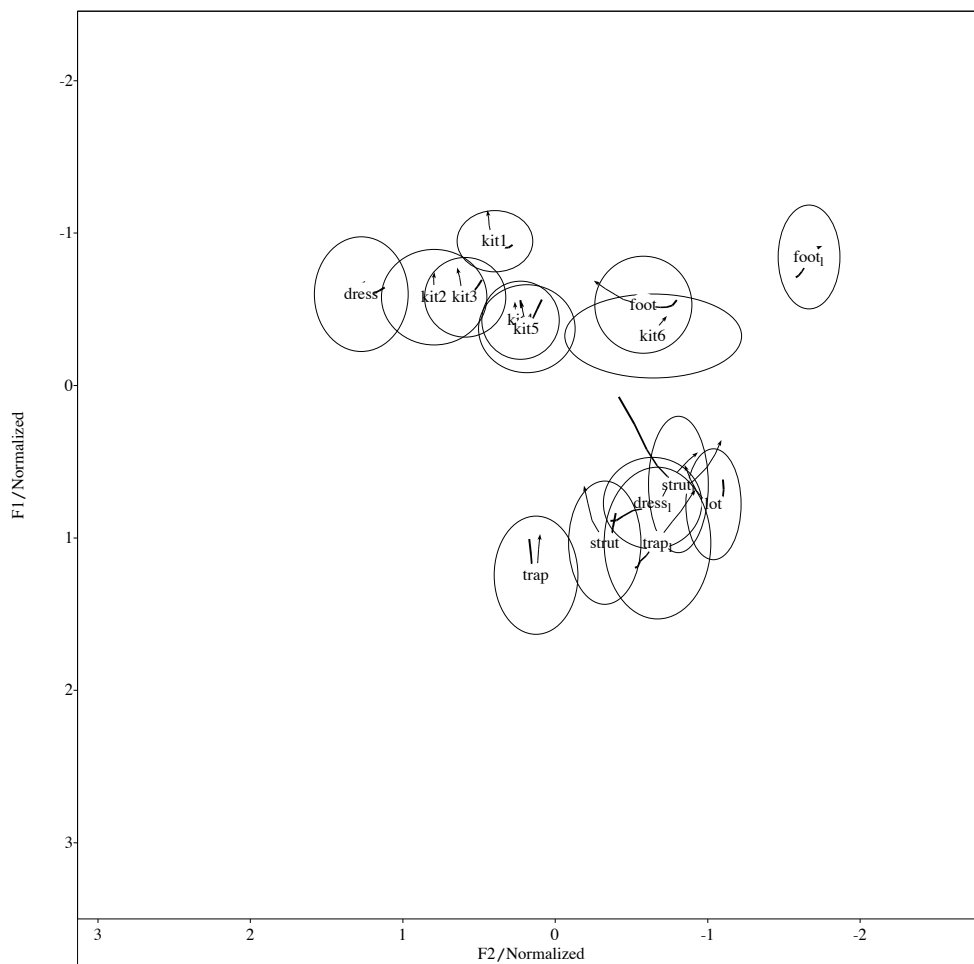


Figure 5.4: Range of variation for short vowels of WSAE in WL style.

Source: Bekker (2009: 268).

Key: Ellipses indicate one standard deviation.

The comparable divisions in the data in this study include most of the allophonic distinctions Bekker (2009) has isolated (discussed in section 5.4 below). KIT5 does not appear in my data, however, because of the phonological exclusions placed on the data before comparative analysis (see section 4.3 for details). KIT2 and KIT3 are represented in my data by three categories: preceding velar consonants, following velar consonants and word-initially. KIT4 is represented in my data as unmarked SIT tokens. KIT6 is comparable with pre-velar-/l/ tokens in the present study, except that it does not include tokens after /w/ as these were also excluded from the analysis.

5.4.2 CSAE Reports

CSAE, like all varieties of SAE, displays a split in the KIT set (Lass 1995: 97; Wood 1987:122-123). The *IT* subset is realised as [ɪ] or [i] word-initially, after /h/, following and preceding velar consonants and before palatal consonants. The *SIT* subset is realised as centralised [ɪ̯], which can be lowered to [ə], in all other environments. While the schwa realisation is restricted to Broad speakers of WSAE (Lass 1995: 97), for CSAE speakers it also occurs amongst more General speakers (Wood 1987: 111). Wood (1987: 122) also notes that use of a raised [i] in the *IT*-subset is a stigmatised feature, used exclusively by Broad CSAE speakers. Before /l/, KIT is typically retracted to [ɣ] (Finn 2004: 970).

5.4.3 CSAE Acoustic Data

a) Overview

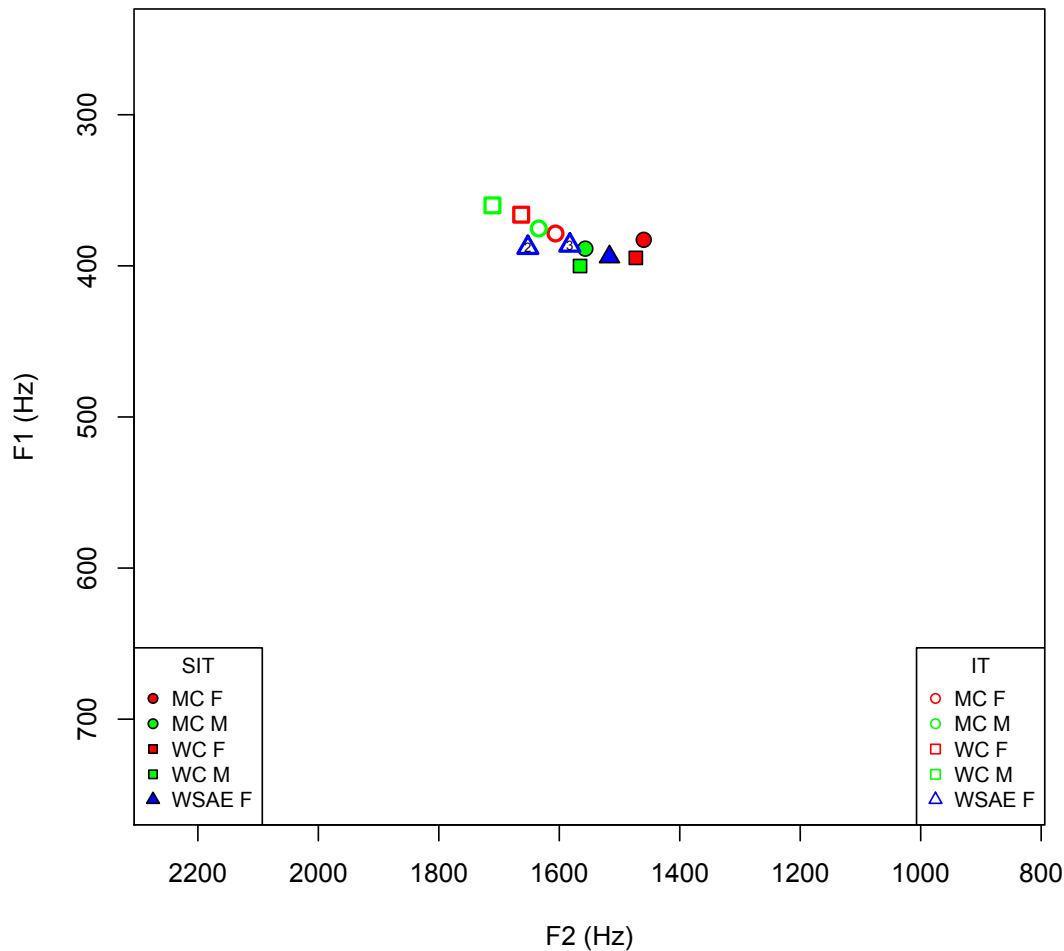


Figure 5.5: Mean values for IT~SIT in Interview Style (IS), disaggregated by class and gender.

The acoustic data for the CSAE speakers in this sample show evidence of the KIT split. In figure 5.5, the IT and SIT subsets are disaggregated by class and gender. One point is plotted for each speaker group, for each of the subsets. A mean value for Bekker's (2009) KIT2 (preceding and following velar environments), KIT3 (preceding alveolar consonants) and KIT4 (unmarked i.e. SIT) tokens are provided as a comparison. The figure shows that the combined IT tokens (i.e. those preceding all palato-alveolar consonants (viz. /ʃ, ʒ, tʃ, dʒ/), tokens preceding and following velar consonants (viz. /k, g, ŋ/), tokens in which the vowel is word-initial, and those following /h/) are realised fronter and higher than those in the remaining phonetic environments for all speaker groups. (The data are deconstructed into the different conditioning environments in the following section.)

The graph (figure 5.5) also shows that female speakers use the most centralised SIT vowel, more so than WSAE speakers in Bekker's (2009) study; male speakers use a

more fronted variant. Working-class male speakers use the highest, frontest values for the IT-subset confirming Wood's (1987) suggestion that working-class speakers tend to use a variant approximating [i] for this set. Wood's (1987) claim that CSAE speakers use a lower variant for the SIT-subset, approximating schwa, compared with WSAE speakers, are not strongly supported by the acoustic data.

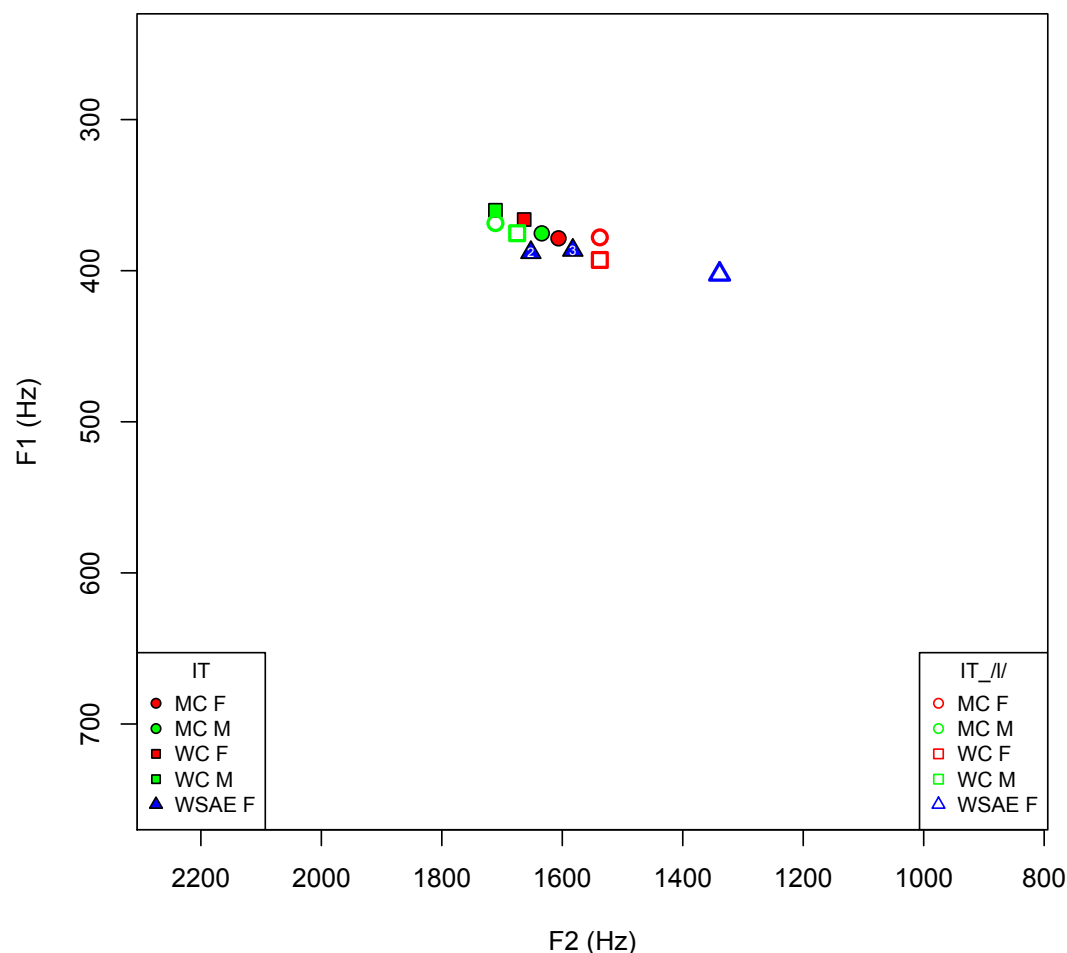


Figure 5.6: Mean values for IT and IT_/l/ in Interview Style (IS), disaggregated by class and gender.

Figure 5.6 shows that IT tokens before velar /l/ are not retracted for CSAE speakers⁴³. Males show some degree of retraction, but it is not nearly as far back as Bekker's (2009) KIT6 tokens (which are made up not only of pre-/l/ tokens, but also post-/w/ tokens).

⁴³ Acoustic readings before /l/ are known to have coarticulatory effects, which is the reason for their exclusion from the primary analysis (see section 4.3 for details of exclusions).

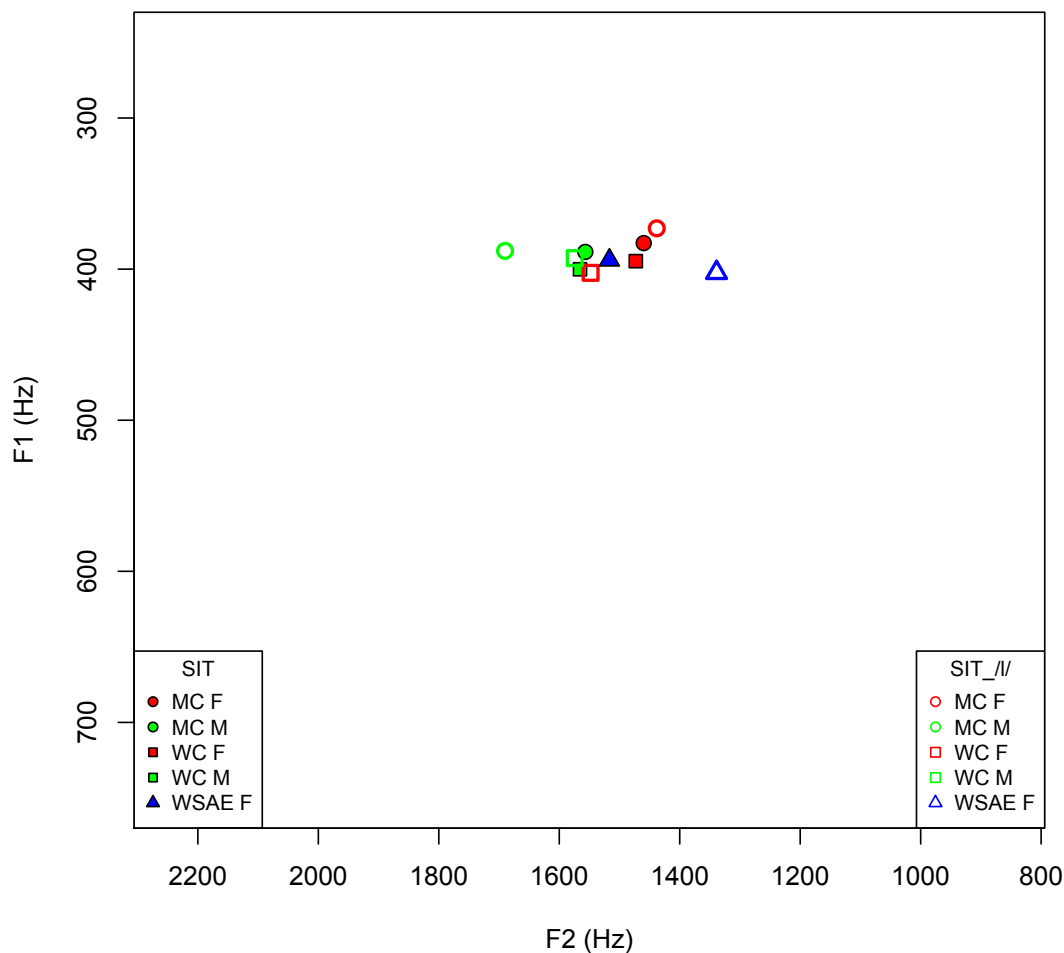


Figure 5.7: Mean values for SIT and SIT_/l/ in Interview Style (IS), disaggregated by class and gender.

For the SIT subset (figure 5.7), there is also no clear retraction of tokens before velar /l/. Therefore based on figures 5.6 and 5.7, there is no acoustic evidence to support Finn’s (2004) claim that CSAE speakers retract the KIT set before velar /l/.

In order to test whether the IT-subset is longer than the SIT-subset, a very small selection of tokens (n=16) were manually measured from four speakers – one male and one female from each of the two social class categories⁴⁴. The results are presented in table 5.3 below:

⁴⁴ Manual measurements were taken rather than using the duration measurements from the Forced Alignment output in order to ensure the accuracy of the readings.

Speaker	SIT ₁	SIT ₂	IT ₁	IT ₂
F16 (MC)	0.039	0.029	0.069	0.598
F13 (WC)	0.049	0.039	0.059	0.189
M6 (MC)	0.029	0.039	0.055	0.062
M15 (WC)	0.039	0.059	0.099	0.082

Table 5.3: Vowel length of two IT and two SIT tokens for four speakers (msec).

Comparing the vowel length (in msec) between the SIT and IT tokens in table 5.3, it is quite clear that IT tokens are indeed longer than SIT tokens, for both middle-class and working-class speakers. Speakers of Received Pronunciation use a very fronted variant for the KIT set (Gimson 1989), as do working-class speakers in this sample, but the variant used by these CSAE speakers does not come across like Received Pronunciation, precisely because the IT-subset is half long, produced as [iː].

To show further patterns, the acoustic data have been disaggregated by formant, class and environment, and following that, by formant, class and gender in the boxplots below.

b) Phonological conditioning of the KIT vowel

In this section, data is provided to ascertain whether the phonological environments generally reported to condition the split in the KIT set hold for the CSAE speakers in this study. In figures 5.8 to 5.11, IS data are disaggregated by environment, dividing the data into five categories, explained in table 5.4.

Environment name	Explanation	Bekker's category
Unmarked	SIT tokens	KIT4
_palatal	Palatal consonant following	KIT3
Velar_	Velar consonant preceding	KIT2
_velar	Velar consonant following	KIT2
Word initial	Word initially	KIT2

Table 5.4: Phonological environments for KIT tokens.

The 'unmarked' category refers to SIT tokens in IS. The remaining four categories represent environments in which SAE speakers typically use a higher, fronter variant (compared with the unmarked tokens). While they are presented separately in the boxplots below, they will be collectively referred to as the IT-subset.

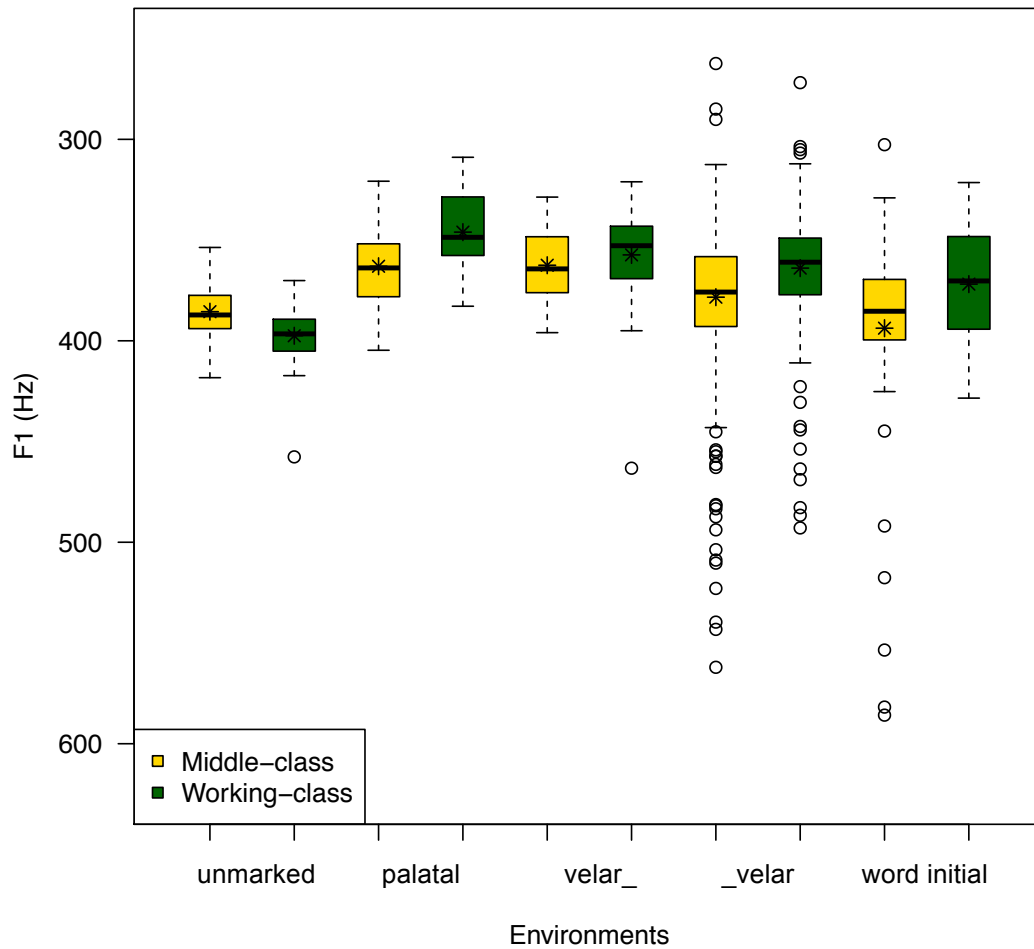


Figure 5.8: Boxplot for KIT F1 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by class.

Figure 5.8 shows that all of the IT-environments have a higher range of values compared to the SIT-subset. Statistically, all of these environments are significantly different to the SIT-subset ($p < 0.05$), save for the ‘word initial’ environment. The difference between middle- and working-class speakers was also significant ($p = 0.0005$), with working-class speakers using a higher range across all IT-environments compared with the middle-class speakers, and a lower range for the SIT tokens.

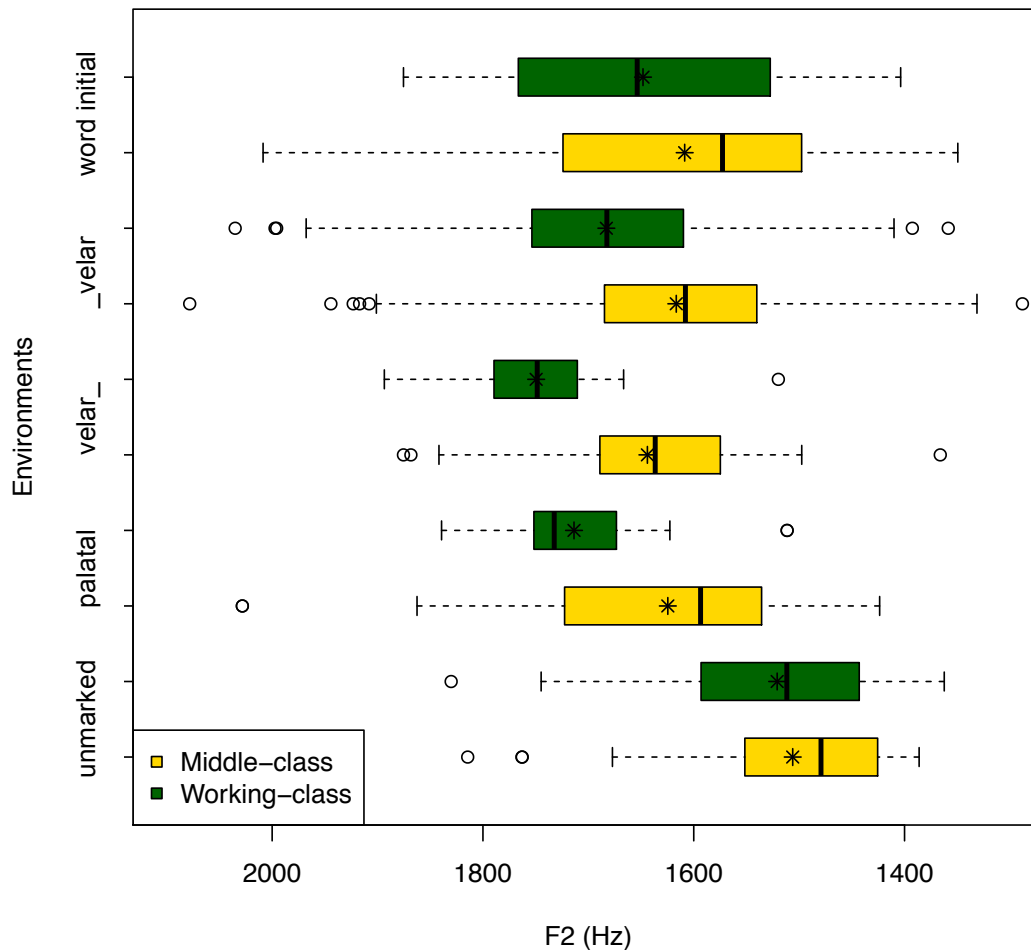


Figure 5.9: Boxplot for KIT F2 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by class.

Similarly for F2, it is clear from the distribution of the data in figure 5.9 that all IT-environments are fronter than the unmarked SIT category, and these findings are supported by the statistical tests ($p < 0.05$ for all environments compared with the SIT tokens). Class was also significantly different on the F2 plane ($p = 0.0055$), showing that working-class speakers generally use a fronter range of values for this set than middle-class speakers.

The two graphs above (figures 5.8 and 5.9) thus show that in terms of class, working-class speakers tend to use a higher and fronter IT-variant than the middle-class speakers. In terms of the conditioning phonological environments, in all cases the IT-environments were significantly higher and fronter than the SIT tokens and a lower SIT value than the middle-class speakers.

The following figures (5.10 and 5.11) disaggregate the data for the phonological environments for KIT by gender in order to ascertain what gender patterning is evident

in the KIT data. Because the data are essentially the same as those presented in figures 5.8 and 5.9, the statistical results in terms of phonological environment are the same as reported above. This information will therefore not be repeated, so only the results for gender are reported. (It remains useful to see the distribution in the data by males and females for the different environments, hence the provision of the boxplots.)

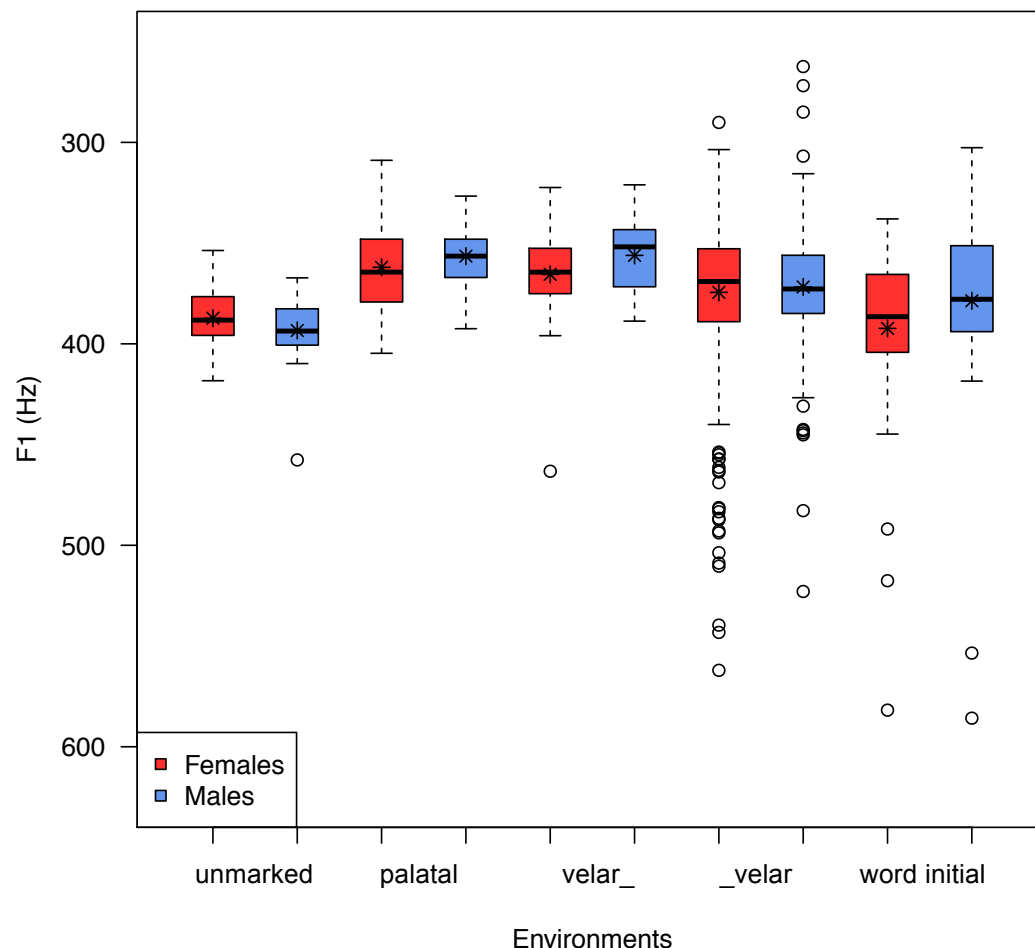


Figure 5.10: Boxplot for KIT F1 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by gender.

For F1, the difference between males and females is not very noticeable in the boxplots (figure 5.10), as males and females have very similar mean values in each of the environments. Not surprisingly then, gender was not a significant factor ($p=0.3851$).

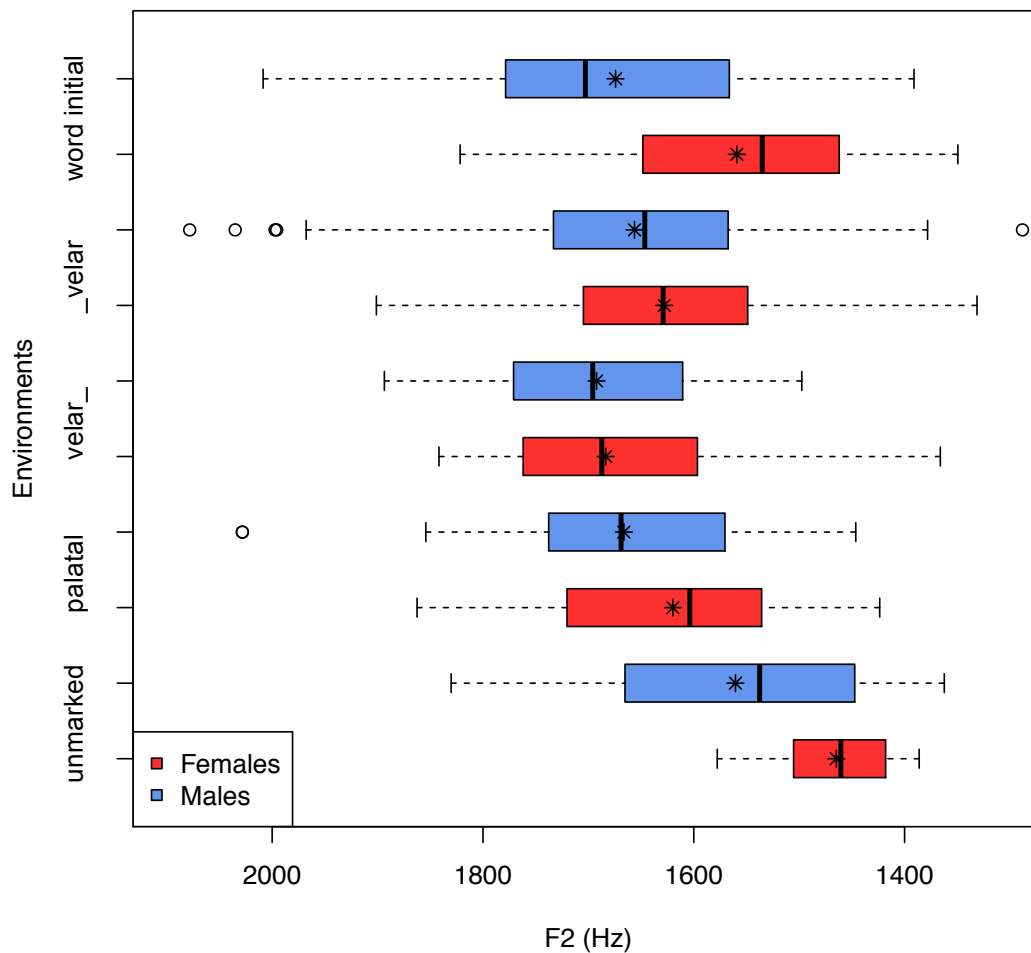


Figure 5.11: Boxplot for KIT F2 values in Interview Style (IS), showing the distribution of SIT (unmarked tokens) vs. the different phonological environments of the IT-subset, disaggregated by gender.

On the F2 plane (figure 5.11), the gender difference borders on significant ($p=0.0555$): male speakers tend to use a slightly fronter range of values than females. This result is based on a linear mixed-effects model for all tokens.

In terms of gender, the difference between male and female speakers is not significant, although the data show that males use a fronter variant than females in all environments and a higher variant in all IT-environments.

c) Class and Style

In this section, the data are disaggregated by formant, class and speech style. Both the SIT and IT-subsets are represented as a point of comparison with the more careful styles, viz. Reading Passage (RP) style and Word List (WL) style. The WL tokens are also split into IT and SIT categories because both sets were represented on the list (by the words *kit* and *bit* respectively). The words making up the tokens in RP style

belong to the IT-subset, so a high, front variant is expected in this style. The words are *bricks* (n=26) and (the first syllable of) *finish(ed)* (n=17). Although *finish*, according to the phonological rule governing the KIT split, should belong to the SIT-subset, it is in fact an exception to the rule (for CSAE speakers in the sample), and realised as part of the IT-subset.

Figures 5.12 to 5.15 depict several boxplots of different categories of the KIT data. These include tokens in IS for both the IT and SIT subsets, referred to as IT in IS and SIT in IS, where necessary. RP tokens are referred to as such, and the two sets of WL tokens are referred to as WL SIT and WL IT.

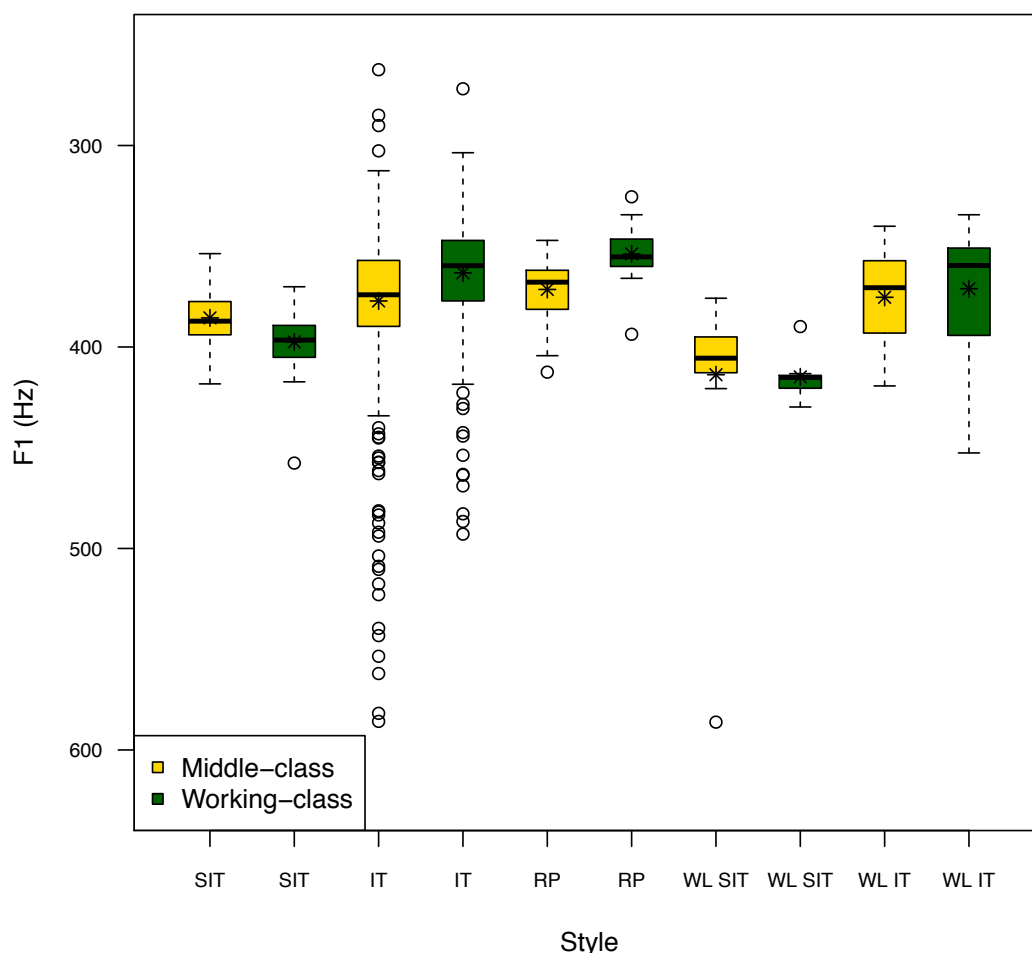


Figure 5.12: Boxplot for KIT F1 values in all speech styles, disaggregated by class.

Figure 5.12 displays the range of different subsets of KIT data for the first formant. The boxplots show that all categories making up the IT-subset (i.e. IT in IS, RP and WL IT) have a higher range of values, compared with the SIT categories (ie. SIT in IS and WL SIT). The statistical tests support these findings: IT tokens are significantly higher than the SIT tokens ($p < 0.001$) as are the WL IT tokens ($p = 0.0059$), and RP

tokens ($p < 0.001$). WL BIT tokens are significantly lower than the SIT in IS tokens, showing a stylistic difference between IS and WL style.

Stylistically, there is no significant difference between RP tokens and IT in IS tokens ($p = 0.193$), nor with WL IT tokens ($p = 0.8748$) (compared to IS). Of course, WL BIT is significantly different to IT tokens in IS ($p < 0.001$). The difference between middle-class and working-class speakers is also significant ($p = 0.0011$), with working-class speakers using a higher range of values for the IT categories, and a lower range for the SIT categories. The KIT split is thus greater for working-class speakers.

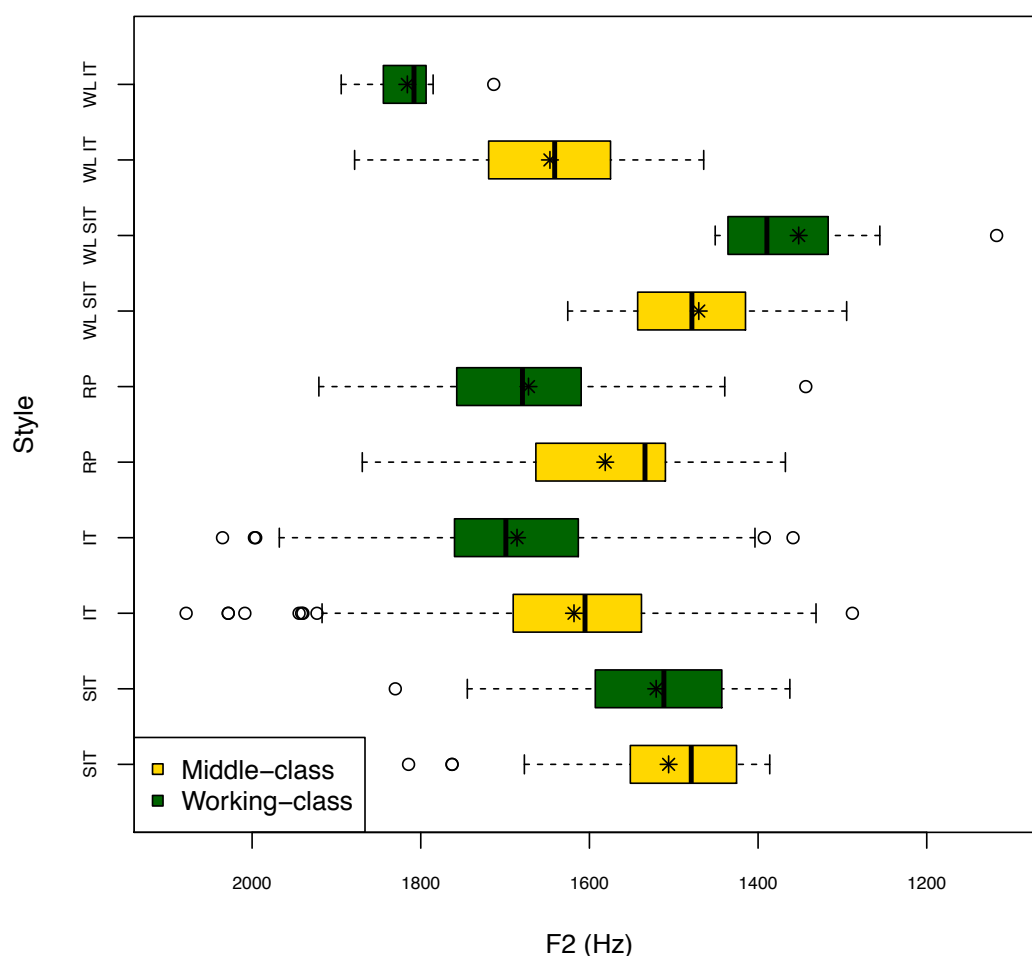


Figure 5.13: Boxplot for KIT F2 values in all speech styles, disaggregated by class.

Figure 5.13 shows the data for F2 values of the KIT tokens, disaggregated into the three different speech styles for each of the class groups. Compared with the SIT tokens in IS, IT tokens are significantly fronter ($p < 0.001$), as are the RP tokens ($p < 0.001$) and WL IT tokens ($p < 0.001$). The WL SIT tokens are significantly backer than the SIT tokens in IS ($p = 0.0175$) as the boxplots show, again proving a stylistic difference between IS and WL style.

Comparing the IT tokens to RP tokens, there is evidence of a stylistic difference between the two categories ($p=0.0306$): in this case RP tokens are backer than the IT tokens in IS, also evident in the boxplot (figure 5.3). WL IT tokens are significantly fronter than the IT tokens in IS ($p=0.0005$), and of course WL SIT tokens are significantly backer ($p<0.001$).

In terms of class, working-class speakers use a fronter range for the IT categories, and a backer range for the SIT categories. The class difference is significant ($p=0.0039$).

d) Gender and Style⁴⁵

In figures 5.14 and 5.15, the data for the KIT set have been disaggregated by formant, gender and style.

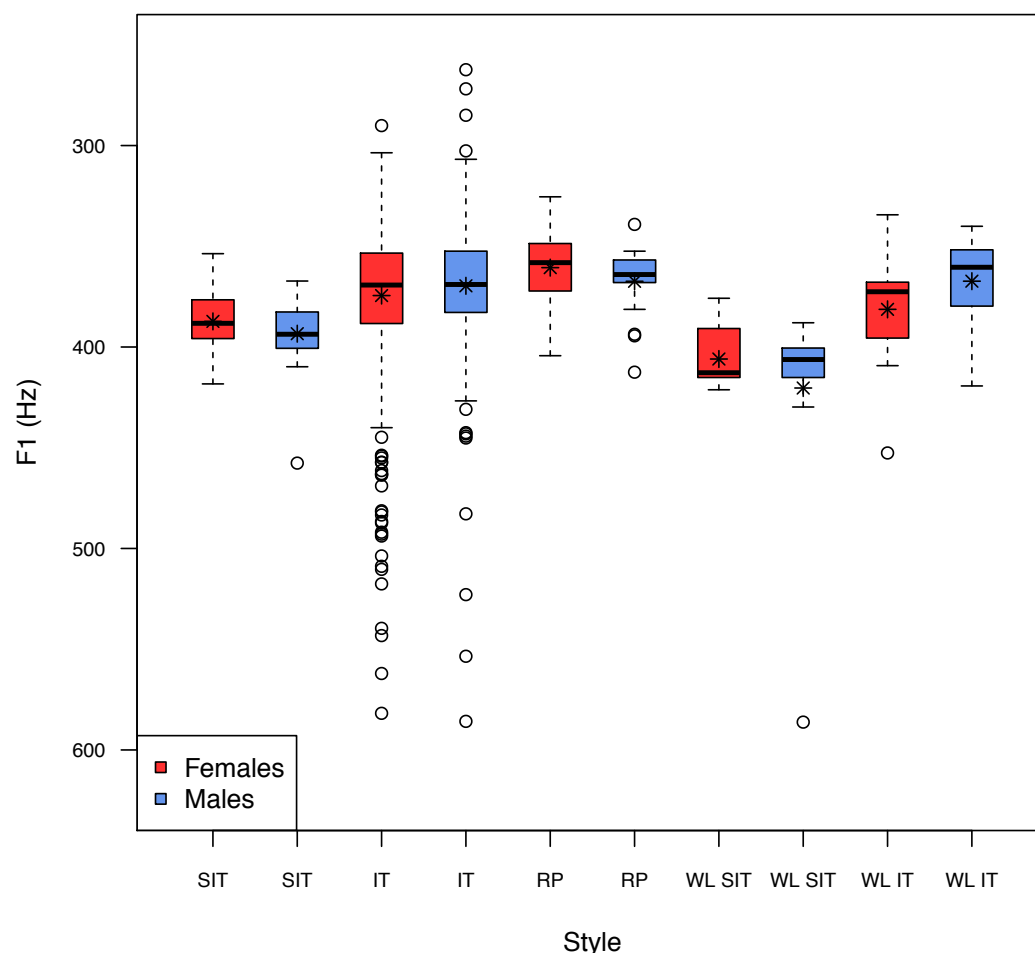


Figure 5.14: Boxplot for KIT F1 values in all speech styles, disaggregated by gender.

⁴⁵ In this section, the statistical results for the stylistic differences are not provided, as they are virtually identical to those reported in the previous section on ‘Class and Style’. The discussion and visual analysis of the graphs does take style into account, however. The same applies to the ‘Gender and Style’ discussion in each of the vowel subsections throughout the chapter.

The gender distinction between the tokens of the KIT set is not significant in terms of vowel height ($p=0.3959$). Figure 5.14 shows that in WL style, male speakers use a higher IT value than females, and a lower SIT value. The SIT tokens show the same trend (males have a slightly lower mean value than females).

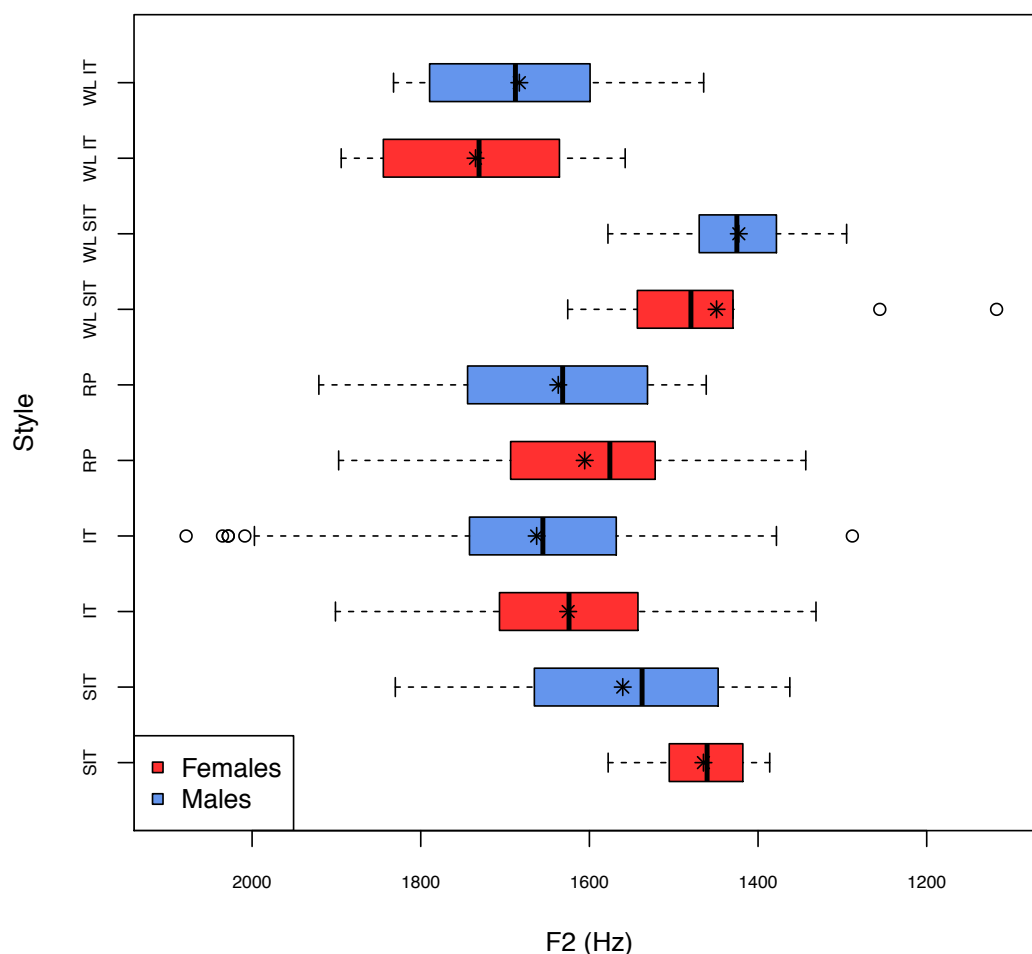


Figure 5.15: Boxplot for KIT F2 values in all speech styles, disaggregated by gender.

In terms of vowel advancement, gender difference is also not significant for the speakers in this sample ($p=0.0805$) across the entire range of data. The boxplots in figure 5.15 show, however, that in WL style, males use a backer value than females for both subsets. In IS, males use a fronter variant for both subsets, and this pattern is also evident in RP style.

5.4.4 Summary: KIT Findings

In summary, the KIT set shows a definite class distinction, with working-class speakers consistently using a higher and fronter variant for the IT-subset compared with their middle-class counterparts. For the SIT-subset, working-class speakers use a backer and lower variant compared with middle-class speakers. These trends hold across IS, RP and WL styles. For this set, the acoustic data do not show any

significant difference between male and female speakers, either in terms of vowel height or advancement, but from the distribution of the data in WL style, it seems that males tend to use a lower and backer SIT token and higher IT token than females.

It was also confirmed from the acoustic data that the KIT set is split according to the phonological rules outlined by Lass (2002): before and after velar consonants, before palatal consonants and word initially, the set is realised as the IT-subset with a higher and fronter variant. In other phonological environments, the SIT-subset is realised with a lower and more centralised vowel quality.

Furthermore, it was found that the IT-subset is produced with a longer vowel than the SIT-subset, suggesting that the vowel used in this subset is half-long.

5.5 DRESS

5.5.1 WSAE Reports

According to Lass (2002: 115), DRESS is not a significant social marker, although he reports that a gender distinction is in evidence for this set: females tend to use a closer variant than the half-close front [e] that males typically use in non-Conservative lects, with their variant overlapping somewhat with the KIT set. In Broad SAE, the pre-/l/ allophones often have preceding [j] initially and after /h/ e.g. *help* [(h)jɛlp]. This vowel lowers and retracts before velar /l/ for some General and Broad speakers, to [ɛ] or [æ]. Bekker's (2009: 228f) acoustic data show that DRESS is a close front vowel which overlaps with KIT2 (i.e. word-initially and preceding and following velars).

5.5.2 CSAE Reports

Aural accounts of CSAE report this set to be realised as [e]. Lowered [ɛ] and [æ] also occur (Finn 2004: 968), although this is mainly before velar /l/. Wood (1987: 122) reports unique realisation of the word *yes* as [jə(:)s] (i.e. a shift in word class to the NURSE set). Two other words that follow this trend, according to my own impressions, are *jelly* and *jealous* (without additional length). The acoustic analysis will deal with these words separately, using IS tokens of *yes* from the interview (which are excluded from the main analysis as were all other words following /j/ - see section 4.3). The words *jelly* and *jealous* were included in the WL in order to facilitate this comparative analysis.

5.5.3 CSAE Acoustic Data

a) Overview

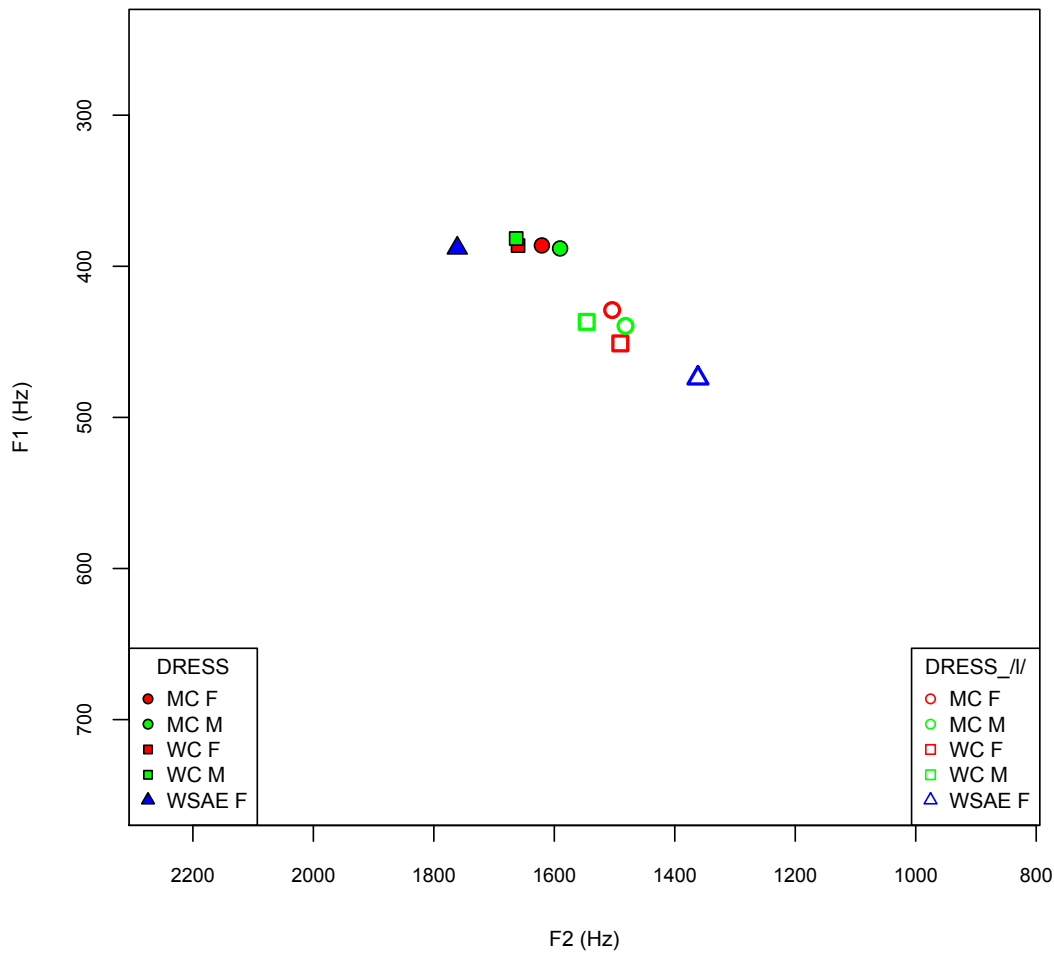


Figure 5.16: Mean values for DRESS and DRESS_// in Interview Style, disaggregated by class and gender.

Figure 5.16 shows that tokens of DRESS seem to be realised quite uniformly by all speaker groups, in terms of vowel height. The variants used by middle-class speakers are backer than those used by working-class speakers. The general realisation of this set by the groups in this sample confirms the earlier accounts as it is in the region of [e], although all CSAE groups realise this set further back than Bekker's (2009) female WSAE speakers, who use a seemingly fully front DRESS vowel. In terms of vowel height, all CSAE speakers use a variant as high as Bekker's (2009) WSAE speakers.

Figure 5.16 also confirms that retraction occurs for this set in the phonological environment preceding tautosyllabic velar /l/, although the retraction is not as pronounced as for the WSAE speakers: the two WSAE mean tokens are, respectively,

the furthest front and furthest back on the graph. All CSAE speaker groups use a lower, centralised vowel before velar /l/.

The following section covers the words in the set that are thought be anomalous.

b) Unique words

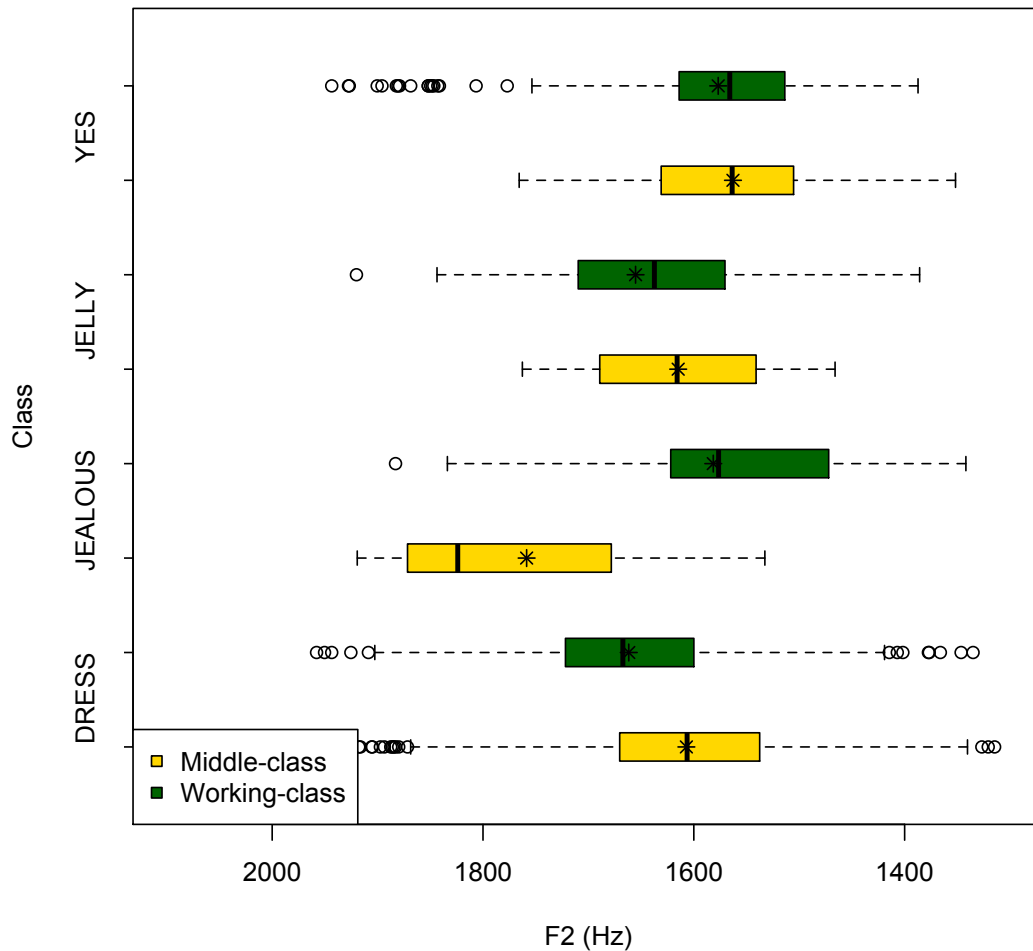


Figure 5.17: Boxplot for DRESS, *jealous*, *jelly* and *yes* F2 values, disaggregated by class.

The boxplots in figure 5.17 show the spread of the data on the F2 plane in order to ascertain whether the speakers in this sample confirm Wood's (1987) use of a more centralised variant for the word *yes*, and my own impressions of the words *jelly* and *jealous*. The latter words were included in the word list in order to check whether this distinction is clear in the acoustic data. The DRESS tokens in figure 5.17 represent all clean tokens of DRESS, not only tokens of the word *dress*, and are used as a point of comparison for the other three words. Statistically, *yes* was produced significantly backer than DRESS words ($p < 0.001$), as was *jealous* ($p = 0.033$). This was not the case for *jelly* ($p = 0.5363$), where middle-class speakers actually used a backer value than

the working-class speakers. For *yes* and *jealous*, however, it is confirmed that CSAE speakers use a more centralised variant than for the rest of the DRESS set.

In the figures that follow, the data for DRESS are disaggregated in terms of formant, style, class and gender to further elucidate trends in the acoustic data for CSAE speakers.

c) Class and Style

In this section, the data are disaggregated by class and speech style.

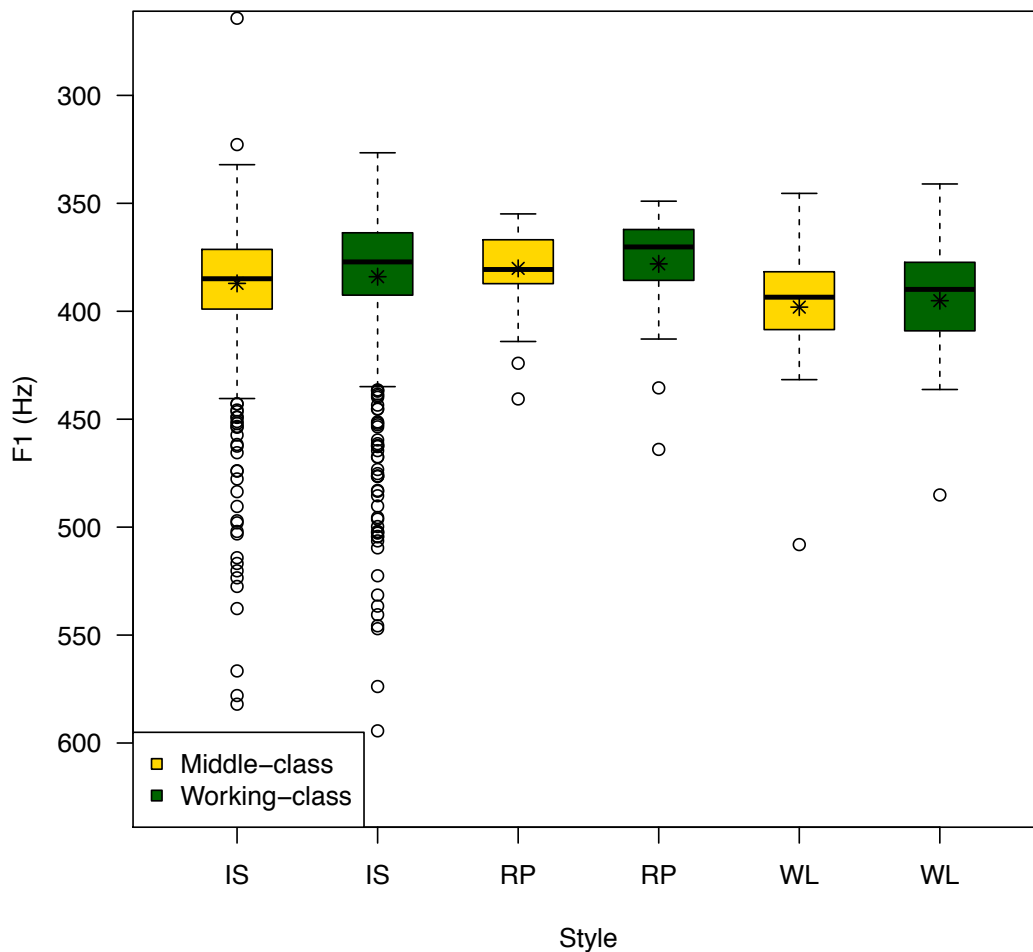


Figure 5.18: Boxplot for DRESS F1 values in all speech styles, disaggregated by class.

In terms of vowel height, the boxplot (figure 5.18) shows that in WL style speakers use a lower range of values compared with IS ($p=0.0002$). The height difference between IS and RP style is not as clear, however, and not statistically significant ($p=0.0622$). The difference between middle- and working-class speakers is also not significant for this set ($p=0.4312$).

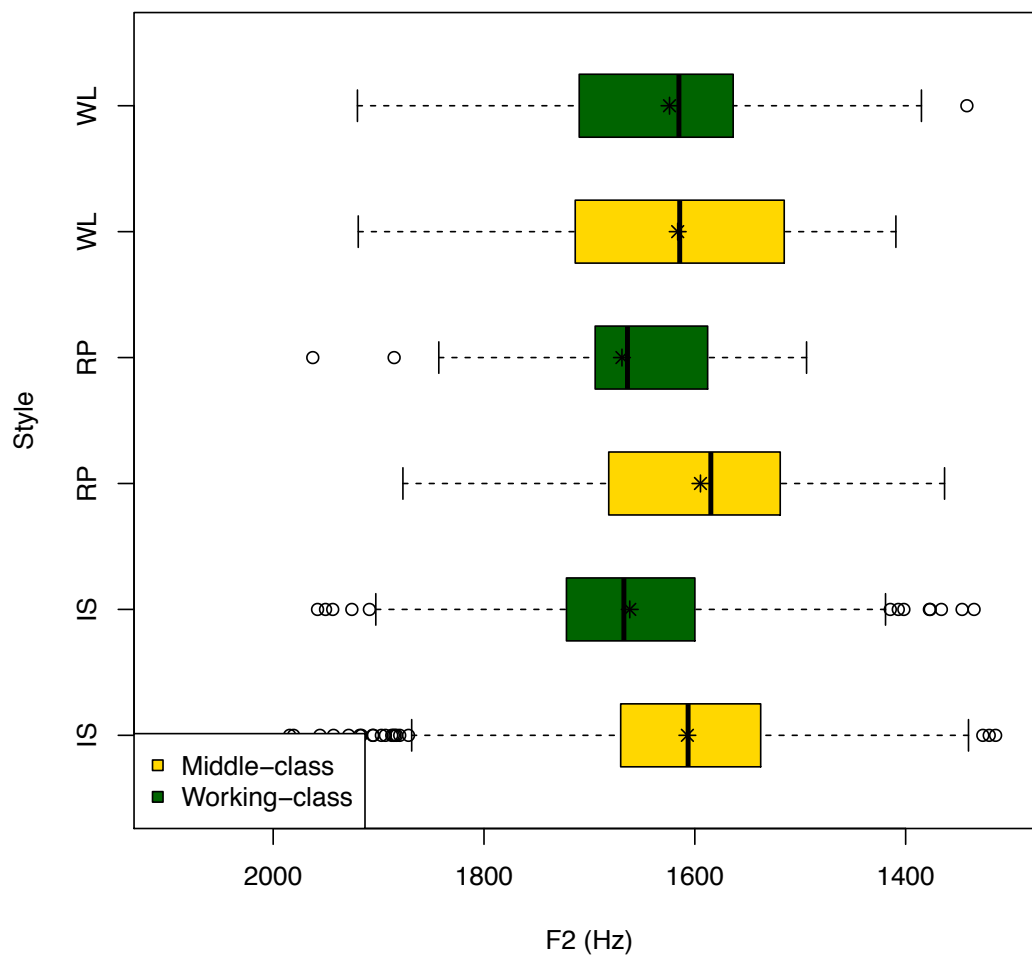


Figure 5.19: Boxplot for DRESS F2 values in all speech styles, disaggregated by class.

In terms of advancement, there is a significant difference between middle- and working-class speakers ($p=0.0142$): in the boxplot (figure 5.19), it is clear in RP style and IS that working-class speakers use a frontier mean range than middle-class speakers do.

Stylistically, RP style is not significantly different to IS ($p=0.1859$) but WL style is ($p=0.0079$). The boxplots do not clearly show how the two styles differ, despite the significant result.

d) Gender and Style

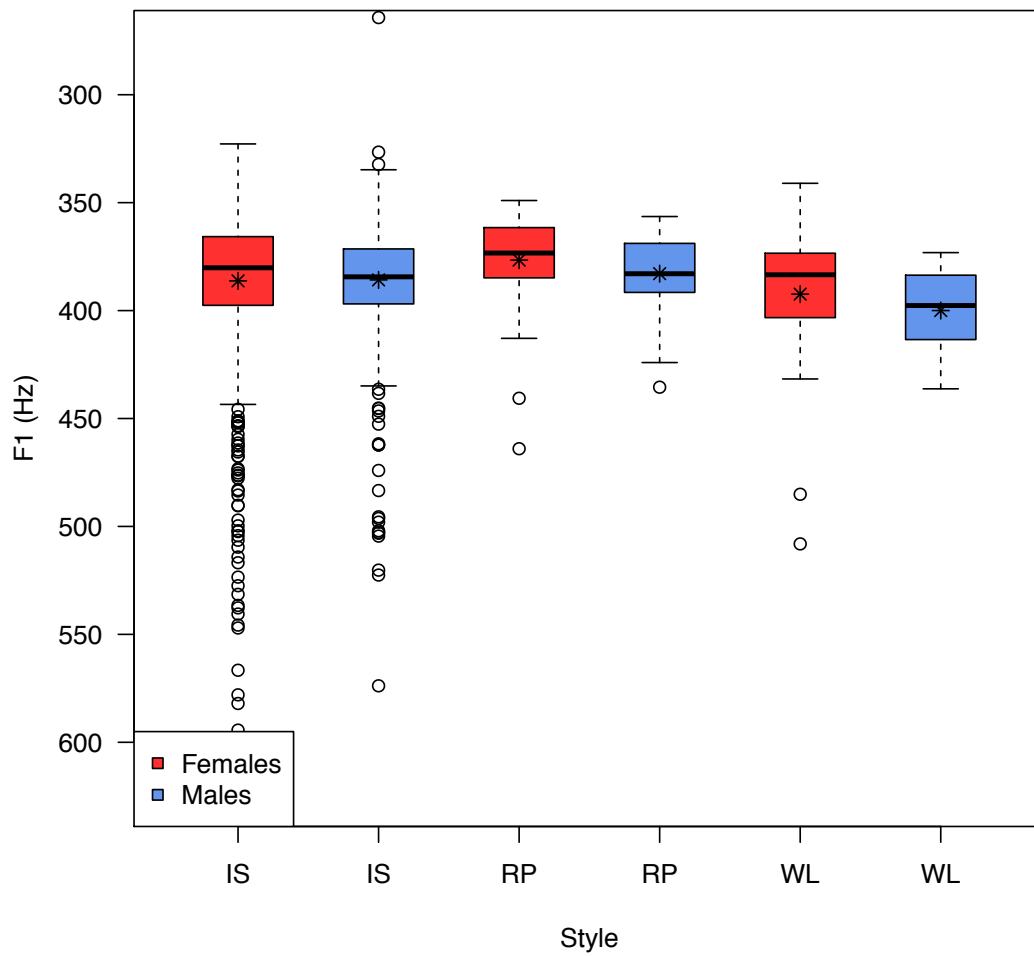


Figure 5.20: Boxplot for DRESS F1 values in all speech styles, disaggregated by gender.

The distinction between male and female speakers on the F1 plane is not significant ($p=0.9567$), although the boxplots (figure 5.20) seem to show that females tend to use a slightly higher mean range of variants in RP and WL styles compared with the male speakers.

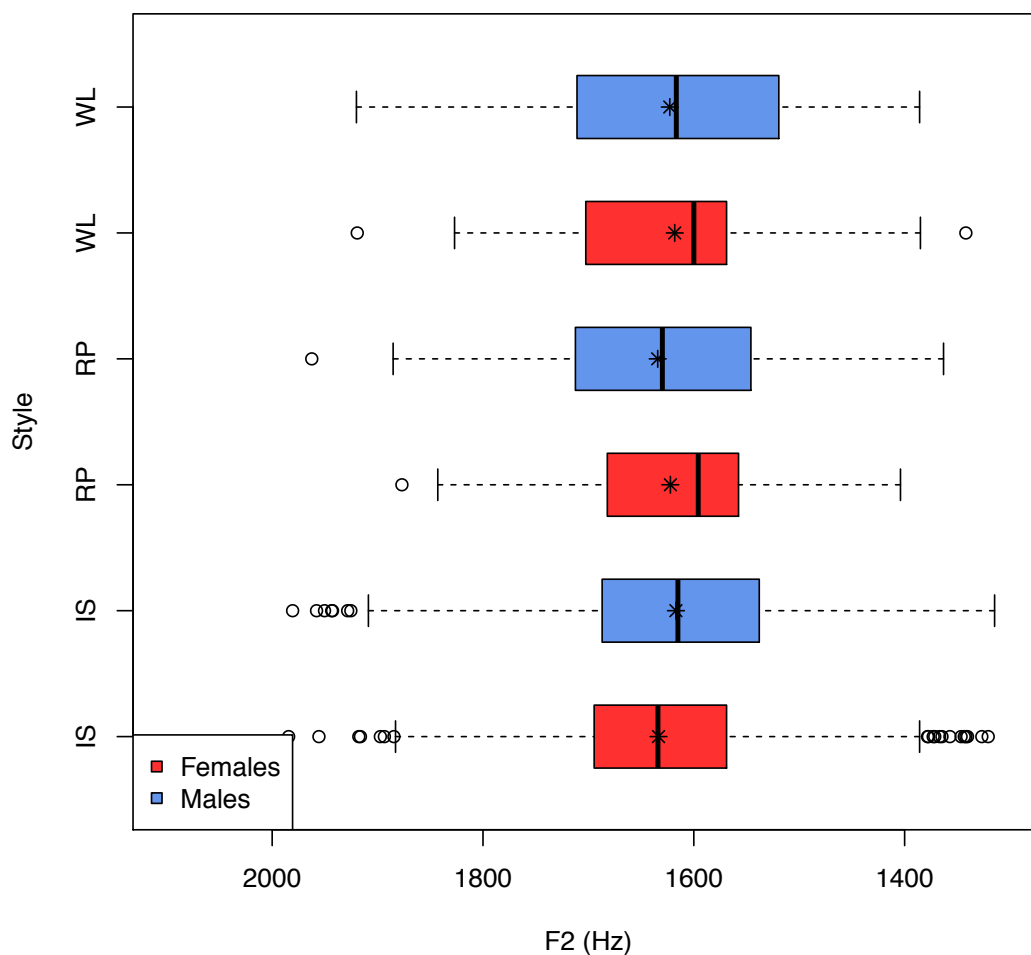


Figure 5.21: Boxplot for DRESS F2 values in all speech styles, disaggregated by gender.

In terms of advancement, generally, the boxplots in figure 5.21 show that males tend to use a wider range of values compared with females across all three speech styles, but the difference between males and females is also not significant ($p=0.2779$).

5.5.4 Summary: DRESS Findings

For this set, the acoustic data do not show a significant gender distinction in terms of vowel height or advancement. In terms of class distinction, the only significant distinction seems to be that working-class speakers use a slightly fronter variant than middle-class speakers. It is possible that the use of this fronted DRESS variant by working-class speakers is structurally related to the higher and fronter variant used in the IT-subset of the KIT vowel by the same group of speakers. Both working- and middle-class speakers use a lower and slightly fronter DRESS vowel in WL style compared with IS, but RP style showed no significant difference compared with IS. The acoustic data also confirmed that the words *yes* and *jelly* are realised further back than regular tokens of DRESS and that these CSAE speakers retract this set before velar /l/.

The acoustic results thus confirm that DRESS is not a very salient social marker in SAE: for CSAE speakers, there was little difference in the data between different groups of speakers and nothing that would suggest any emerging trends in this set, although it would be expected that DRESS will begin lowering, as is evident in the KIT and TRAP sets (see sections 5.4 and 5.6 respectively), as part of the reverse vowel shift.

5.6 TRAP

5.6.1 WSAE Reports

TRAP is an important social marker in SAE as it distinguishes the lects of the SAE hierarchy, and is showing interesting patterns of change, detailed below. Lass (2002: 115) reports the traditional value for this set in WSAE to be [æ] in Conservative and General SAE, while Broad SAE has a more [ɛ]-like quality. Use of the higher variant is highly stigmatised. It has been argued that raised TRAP is a relic of the input variety of SAE spoken by the 1820 Settlers (Trudgill 2004), supported by the evidence in the chronicle of one of the Settlers (discussed in section 5.2 above) cited by Lass (2002: 114).

A lower variant is more prestigious in SAE. In Received Pronunciation, this vowel is lowering (as is DRESS and KIT), and lowered TRAP also features in Modern Cockney accents as well as other accents within mainland England, as a starting point for the SECS-shift (Bekker 2009: 192). Bekker (2009: 201f) reports that TRAP in SAE is lowering, based on his citation-style acoustic data of adolescent White females, who use a value that is slightly higher than TRAP in Received Pronunciation. In WSAE, this set lowers and retracts before velar /l/, which is very clear in Bekker's (2009: 196) data.

5.6.2 CSAE Reports

Impressionistic accounts of CSAE have this set raised towards [ɛ] (Wood 1987: 122), although [æ] and slightly diphthongised [æ^ə] are reported to occur. TRAP purportedly does not retract before /l/ in CSAE (Finn 2004: 970).

5.6.3 CSAE Acoustic data

a) Overview

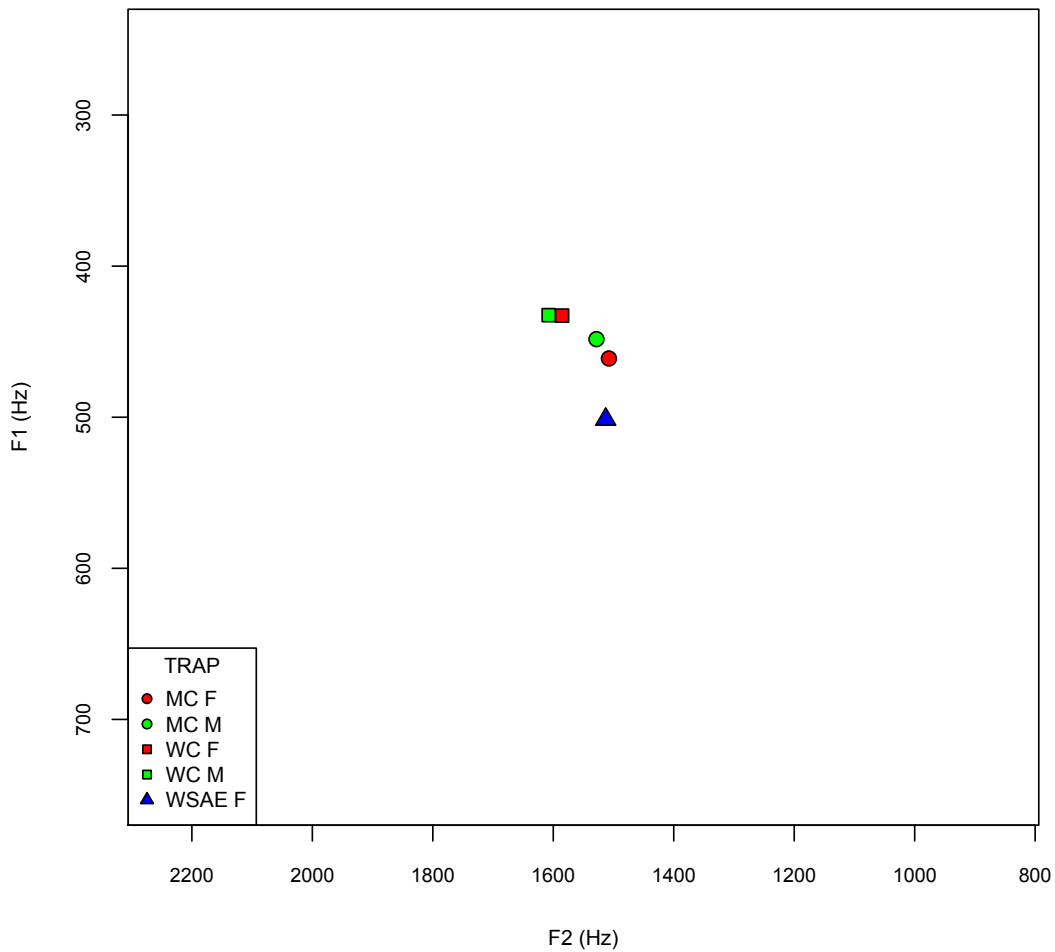


Figure 5.22: Mean values for TRAP in Interview Style, disaggregated by class and gender.

The overview of the TRAP set in figure 5.22 shows that while none of the CSAE speakers use a variant as low as Bekker's WSAE speakers, middle-class males and females use a lower variant than do their working-class counterparts.

This finding is unsurprising given the stigma attached to the use of raised TRAP. As such, it is associated with working-class speakers of all varieties of SAE (except Indian SAE). The use of the lower variant middle-class speakers is likely to be due to the speakers' exposure to middle-class WSAE in their educational, and in some cases, residential environments.

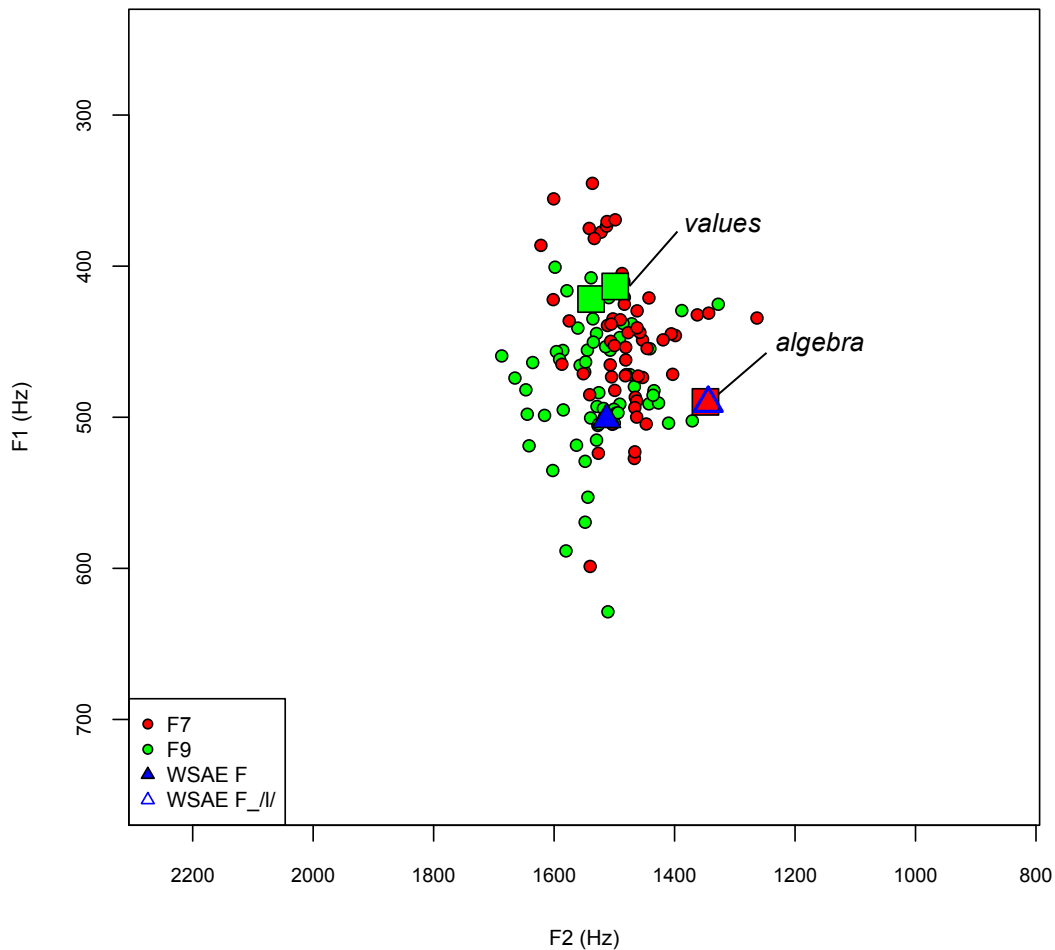


Figure 5.23: TRAP tokens for speakers F7 and F9, showing the position of tokens before velar /l/.

With reference to the retraction of TRAP before velar /l/, Finn (2004: 970) suggests that it does not occur in CSAE. Bekker's (2009) data clearly show that it is indeed a characteristic of WSAE when in syllable-final position i.e. tautosyllabic with the preceding vowel. In Bekker's dataset, the word *pal* provides the data that shows the retraction of TRAP in this environment. In the present dataset, TRAP only occurs in this environment in three occurrences of the vowel for which formant measurements were extracted. Two tokens were produced by speaker F9 in the word *values* and the other token by another speaker (F7) for the word *algebra*. Figure 5.23 shows the placement of these tokens in relation to the other tokens for these two speakers. The word *values* is not retracted in relation to speaker F9's other tokens of TRAP. The word *algebra* is backer than most of speaker F7's other TRAP tokens, and as back as the mean value for Bekker's (2009) retracted TRAP set. The small number of tokens with the relevant environments makes it impossible to draw acoustic conclusions about whether or not CSAE speakers in fact retract this set before tautosyllabic, velar /l/.

b) Class and Style

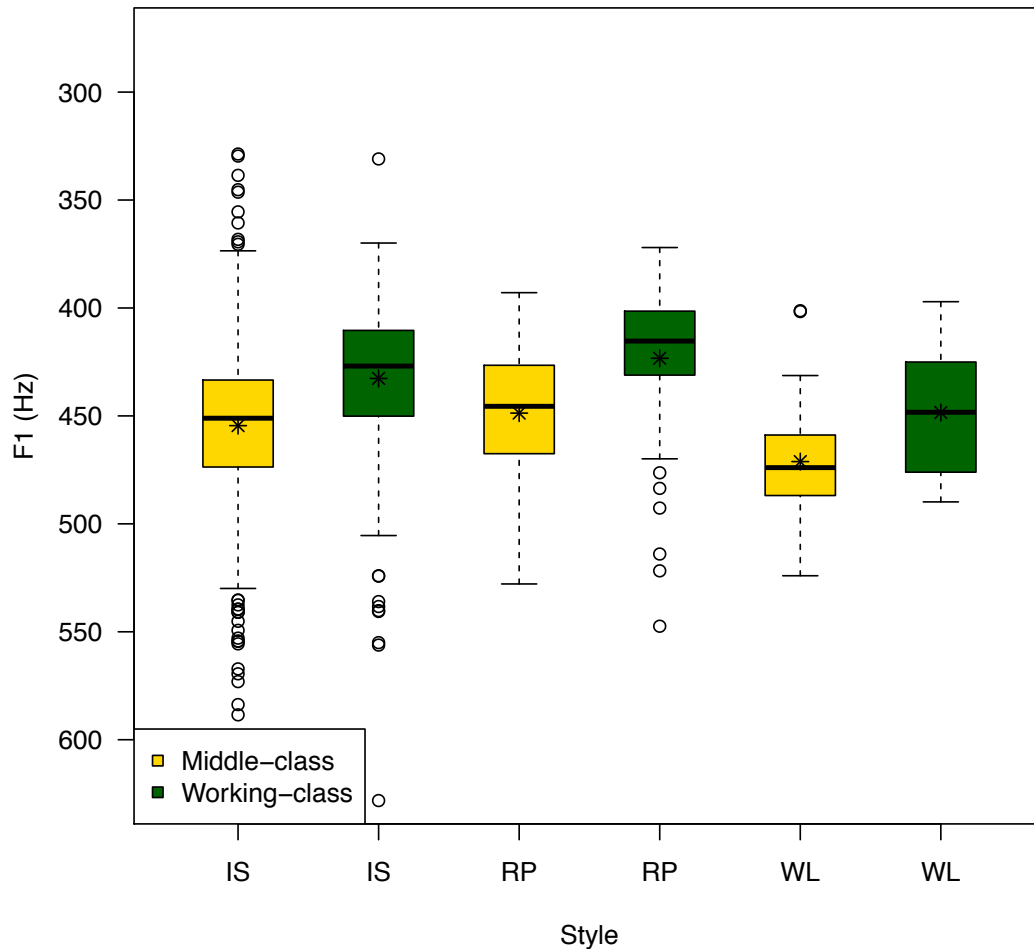


Figure 5.24: Boxplot for TRAP F1 values in all speech styles, disaggregated by class.

Figure 5.24 shows that in terms of vowel height, working-class speakers use a significantly higher TRAP vowel than middle-class speakers ($p < 0.001$) in all styles. Both RP and WL styles are significantly different to IS ($p = 0.0464$ and $p = 0.0011$ respectively): the boxplots in figure 5.24 show the difference between WL style and IS for both classes more clearly, where it is evident that the range of WL style tokens is lower than those produced in IS. This is unsurprising given that the higher prestige variant is expected in more careful speech styles. TRAP in RP style is generally realised higher than in IS.

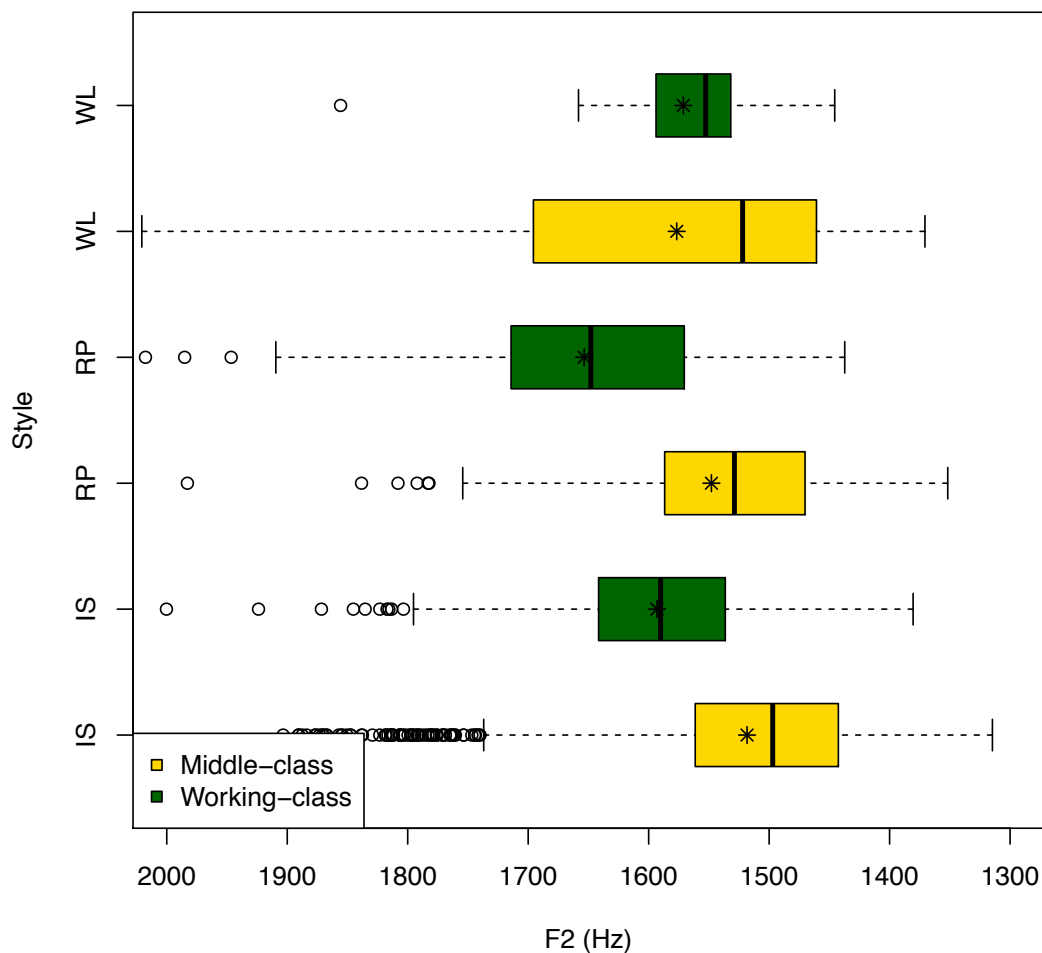


Figure 5.25: Boxplot for TRAP F2 values in all speech styles, disaggregated by class.

Figure 5.25 also shows a clear class distinction in terms of F2: working-class speakers use a frontier range of values than middle-class speakers ($p < 0.001$), most evident when comparing the boxplots for RP style and IS. Stylistically, RP style is significantly frontier than IS ($p = 0.0001$), but the distinction between IS and WL style is not significant ($p = 0.1644$). The boxplots show that in WL style, despite a similar mean value for middle- and working-class speakers, the former group use a much wider range of values compared with the latter.

c) Gender and Style

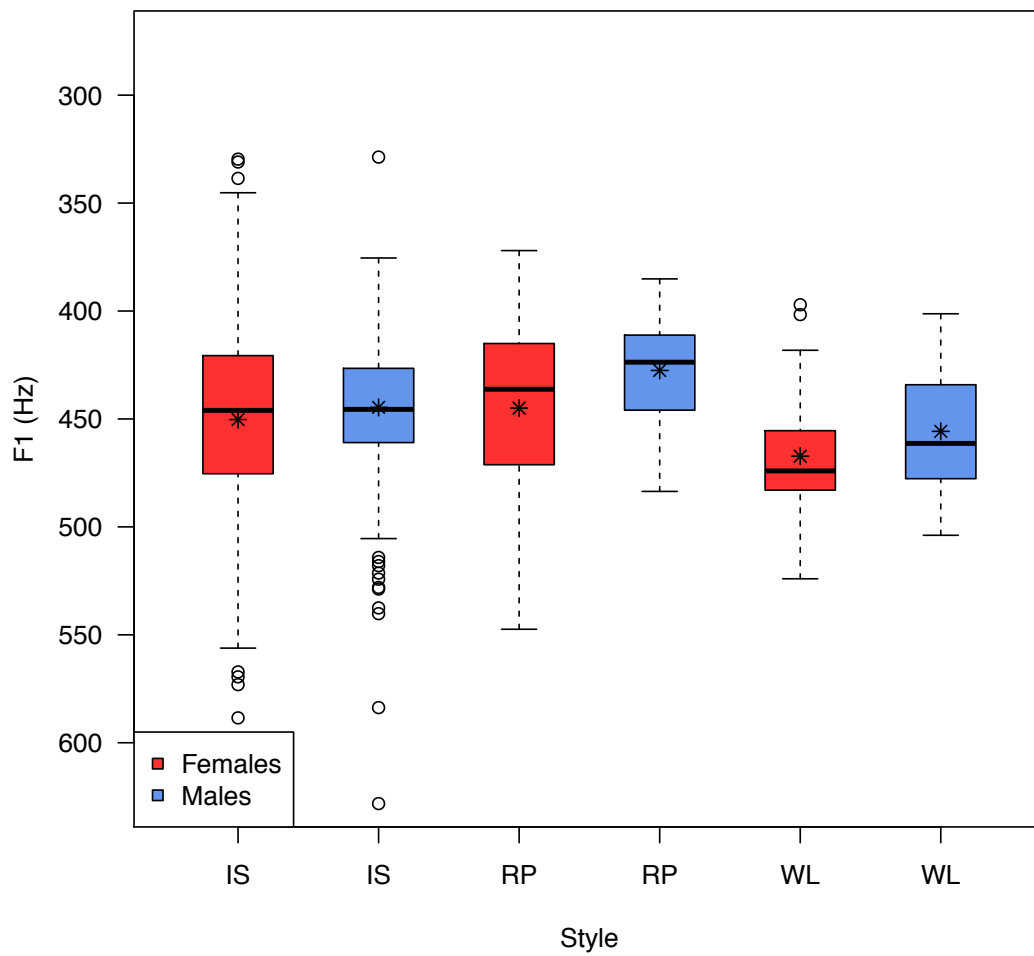


Figure 5.26: Boxplot for TRAP F1 values in all speech styles, disaggregated by gender.

The gender difference for the TRAP set is not significant on the F1 plane ($p=0.1196$), although the boxplots in figure 5.26 show that in IS and RP style, female speakers use a wider range of values than do the male speakers. In all three styles, the boxplots also show that females use a slightly lower mean value compared with male speakers.

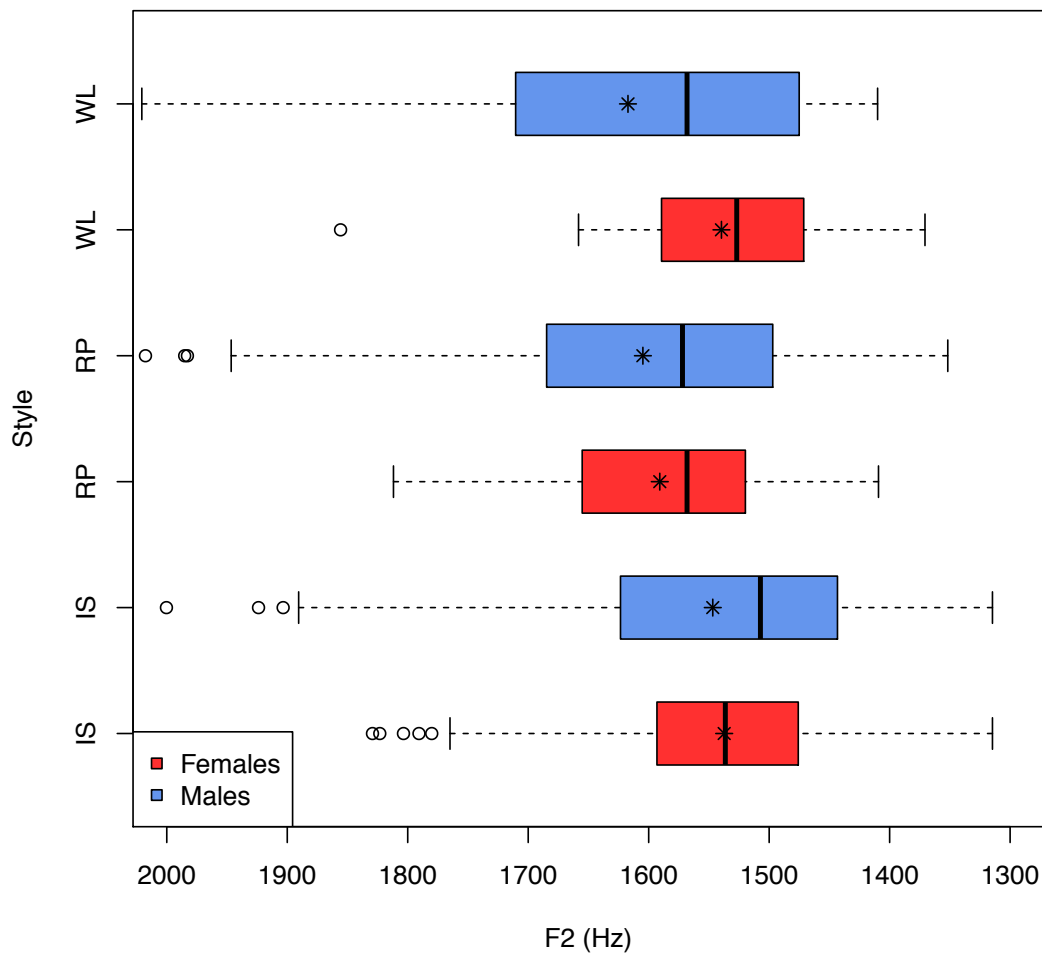


Figure 5.27: Boxplot for TRAP F2 values in all speech styles, disaggregated by gender.

Similarly, on the F2 plane, the distinction between tokens produced by males and females is not significant ($p=0.6249$). The boxplots in figure 5.27 show that in WL style, male speakers use a wider range with a fronter mean value than do females. In the other two styles, the males' mean value is also slightly fronter than the females' value.

5.6.4 Summary: TRAP Findings

On the whole, the data show that middle-class speakers tend to use a lower, more centralised TRAP vowel, while working-class speakers use a higher, fronter variant. In the more careful styles (RP and WL style), speakers use a lower vowel than in IS. This is in keeping with the position that a lower TRAP vowel is a higher prestige variant, while higher TRAP is stigmatised. Middle-class speakers' TRAP vowel approximates the low value used by Bekker's (2009) female speakers, suggesting that they are following the lowering trend of WSAE speakers. It remains to be seen whether middle-class CSAE speakers will lower this vowel further, following the

apparent trend in WSAE. There is no significant distinction between male and female speakers for this set.

It was not possible to confirm acoustically whether or not CSAE speakers in fact retract this set before velar /l/ because of the small number of pre-/l/ tokens in the relevant phonological environment.

5.7 LOT

The remaining short vowels, LOT, STRUT and FOOT are not subject to the chain shift which affects the front vowels hitherto discussed, and thus seem more stable in their traditional positions.

5.7.1 WSAE Reports

Lass (2002: 115) describes this set as a short, open and weakly rounded back vowel, which can be centralised to [ɔ̞]. He further reports that certain younger General speakers produce raised, less rounded variants, approximating central [ʌ]. This is not attested in Bekker's (2009) data: he reports that LOT occupied a standard low back position for WSAE females (2009: 342) in his citation-form data.

5.7.2 CSAE Reports

Wood's (1987: 122) account of CSAE records L2 speakers typically using either higher, rounded [ɔ] or low, unrounded [ɒ], whereas L1 speakers use [ɒ] more consistently. A lengthened and slightly diphthongised [ɒ(:)ʰ] is also reported to occur. A WANT subset is realised with [ʌ] in words like *want* and *non-* (e.g. *non-smoking*). These words will be dealt with separately in the ensuing acoustic analysis. Retraction before /l/ is not reported to occur in this set.

5.7.3 CSAE Acoustic Data

a) Overview

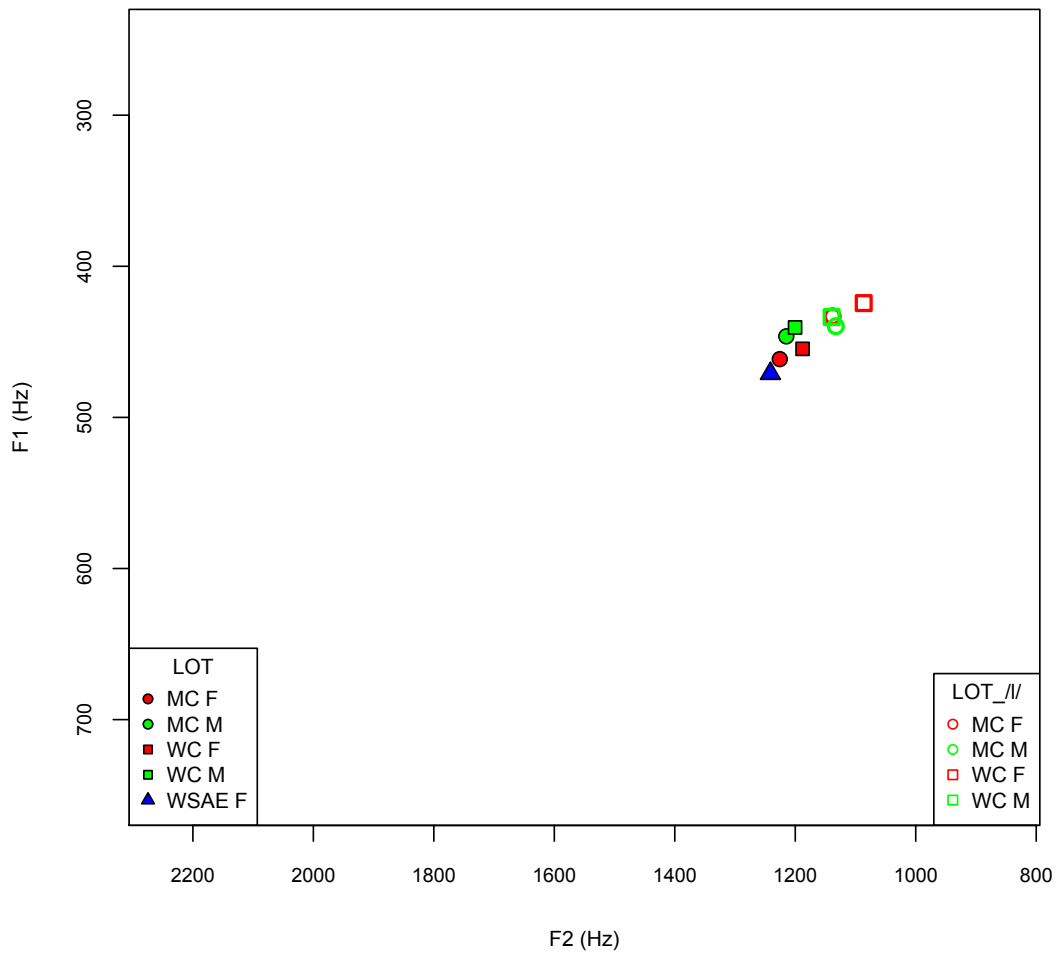


Figure 5.28: Mean values for LOT and LOT_/l/ in Interview Style, disaggregated by class and gender.

Figure 5.28 shows that male speakers tend to use a slightly higher variant than females and working-class speakers use a backer mean value than middle-class speakers. The CSAE speakers use a slightly higher variant compared with Bekker's (2009) WSAE speakers.

For LOT, retraction before velar /l/ is not reported by Wood (1987), nor by Bekker (2009) for his data. This is probably because LOT is a back vowel to begin with. In my data, however, I found a degree of retraction before velar /l/, with the difference to the regular IS tokens being highly significant ($p < 0.001$). So for CSAE speakers, there does seem to be retraction for the LOT set before velar /l/ by all speaker groups, who realise this subset even backer than regular LOT tokens.

b) Unique words

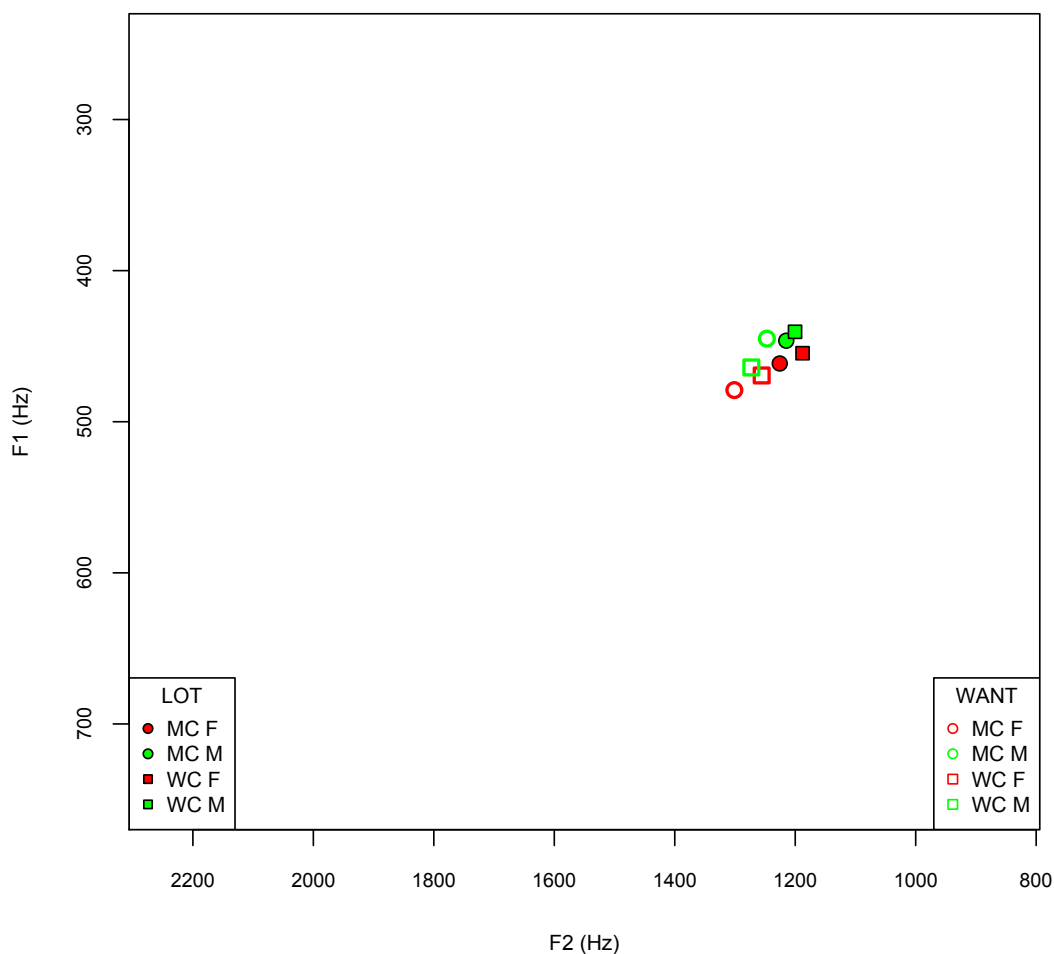


Figure 5.29: Mean values for LOT and *want* tokens, disaggregated by class and gender.

The words identified by Wood (1987) as having pronunciations which deviate from the norm for this set, were *want* and *non-* (as in *non-stop* or *non-smoking*). Figure 5.29 shows a mean value for the tokens of *want* in the dataset, indicating that *want* is realised as more centralised than other tokens of LOT by all groups of speakers, confirming Wood's finding. Figure 5.30 (overleaf) shows that for *non-* words, the findings are not as consistent: Middle-class females produced this value with a backer mean value than other LOT tokens, working-class males used a slightly more centralised mean value, and middle-class males used a mean value that was not much different to the other LOT tokens for their grouping, as the points lie almost on top of one another on the graph. The working-class females in the sample did not produce any tokens of *non-*. These data are based on very few observations of *non-* (n=14), which is perhaps why the data are slightly inconclusive.

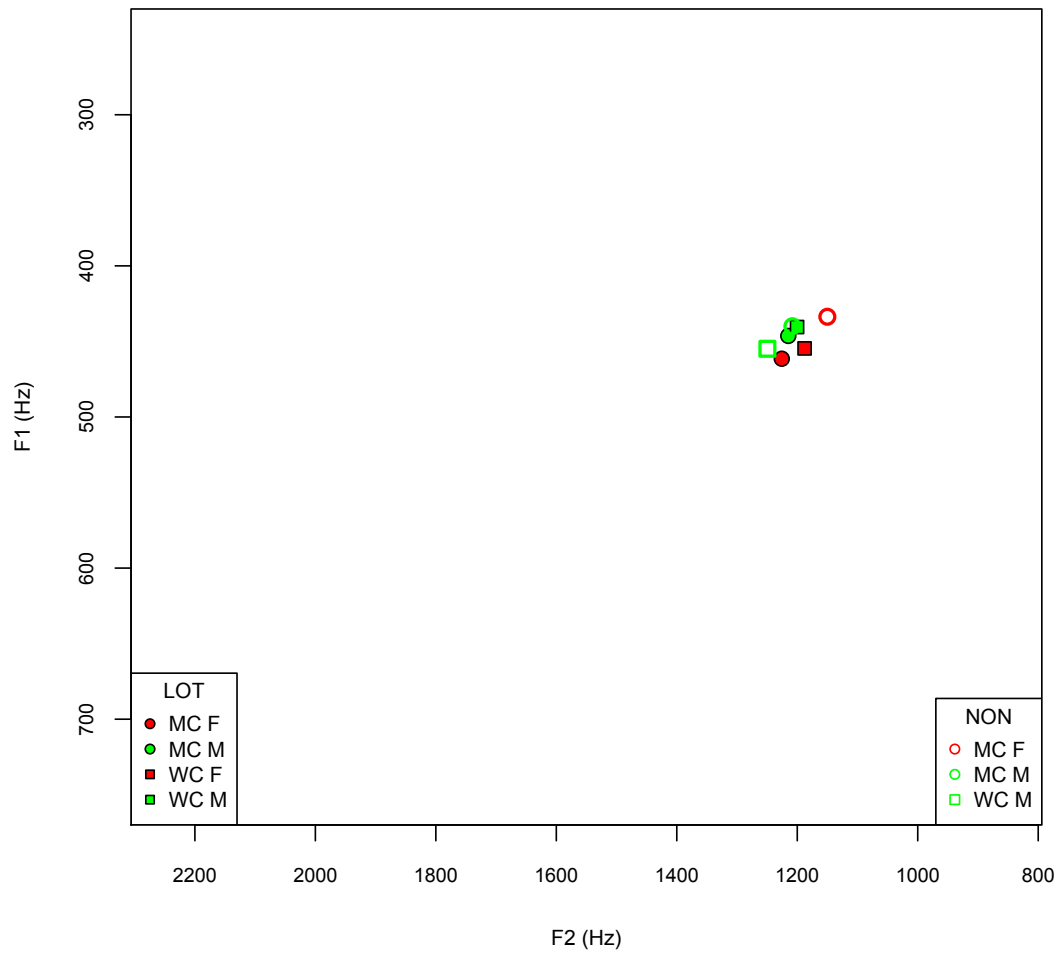


Figure 5.30: Mean values for LOT and *non*- tokens, disaggregated by class and gender.

c) Class and Style

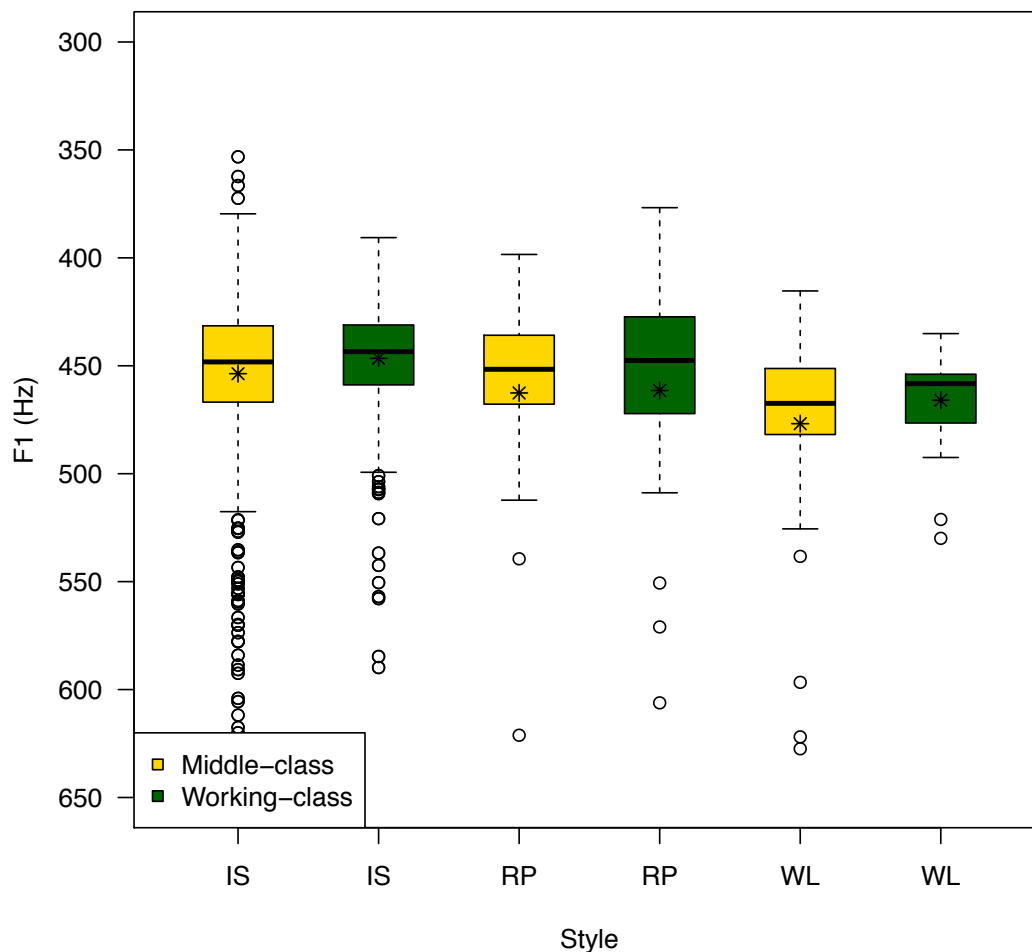


Figure 5.31: Boxplot for LOT F1 values in all speech styles, disaggregated by class.

For this set, in terms of F1, the class distinction is not significant ($p=0.2048$), although the boxplots in figure 5.31 show that working-class speakers use a slightly higher mean value in all three speech styles compared with middle-class speakers. Stylistically, however, both RP and WL styles are significantly lower than IS ($p=0.0033$ and $p<0.001$ respectively) as the boxplots in figure 5.31 show, particularly comparing the means.

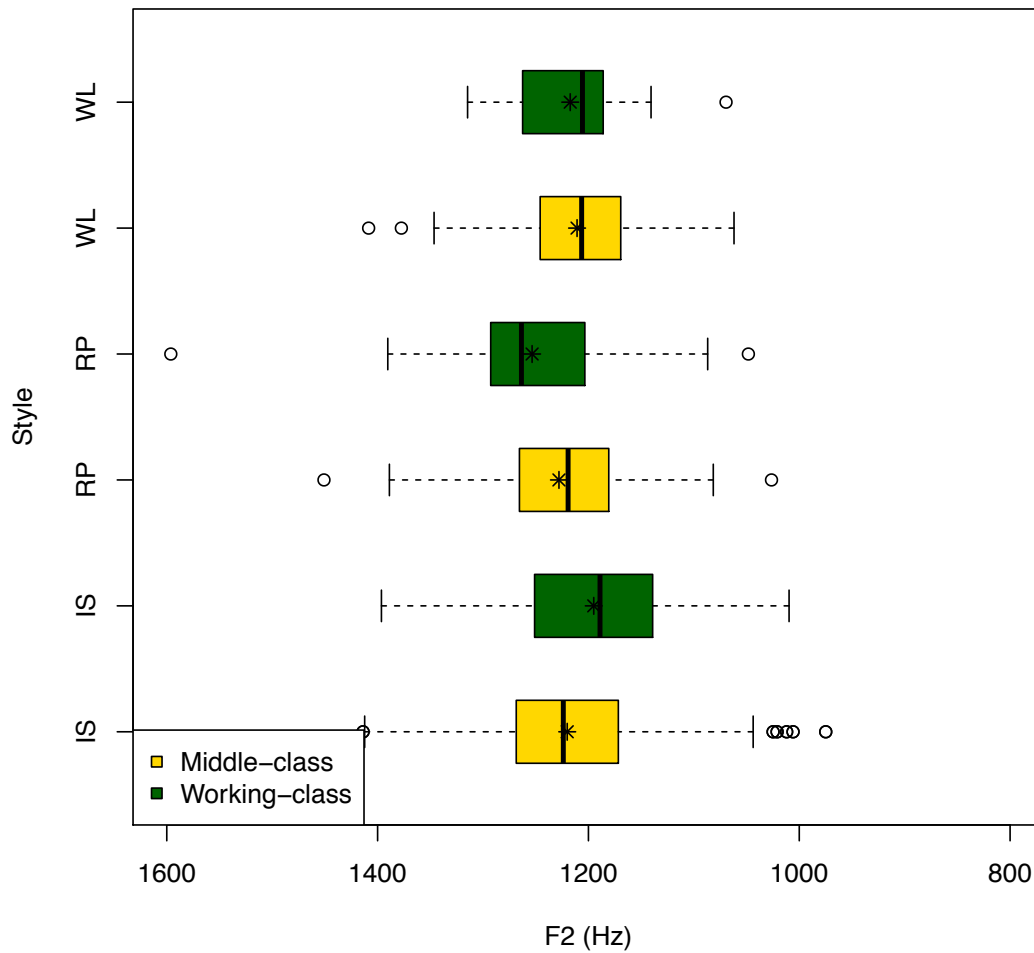


Figure 5.32: Boxplot for LOT F2 values in all speech styles, disaggregated by class.

Although the graphical evidence in figure 5.32 does not clearly show a difference between working- and middle-class speakers in terms of advancement for this set, the statistical test shows that there is indeed a significant difference ($p=0.042$). Since most of the tokens are in IS, this result is likely to be as a result of the backer range of values used by working-class speakers in IS. RP style is realised frontier than IS ($p=0.0001$), but the difference between IS and WL style is not significant (0.9549).

d) Gender and Style

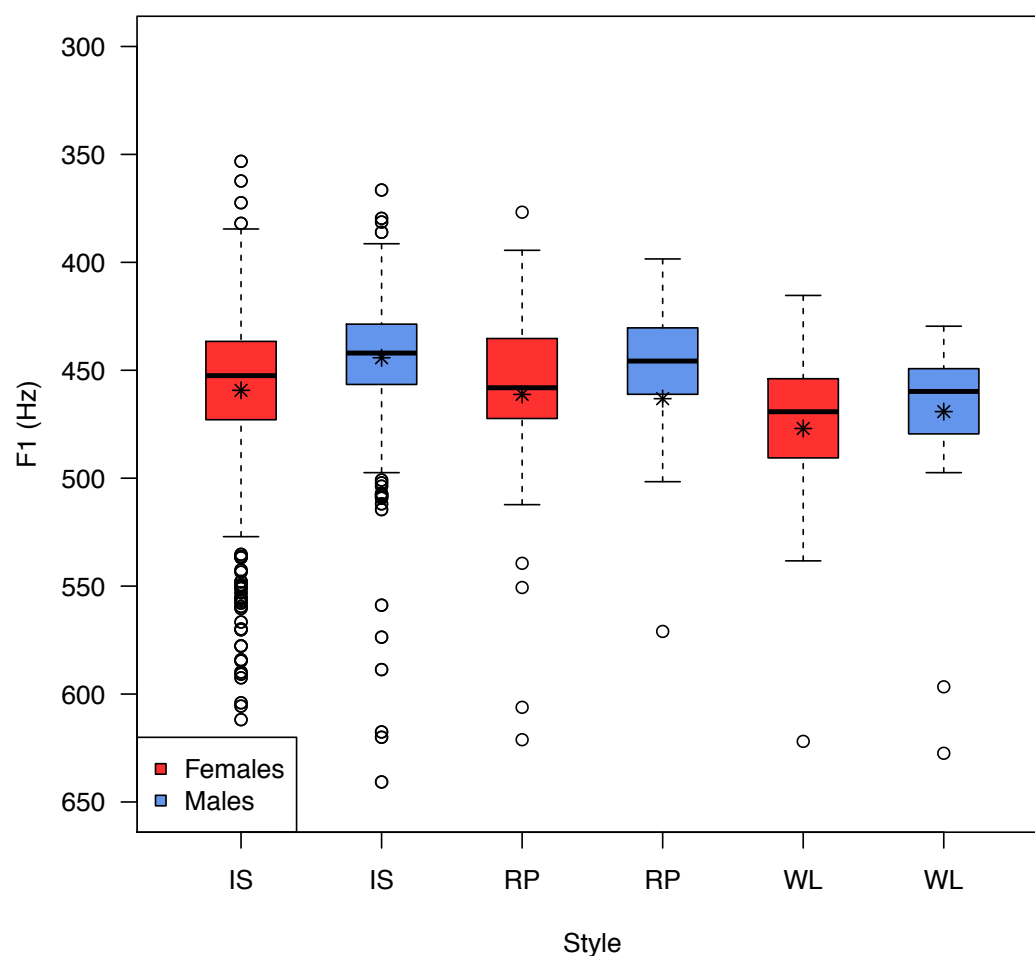


Figure 5.33: Boxplot for LOT F1 values in all speech styles, disaggregated by gender.

There is a significant gender difference between male and female speakers in terms of vowel height ($p=0.0247$): figure 5.33 shows that in IS and WL style, female speakers use a lower mean value than males.

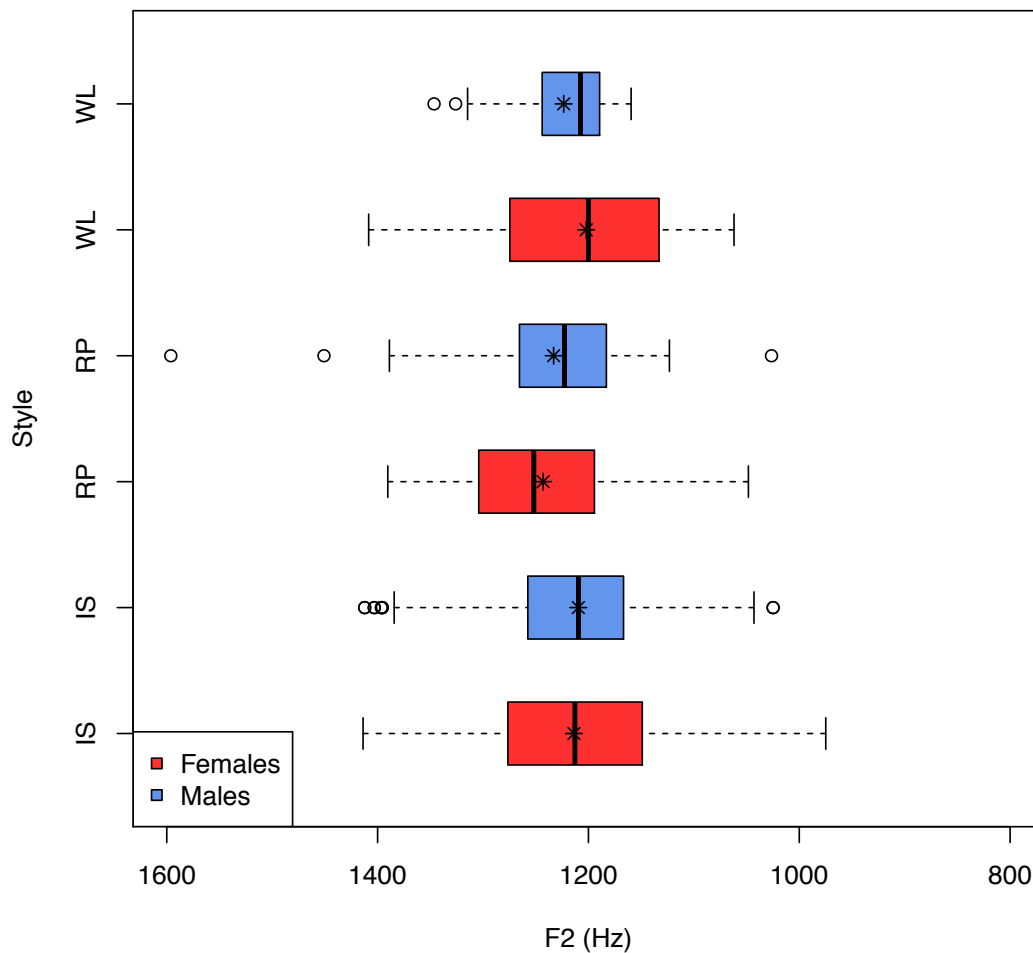


Figure 5.34: Boxplot for LOT F2 values in all speech styles, disaggregated by gender.

There is no significant gender distinction evident on the F2 plane ($p=0.8912$), although the boxplots in figure 5.34 show that females use a wider range of values in all three speech styles compared with the male speakers, despite similar means.

5.7.4 Summary: LOT Findings

The acoustic data show that the middle-class speakers use a slightly more centralised variant in IS than working-class speakers do. In the more careful speech styles, speakers use a lower variant than in IS. There is also evidence of female speakers using a lower variant than males. This contrasts with the findings of Wood (1987: 127), who suggested more of a height difference for this set. There is retraction in this set before velar /l/, even though it was not reported in the earlier literature on this variety, and the word *want* is indeed realised as more centralised (approximating [ʌ]) as Wood suggested. Findings for ‘non-’ words were inconclusive.

5.8 STRUT

5.8.1 WSAE Reports

Lass (2002: 115) reports STRUT to be a weak social marker, realised normally as central open [a] to [ä]. Backer and opener values are associated with Conservative and older General speakers, while higher and fronter values (as far front as [ɛ]) occur amongst younger, female General speakers. The nineteenth century input variety for SAE had a retracted STRUT vowel, as the fronting of this set only reached completion in England in the middle of the twentieth century. It therefore fronted in SAE as an endogenous development (Trudgill 2004: 135). Bekker (2009: 346) provides acoustic evidence of a fronted STRUT set for WSAE, and states that this vowel has not retracted as one might have expected, on account of the lowered TRAP vowel which he reports in his data. This is an expected retraction if in fact a SECS-like shift is in motion in SAE. There is thus some overlap between TRAP and STRUT in his data.

5.8.2 CSAE Reports

In CSAE, L2 speakers are reported to use either [a] or [ɐ], while L1 speakers use [ɐ] more consistently, with [ʌ] occurring sporadically (Finn 2004: 970; Wood 1987: 122). The *ONE*-subset (e.g. (-)one, once) is variably realised as [a] for L2 speakers and [ɐ] or [ɒ] for L1 speakers (Finn 2004: 970).

5.8.3 CSAE Acoustic Data

a) Overview

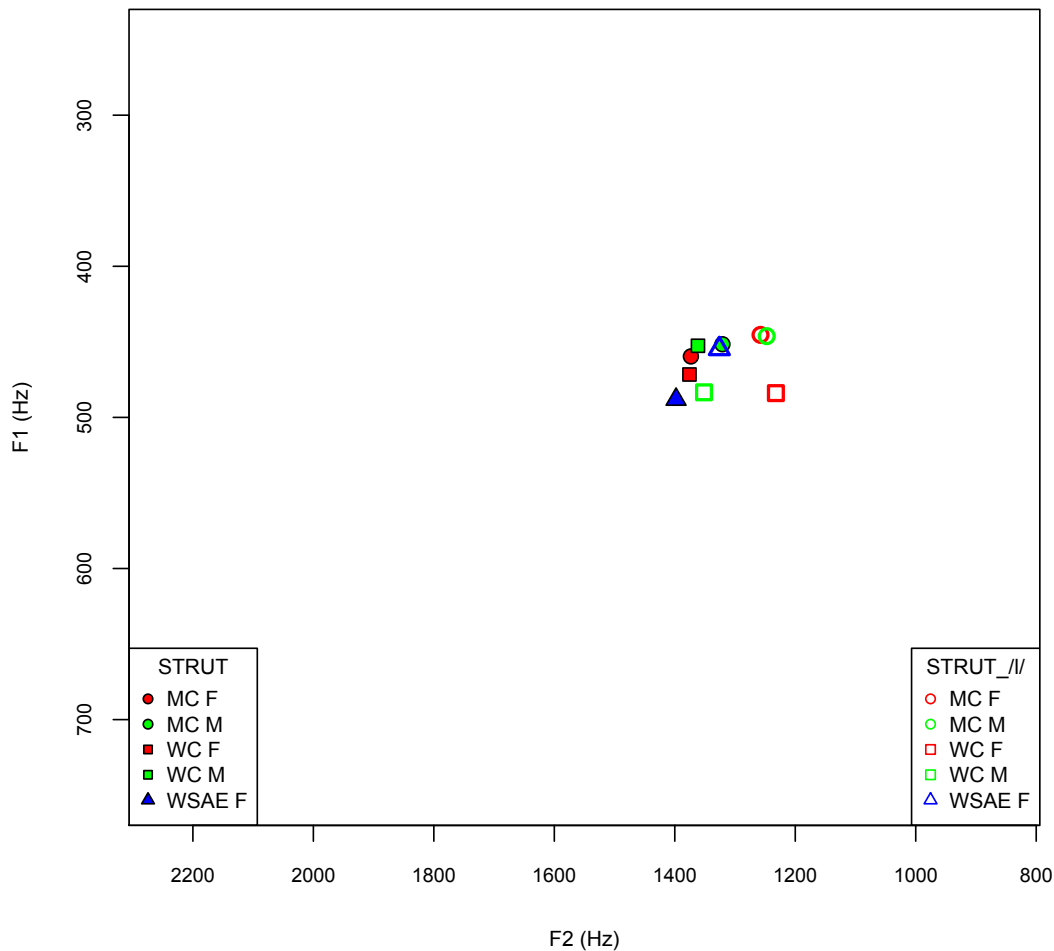


Figure 5.35: Mean values for STRUT and STRUT_/l/ in Interview Style, disaggregated by class and gender.

For the STRUT set, male speakers use a higher mean value than female speakers. The scatterplot (figure 5.35) shows that all CSAE speakers use a higher and slightly backer variant than Bekker's (2009) WSAE speakers. Generally it seems that this set is realised in the region of centralised [ɐ].

While the impressionistic accounts of CSAE do not mention retraction of the STRUT set before velar /l/, figure 5.35 shows some evidence of such retraction, but only for middle-class speakers. The difference between clean, IS style tokens and pre-/l/ tokens is significant on the F2 plane for middle-class speakers ($p < 0.001$), and borders on significant for working-class speakers ($p = 0.0576$). Working-class males produce a lower mean value pre-/l/, based on a rather small number of tokens ($n = 7$), so it is possible that the data do not reflect an accurate portrayal of the patterns of retraction

for this set before velar /l/. WSAE speakers use a higher and slightly retracted variant in pre-/l/ environment, as shown on figure 5.35.

b) Unique words

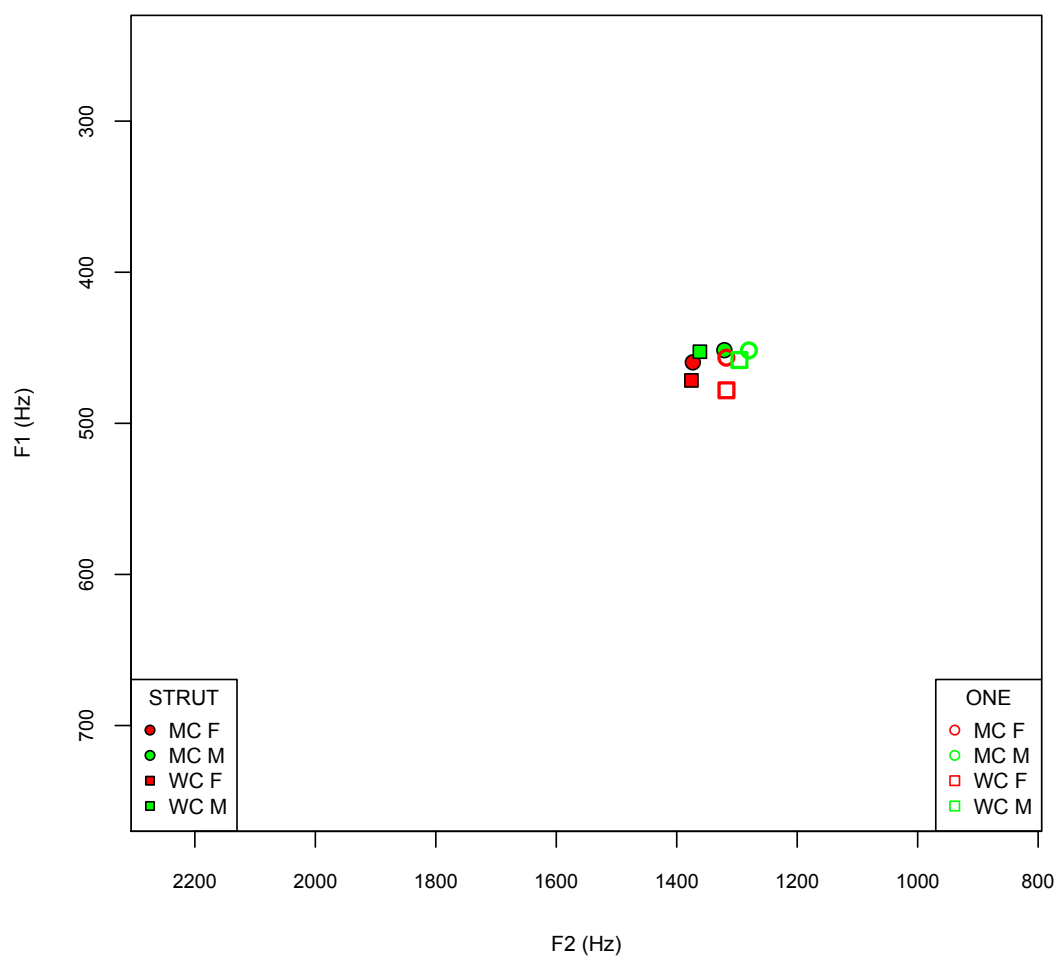


Figure 5.36: Mean values for STRUT and ONE in Interview Style, disaggregated by class and gender.

Figure 5.36 shows that as expected based on the report from Wood (1987), ONE words (which include 587 combined instances of the words *one* and *once*) are realised slightly backer than clean STRUT tokens for all speaker groups. This subset is not fronted to [a] or lowered to [ɒ] as Finn (2004: 970) reports.

c) Class and Style

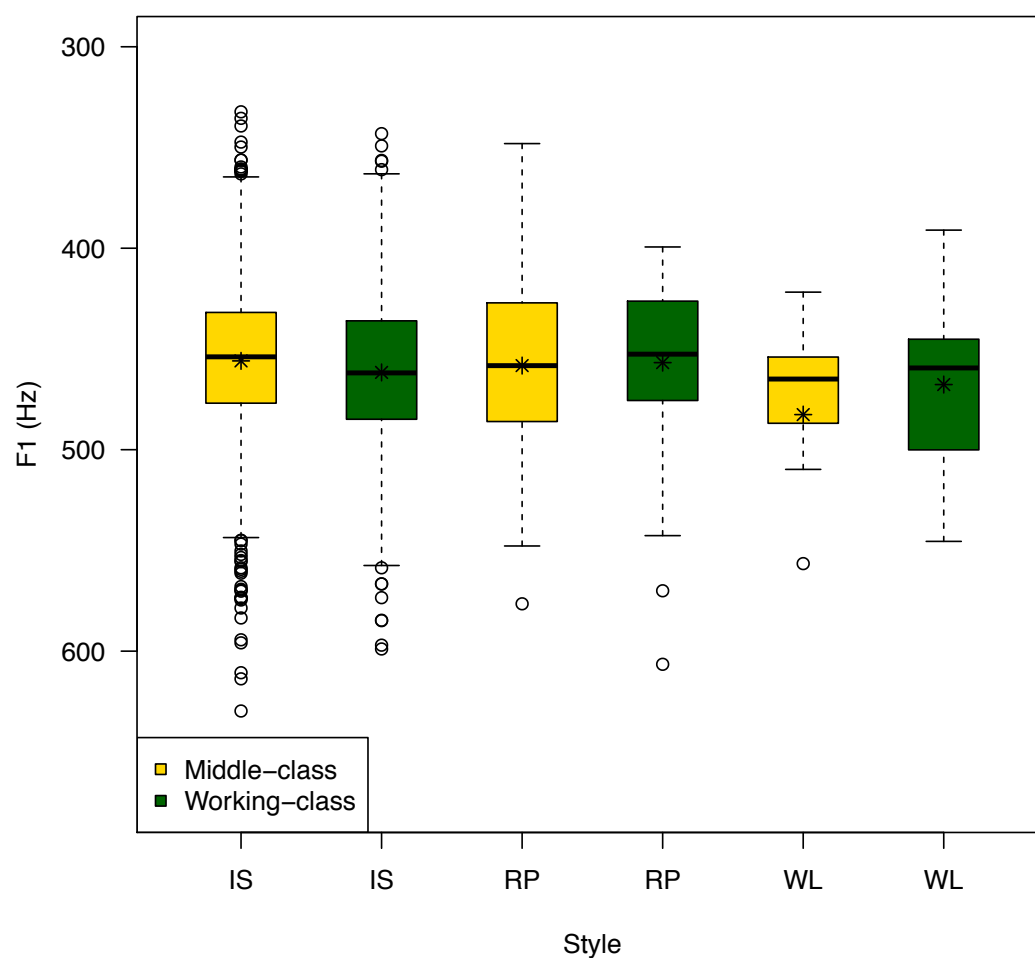


Figure 5.37: Boxplot for STRUT F1 values in all speech styles, disaggregated by class.

In terms of F1, figure 5.37 shows no significant distinction between speakers belonging to the different class categories ($p=0.2982$). Stylistically, WL style is realised slightly lower than IS ($p=0.0041$), but RP style shows no significant difference to IS ($p=0.8692$).

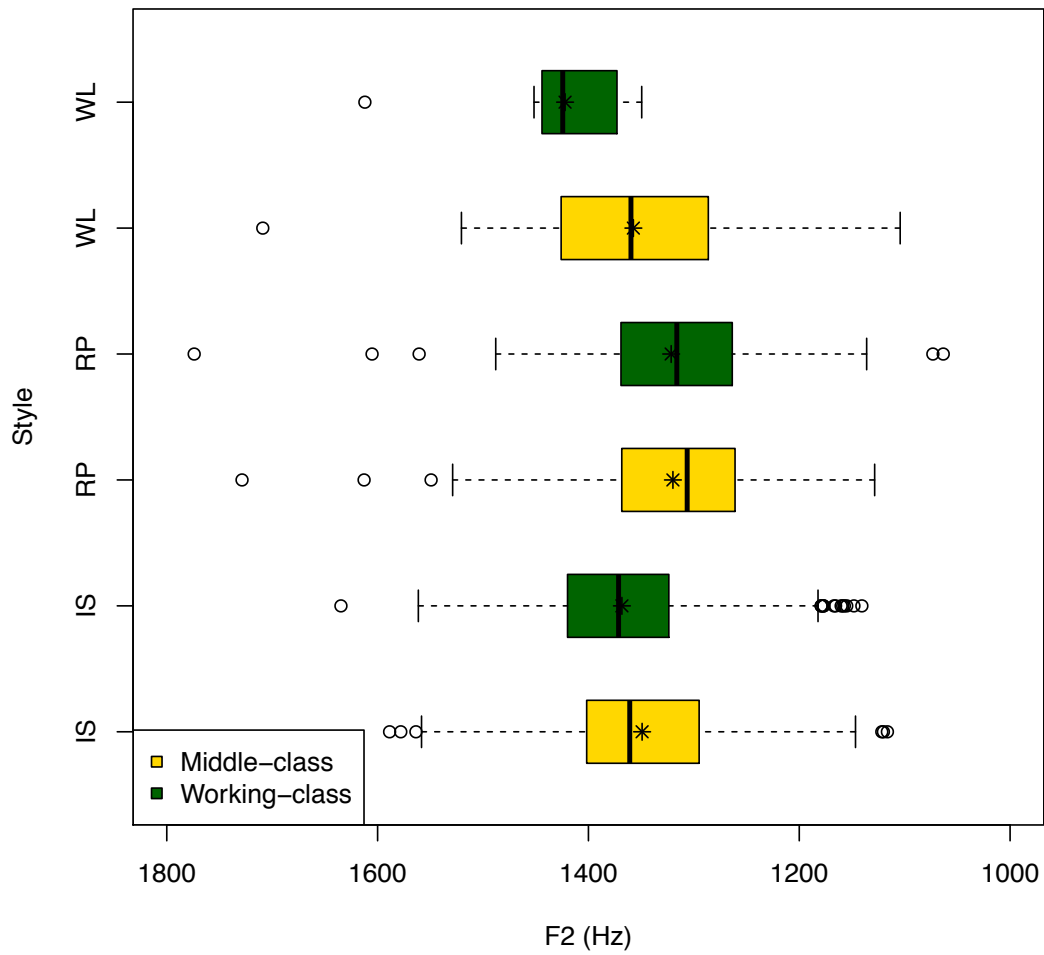


Figure 5.38: Boxplot for STRUT F2 values in all speech styles, disaggregated by class.

In terms of advancement, there is also no significant class distinction ($p=0.1538$). On this plane, however, RP style is realised significantly backer than IS ($p<0.001$). WL style is realised fronter than IS ($p=0.0219$) (see figure 5.38).

d) Gender and Style

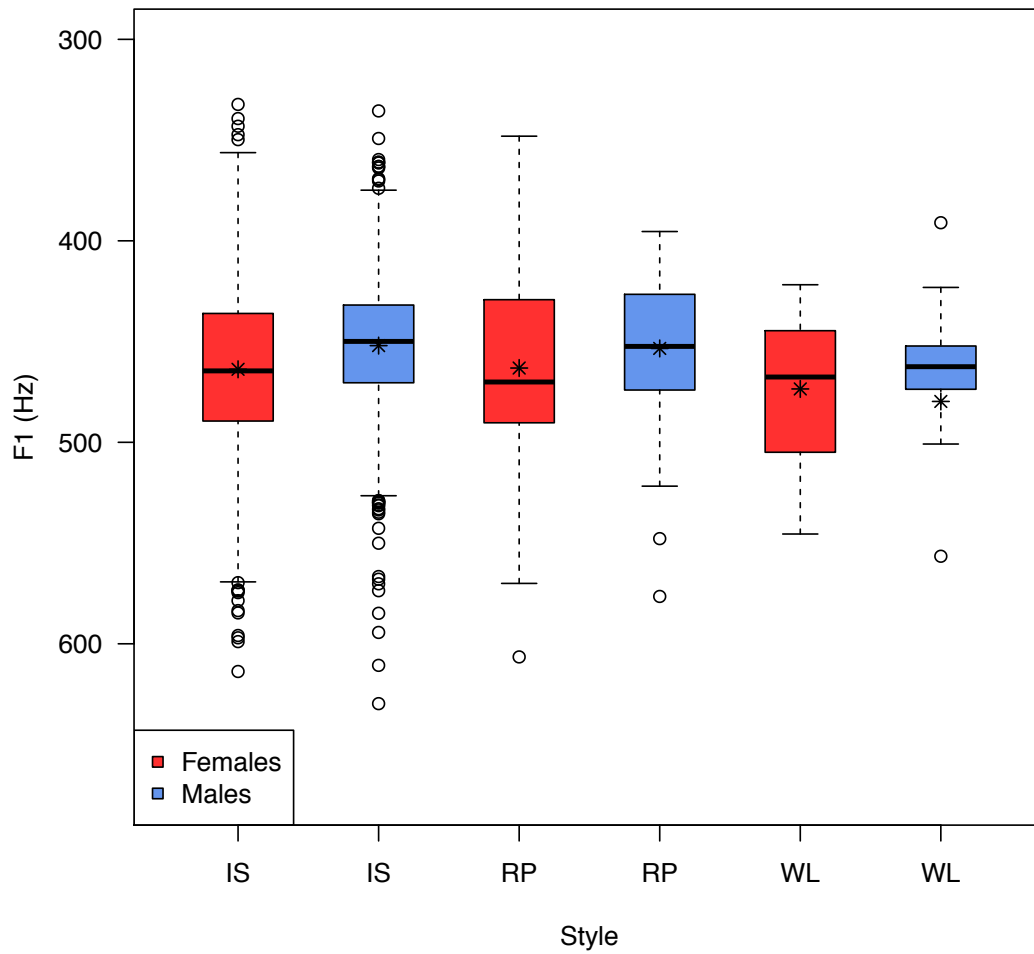


Figure 5.39: Boxplot for STRUT F1 values in all speech styles, disaggregated by gender.

This set displays a significant gender distinction for F1: males use a higher mean value than females ($p=0.0053$), most evident in the boxplots (figure 5.39) for IS and RP styles.

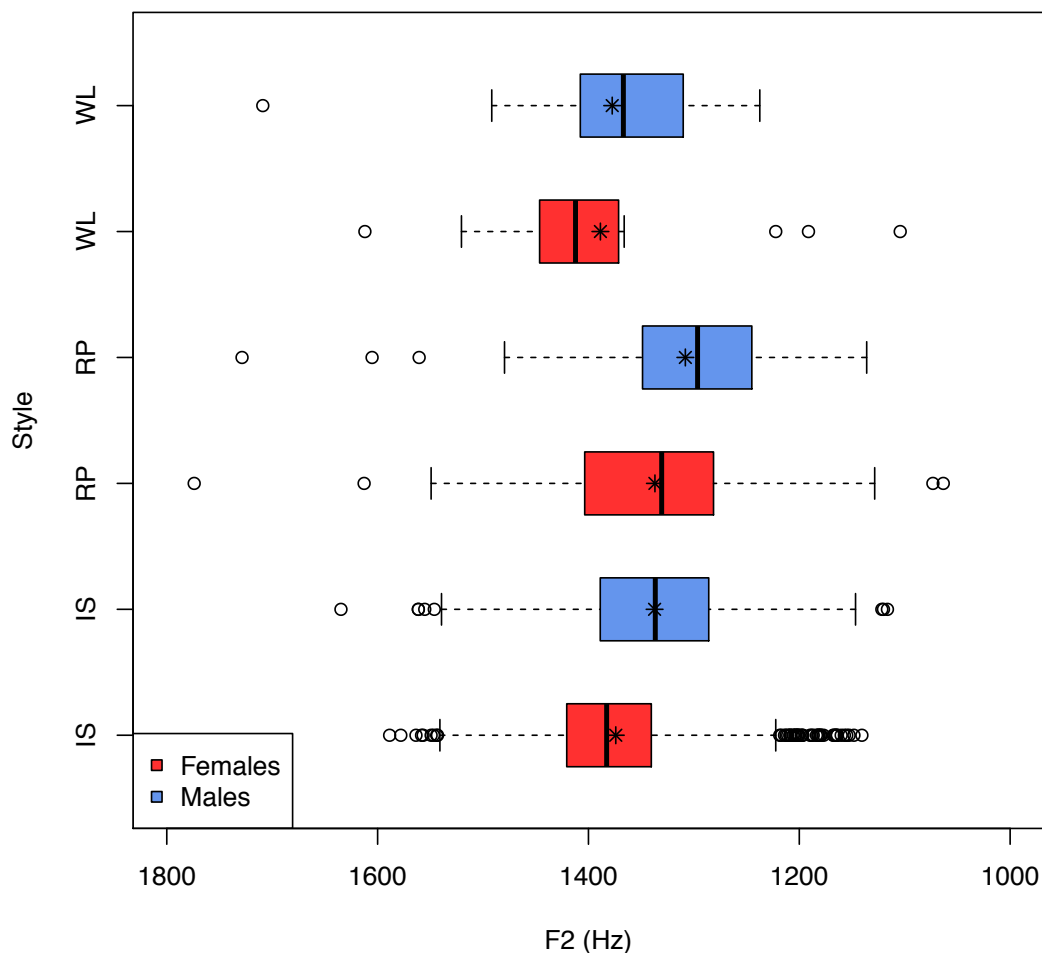


Figure 5.40: Boxplot for STRUT F2 values in all speech styles, disaggregated by gender.

Similarly on the F2 plane, there is a significant gender distinction: female speakers use a more centralised variant than males ($p=0.0067$), as figure 5.40 shows.

5.8.4 Summary: STRUT Findings

For this set, gender is a significant factor: females use a fronter and lower variant than male speakers, which corresponds with the lower, more centralised variant used by Bekker's (2009) WSAE female speakers. There is no clear class distinction coming through in the acoustic data. Stylistically, WL style is realised slightly lower and fronter than IS, which supports the general trend that higher prestige variants are used in more formal speech styles. RP style is significantly backer than IS, which is an anomalous finding.

The acoustic data for middle-class speakers shows some retraction before velar /l/, but the tokens available for a comparative analysis for working-class speakers (particularly females) make the findings inconclusive. A subset of *one* words are realised backer than clean tokens of the STRUT set.

5.9 FOOT

5.9.1 WSAE Reports

Lass (2002: 115) reports this set to be realised as a centralised, high-mid back vowel [ʊ], or a variant slightly fronter than this. Younger female General speakers may realise this set with a fronter, lower value, approximating [ɥ]. Bekker (2009: 253) reports a fronted FOOT set, which is in keeping with Lass' (2002) findings for young female speakers, as it is precisely these speakers upon which Bekker's acoustic account is based. Bekker (2009: 355) further reports a diphthongal quality to this set, with the glide moving to a fronted position.

5.9.2 CSAE Reports

Earlier accounts of CSAE report that this set is typically realised in the region of [u] i.e. very back and rounded. L1 speakers use a centralised variant [ɥ] and also [ɤ] (Finn 2004: 970). This set is not reported to be affected by following velar /l/.

5.9.3 CSAE Acoustic Data

a) Overview

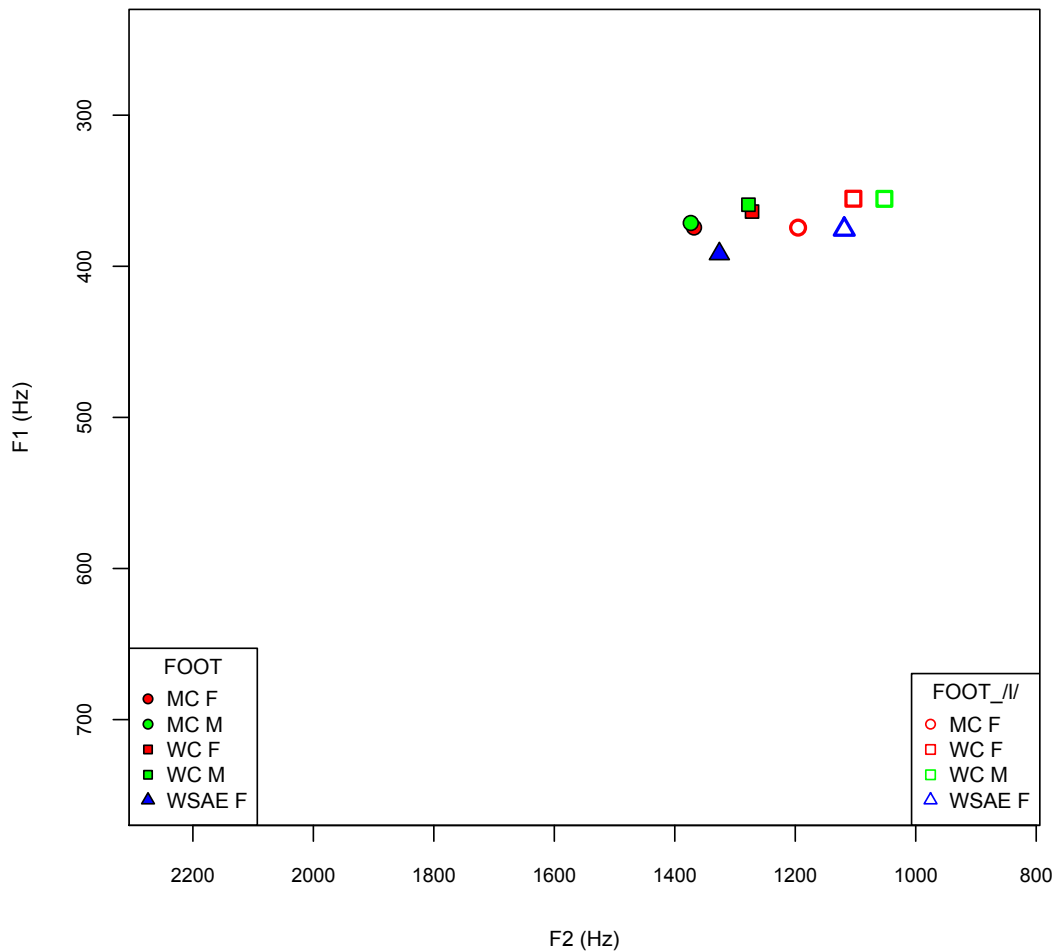


Figure 5.41: Mean values for FOOT and FOOT_/l/ in Interview Style, disaggregated by class and gender.

The general picture for FOOT (figure 5.41) seems to show that working-class CSAE speakers use a backer and slightly higher variant than middle-class speakers. This supports Wood's (1987) findings regarding the back, rounded variant historically used in the Coloured community. Middle-class speakers use a variant that is higher than Bekker's (2009) WSAE speakers, and also slightly more centralised, which supports Finn's (2004) findings regarding the use of [ɥ], although a lower [ɣ] does not seem to be as prevalent as he suggests.

Despite reports from the CSAE scholars that this set is not affected by following velar /l/, the acoustic data suggest otherwise (see figure 5.41). The pre-/l/ values are based on very few tokens (n=7), however, and there were none produced by middle-class males. The small number of tokens shows clear retraction for working-class speakers,

as well as middle-class females. This suggests that as for Bekker's (2009) WSAE speakers, retraction for the FOOT set also occurs before velar /l/ in CSAE.

b) Class and Style

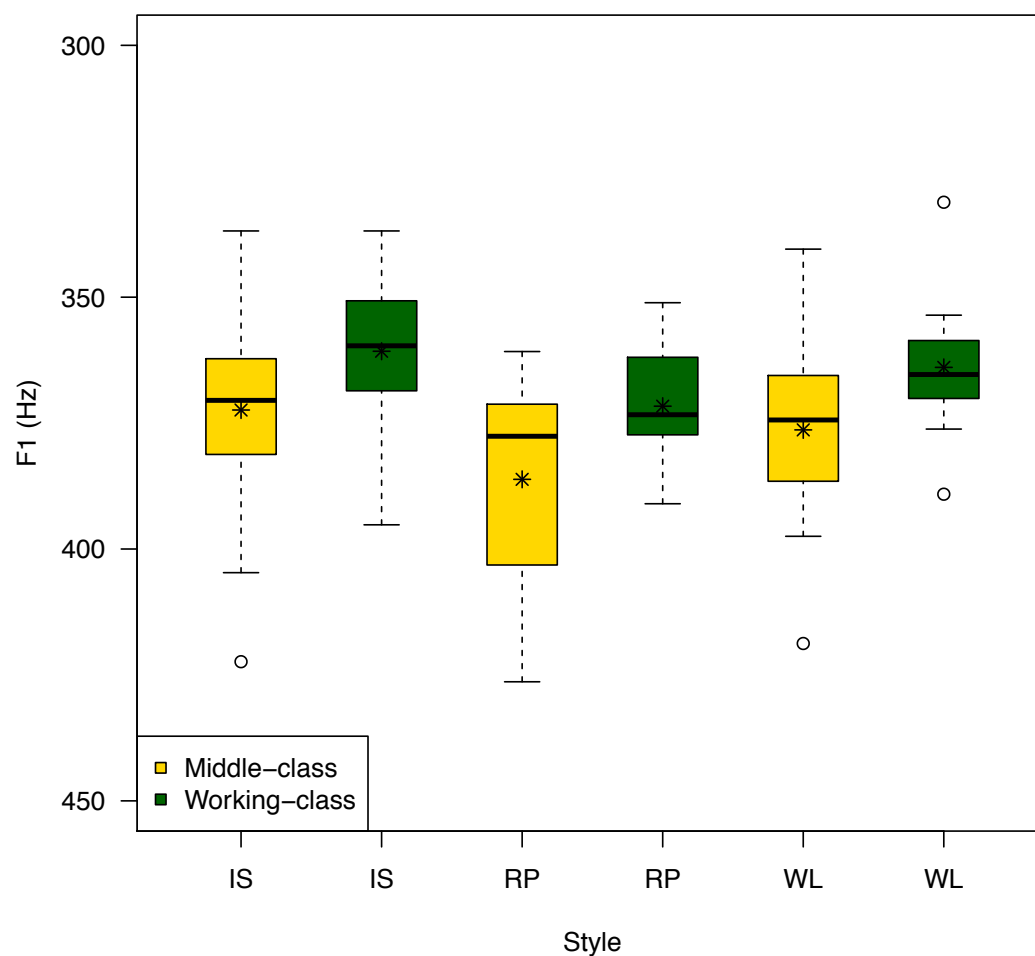


Figure 5.42: Boxplot for FOOT F1 values in all speech styles, disaggregated by class.

For the FOOT set, working-class speakers produce significantly higher values than do middle-class speakers ($p=0.0006$), as is evident in the boxplots in figure 5.42. RP style is realised lower than IS tokens ($p=0.0001$), but there is no significant difference between WL style and IS in terms of vowel height ($p=0.2739$).

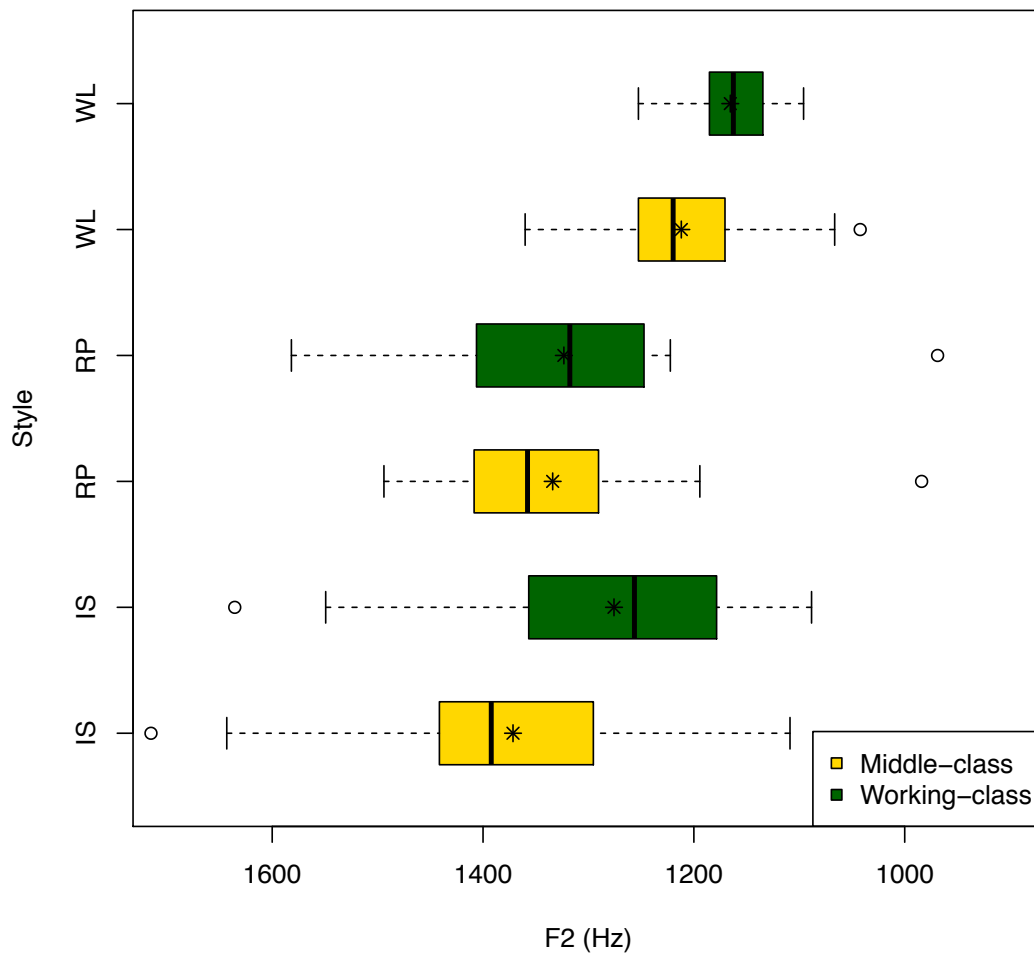


Figure 5.43: Boxplot for FOOT F2 values in all speech styles, disaggregated by class.

Class is also a significant factor in terms of advancement, as is evident on the boxplots in figure 5.43: working-class speakers use a significantly backer range of values than middle-class speakers for this set ($p=0.0004$). This is most evident in IS, which makes up the bulk of the tokens, but in the two more careful styles, the mean values for working-class speakers are also backer than for middle-class speakers. On the F2 plane, RP style is not significantly different to IS ($p=0.9217$) but WL style is produced significantly backer than the IS tokens ($p<0.001$).

c) Gender and Style

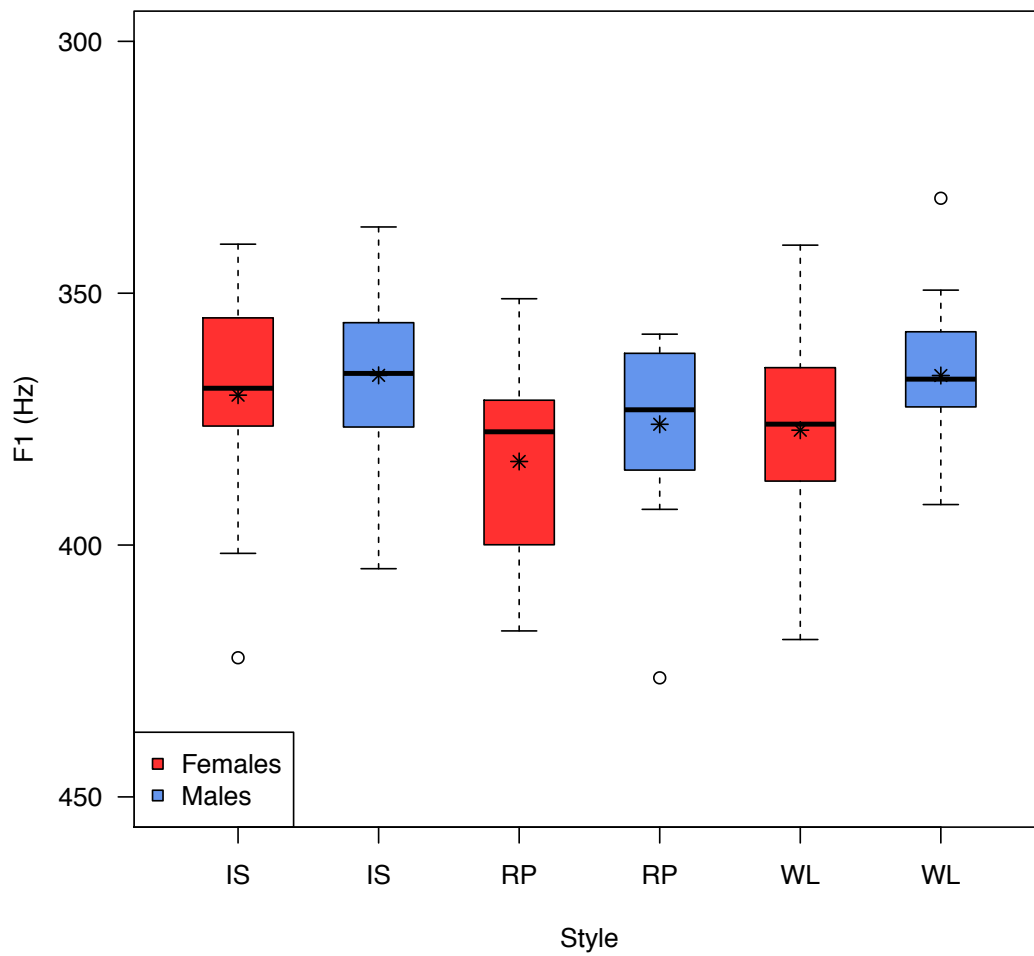


Figure 5.44: Boxplot for FOOT F1 values in all speech styles, disaggregated by gender.

In terms of vowel height, there is no significant gender distinction between male and female speakers for this set ($p=0.1052$), although the boxplots in figure 5.44 show that in RP and WL styles, females use a lower mean value and a higher range of values than male speakers.

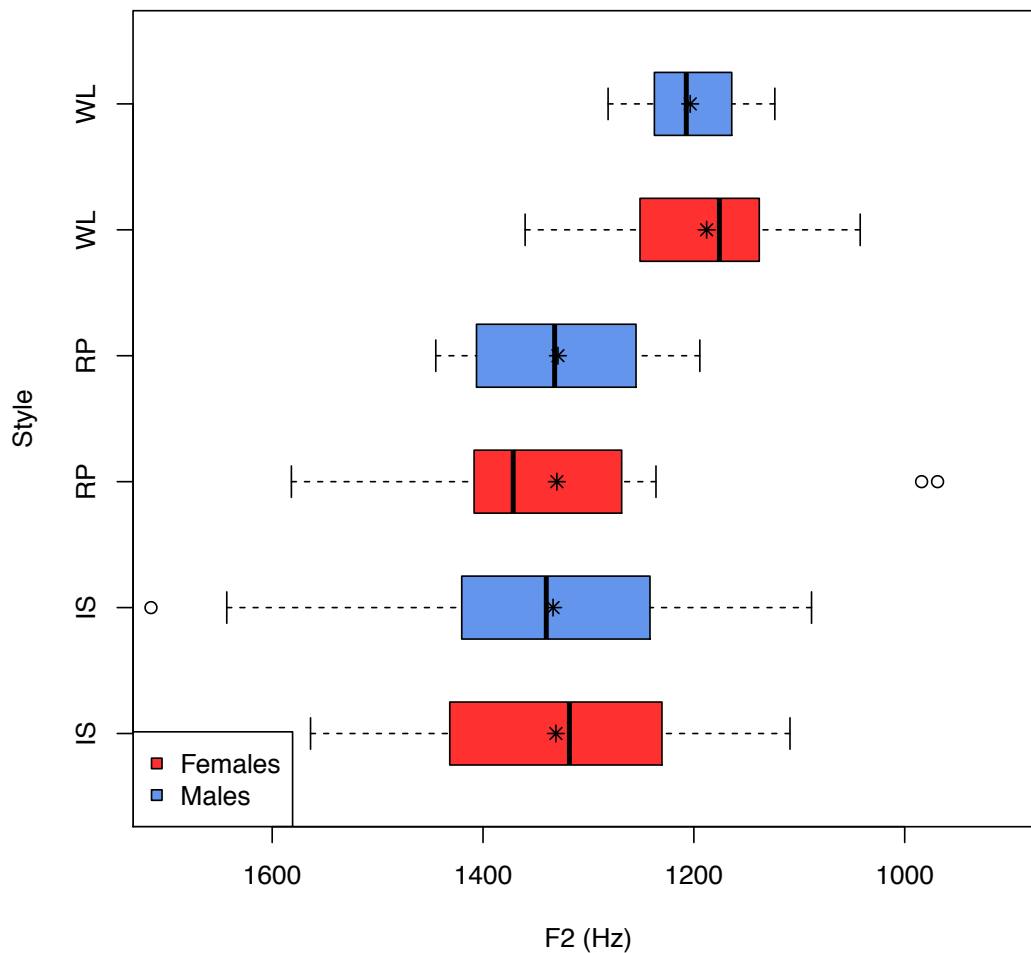


Figure 5.45: Boxplot for KIT F2 values in all speech styles, disaggregated by gender.

Similarly, on the F2 plane, there is no significant gender distinction ($p=0.7889$), with very similar mean values for males and females in each speech style (see figure 5.45).

5.9.4 Summary: FOOT Findings

The acoustic data show that working-class speakers use a higher and backer variant than middle-class speakers. The use of the more centralised variant by middle-class speakers mirrors the trend amongst WSAE speakers (Bekker 2009). In RP style, speakers use a lower vowel, and in WL style, they tend to use a backer vowel than in IS. This set does not display a significant gender distinction for either vowel height or advancement.

5.10 Conclusion

An analysis of the six short vowels were presented in this chapter, comparing aural and acoustic accounts of WSAE and traditional aural accounts of CSAE to the acoustic data from the present study. A summary of the significant findings are presented in table 5.5 below.

	Class	Gender	Speech Style	Velar /l/	IPA value*
KIT	Working-class speakers use higher and fronter IT-subset, and lower SIT-subset.	No significant gender distinction.	WL style realised backer and lower for the SIT-subset compared with IS.	No retraction in either the SIT or IT subsets.	IT: [iː]~[ɪ] SIT: [ɪ]~[ə]
DRESS	Working-class speakers use a fronter variant than middle-class speakers.	No significant gender distinction.	WL style realised lower and fronter than IS.	Clear retraction.	[e]
TRAP	Middle-class speakers use a lower variant. Working-class speakers retain a traditional raised variant.	No significant gender distinction.	WL style realised lower than in IS.	Inconclusive whether or not retraction occurs in this set.	[ɛ]~[æ]
LOT	Working-class speakers use a backer variant than middle-class speakers.	Females use a lower vowel than males.	WL and RP styles realised lower than IS.	Retraction occurs.	[ɔ]
STRUT	Class difference was not significant for this set.	Males use a higher and backer vowel than females.	WL style realised lower and fronter than IS; RP style realised backer than IS.	Middle-class speakers retract this set.	[ʊ]
FOOT	Working-class speakers use a higher and backer vowel than middle-class speakers.	No significant gender distinction.	WL style realised backer than IS; RP style realised lower than IS.	Retraction occurs.	[ʊ]~[ö]

Table 5.5: Summary of acoustic findings for the short vowels of CSAE.

* This value is approximate.

Stylistically, the results for RP style were not what one would expect: as a more careful style than IS, but less so than WL style, RP style is generally expected to be realised somewhere between these other two styles. In the acoustic data, RP style was often not statistically significantly different to IS, and in a few cases, displayed the opposite trend to WL style. Much of the acoustic data supported the findings of the earlier CSAE scholars.

In the following chapter, a similar detailed analysis is provided for the long monophthongs: BATH, NURSE, FLEECE, THOUGHT, GOOSE and the monophthongised diphthong SQUARE.

Chapter Six

Long Monophthongs

6.1 Introduction

The analysis of the five long monophthongs, viz. BATH, NURSE, FLEECE, THOUGHT and GOOSE are presented in this chapter. In addition, a monophthongised diphthong, SQUARE, is presented as the only vowel in CSAE and SAE more generally that is not considered a traditional monophthong in English studies, and Received Pronunciation specifically. The reasons for the inclusion of SQUARE are discussed below. The chapter is structured as follows: section 6.2 provides an overview of all the long vowels that are discussed in the chapter, indicating their position relative to one another in the vowel space. Sections 6.3 through 6.7 discuss, in turn, each of the long monophthongs and section 6.8 provides a description of SQUARE. A tabular summary of the findings is presented in section 6.9, followed in section 6.10 by an overview of the full simple vowel system of CSAE (i.e. long and short vowels), based on the acoustic data in this study.

As in the previous chapter, each subsection dealing with a lexical set (i.e. 6.3 to 6.8), is structured as follows: Firstly, an account of the trends for each vowel in WSAE is provided, based on reports by Lass (2002) and Bekker (2009). The latter is an acoustic account of young female WSAE speakers. This is followed by a summary of the aural reports on CSAE (detailed in chapter two), and the acoustic analysis of the present sample of speakers. The acoustic analysis includes an overview of the data, and then a more detailed disaggregation of the data into formant, class, gender and speech style, in which statistical results are provided. Each section is concluded with a discussion of the findings in relation to the WSAE and CSAE reports.

6.2 Acoustic Overview of the Long Monophthongs of CSAE

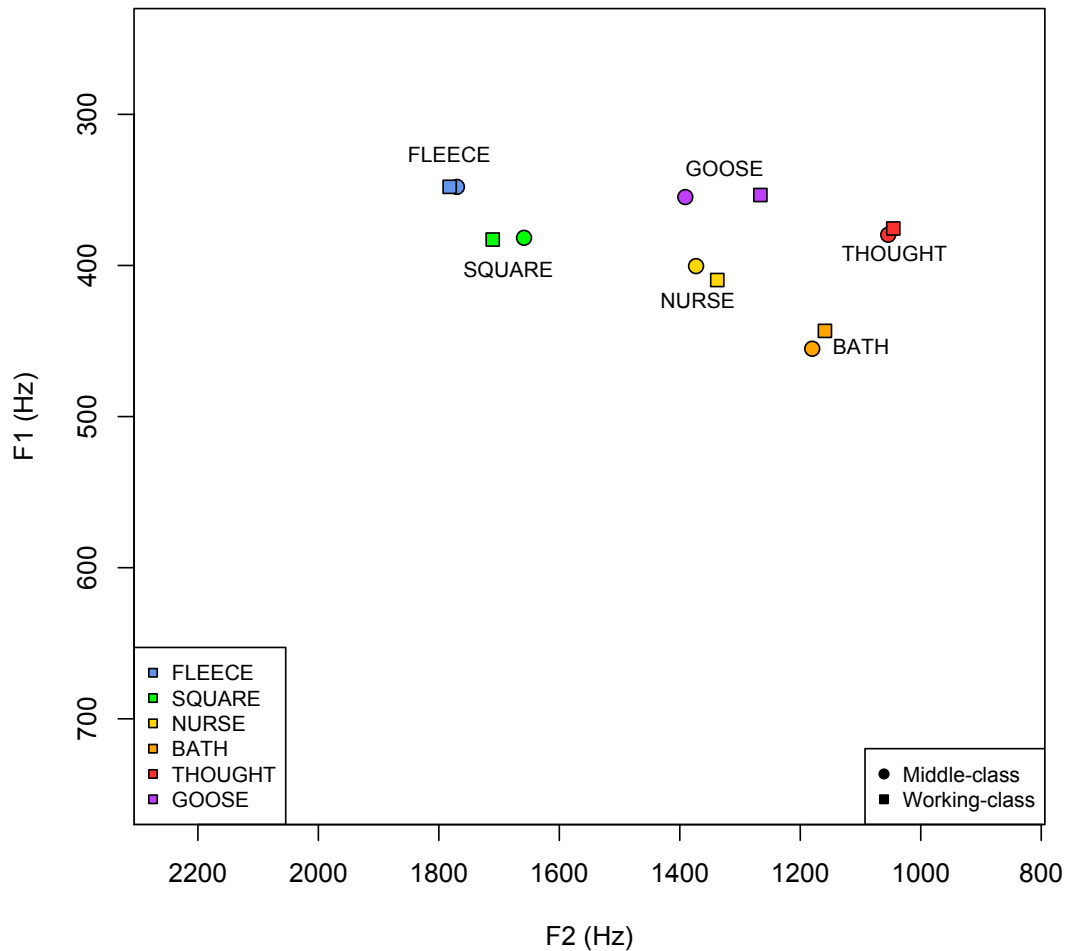


Figure 6.1: An acoustic overview of the long vowel system of CSAE, showing mean values for middle- and working-class speakers.

In order to ascertain the position of each of the long vowels relative to one another, an overview of the acoustic data for the CSAE speakers in this study is provided here. Figure 6.1 provides two mean values for each of the long vowels in the CSAE system, one each for middle- and working-class speakers. The F1 axis runs from 250Hz to 750Hz, which is the range of the normalised data on this plane, while the F2 axis runs from 850Hz to 2,250Hz. The mean values fall within a much smaller area of the plot, however: for F1 all points fall between 300Hz and 500Hz, and between 900Hz and 2,000Hz for F2. For each vowel, two points are plotted: one for each social class category, i.e. middle-class and working-class. Some class variation is evident for all of the lexical sets, save for FLEECE and THOUGHT, but the distinction is most notable in the GOOSE set, where middle-class speakers use a more centralised variant than the working-class speakers. Class differences are discussed in the subsection relevant to each vowel. Each vowel seems to occupy a unique position within the vowel space.

In the sections that follow, each of the long monophthongs will be analysed in detail, beginning with the FLEECE vowel.

6.3 FLEECE

6.3.1. WSAE Reports

Lass (2002: 116) reports that speakers of all lects of WSAE realise this set with a long, close [i:] with no social variation. There is no evidence of diphthongisation, as occurs in other Southern Hemisphere varieties like Australian English (AusE) and New Zealand English (NZE). Bekker (2009: 235) concurs, showing acoustically that FLEECE for his speakers is consistently realised as a high, front monophthong.

6.3.2 CSAE Reports

The reports on CSAE indicate some variation for this set, however. While [i:] dominates the distribution (with optional length), diphthongised variants [iə] and [iʊ] (with an optionally lengthened first element) are reported to occur (Finn 2004: 971), although their distribution is not described. Wood (1987: 128) suggests that FLEECE may become weakly diphthongised before velar /l/.

6.3.3 CSAE Acoustic Data

a) Overview

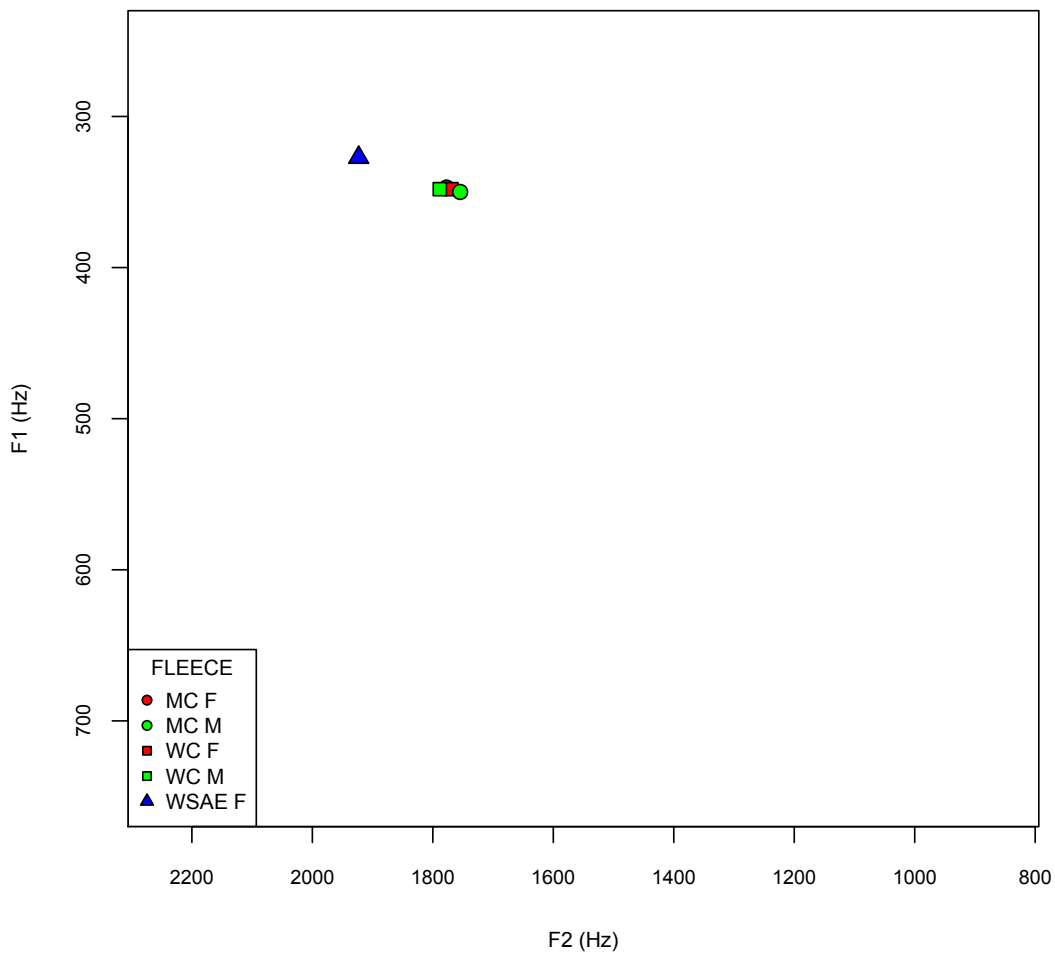


Figure 6.2: Mean values for FLEECE in Interview Style (IS), disaggregated by class and gender.

The acoustic data in figure 6.2 shows a mean value for each group of CSAE speakers, as well as a mean value for Bekker's (2009) WSAE speakers for FLEECE. The set shows almost no variation for the different CSAE speaker groups, who use a significantly backer and slightly lower mean value compared with the WSAE speakers. This is likely to be due to the fronting GOOSE set amongst WSAE speakers, which has pushed FLEECE fully front. GOOSE is not as fronted for CSAE speakers (see section 6.5 below), which could account for why FLEECE is not fully fronted.

	F1	F1_/l/	Difference in F1	F2	F2_/l/	Difference in F2
MC F	347	353	6	1777	1732	45
WC F	348	352	4	1769	1735	34
MC M	349	354	5	1754	1725	29
WC M	348	354	6	1788	1798	10
WSAE F	327	351	44	1923	1808	115

Table 6.1: Comparative mean values for F1 and F2 in regular and pre-/l/ tokens of FLEECE (Hertz).

With regards to retraction of this set before velar /l/, the mean F2 values for regular and pre-/l/ tokens of FLEECE, showing one value for each speaker group (table 6.1) shows that the set, in fact, does not retract⁴⁶. For all CSAE speaker groups, the difference between the F2 values is negligible (less than 50Hz), suggesting that retraction does not occur in this set. By contrast, the WSAE speakers in Bekker's (2009) study show some retraction, with a difference of 115Hz on the F2 plane between regular and pre-/l/ tokens.

While the particular methodology employed to extract formant measurements does not allow for any analysis of the trajectory of the vowel, in order to provide preliminary acoustic findings with regards to diphthongisation of this lexical set, six tokens of FLEECE were randomly selected from four of the speakers: one male and one female from each of the two social class categories. Two manual readings were taken from each token – one for the nucleus of the vowel, and the other for the (would-be) glide. The readings are provided in table 6.2 below.

⁴⁶ Tables have been provided for the lexical sets that do retract before velar /l/ to show the difference in the mean values for regular and pre-/l/ tokens. Because the values were so similar for regular and pre-/l/ tokens, their representation on a scatterplot was unclear because many of the points overlapped.

	F16 (MC)			F13 (WC)			M6 (MC)			M15 (WC)		
	1	2	D	1	2	D	1	2	D	1	2	D
F1	607	594	-7	480	448	-32	296	286	-10	376	378	2
F2	2960	2853	-	3045	2986	-59	2003	1979	-24	1836	1877	41
			107									
F1	439	469	30	521	556	35	256	300	44	330	338	8
F2	2683	2652	-31	2829	2783	-46	2017	2024	7	2060	1983	-77
F1	419	397	-22	506	456	-50	364	340	-24	355	360	5
F2	2734	2695	-39	2914	2853	-61	2040	1868	-	1911	2010	99
									172			
F1	341	383	42	439	420	-19	365	376	11	347	371	24
F2	2844	2884	40	2454	2567	113	1885	1916	31	2111	1905	-
												206
F1	470	480	10	409	435	26	327	340	13	337	368	31
F2	2652	2598	-54	2947	2985	38	1943	1993	50	2069	2049	-20
F1	454	453	1	429	420	-9	265	257	-8	359	393	34
F2	2654	2711	111	2920	2906	-14	2025	1958	-67	2160	2073	-87

Table 6.2: Comparative formant values of six tokens of FLEECE for four speakers (Hertz; unnormalised).

Key: 1=nucleus reading; 2=glide reading; D=difference.

The reported diphthongisation of this set (Finn 2004: 971) suggests that the direction of the trajectory is towards [ə] or [ʊ], which corresponds with higher F1 and lower F2 values in the case of [ə], and lower F2 values in the case of [ʊ] being the endpoint of the glide. None of the readings for the second F1 values had a difference of more than 100Hz, which would have been considered significant. Where the values for the second F2 readings are more than 150Hz less than the first reading, they are underlined to indicate that there is a significant glide. Only two of the 24 tokens (eight percent) had a slight backwards trajectory, which leads to the conclusion that this set is generally monophthongal, with little acoustic evidence to support the notion that diphthongisation occurs. Of course, this can only be confirmed with an acoustic study which investigates this phenomenon with significantly more tokens than have been analysed presently.

b) Class and Style

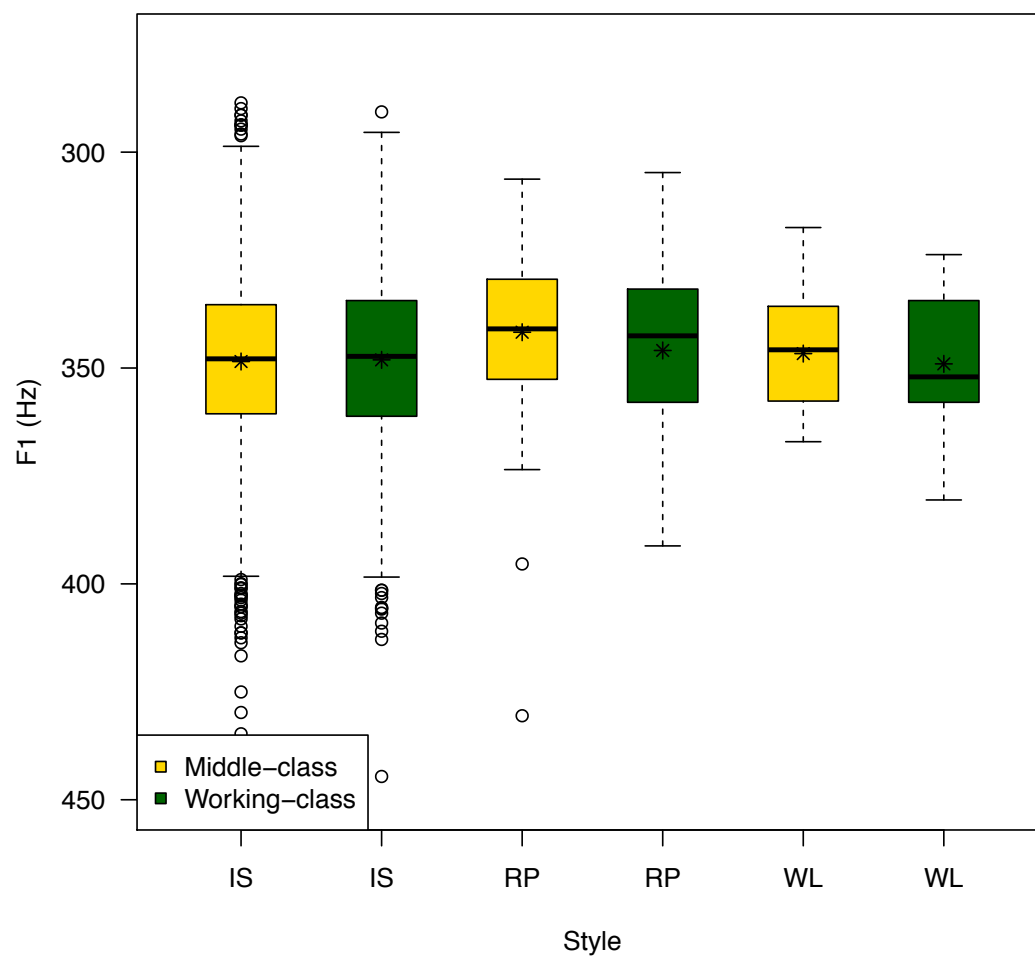


Figure 6.3: Boxplot for FLEECE F1 values in all speech styles, disaggregated by class.

On the F1 plane, there is no clear distinction between middle- and working-class speakers ($p=0.8803$), as figure 6.3 shows. RP style is realised significantly higher than IS ($p=0.0022$) but WL style is not ($p=0.8669$).

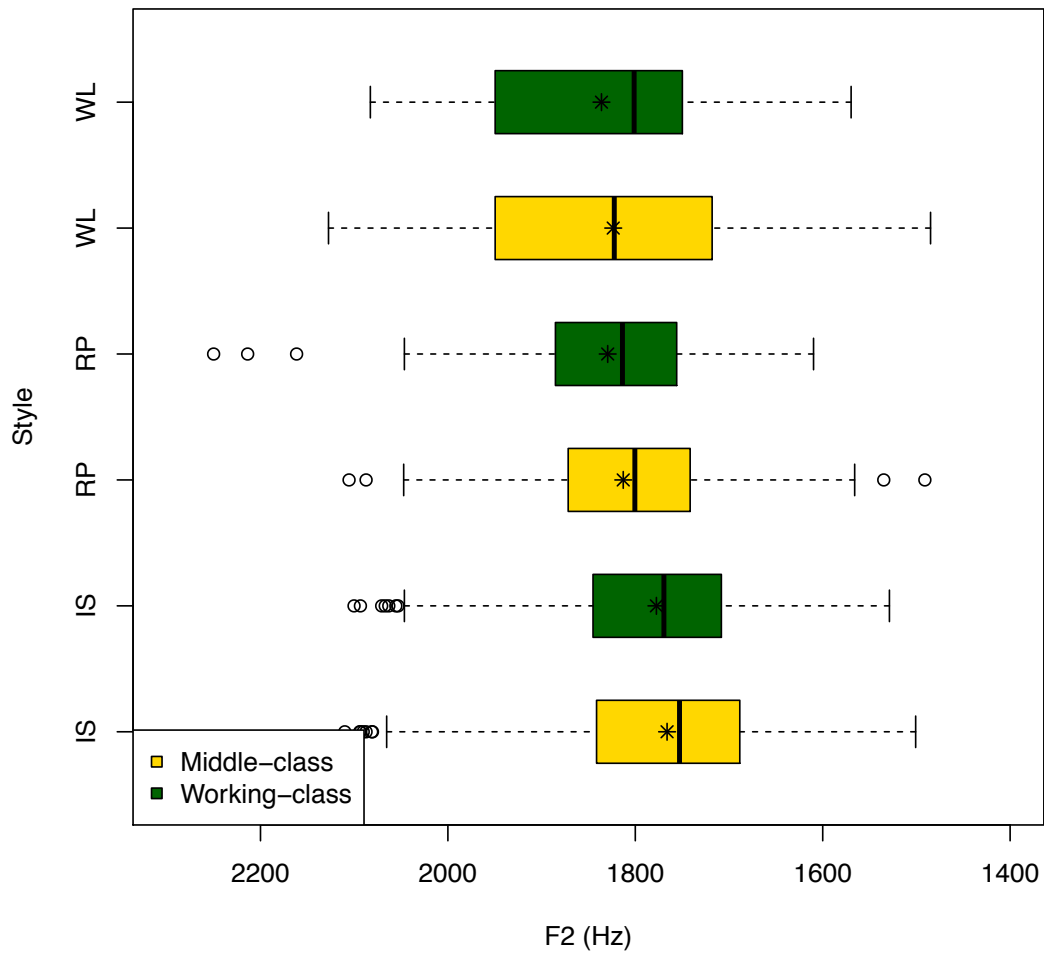


Figure 6.4: Boxplot for FLEECE F2 values in all speech styles, disaggregated by class.

The class distinction is also not significant on the F2 plane ($p=0.3651$). In terms of stylistic differences, RP and WL styles are significantly fronter than IS ($p<0.001$ and $p=0.0003$ respectively), as can be seen in figure 6.4, particularly when inspecting the mean and median values.

c) Gender and Style⁴⁷

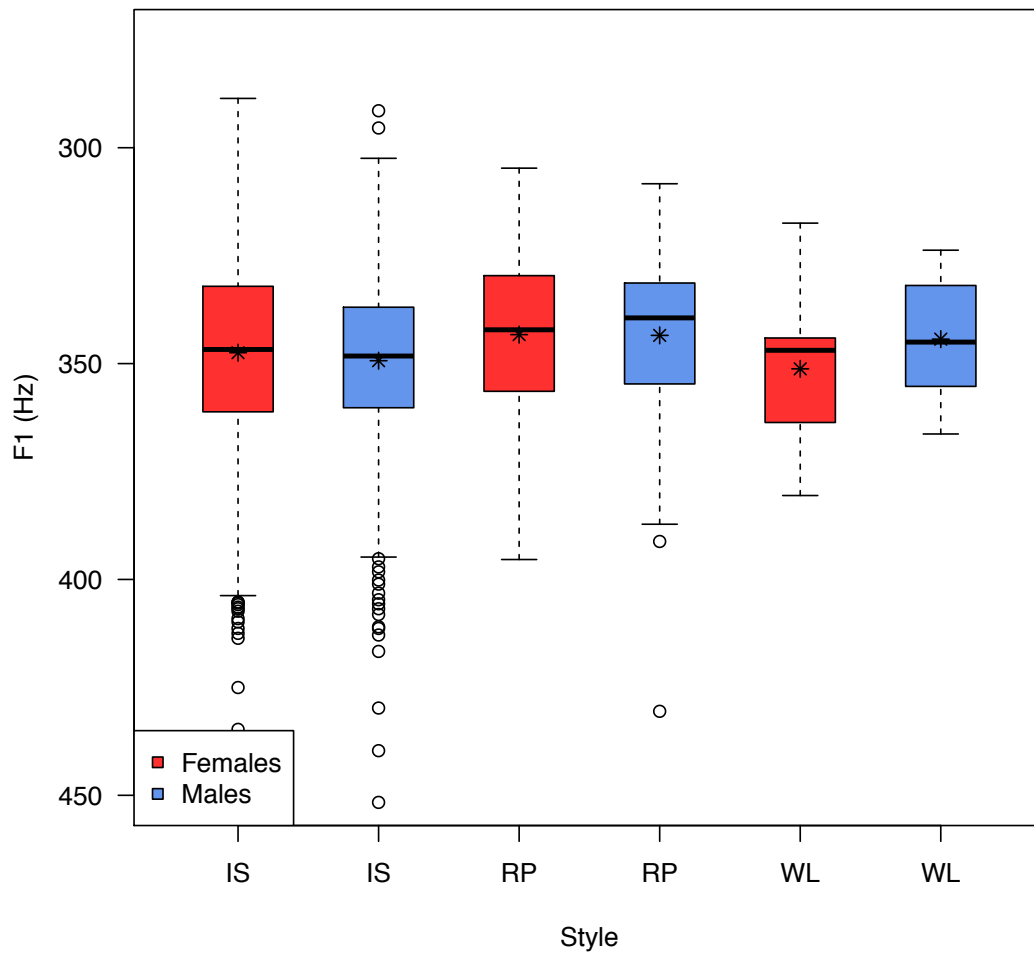


Figure 6.5: Boxplot for FLEECE F1 values in all speech styles, disaggregated by gender.

There is no significant difference between male and female speakers for this set ($p=0.6572$). The boxplots in figure 6.5 show that the mean values in IS and RP style are very similar for males and females, with only WL style showing males using a higher mean value compared with females.

⁴⁷ In this section, the statistical results for the stylistic differences are not provided, as they are virtually identical to those reported in the previous section on 'Class and Style'. The discussion and visual analysis of the graphs does take style into account, however. The same applies to the 'Gender and Style' discussion in each of the vowel subsections throughout the chapter.

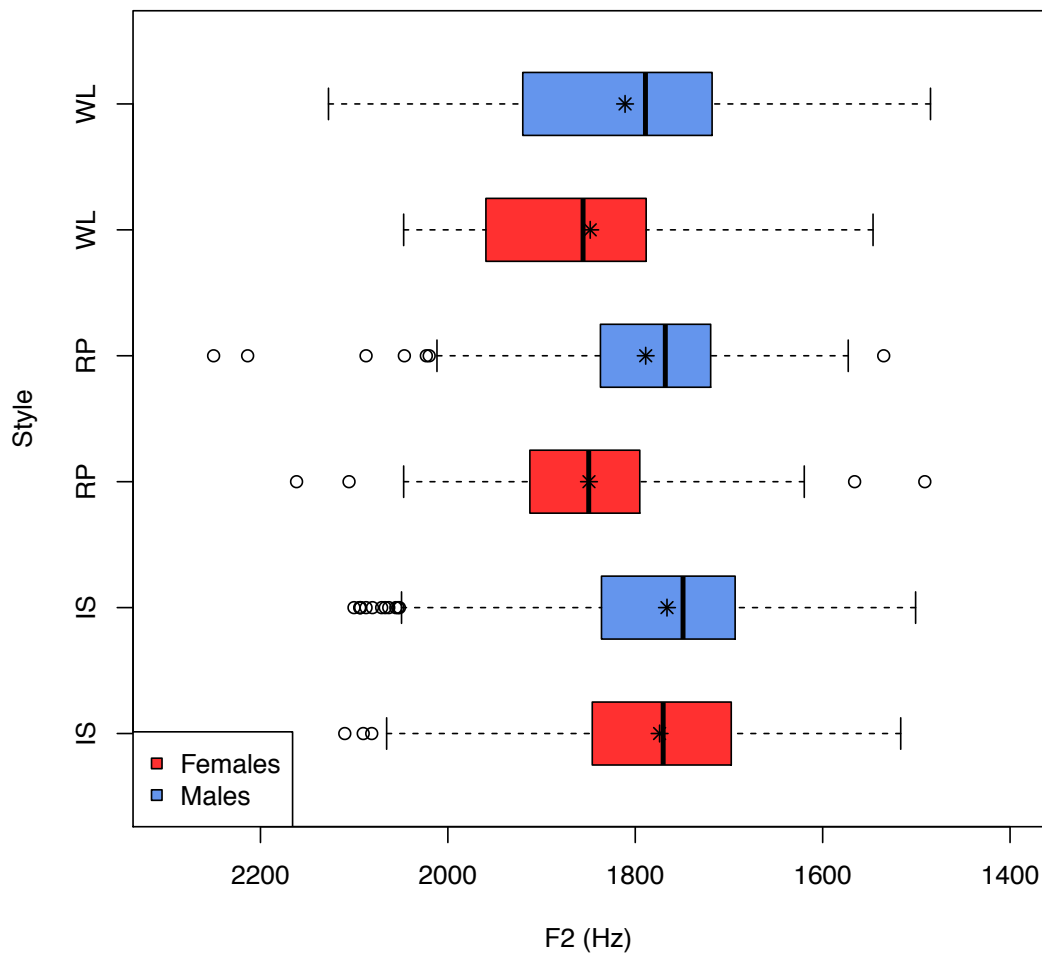


Figure 6.6: Boxplot for FLEECE F2 values in all speech styles, disaggregated by gender.

Gender is also not a significant factor in terms of advancement ($p=0.3004$). Even though the effect is not significant, the boxplots (figure 6.6) show that in the more careful speech styles, male speakers use a slightly backer range of values compared with females.

6.3.4 Summary: FLEECE Findings

This set is realised quite uniformly by all groups of CSAE speakers as a high, front vowel, although not as fronted as the WSAE speakers in Bekker's (2009) study. Stylistically, the more careful styles are realised fronter than IS, confirming that the fronted variant is more prestigious in General SAE.

6.4 NURSE

6.4.1. WSAE Reports

This set is reported to distinguish Conservative speakers from General and Broad speakers (Lass 2002: 116). Conservative speakers use RP-like mid-central unrounded [ɜ:], whereas in other lects it is rounded, half-close and front in the region of [ø:] or slightly lower. Bekker (2009: 378) suggests that rounding of this vowel is filtering into higher classes, and that the General SAE value, as used by the speakers in his study, is a relatively raised, rounded and fronted monophthong, confirming Lass' value of [ø:] for General speakers.

6.4.2. CSAE Reports

Finn (2004: 971) reports that CSAE speakers use an optionally diphthongised variant viz. [ø:(ə)], with [ɜ:], [ɝ:(ə)] and [ə] also occurring. The set is retracted to [o:] before /l/. Wood (1987: 127) records [ɜ] in unstressed syllables.

6.4.3 CSAE Acoustic Data

a) Overview

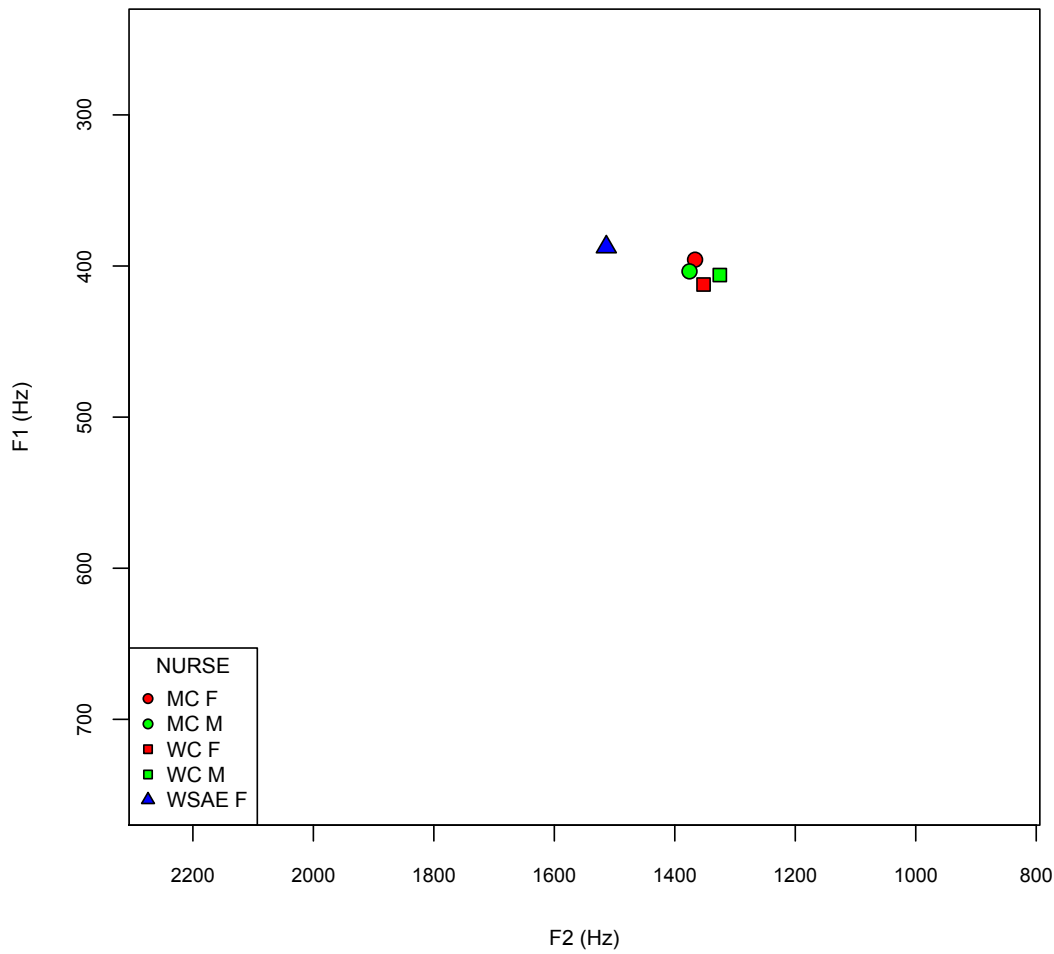


Figure 6.7: Mean values for NURSE in Interview Style (IS), disaggregated by class and gender.

Compared with Bekker’s (2009) WSAE speakers, all groups of CSAE speakers use a backer variant for the NURSE set. While the mean values plotted in figure 6.7 do not seem to show much of a difference between the speaker groups, we see that working-class speakers use a slightly lower and backer variant than the middle-class speakers do.

	F1	F1_/l/	Difference in F1	F2	F2_/l/	Difference in F2
MC F	395	388	7	1366	1341	25
WC F	412	409	3	1352	1271	81
MC M	403	408	5	1375	1362	13
WC M	405	406	1	1325	1323	2
WSAE F	387	393	6	1513	1347	166

Table 6.3: Comparative mean values for F1 and F2 in regular and pre-/l/ tokens of NURSE (Hertz).

The acoustic data suggest that there is no retraction before velar /l/ for CSAE speakers, as is evidenced by the mean values provided for each of the speaker groups in table 6.3. The biggest difference is for working-class females, with a 81Hz difference, but all other CSAE speaker groups had less than a 30Hz difference, which is negligible. There seems to be slight retraction for WSAE speakers, with a difference of 166Hz in the mean values for regular and pre-/l/ tokens on the F2 plane.

b) Class and Style

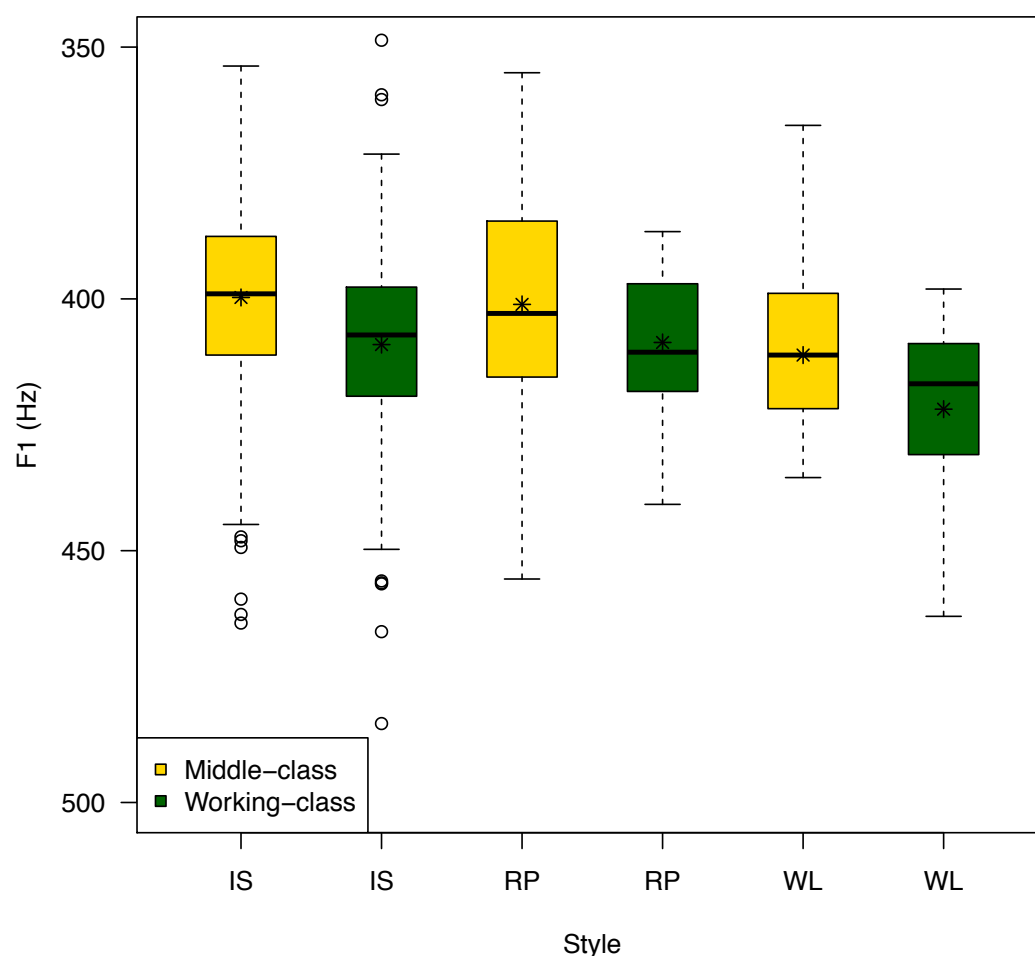


Figure 6.8: Boxplot for NURSE F1 values in all speech styles, disaggregated by class.

The distinction between middle- and working-class speakers is clear from the boxplots in figure 6.8: working-class speakers use a significantly lower range of values for NURSE in all speech styles ($p=0.0044$). In terms of stylistic difference, WL style is significantly lower than IS ($p=0.00017$), but there is no significant difference between RP style and IS ($p=0.9743$).

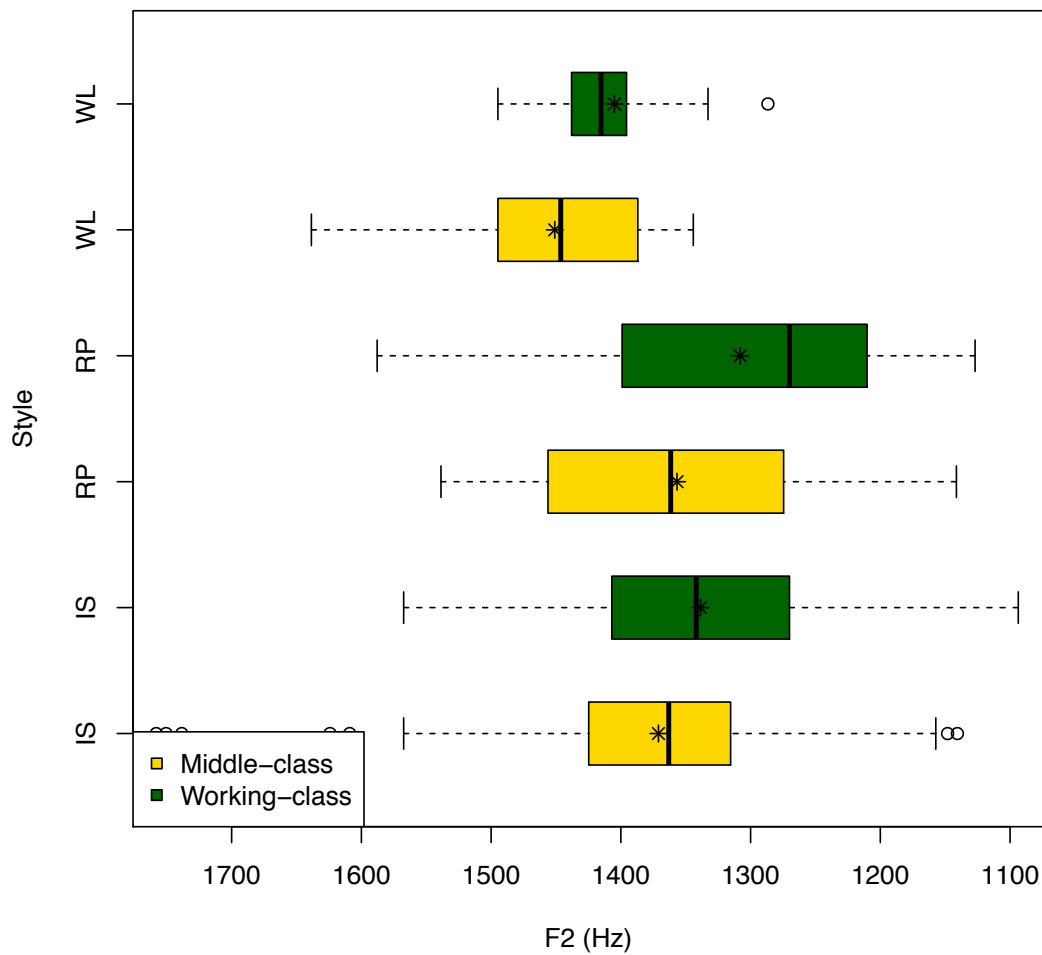


Figure 6.9: Boxplot for NURSE F2 values in all speech styles, disaggregated by class.

Class difference is also significant on the F2 plane: working-class speakers use a backer variant than their middle-class counterparts ($p=0.0066$). This is evident in figure 6.9. Stylistically, as for F1, WL style shows significant variation from IS ($p<0.001$), produced fronter in the more careful style, while RP style shows no significant variation ($p=0.0739$).

c) Gender and Style

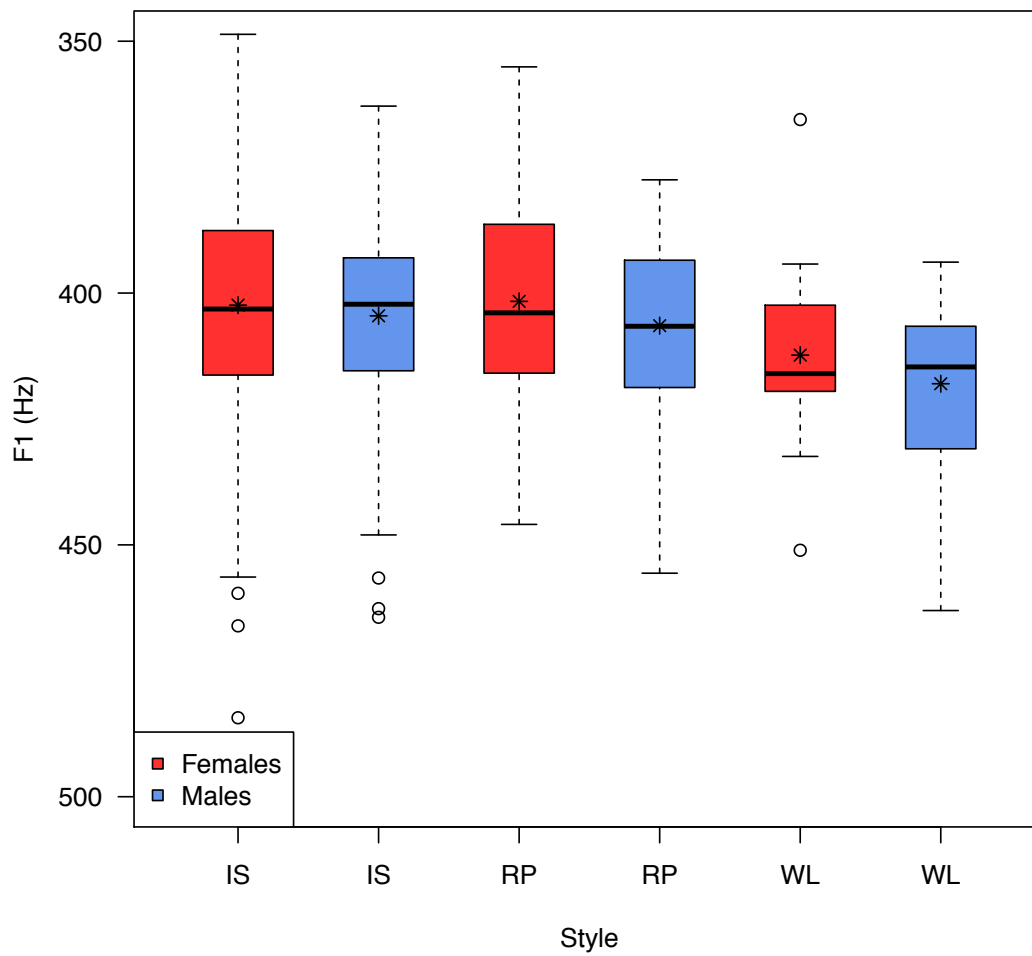


Figure 6.10: Boxplot for NURSE F1 values in all speech styles, disaggregated by gender.

There is no significant gender difference between speakers on the F1 plane ($p=0.5293$), although figure 6.10 shows that in all speech styles, males use a lower mean value compared with females. This is most marked in WL style.

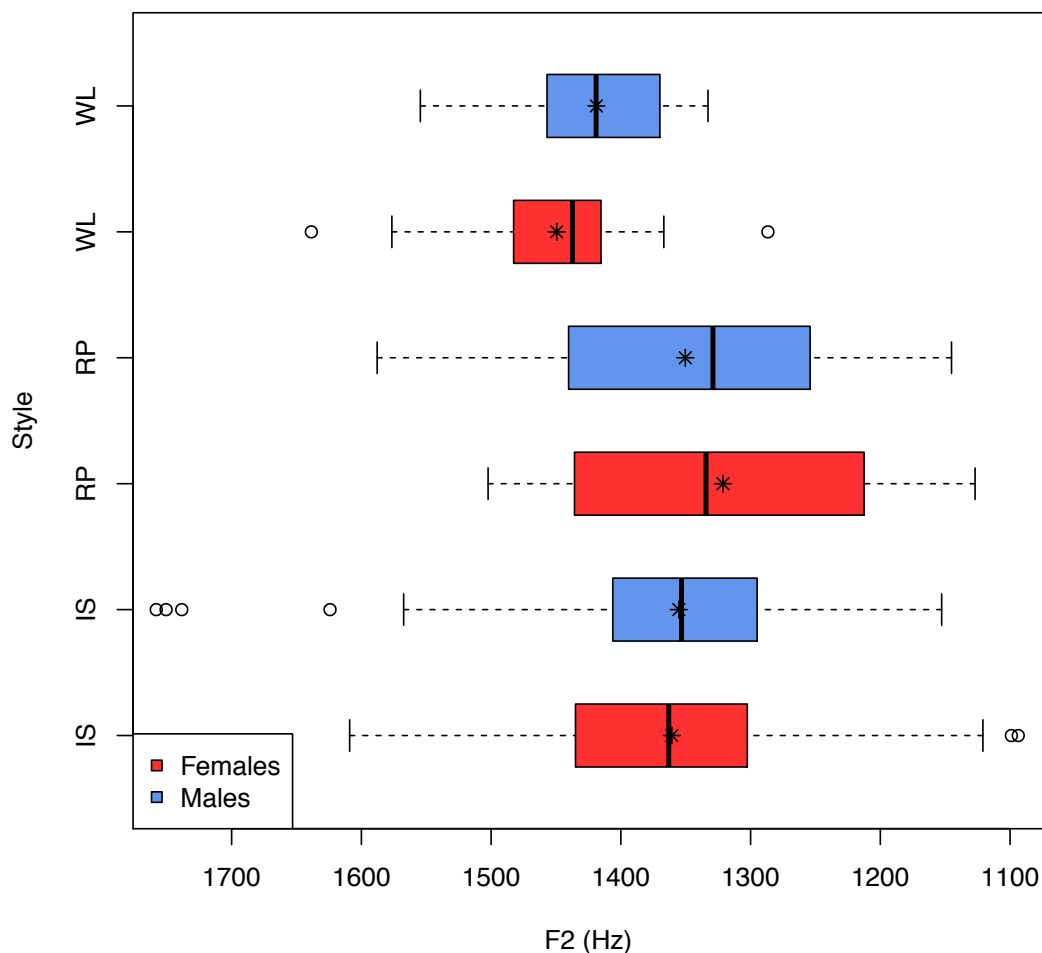


Figure 6.11: Boxplot for NURSE F2 values in all speech styles, disaggregated by gender.

Figure 6.11 shows very similar mean and median values for both genders in IS, but in the more careful styles the findings are conflicting: males use a fronter range than females in RP style, but a backer range in WL style. None of these results are significant, however, as there are no significant differences between male and female speakers on the F2 plane ($p=0.9472$).

6.4.4 Summary: NURSE Findings

The acoustic data for this set show that working-class speakers use a backer and lower NURSE vowel than the middle-class speakers in the sample. The middle-class value is thus closer to the WSAE value, [ø:], which is expected due to the contact between WSAE speakers and middle-class CSAE speakers. The lower and backer variant approximates [ɜ:], as Wood (1987) reports. There was no significant difference between male and female speakers with regards vowel height or advancement. Stylistically, WL style was produced fronter and lower than regular IS tokens, confirming the use of the more prestigious variant in more formal speech styles (at least in terms of advancement).

6.5 GOOSE

GOOSE is an important social variable in SAE, which displays various degrees of fronting based on social and phonological factors. The social groups who traditionally front this vowel most are discussed in the following section. Phonologically, a preceding coronal consonant (i.e. /t, d, n, l, r, s, z, j, ʃ, ʒ, tʃ, dʒ/) produces a fronter vowel than non-coronal environments (i.e. following /k, g, h, m, p, b, f, v/) (Labov et al. 2006: 154; Mesthrie 2010: 10). Preceding /j/ is a particularly fronting environment, so much so that Mesthrie (2010) analyses this as a separate class. No post-/j/ tokens remain in IS for the current dataset – they were removed along with other tokens in phonological environments known to have co-articulatory effects (see section 4.3 for details). The acoustic analysis of GOOSE thus separates the IS tokens into coronal and non-coronal environments.

6.5.1 WSAE Reports

Lass (2002: 116) reports that Conservative speakers use a backish (but not fully back) vowel in the region of [u:], and for other speakers it is more centralised i.e. [ʊ:]. He suggests that younger General speakers, females particularly, have completely fronted this vowel, using [y:]. This fronted variant is apparently regarded as a marker of ‘whiteness’, although as Mesthrie (2010) shows, it is no longer only White speakers who use a fronted variant for this vowel: Black speakers use an equally fronted variant. Bekker’s (2009: 308) acoustic analysis of young females confirms that WSAE speakers of this demographic are indeed using a fronted variant, noting that /j/ in the preceding segmental environment results in further fronting as a co-articulatory effect.

6.5.2 CSAE Reports

Earlier accounts of CSAE confirm that this variety traditionally has a back, rounded [u:] (Wood 1987: 136-137), although a more centralised variant is also reported to occur (Finn 2004; Dennis 2008).

6.5.3 CSAE Acoustic Data

a) Overview

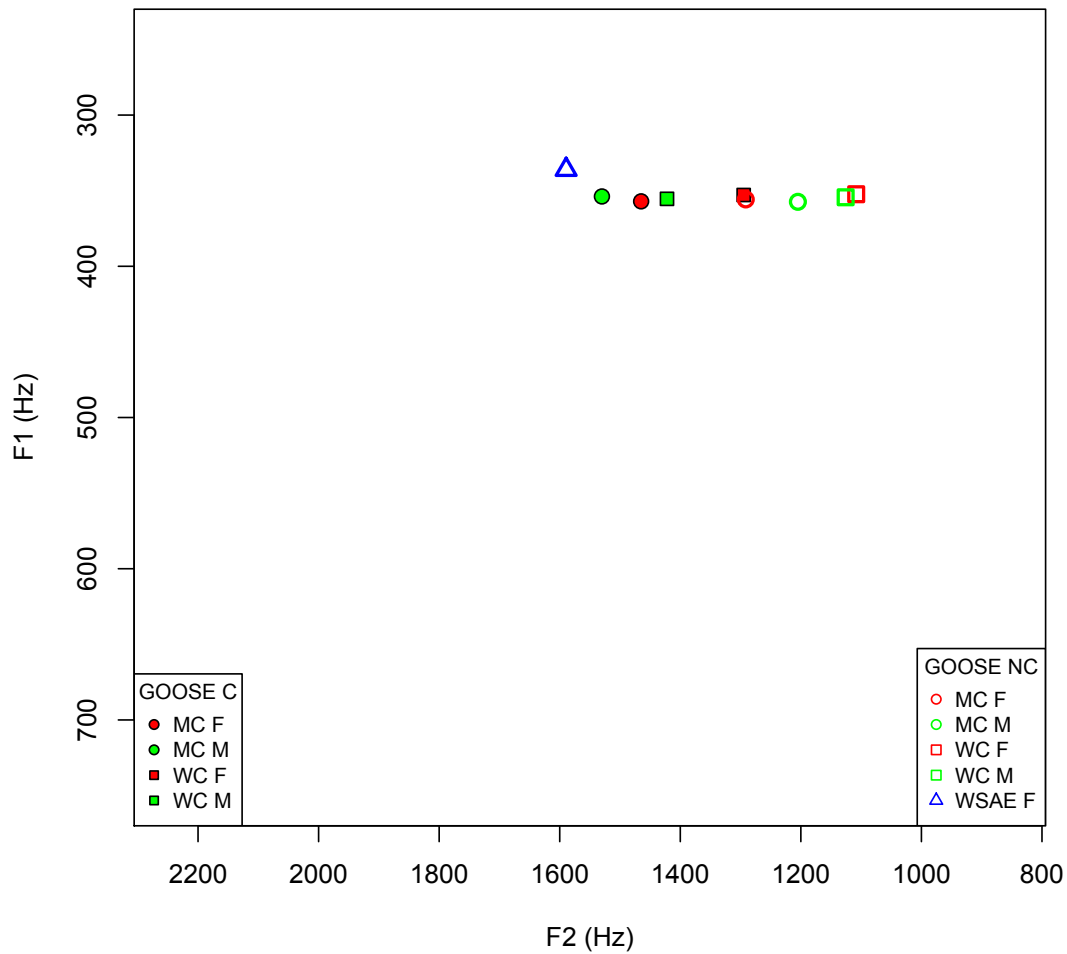


Figure 6.12: Mean values for GOOSE in Coronal and Non-Coronal environments in Interview Style (IS), disaggregated by class and gender.

Figure 6.12 shows a clear distinction between the two phonological environments dividing the GOOSE set: for all speaker groups, GOOSE occurring after coronal consonants are produced further than in non-coronal environments. The mean value for Bekker's (2009) WSAE speakers represents tokens produced with preceding non-coronal consonants (*who'd, boot, food*), and this value is further than for the CSAE speakers' mean values in the coronal environment. Bekker's Word Lists did not contain any tokens of GOOSE with a preceding coronal environment, but it is reasonable to assume that these speakers would use an even more fronted value in non-coronal environments.

Middle-class speakers use a further variant than working-class speakers in both phonological environments, although surprisingly, male speakers use a more fronted mean value in the coronal environment than their respective female counterpart group.

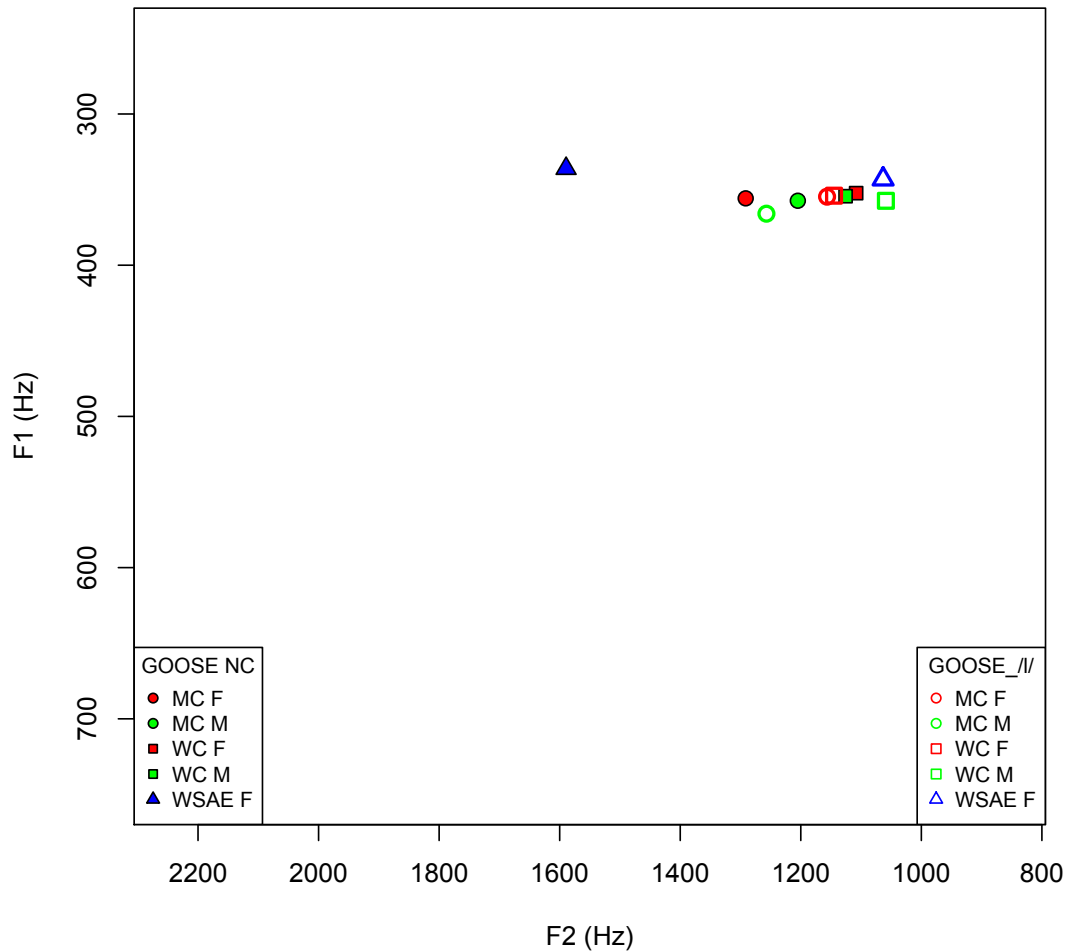


Figure 6.13: Mean values for GOOSE Non-Coronal environments and GOOSE_/l/ in Interview Style (IS), disaggregated by class and gender.

To ascertain whether the GOOSE set retracts before velar /l/, pre-/l/ tokens were plotted on the same graph as the regular GOOSE tokens in the non-coronal environment in figure 6.13. The graph shows that for middle-class females and working-class males, the pre-/l/ tokens are backer still than those in the non-coronal environment, but the middle-class males and working-class females produce a fronter mean value for their pre-/l/ tokens. These acoustic data are thus inconclusive regarding the retraction of GOOSE before velar /l/. The graph also shows that there is significant retraction of this set for Bekker's (2009) WSAE female speakers, who use a very centralised GOOSE vowel in non-coronal environments.

b) Class and Style

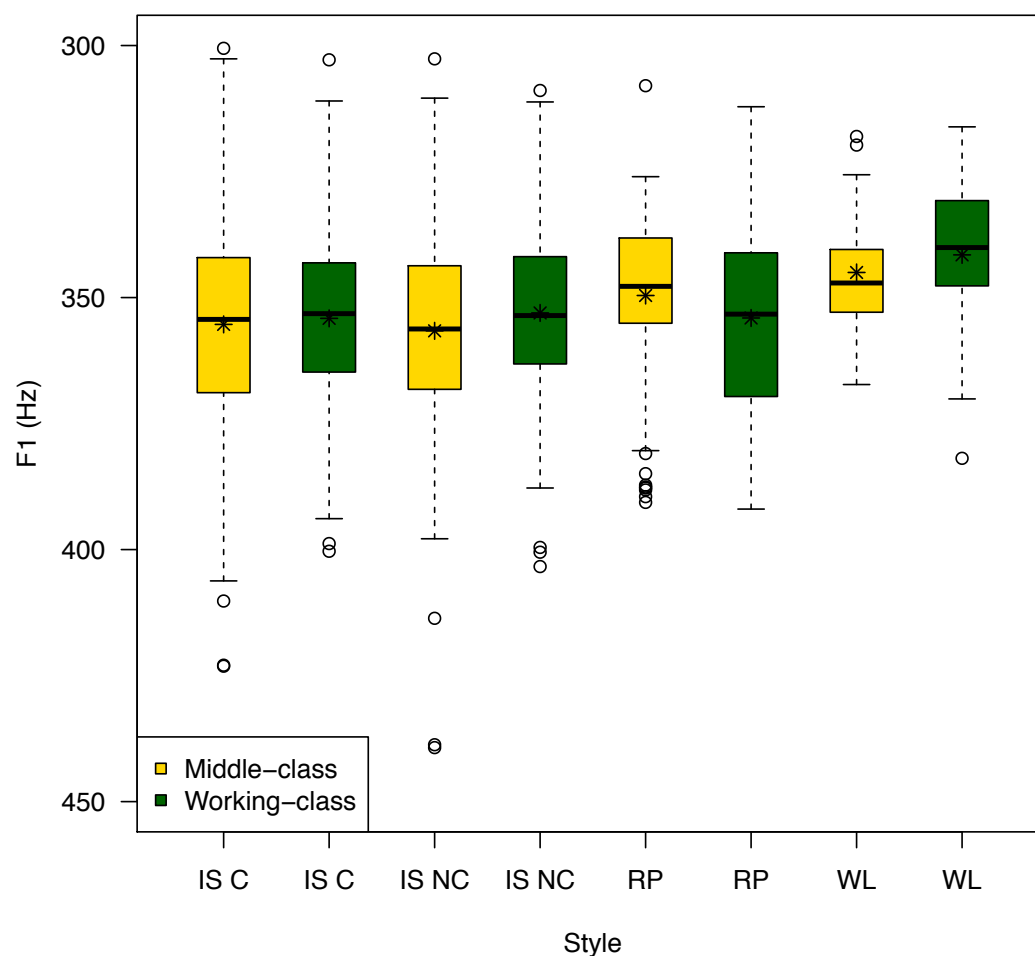


Figure 6.14: Boxplot for GOOSE F1 values in all speech styles, disaggregated by class.

In terms of F1, there is no significant class distinction for this set ($p=0.8858$), nor is there a distinction between IS tokens in the coronal and non-coronal environments ($p=0.2449$). Stylistically, however, the boxplots (see figure 6.14) show that WL style is produced higher than IS ($p=0.0004$), but RP style is not significantly different to the IS tokens ($p=0.2661$).

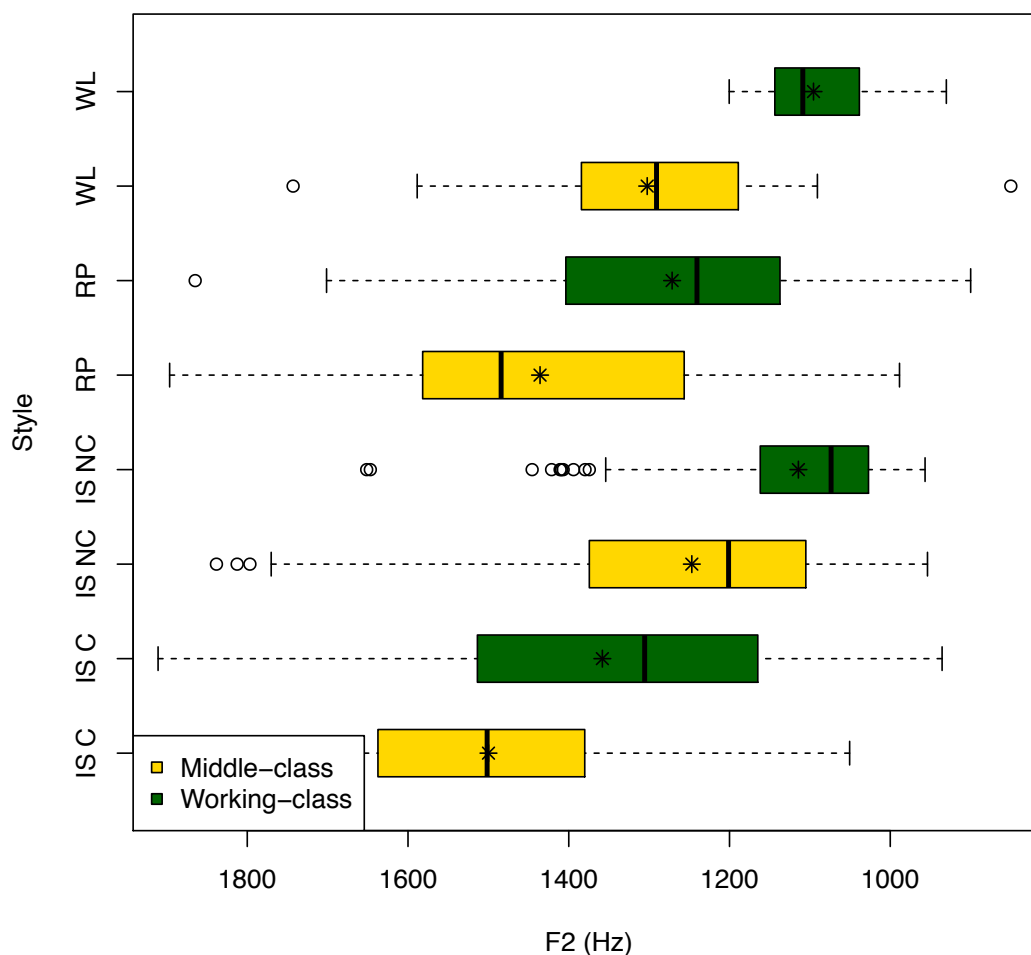


Figure 6.15: Boxplot for GOOSE F2 values in all speech styles, disaggregated by class.

Class distinction is very significant on the F2 plane ($p < 0.001$), as is expected for the GOOSE set: working-class speakers use a backer variant in all speech styles than do the middle-class speakers, as figure 6.15 shows. The boxplots also show that tokens in the coronal environment are produced significantly fronter than those in the non-coronal environment ($p < 0.001$). Both RP and WL styles show significant variation compared with IS ($p = 0.0006$ and $p < 0.001$ respectively), but with divergent trends: RP style is fronter than IS, while WL style is backer than IS. This is attributable to the words making up the RP and WL style tokens: all of the RP tokens are in fronting environments i.e. after /j/⁴⁸ (*you, use, beauty*) and after coronal consonants /t/ (*two*). There was only one word in the RP which occurred after a non-coronal consonant i.e. /f/ in *food*. Conversely, the WL style token is *goose* for all speakers, which is expected to be backer because it follows a non-coronal consonant /g/.

⁴⁸ Words were not excluded from the Reading Passage or Word List because of their phonological environments, hence the inclusion of these words in the analysis. For a list of exclusions in Interview Style tokens, see section 4.3).

c) Gender and Style

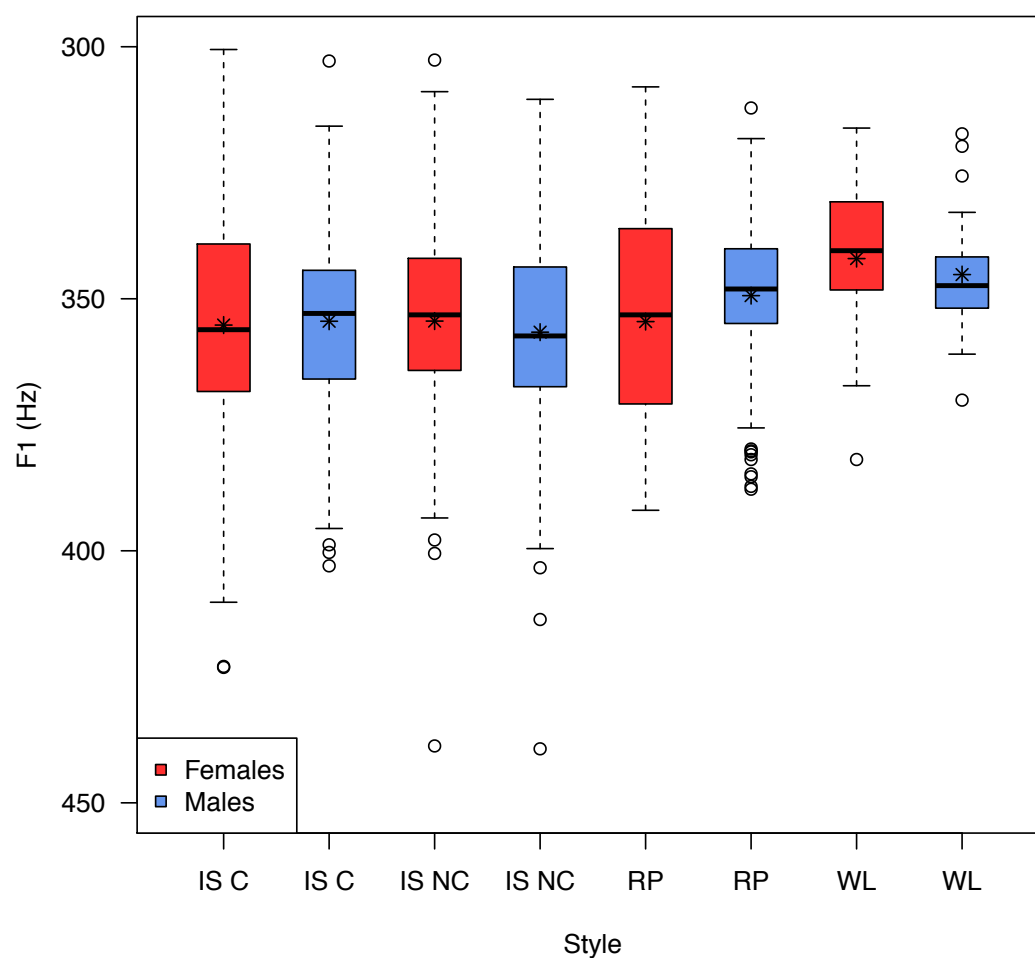


Figure 6.16: Boxplot for GOOSE F1 values in all speech styles, disaggregated by gender.

There is no significant gender distinction for this set in terms of vowel height ($p=0.9342$). The boxplots in figure 6.16 show that females use a wider range of values in all three speech styles compared with male speakers.

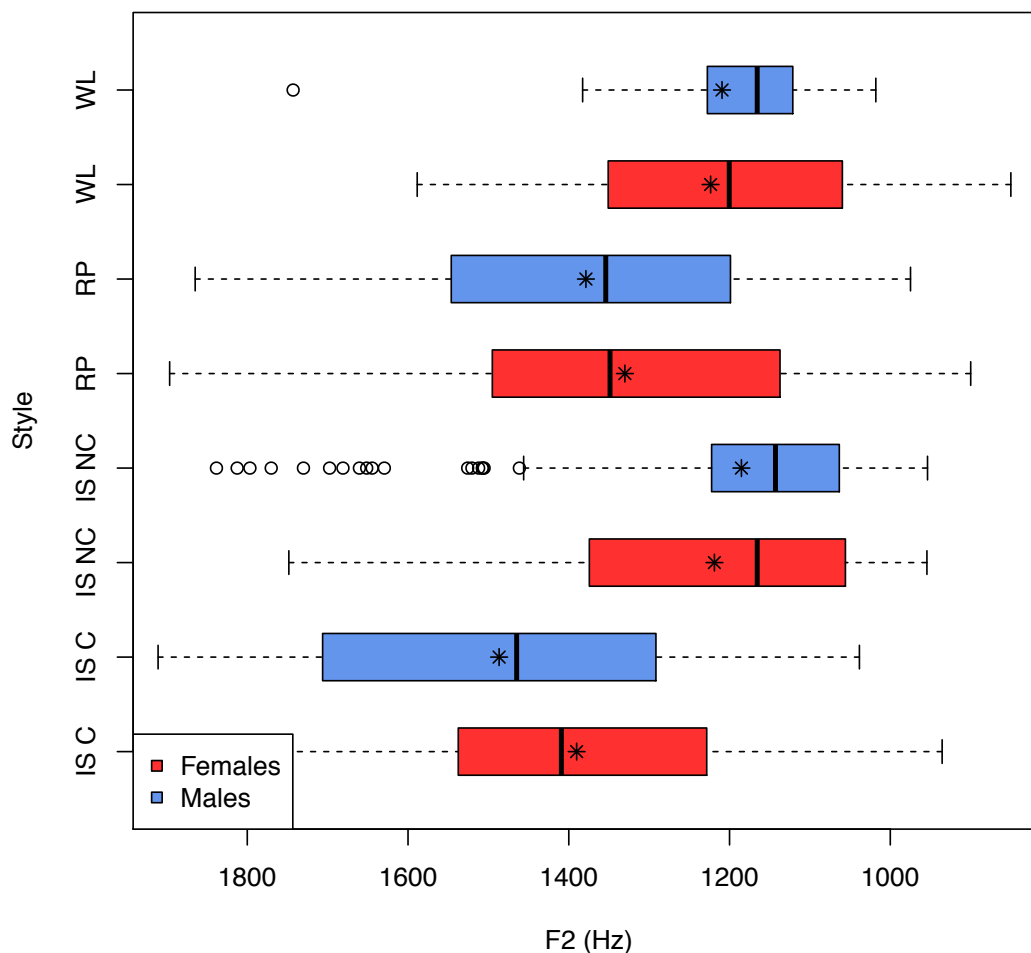


Figure 6.17: Boxplot for GOOSE F2 values in all speech styles, disaggregated by gender.

Gender is also not a significant factor in terms of F2 ($p=0.4857$), although figure 6.17 shows that in the fronting environments (IS Coronal and RP style), males use a fronter mean and median value, and a backer mean and median value in the non-fronting environments (IS Non-coronal and WL style).

6.5.4 Summary: GOOSE Findings

Middle-class speakers use a significantly fronter GOOSE vowel than working-class speakers – following the trend towards fronting which is evident in the speech of the WSAE speakers in Bekker’s (2009) study. The use of the centralised variant by middle-class speakers is thus an interesting, if unsurprising, finding. A preceding coronal consonant is confirmed to be a fronting environment, as tokens produced in this environment were much fronter than those in non-coronal environments.

In terms of stylistic differences, WL style was produced higher and backer than IS, while RP style was realised fronter than IS.

6.6 THOUGHT

6.6.1 WSAE Reports

Lass (2002: 116) reports that this set also separates Conservative speakers from the other lects of SAE. Conservative speakers use an opener vowel in the region of [ɔ:] while General and Broad speakers use a closer [o:]. When followed by a voiceless fricative, this subset is variously realised with either the THOUGHT or LOT vowels, and is labelled by Wells (1982) as the CLOTH set. The more Conservative lects favour the short vowel over long THOUGHT. Bekker (2009) finds that his General speakers produce THOUGHT with [o:], and that the contrast with LOT consists only of a difference in length – the vowels are qualitatively similar.

6.6.2 CSAE Reports

Wood (1987: 122) reports that Broad speakers typically use [ɔ], while Finn (2004: 971) finds that higher [o:] is more common. THOUGHT is not reported to be affected by following /l/.

6.6.3 CSAE Acoustic Data

a) Overview

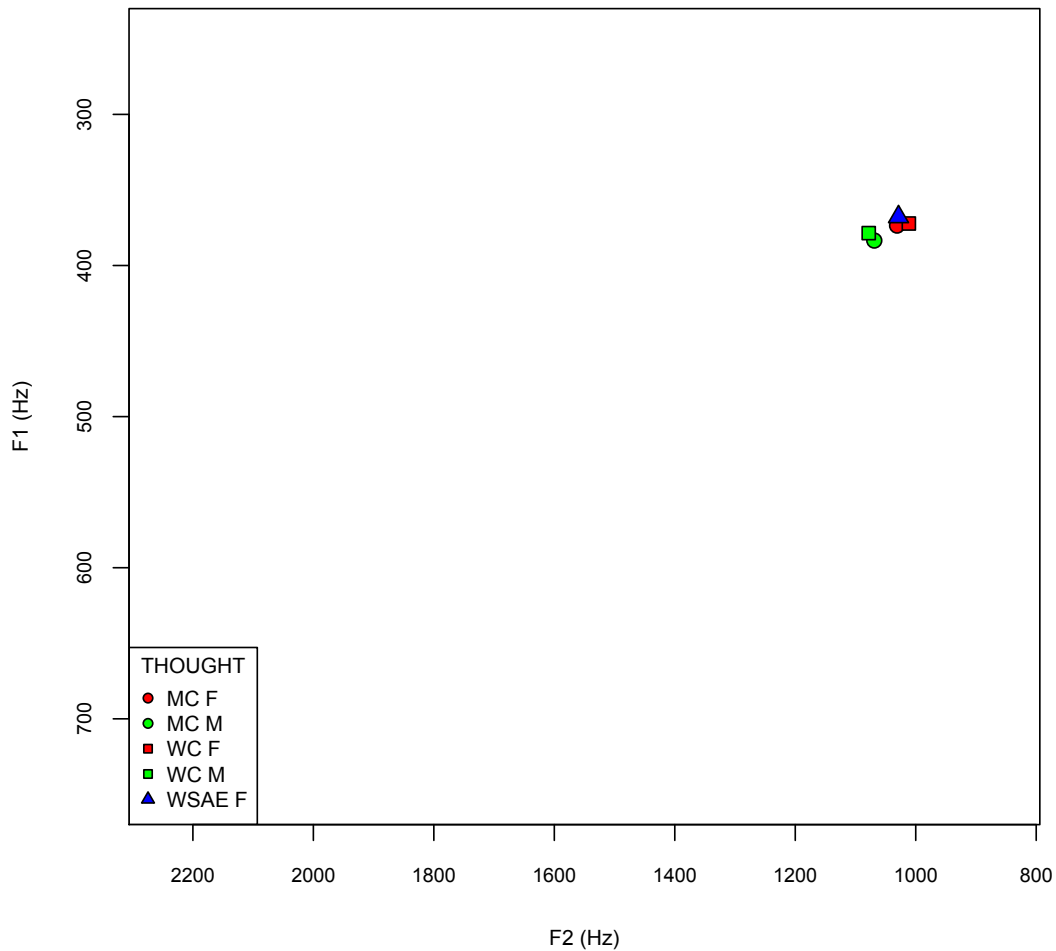


Figure 6.18: Mean values for THOUGHT in Interview Style (IS), disaggregated by class and gender.

The THOUGHT set is realised as a very back vowel by all speakers (see figure 6.18). Female speakers use a mean value which is equally high and back as Bekker's (2009) WSAE female speakers, while male speakers have a slightly fronter mean value. This suggests that CSAE speakers are also using a value in the region of [o:] as Finn (2004) reports.

	F1	F1_/l/	Difference in F1	F2	F2_/l/	Difference in F2
MC F	373	382	9	1030	1055	25
WC F	372	382	10	1011	1040	29
MC M	383	386	3	1068	1090	22
WC M	378	388	10	1078	1102	24
WSAE F	367	361	6	1028	1024	4

Table 6.4: Comparative mean values for F1 and F2 in regular and pre-/l/ tokens of THOUGHT (Hertz).

This set, being realised as far back as it is, does not retract before velar /l/. The mean values presented in table 6.4 show this, with negligible difference in F2 values (less than 30Hz) between regular and pre-/l/ tokens for all speaker groups, including WSAE speakers (Bekker 2009).

b) Class and Style

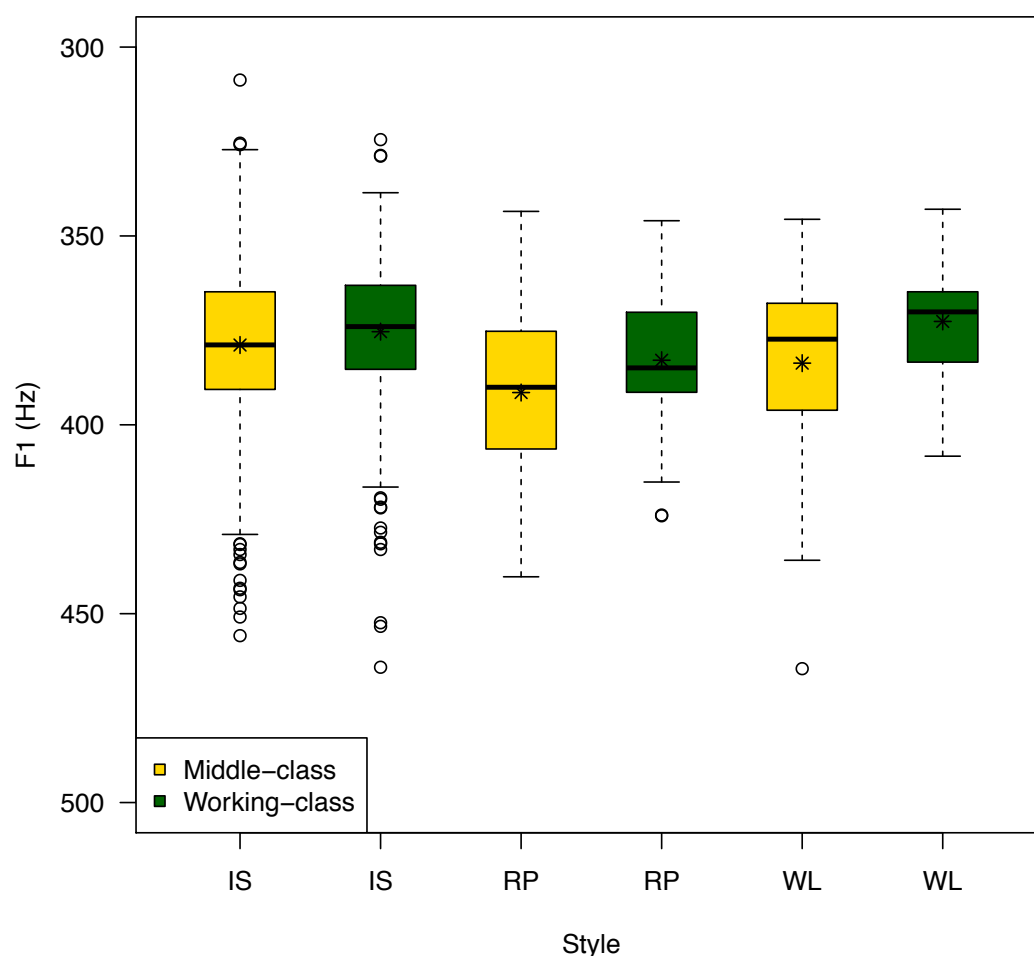


Figure 6.19: Boxplot for THOUGHT F1 values in all speech styles, disaggregated by class.

The class difference for this lexical set is not significant in terms of vowel height ($p=0.209$). Figure 6.19 shows that in the two more careful speech styles, viz. WL and RP style, working-class speakers use a higher range of values compared to middle-class speakers, and it is likely that the insignificant statistical result is due to the larger number of tokens in IS compared with the more careful styles.

Stylistically, only RP style is significantly lower than IS ($p<0.001$), but there is no significant difference between WL style and IS ($p=0.17$).

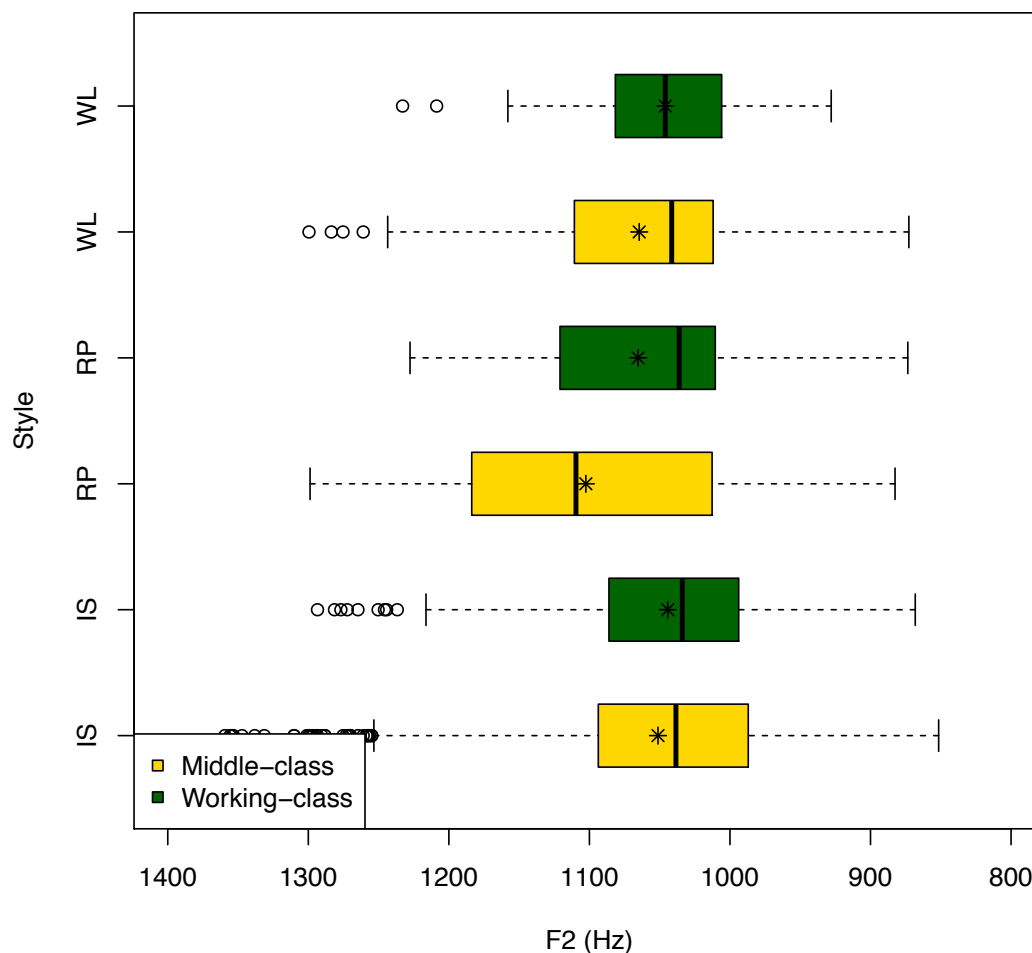


Figure 6.20: Boxplot for THOUGHT F2 values in all speech styles, disaggregated by class.

In terms of vowel advancement, there is also no significant class difference ($p=0.3163$), although the boxplots (figure 6.20) show that in all three speech styles, middle-class speakers have a fronter mean value than do working-class speakers. As for F1, RP style is significantly different to IS tokens ($p=0.0001$), being realised fronter in the more careful style, but WL style shows no significant difference to IS ($p=0.3867$).

c) Gender and Style

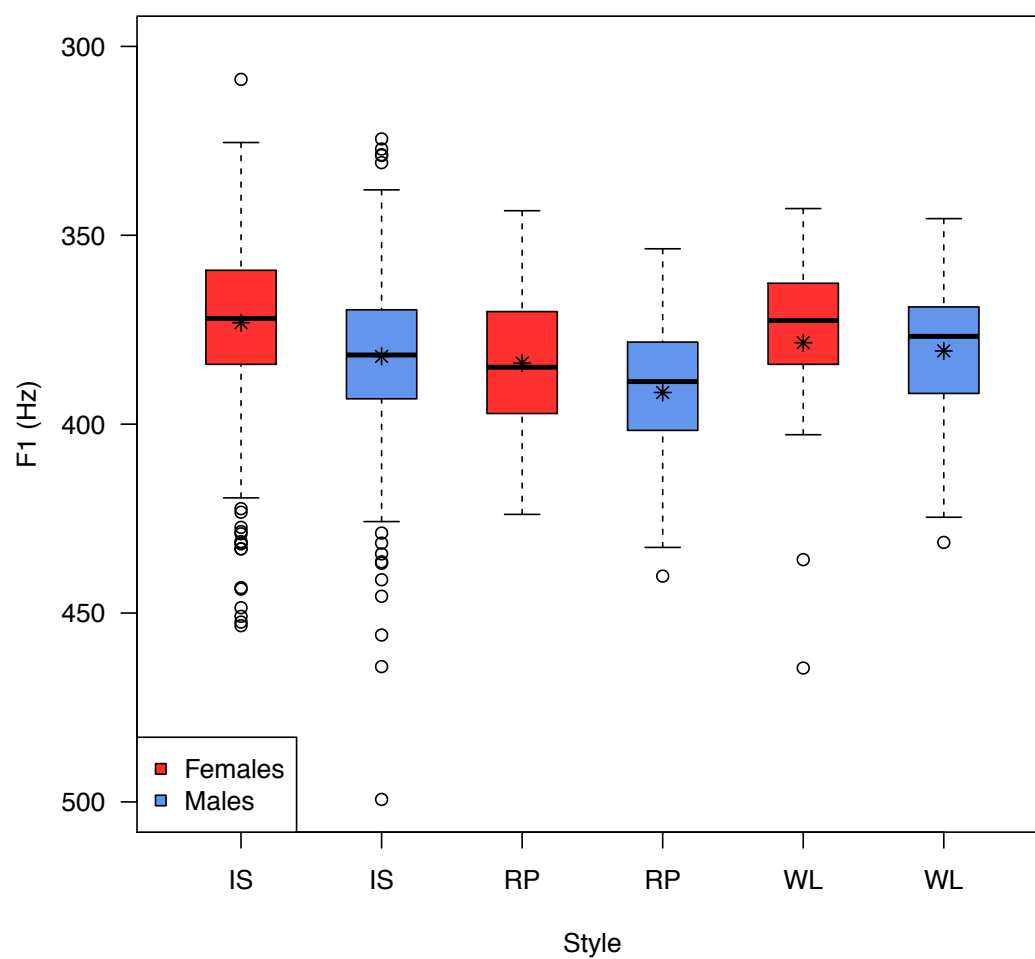


Figure 6.21: Boxplot for THOUGHT F1 values in all speech styles, disaggregated by gender.

For this set, males realise a significantly lower range of values compared with female speakers ($p=0.0024$) in all speech styles. This is very clear in the boxplots in figure 6.21.

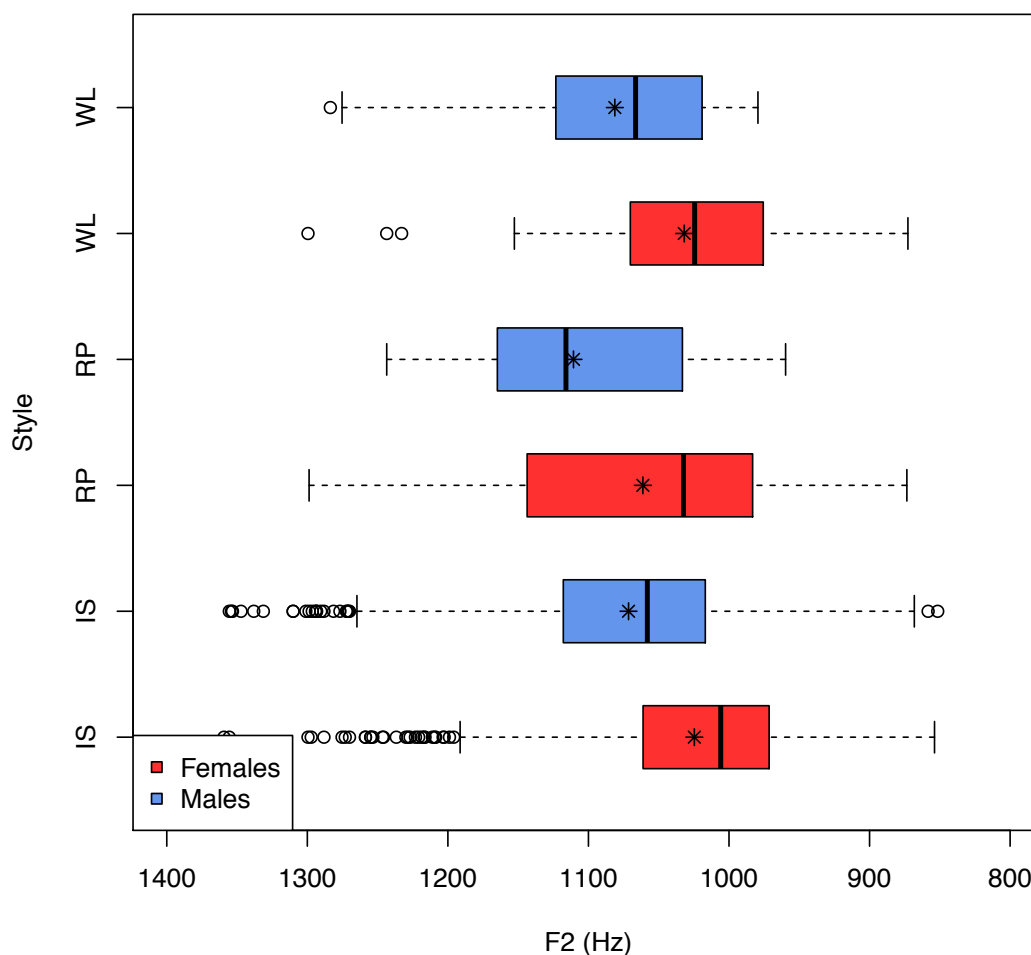


Figure 6.22: Boxplot for THOUGHT F2 values in all speech styles, disaggregated by gender.

Gender is a very significant factor on the F2 plane ($p < 0.001$), with male speakers using a frontier range of values compared with females in all three speech styles, as figure 6.22 shows.

6.6.4 Summary: THOUGHT Findings

While the distinction between middle- and working-class speakers is not significant, gender is a highly significant distinguishing factor, with male speakers using a lower and frontier variant than female speakers. Females thus seem to realise this set in the region of [o:] (which is similar to Bekker's (2009) female speakers), while the value for males is slightly lower and frontier. Wood (1987: 122) suggests that Broad speakers use a lower variant, and generally, women do tend to use more high prestige variants, thus the use of the lower variant by males could be explained this way. The earlier CSAE scholars do not note any difference in this set in terms of vowel advancement, so the use of the frontier variant by males is a new finding. Stylistically, RP style is realised lower and frontier than IS – this result is anomalous with the suggestion that the higher variant is more prestigious, however.

6.7 BATH

6.7.1 WSAE Reports

Another socially significant long vowel is BATH (Lass 2002: 116-7). Conservative SAE has a centralised back [ä:], or even central [a:] in more posh varieties. General and Broad lects produce [ɑ:] – a much backer variant (which can be fully back). Younger male General SAE speakers are reported to have the backest variant. Broad speakers may round to [ɔ:] and sometimes raise to [ɔ:]. Bekker (2009) found that his young female General speakers use a mid back vowel, which suggests that for this demographic, BATH raising may have entered the General SAE lect.

6.7.2 CSAE Reports

Wood (1987) and Finn's (2004) impressionistic accounts of this set find that [a] covaries with [ɑ]. Both variants can be lengthened and sometimes diphthongised with a following schwa i.e. [a:(ə)] ~ [ɑ:(ə)] (Finn 2004: 970). A rounded variant [ɔ:(:ə)] is also reported to occur. There is a subset which is fronted for Broad CSAE speakers: the words *dance* and *chance* are realised with [æ:] (Wood 1987: 123).

6.7.3 CSAE Acoustic Data

a) Overview

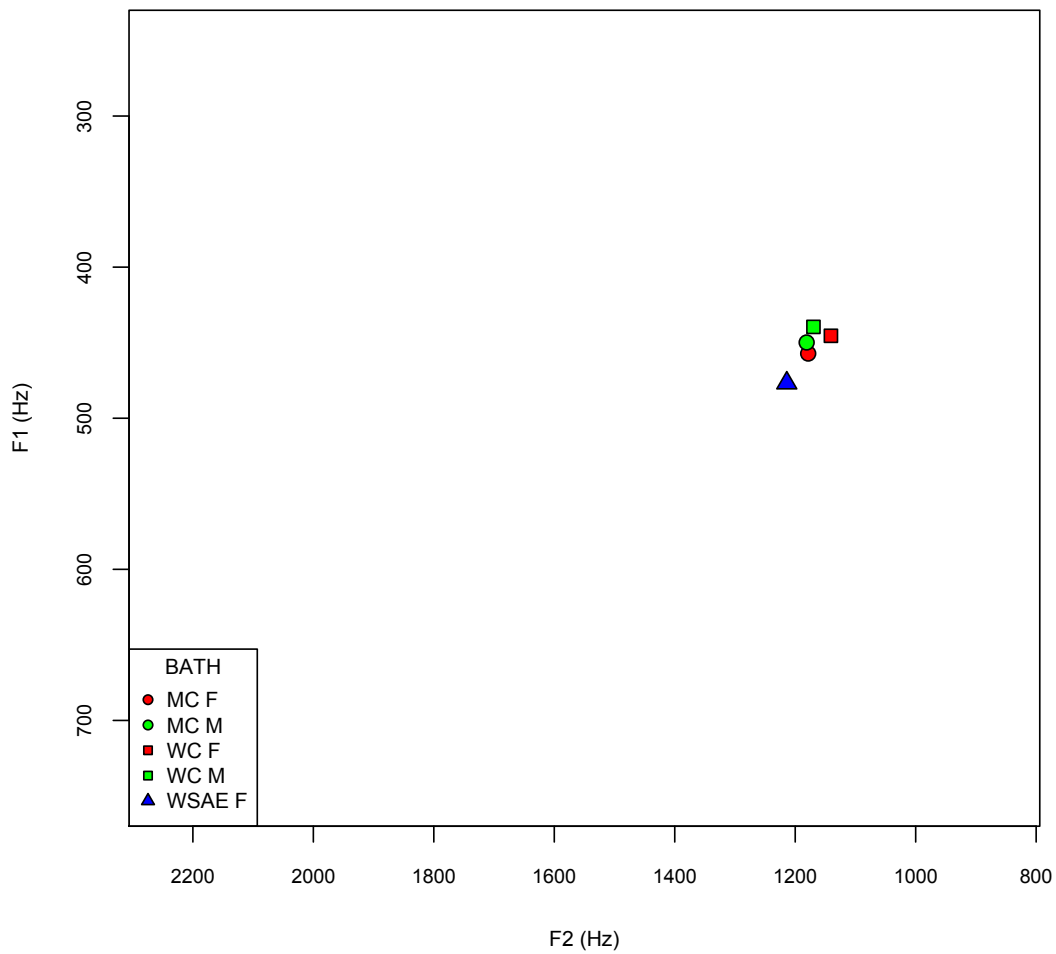


Figure 6.23: Mean values for BATH in Interview Style (IS), disaggregated by class and gender.

Based on the mean values plotted in figure 6.23, CSAE speakers seem generally to use a higher and backer variant compared with Bekker's (2009) female speakers. There does not seem to be much variation between the different CSAE speaker groups, although the mean value used by working-class speakers is slightly higher than that of the middle-class speakers.

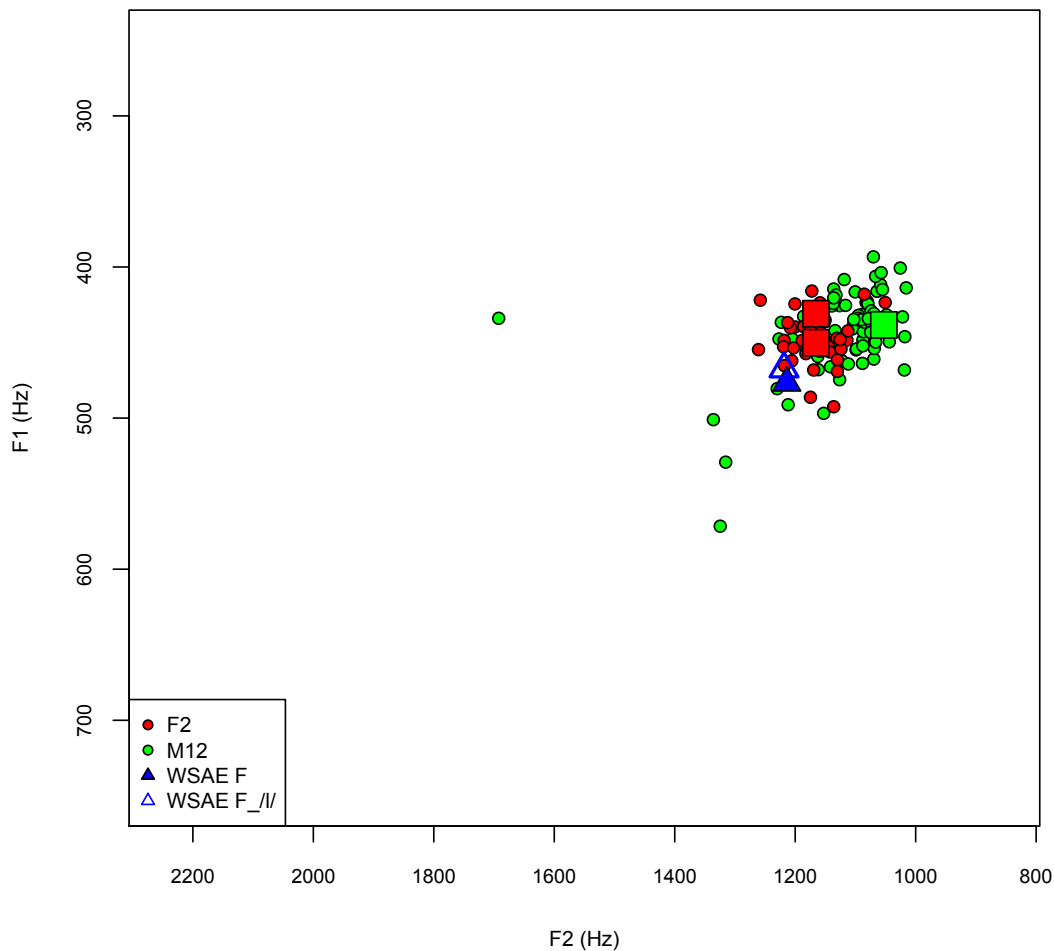


Figure 6.24: BATH tokens for speakers F2 and M12, showing the position of tokens before velar /l/.

Only three tokens of BATH were produced in a tautosyllabic, pre-/l/ position by two of the speakers in the sample. Speaker F2, a working-class female, uttered the word *Paarl* (the name of a town), and speaker M12, a middle-class male, produced two utterances of the name *Charles*. These three tokens are depicted on figure 6.24 with all regular tokens of BATH for the two speakers, as well as a mean value of regular and pre-/l/ BATH tokens for Bekker’s (2009) WSAE speakers. Based on the placement of the three tokens, it seems that BATH is not retracted before velar /l/, although due to the limited number of tokens in the acoustic data, it is not possible to state this conclusively. It seems that WSAE speakers also do not retract this set before velar /l/, based on Bekker’s (2009) mean values.

b) Unique words

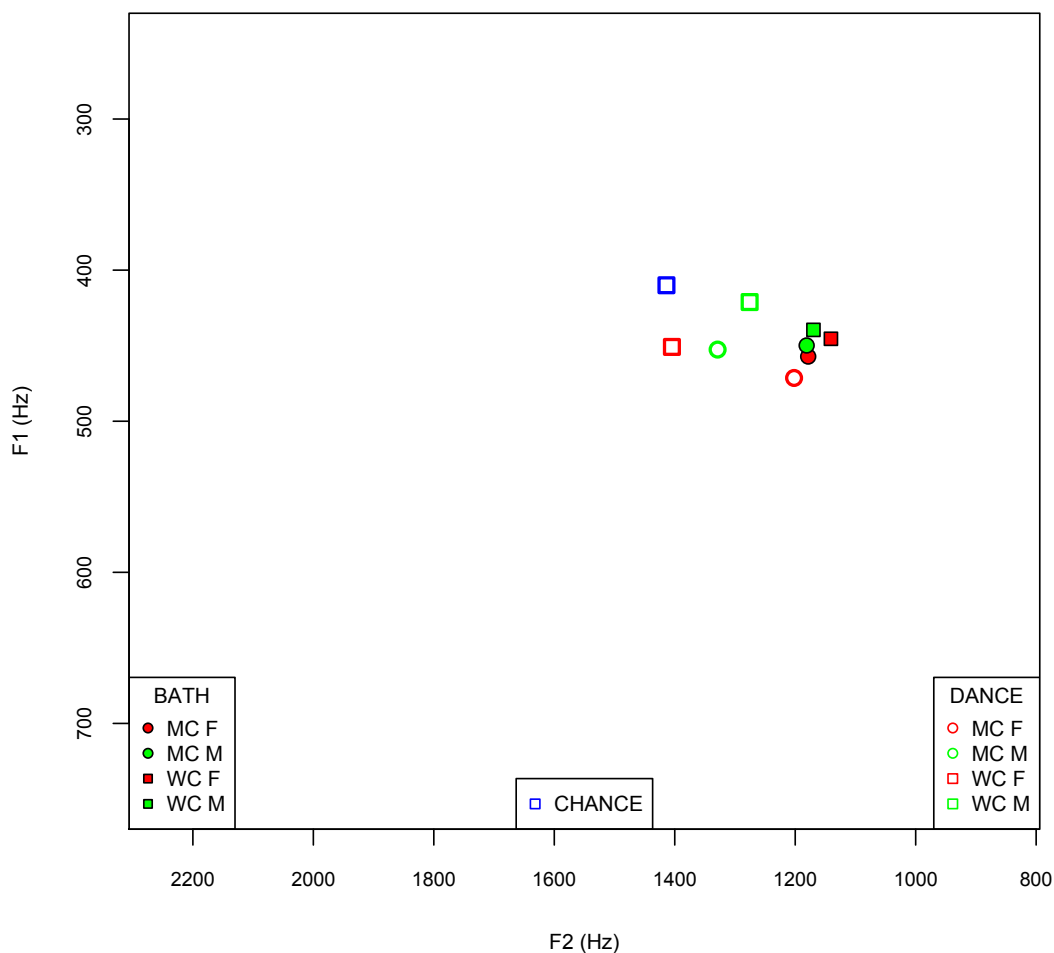


Figure 6.25: Mean values for BATH and DANCE in Interview Style (IS), disaggregated by class and gender, and single token of *chance* by speaker M15.

Figure 6.25 shows that the mean value for DANCE tokens is generally fronter than the mean of regular tokens of the BATH set, especially for male speakers and working-class females. Tokens of the word *dance* do not form part of the set of regular tokens, excluded because they occur before a nasal consonant (see section 4.3 for details). There was only one instance of the word *chance*, produced by a working class male speaker, M15. This token is plotted on figure 6.24, and was produced even fronter and slightly higher than the DANCE tokens for working-class males. The claims that this subset is fronted is thus confirmed in the acoustic data.

c) Class and Style

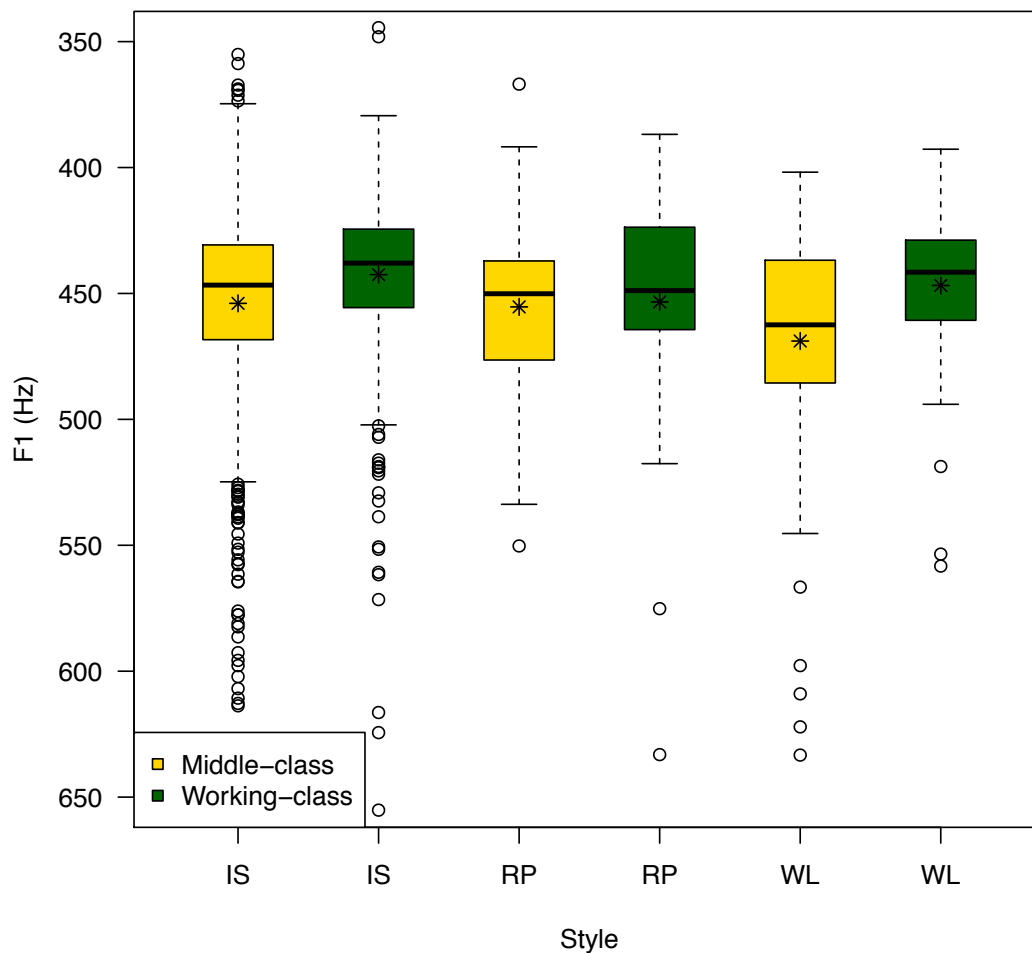


Figure 6.26: Boxplot for BATH F1 values in all speech styles, disaggregated by class.

Working-class speakers produce this set significantly higher compared with the middle-class speakers ($p=0.0099$). This is evident in figure 6.26. Only WL style is significantly different to IS ($p=0.0003$), with the more careful style produced lower than IS tokens. RP style does not show a statistically significant difference to IS ($p=0.2049$).

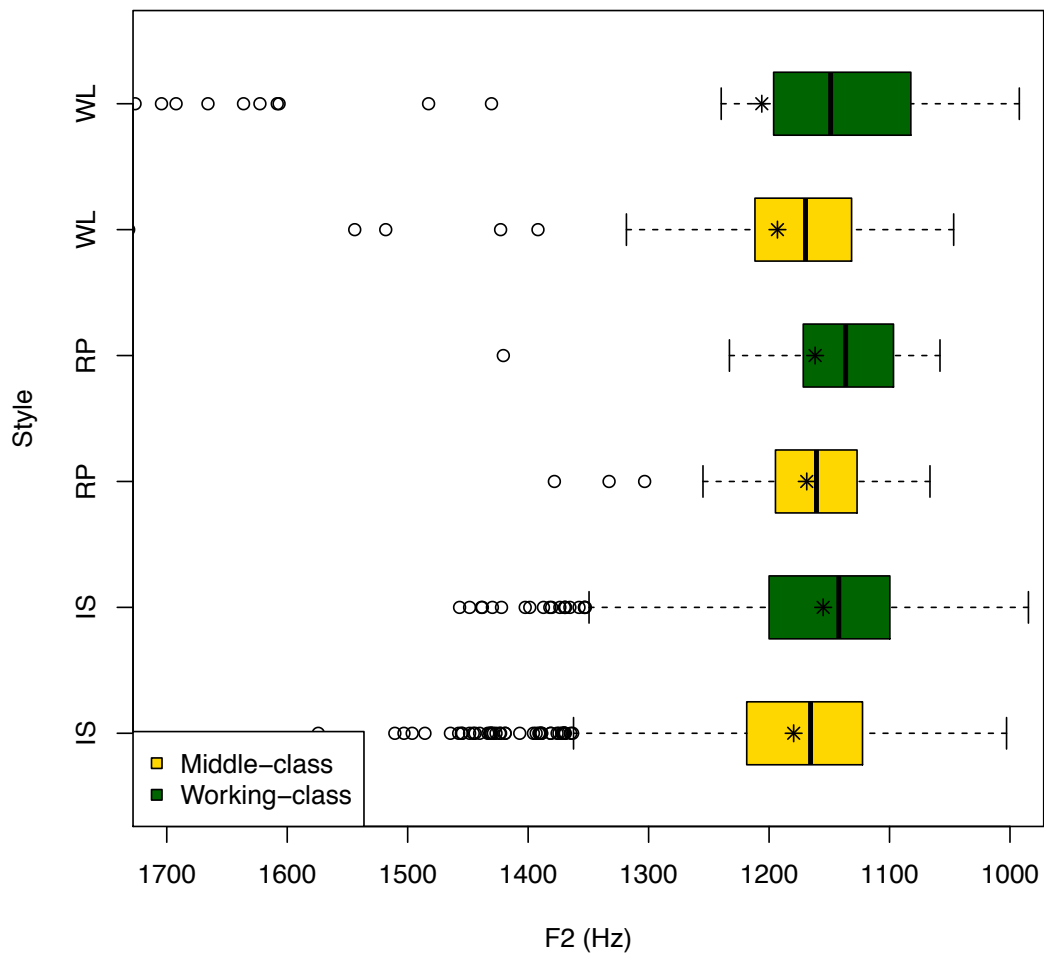


Figure 6.27: Boxplot for BATH F2 values in all speech styles, disaggregated by class.

Class is not a significant factor in terms of vowel advancement for the BATH set ($p=0.1081$). Stylistically, WL style is produced significantly fronter than IS ($p=0.0001$), but there is no significant difference between RP style and IS ($p=0.8256$). This is evident from the boxplots in figure 6.27.

c) Gender and Style

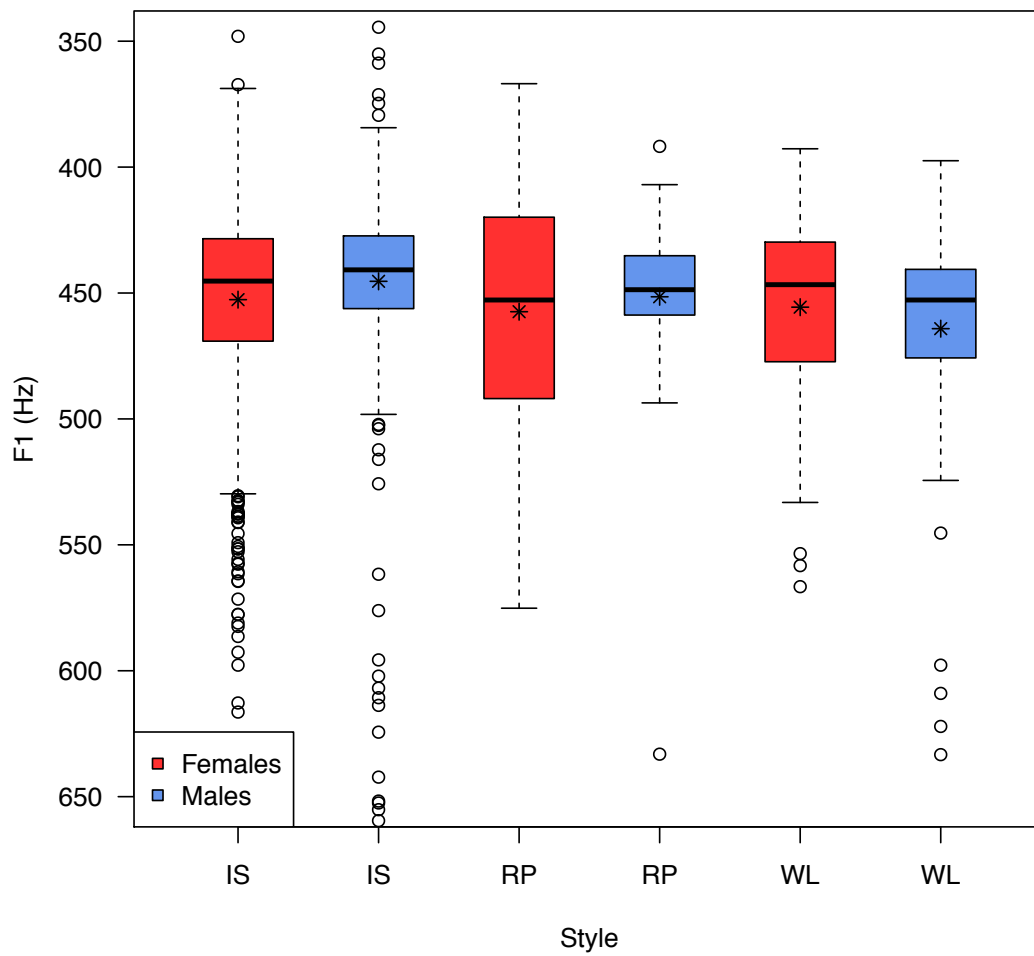


Figure 6.28: Boxplot for BATH F1 values in all speech styles, disaggregated by gender.

With reference to vowel height, there is no significant gender difference ($p=0.3712$). The boxplots in figure 6.28 show that female speakers use a much wider range of values compared with males. In IS and RP style, the means and medians for males are higher than those of the female speakers, while the opposite trend is apparent in WL style.

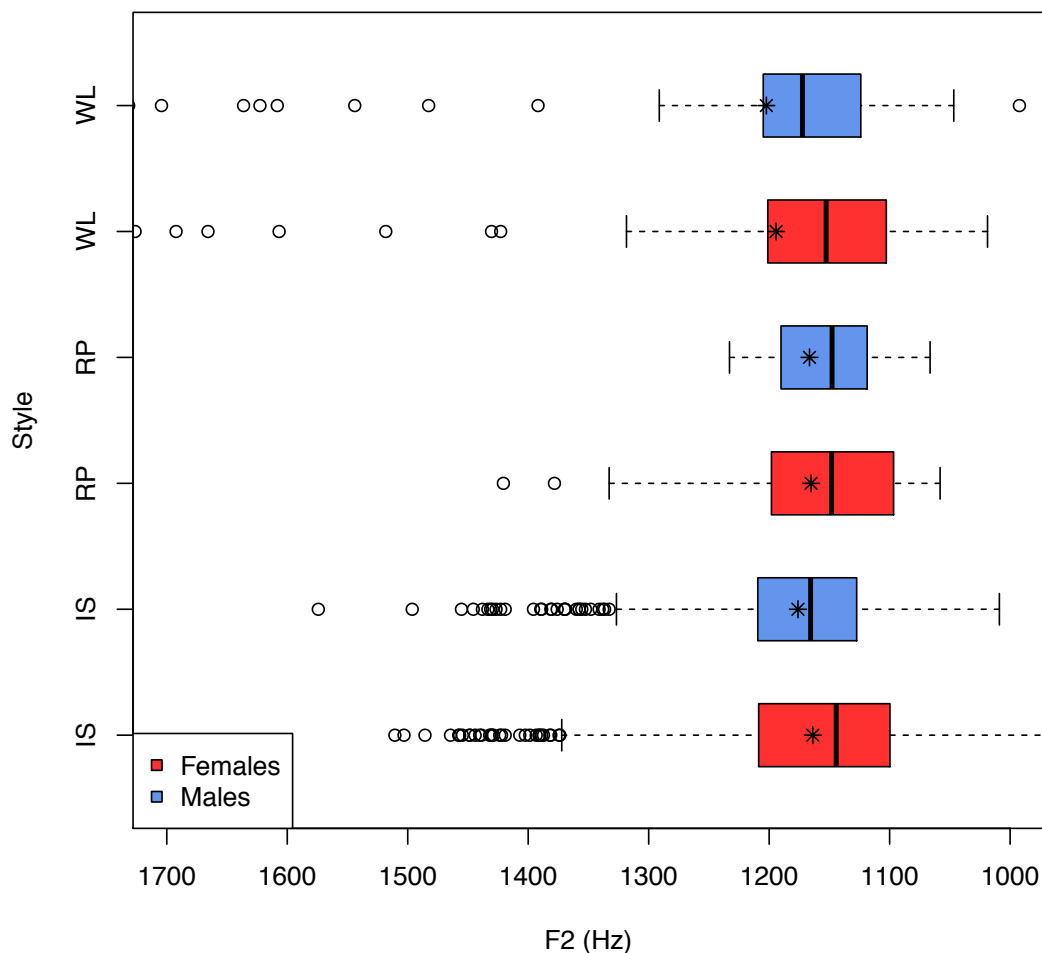


Figure 6.29: Boxplot for BATH F2 values in all speech styles, disaggregated by gender.

In terms of vowel advancement, there is also no significant gender distinction ($p=0.1722$), as figure 6.29 shows.

6.7.4 Summary: BATH Findings

The acoustic data show that working-class speakers use a higher BATH vowel than middle-class speakers do. Gender distinction was not significant, and stylistically, WL style was realised lower and fronter than IS, suggesting that the variant used by the middle-class speakers is more prestigious than the raised variant. Since Bekker's (2009) WSAE speakers also use a lower variant, this would make sense within the SAE system as a whole.

A subset which includes the words *dance* and *chance* are produced with a fronter variant than regular BATH tokens.

6.8 SQUARE: A Rogue Monophthongal Diphthong

Traditionally and historically a diphthong, the inclusion of SQUARE in a chapter entitled ‘Long Monophthongs’ might seem out of place. In General and Broad varieties of SAE, however, this vowel is largely monophthongal, more often in the latter than the former lect (Lass 2002: 118). Lass claims that the monophthongised realisation of SQUARE is stigmatised, even amongst General speakers who use the monophthong, though its use has become very commonplace in SAE generally.

In a very early report of SAE, the absence of the glide for this set is documented (Hopwood 1928: 19), and use of the glide-weakened or glideless variant has been confirmed by many scholars subsequently (e.g. Lanham 1967: 63; Lass 1990: 277). For this reason, an analysis of SQUARE is presented here, in the same manner as the traditional long monophthongs that have been presented in previous sections.

6.8.1 WSAE Reports

Broad speakers tend to use a closer vowel – in the region of [e:] – than General speakers, who use [ɛ:]. The diphthongal variant in use primarily by Conservative speakers is [ɛə].

6.8.2 CSAE Reports

According to the earlier accounts of CSAE, [e:] occurs most commonly (Finn 2004: 973), while a slightly diphthongised variant [e:°] is also reported to occur. Wood (1987: 126) reports a lower [ɛ:] for his speakers.

While acoustic methodology does not allow me to make claims as to whether the speakers in my sample ever use a diphthongal variant for this set, I can confirm that the preponderance is towards the monophthongised variant, based on the preliminary acoustic results presented in the following section.

6.8.3 CSAE Acoustic Data

a) Overview

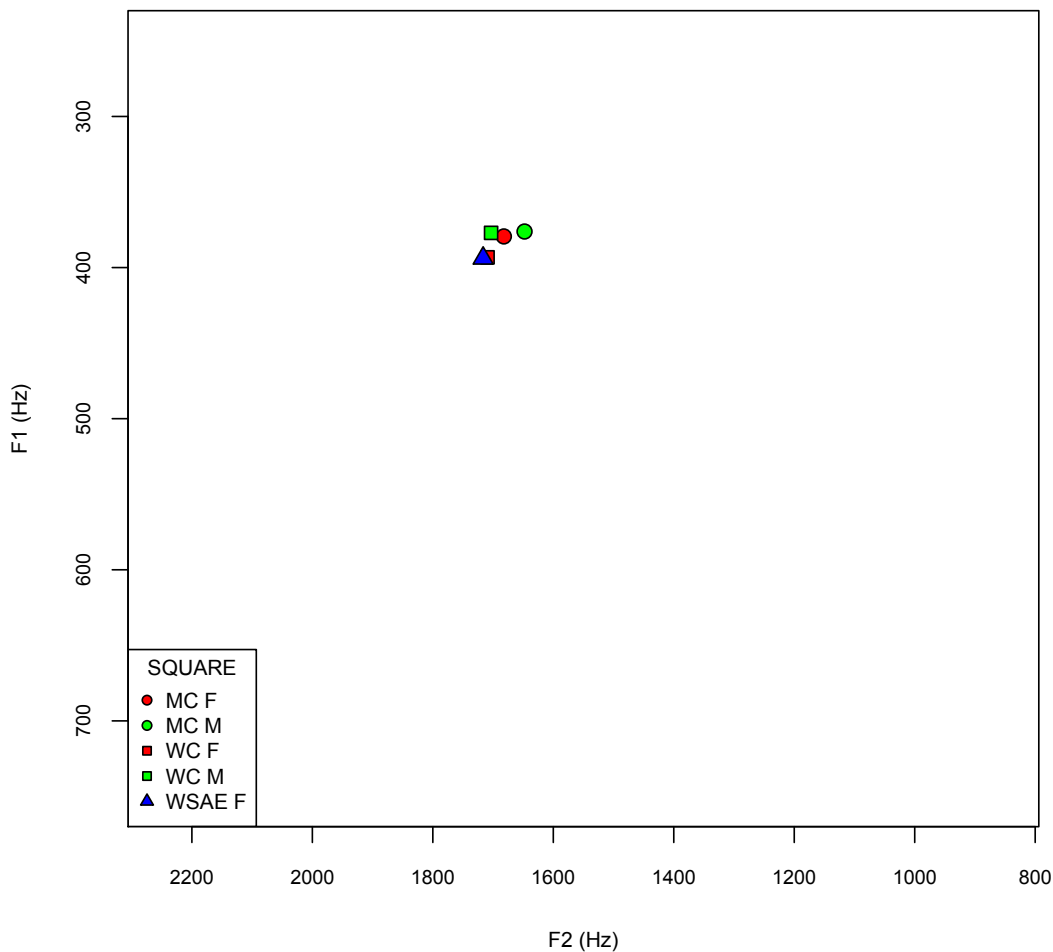


Figure 6.30: Mean values for SQUARE in Interview Style (IS), disaggregated by class and gender.

Based on the plotted means in figure 6.30, there does not seem to be much variation between the different speaker groups. Working-class females produce a slightly lower mean, but otherwise generally in the region of [e:], this set is realised in the same region as Bekker's (2009) WSAE speakers. There were no tokens of SQUARE before tautosyllabic /l/ in the acoustic dataset, so it cannot be confirmed whether or not retraction occurs in this set for this group of speakers, but no retraction is mentioned in the reports by Finn (2004) or Wood (1987).

In order to provide preliminary acoustic results for whether or not this vowel is indeed monophthongised amongst the CSAE speakers in the sample, manual measurements were taken for six randomly selected tokens of SQUARE for four speakers – one male and one female from each of the class groups. Two measurements were taken per

token: the first measuring the nucleus and the second, the glide. The formant values are provided in table 6.5 below.

	F16 (MC)			F13 (WC)			M6 (MC)			M15 (WC)		
	1	2	D	1	2	D	1	2	D	1	2	D
F1	491	496	5	603	592	-11	483	453	-30	423	445	22
F2	2336	2225	-	2090	2035	-55	1722	1700	-22	1906	2032	124
			111									
F1	512	566	54	753	741	-12	462	478	16	440	418	-22
F2	2197	2115	-82	1598	1566	-32	1861	1886	25	1797	1730	-67
F1	508	498	-10	484	466	-18	455	494	39	383	430	47
F2	2505	2240	-	2607	2614	7	1744	1777	33	2047	1995	-52
			265									
F1	525	556	31	410	617	204	421	446	25	474	479	5
F2	2514	2519	5	2569	2334	-	1958	1807	-	2000	1810	-
						235			151			190
F1	505	526	21	550	565	15	460	449	-11	397	428	31
F2	2478	2478	0	2093	1881	-	1774	1766	-12	1985	2005	20
						212						
F1	572	583	11	551	491	-60	449	425	24	414	431	17
F2	2745	2723	-22	2704	2657	-47	1661	1705	44	1914	1574	-
												340

Table 6.5: Comparative formant values for six tokens of SQUARE for four speakers (Hertz; unnormalised).

Key: 1=nucleus reading; 2=glide reading; D=difference.

As the traditional diphthongal value for SQUARE is [eə], any tokens that have diphthongal qualities could have a slightly higher F1 value, but generally a glide would be indicated by a lower F2 value in the measurement of the glide. In table 6.5, differences in values between the F1 readings for the nucleus and glide are underlined (indicating significance) if the second reading has a value which is higher by more than 100Hz. For the F2 readings, if the second reading has a value that is lower by more than 150Hz, this value is underlined to indicate that there is a significant difference. Using this measure, six of the 24 tokens (25 percent) seem to have something of a glide, which suggests that while this set is predominantly realised as a monophthong, there is some relic of the the diphthongal quality that historically

characterised this set. A more detailed acoustic analysis of a greater number tokens is required to validate these findings, however.

b) Class and Style

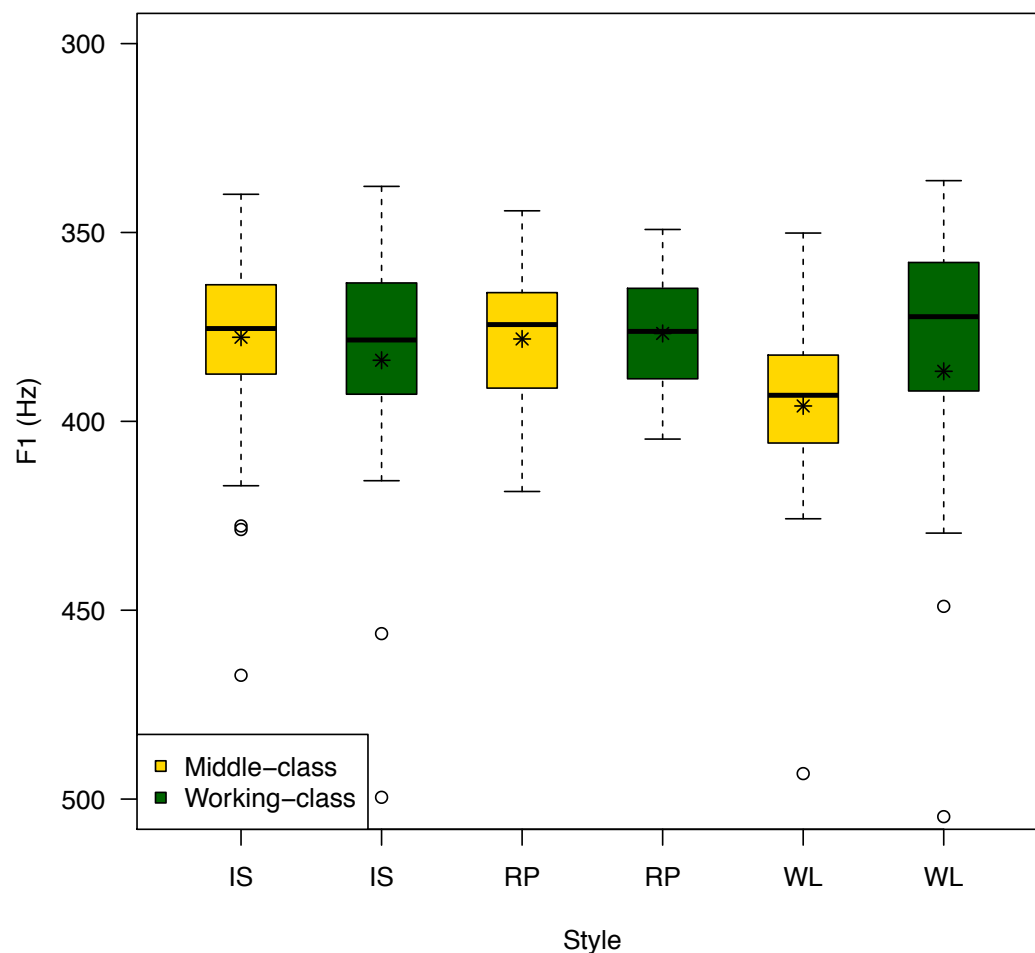


Figure 6.31: Boxplot for SQUARE F1 values in all speech styles, disaggregated by class.

This set does not show significant variation between middle- and working-class speakers ($p=0.8169$), as figure 6.31 shows, particularly in IS and RP style. There is a significant difference between WL and IS tokens ($p=0.0184$), with WL realised lower than IS. RP style shows no significant distinction from IS, however ($p=0.6191$).

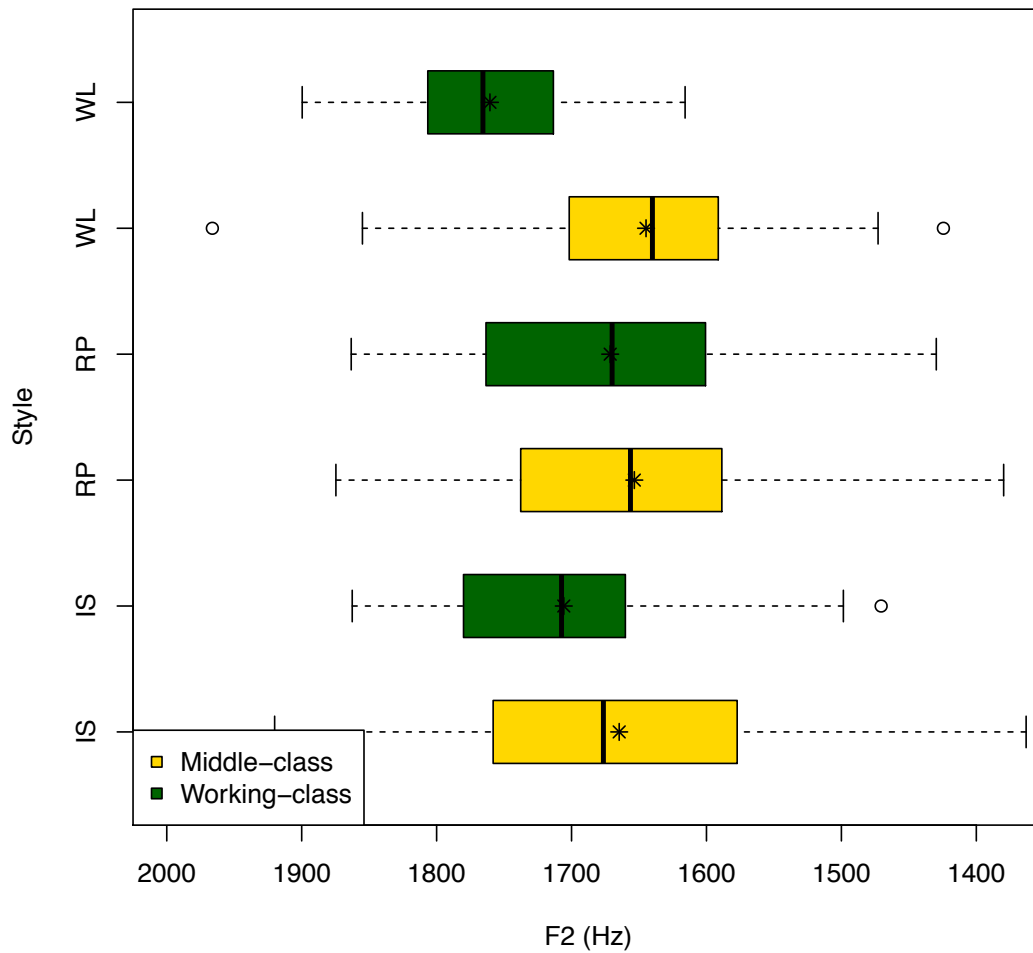


Figure 6.32: Boxplot for SQUARE F2 values in all speech styles, disaggregated by class.

On the F2 plane, the class distinction borders on significant ($p=0.0464$), with working-class speakers using a frontier range than middle-class speakers. Stylistically, neither WL nor RP styles are significantly different to IS ($p=0.4137$ and $p=0.6282$ respectively), as figure 6.32 shows.

c) Gender and Style

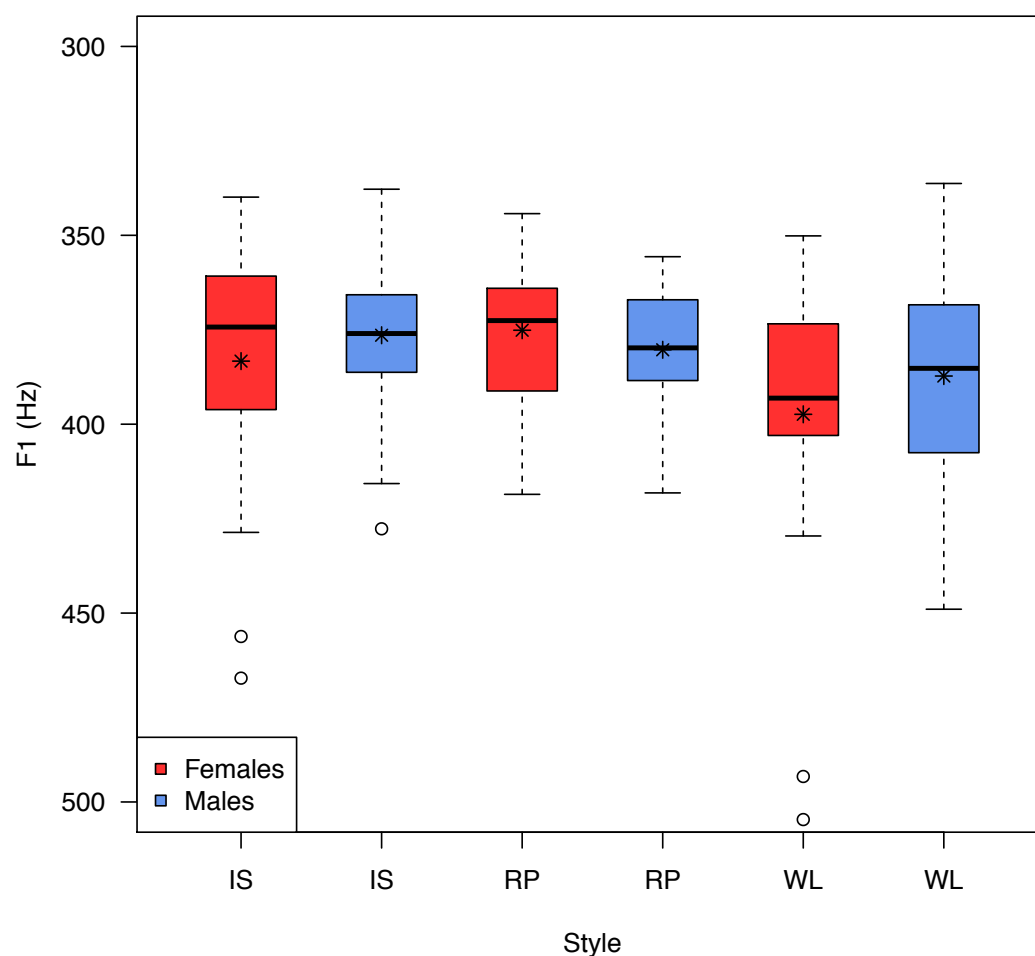


Figure 6.33: Boxplot for SQUARE F1 values in all speech styles, disaggregated by gender.

For this set, there is no significant gender difference with regards to vowel height ($p=0.3473$). In IS and WL style, females use a lower mean value than males, but the opposite trend is apparent in RP style, as can be seen in figure 6.33.

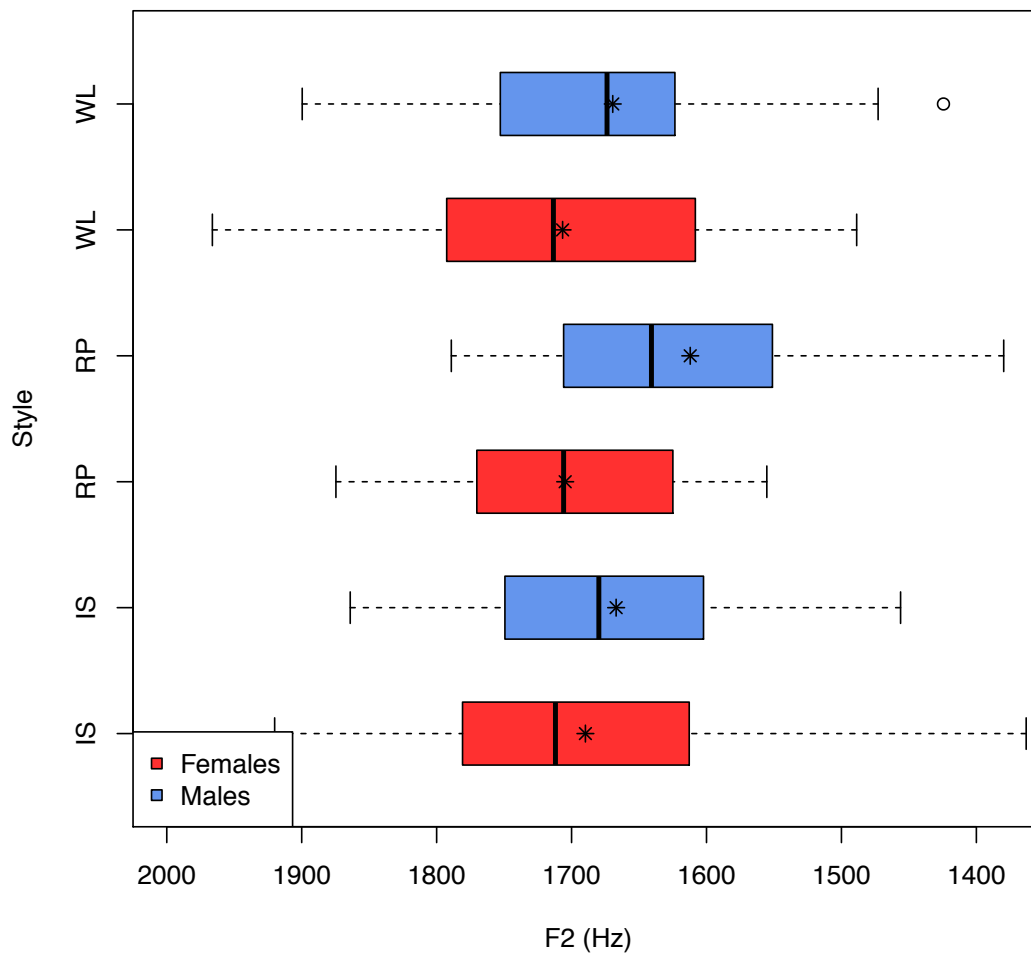


Figure 6.34: Boxplot for SQUARE F2 values in all speech styles, disaggregated by gender.

Similarly on the F2 plane, the distinction between males and females is not significant ($p=0.0753$). This is probably because most tokens upon which the statistical results are based are IS tokens, and the difference here is perhaps not significantly different, but the boxplots in figure 6.34 show clearly that males use a backer mean value in all three speech styles, and generally a backer range of values in RP and WL styles.

6.8.4 Summary: SQUARE Findings

Working-class speakers use a slightly fronter variant than middle-class speakers, but there were no other social factors that significantly influence the distribution of this set. This set is generally produced in the region of [e:]. Preliminary acoustic readings of a small number of tokens suggest that there is some relic of the diphthongal quality that historically characterised this vowel.

6.9 Conclusion

In this chapter, the results of the acoustic analysis of the long monophthongs of CSAE were presented. Table 6.6 provides a summary of the significant findings in the acoustic data.

	Class	Gender	Speech Style	Velar /l/	IPA value*
FLEECE	No significant class distinction.	No significant gender distinction.	RP higher and fronter than IS; WL fronter than IS.	No retraction.	[i:]
NURSE	Working-class lower and backer than middle-class speakers.	No significant gender distinction.	WL lower and fronter than IS.	No retraction.	[ɜ:]
THOUGHT	No significant class distinction.	Males lower and fronter than females.	RP lower and fronter than IS.	No retraction.	[o:]
BATH	Working-class higher than middle-class speakers.	No significant gender distinction.	WL lower and fronter than IS.	No retraction.	[ɔ:]
GOOSE	Middle-class fronter than working-class.	No significant gender distinction.	WL higher than IS.	Data inconclusive.	[u:] ~ [ʊ:]
SQUARE	Working-class fronter than middle-class.	No significant gender distinction.	WL lower than IS.	No data for analysis.	[e:]

Table 6.6: Summary of acoustic findings for the long vowels of CSAE.

* This value is approximate.

The acoustic data generally support the findings of Finn (2004) and Wood (1987), although the gender distinction in the THOUGHT set had not been recorded by either of the scholars.

6.10 CSAE Vowel System

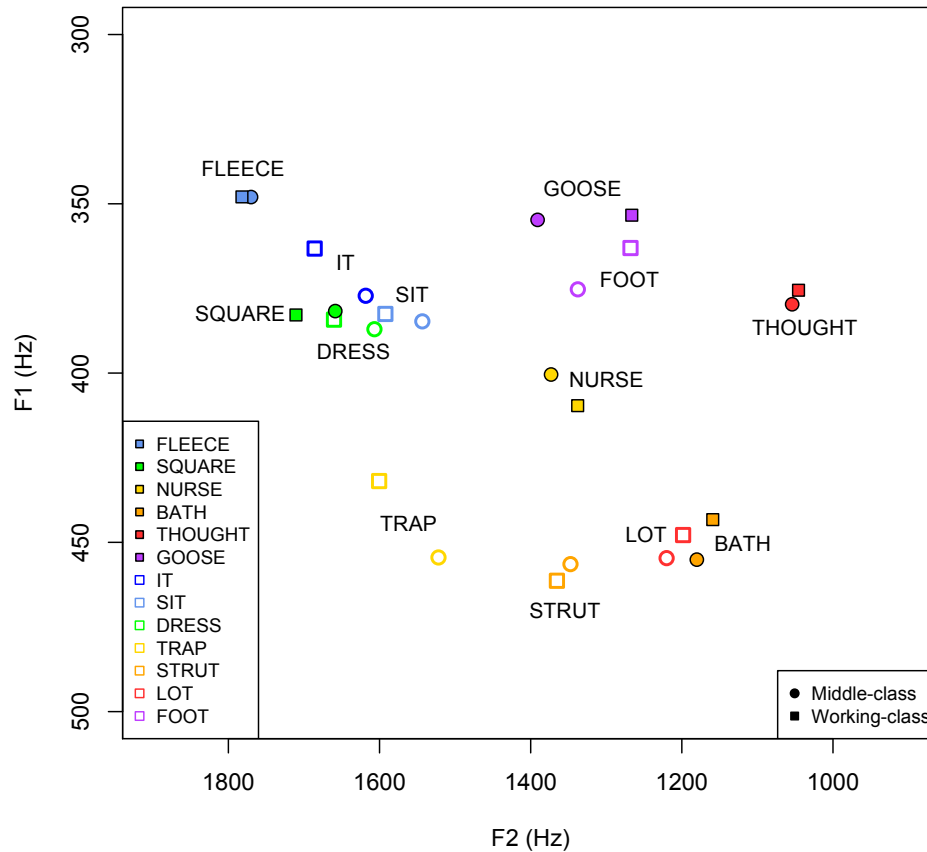


Figure 6.35: An acoustic overview of the simple vowel system of CSAE. Mean values for middle- and working-class speakers.

The full simple vowel system of CSAE is represented in figure 6.35, which shows two points for each vowel: one each for middle- and working-class speakers. The axes are not the same as in the scatter plots presented previously: the range has been reduced so that the distinction between points is clearer. The long vowels are represented with solid points, and the short vowels with outlined shapes.

The graph shows that FLEECE is the highest and frontest vowel, equally so for both working- and middle-class speakers. The vowel quality for this set is [i:]. In relation to FLEECE, the high, fronted subset of the KIT vowel (labelled IT), is lower and backer – with a clear class distinction. Working-class speakers use a higher and slightly fronter variant, in the region of raised [ɪ] (approximating [i]), than the middle-class speakers for this subset, who seem to use a slightly lowered [ɪ]. The centralised subset of the KIT set (labelled SIT), is indeed slightly centralised when compared with the IT subset, but is not as central as NURSE, which is realised with the quality [ɜ:]. The KIT set is showing evidence of lowering as part of the reverse vowel shift, as it overlaps somewhat with DRESS in these acoustic data.

DRESS and SQUARE, the monophthongised diphthong, seem to share vowel quality – [e] – and differ primarily in length, although SQUARE mean values are slightly fronter than the DRESS tokens, and working-class speakers produce fronter mean values than the middle-class speakers for both sets.

TRAP occupies a unique position in the vowel space in a low, front position. This set shows a clear class distinction: middle-class speakers use a much lower mean value compared with working-class speakers. This lowered value approximates the lowered TRAP used by WSAE speakers (Bekker 2009). STRUT is realised as a low, central vowel – [ɐ] – with little class differentiation by the CSAE speakers.

LOT and BATH are qualitatively very similar, differing predominantly in length. Both sets are realised low and back in the region of [ɔ] for working-class speakers, and a slightly lower variant for middle-class speakers. THOUGHT is the backest vowel in the CSAE vowel space, realised as [o:] by both speaker groups in the sample.

Finally, the figure shows similar patterns for the GOOSE⁴⁹ and FOOT lexical sets: middle-class speakers use a more centralised value than working-class speakers. The trend for GOOSE-fronting is thus mirrored in the short vowel FOOT, which occupies a similar, though slightly lower, position in the vowel space. The working-class speakers do not use a fully back vowel for either set, with GOOSE realised as fronted [u:] for these speakers, and central [ɯ:] for the middle-class speakers.

This acoustic study of CSAE thus confirms many of the findings of the impressionistic scholars (viz. Finn (2004) and Wood (1987)), and shows clearly that there is a class distinction between middle-class and working-class speakers for many of the lexical sets. The significance of these findings in terms of the historical and sociopolitical landscape in which this study is located will be discussed in the concluding chapter.

⁴⁹ Tokens in coronal and non-coronal preceding environments have been combined for the purpose of providing an overall picture of the long vowel system.

Chapter Seven

Summary and Conclusions

7.1 Overview of Research Aim

The aim of this thesis was to provide a detailed acoustic description of the phonetic variation and changes evident in the monophthongal vowel system of CSAE speakers in Cape Town, highlighting the most salient changes compared with traditional aural reports of the variety and indicating the extent of the change amongst both working-class and middle-class speakers.

7.2 Research Objectives Recapped

The research was guided by the following objectives:

- i. To produce a complete acoustic analysis of the monophthongal vowel system of CSAE, both middle-class and working-class, using modern methods of acoustic analysis, including AVM.
- ii. To highlight the most salient changes in the phonetic system of middle-class speakers, compared with earlier accounts of the variety.
- iii. To investigate whether changes affect only middle-class speakers, or whether changes have affected the working-class system.

7.3 Research Questions Recapped

The following research questions were proposed in order to maintain a clear focus on the aim and objectives:

- i. How does an acoustic phonetic analysis of present day CSAE differ from the existing accounts of CSAE phonetic and phonological systems, if at all?
- ii. What are the most salient changes that have arisen in the phonetic system of middle-class CSAE speakers in a post-apartheid social setting?
- iii. To what extent are such changes evident in the phonetic system of working-class CSAE speakers?

7.4 Research Questions Answered

Here, a brief answer for each of the research questions is provided, and a detailed summary of the findings is presented below.

- i. The acoustic account confirms many of the findings of scholars who studied this variety using traditional aural methodologies. The changes that

are evident can be attributed to socio-educational change in the post-apartheid setting.

- ii. The most salient changes affect the following lexical sets: TRAP, GOOSE and FOOT (details in section 7.7 below).
- iii. Working-class speakers were found to have maintained the monophthongal vowel system described by earlier scholars of CSAE i.e. changes affecting the middle-class variety have not affected the quality of vowels produced amongst working-class speakers.

7.5 Significance of the research

The original contribution to knowledge provided in this thesis is primarily that it is the first acoustic account of an understudied variety of English in South Africa, viz. CSAE, as previous accounts of CSAE have all been done using non-acoustic methodology. In addition, the methodology employed to automatically time-align the audio recordings with orthographic transcriptions and extract formant measurements, providing the raw data for the acoustic analysis, was applied for the first time to any variety of SAE. This required that P2FA and extractFormants programs be adapted in various ways in order to accurately analyse CSAE. This paves the way for future scholars of SAE to use methods of AVM, which should increase the volume of data that can be analysed within temporally economic time frames.

The study also highlights changes in the vowel system that have occurred as a result of South Africa's post-apartheid socio-educational reform within the Coloured community, which has not been documented sociophonetically before now.

7.6 A Methodological Note

While the implementation of Automatic vowel measurement on CSAE has provided documentation of the present state of the monophthongal vowels in this variety, its implementation for CSAE and other varieties of SAE remain less than ideal. Before future work is undertaken using AVM methodology, the tools need to be optimised for the variety under investigation. Ironically, a process which is designed to be time-saving became rather time consuming for the present researcher, because of the editing and adaptations required at almost every step of the automatic alignment process (details in chapter four). Pioneering this methodology on SAE, however, it was necessary to use a trial and error approach to setting up the programs and running

the SAE data through them, and of course, this effort was not in vain, as it has provided an acoustic account of the current-day vowel system of CSAE.

As mentioned in a footnote in chapter four, the use of FAVE (Rosenfelder et al. 2011) – a program suite which offers the current versions of P2FA and extractFormants as FAVE-align and FAVE-extract respectively – is a promising option which has been found to perform well on non-US varieties of English (e.g. MacKenzie and Turton 2013). Even using FAVE, however, it would still be necessary to amend the CMU Pronouncing Dictionary so that its phonemic transcriptions reflect the variety of English being investigated, rather than NAE. Ideally, phonemes such as OH and EE (for the lexical sets LOT and SQUARE respectively) would need to be added to the dictionary, and to the acoustic models accepted by FAVE, in order to accommodate the phonemic distinctions in varieties of English other than NAE.

7.7 Summary of Findings

This section provides a summary of the findings in chapters five and six for each of the lexical sets.

KIT

The acoustic data showed that working-class speakers have a greater distinction between the two subsets which make up this lexical set: they use a higher, half long [iː] for the IT-subset, and a more centralised [ə] for the SIT subset. Middle-class speakers have less of a distinction, using [ɪ] and [i] for IT and SIT respectively. Stylistically, WL style was realised backer and lower for the SIT-subset compared with IS speech. No significant gender distinctions were found in this set, and no retraction before velar /l/ was found for either subset. This set overlaps with the DRESS set, having lowered slightly as part of the reverse vowel shift which is in process in SAE.

DRESS

Working-class speakers use a fronter variant for this set compared with middle-class speakers, with a quality of [e], and slightly centralised [ɛ] for middle-class speakers. The set shows clear retraction before velar /l/ for all speaker groups. With regards to stylistic differences, WL style is realised lower and fronter than IS. No significant gender differences were recorded for this set.

TRAP

This vowel class is also affected by the reverse vowel shift, as it displays a trend towards lowering, particularly by middle-class speakers (who are following the trend which is more pronounced amongst WSAE speakers – see Bekker (2009)). The value used by working-class speakers is in the region of [ɛ], while middle-class speakers use a lower [æ]. In WL style, speakers use a lower variant than in IS, confirming that the lower variant carries prestige in SAE. This set also did not show a significant distinction between male and female speakers. The acoustic data did not provide conclusive evidence to show whether or not retraction before velar /l/ is a feature of CSAE for this vowel.

LOT

The acoustic data show that for this set, working-class speakers use a backer variant than middle-class speakers do, and that females produce a lower variant than male speakers. The set is realised in the region of [ɔ]. Each of the more careful speech styles was realised lower than IS, suggesting that the lower variant is more prestigious. This is in keeping with the finding that females (including Bekker's (2009) WSAE females) use a lower variant, as females generally tend to use more prestigious variants than male speakers (Chambers 2009: 114). It was confirmed that retraction does occur before velar /l/ for this set.

STRUT

The STRUT set did not show a significant difference between the two social classes, but male speakers were found to produce a higher and backer variant than females and the general value for this set is [ʊ]. With reference to stylistic differences, the two careful speech styles provide conflicting results: WL style is produced lower and fronter than IS, while RP style was produced backer than IS. Retraction before velar /l/ was found amongst middle-class speakers only. While the data show that the TRAP set is lowering, especially amongst middle-class speakers, the STRUT set remains the lowest vowel in the CSAE vowel system.

FOOT

Working-class speakers use a variant that is quite high and back, in the region of [u] (although this set was not realised as back as the THOUGHT set). Middle-class speakers use a more central and lower [ʊ]. This mirrors the trend that is evident in the GOOSE vowel for these speakers. In WL style, speakers used a backer variant than in IS,

while RP style was produced lower than IS. Retraction before velar /l/ was found to occur in this set. The distinction between male and female speakers was not significant.

FLEECE

This set is realised in the region of [i:] for all speaker groups, and is the highest and frontest vowel in the CSAE vowel space. There are no significant distinctions in terms of social class or gender for this vowel class. Stylistically, the data show that both WL and RP styles are produced fronter than IS, and RP style is also produced higher, suggesting that use of a fronter variant is most prestigious. Indeed, Bekker's (2009) WSAE speakers used a fronter variant than the CSAE speakers, confirming that this variant is more prestigious.

NURSE

The NURSE set shows a significant distinction between working- and middle-class speakers: the former group use a lower and backer variant than the latter, in the region of [ɜ:]. WL style was shown to be produced lower and fronter than IS. No retraction before /l/ was evident in this set, nor was there any significant distinction between male and female speakers.

THOUGHT

Class was not a distinguishing factor for this vowel set, but the data show that males use a lower variant than female speakers. The set is produced in the region of [o:]. Stylistically, RP style was produced lower and fronter than IS. The data did not show any retraction before velar /l/. THOUGHT is the backest vowel in the CSAE vowel space.

BATH

Working-class speakers produced a higher variant than their middle-class counterparts, realising the set as [ɔ:]. There was no significant gender distinction, nor did the set show retraction before velar /l/. In terms of stylistic variation, WL style was produced lower and fronter than IS, more in keeping with the middle-class variant. This set is almost identical to the LOT set in terms of vowel quality, and differs only in length.

GOOSE

For this vowel class, there is a significant distinction between middle- and working class speakers: middle-class speakers use a more centralised variant for this set – [u:], while working-class speakers maintain use of a high, back [u:]. There was no significant difference between male and female speakers, but it was found that retraction occurs before velar /l/. WL style was produced higher than IS, suggesting that a higher variant is regarded as more prestigious.

SQUARE

Traditionally a diphthong, this set has been included in this analysis due to its long history of being produced as a glideless or glide-weakened vowel by SAE speakers (see Hopwood 1928: 19). The acoustic analysis shows that working-class speakers use a fronter variant than middle-class speakers, and that WL style is produced lower than IS. The set is produced in the region of [e:], differing from the DRESS set only in length. There were no data available to test whether or not the set retracts before velar /l/, and there was no significant distinction between male and female speakers.

7.8 Relevance of the Findings

The results of the acoustic study show social class to be a very significant distinguishing factor in terms of the production of the monophthongs of CSAE. All of the lexical sets, save for STRUT, FLEECE and THOUGHT, displayed significant differences between middle- and working-class speakers. Where there are differences, the working-class speakers have maintained the vowel system traditionally used by CSAE speakers, as described by Wood (1987), while the changes to the system are being introduced by middle-class speakers, and the trend for these changes is in the direction of WSAE. This is unsurprising given that the social changes that resulted from the political transition from apartheid to democracy (which began in the early 1990s) have effectively redistributed access to opportunities from a race-based system to a class-based system i.e. middle-class South Africans from all race groups are now privy to the resources that the apartheid system reserved exclusively for Whites.

Deracialisation of former White residential and educational sites has progressed steadily over the 20 year period since the introduction of democratic system of governance to South Africa. It is in these deracialised sites that the middle-class speakers in this sample lived and were educated, and where they came into contact with White peers, whose vowel system they approximate in the changes that are being

introduced into the CSAE system. The changes are most evident in two lexical sets: TRAP and GOOSE. WSAE speakers are lowering this set, producing a variant even lower than STRUT (Bekker 2009: 201) – evidence of the reverse vowel shift which is in place in SAE. Middle-class CSAE speakers use a much lower TRAP vowel than working-class speakers, who have retained use of [ɛ] – a stigmatised variant used by Broad SAE speakers across the racial groups. The lowered variant used by middle-class CSAE speakers is not as low as Bekker's (2009) WSAE speakers, however, suggesting firstly that the change is perhaps still in progress, but also that the middle-class speakers, despite their contact with White peers, have retained an intermediate position in terms of their vowel system, placing them in between working-class CSAE speakers and WSAE speakers.

The same trend is evident for the GOOSE set: working-class CSAE speakers have retained the use of the traditional high, back value, while middle-class speakers are using a much more central variant. Bekker's (2009) WSAE speakers use almost a fully front [y:] for this set, so again it is evident that middle-class CSAE speakers position themselves intermediately between the other two speaker groups.

While only thorough sociological analysis could account for why these middle-class speakers do not fully approximate the WSAE vowel system, nor retain the quality of the traditional CSAE system, I conjecture that these speakers feel as though they occupy an intermediate space within the complex social system that is post-apartheid South Africa. Revitalisation of a Coloured identity is evident from the interviews with middle-class speakers, one of whom (M16) stated that 'My friends ... taught me that it's cool to be Coloured ... and to be proud of it'. This contrasts starkly with how previous generations of Coloured people viewed their designation as 'Coloured': something undesirable and from which families and individuals would try to dissociate themselves (Ridd 1981:187-189 cited in Wood 1987: 38-40).

7.9 Directions for Future Research

The stated focus of this dissertation was to provide a thorough and detailed acoustic analysis of the short vowels and long monophthongs of CSAE using novel methodology never before applied to this – or any other – variety of SAE. There remains much room for further work on CSAE, as it is an under-studied variety of English; the last detailed account of the phonetics, morphology, syntax, lexical and discourse features was Tahir Wood's Masters thesis, a work completed more than 25

years ago (in 1987). It is indubitable that features of CSAE have changed in this time, following the political and social changes that have characterised South Africa, as well as the consideration that languages naturally change over time. Future research on this variety could also document change in the diphthongal vowels, such as the extent to which PRICE and MOUTH sets have become glide-weakened following the trend reported by Bekker (2009: 186) in his citation-style data for young, White females.

CSAE also exhibits dissimilarities from other varieties of SAE in its consonantal system, e.g. use of obstruent /r/ and dentalised /t/. Further research into such and other features is also overdue. A follow-up study in the vein of Wood's (1987) phonetic and phonological portion of his dissertation, obviously taking methodological advancement into consideration, would be a very useful research endeavour to document a lively variety of SAE.

An up-to-date account of the morpho-syntactics of CSAE as a uniquely non-standard variety of SAE would also be an interesting research enterprise. It seems, from personal experience, that the morpho-syntactic accounts of CSAE to date (e.g. McCormick 2004; Malan 1996) have not included all non-standard features of the variety e.g. use of *then* as a displaced conjunction in *I then told you that yesterday* (meaning: 'But I told you that yesterday'). Further research might also be conducted with an ethnographic bent, looking perhaps at a single social network in order to ascertain more thoroughly the role of social networks in carrying changes throughout the community. Due to the relative dearth of studies focussing on this variety, any research conducted that includes analysis of features of CSAE would be a useful contribution to the literature on this topic.

7.10 Conclusion

In conclusion, the acoustic reports of CSAE as documented in this dissertation largely support the findings of earlier scholars on this variety, although there are definite changes filtering into the middle-class vowel system as a result of contact with WSAE speakers in deracialised socio-educational settings. Many features of modern CSAE remain to be documented, and it is hoped that the present work provides a basis for future work on this socially significant minority variety of SAE.

Appendix A

Demographic Data

University of Cape Town Linguistics Section

Principal Researcher: Tracey Toefy

Name of Participant:

Date of Birth:

Primary School(s):

High School(s):

Occupation:

Race:

Address:

Contact no:

Email address:

First language:

Other languages:

Do I have your consent to record this interview?

Conditions

- I agree to participate in this research project
- I have read this consent form and the information it contains and had the opportunity to ask questions about them
- I agree to my responses being used for education and research on condition my privacy is respected, subject to the following:
I understand that my personal details will be used in aggregate form only, so that I will not be personally identifiable
- I understand that I am under no obligation to take part in this project
- I understand I have the right to withdraw from this project at any stage

Name of Participant:

Signature of Participant:

Signature of researcher:

Date:

Appendix B

Word List

1. KIT
2. BIT
3. DRESS
4. TRAP
5. LOT
6. STRUT
7. FOOT
8. BATH
9. CLOTH
10. NURSE
11. FLEECE
12. FACE
13. PALM
14. THOUGHT
15. GOAT
16. GOOSE
17. PRICE
18. CHOICE
19. MOUTH
20. NEAR
21. SQUARE
22. START
23. NORTH
24. FORCE
25. CURE
26. DANCE
27. JELLY
28. JEALOUS

Appendix C

Reading Passage

Two cats were having a conversation. ‘How can I hoist this load of bricks to the top of that building?’ said one. ‘Use mice,’ said the other. ‘But where can I find mice?’ asked the first cat. ‘Look, you should try over there at the construction site,’ said the second one. ‘They use them as cheap labour.’ A third cat joined the party: ‘I saw a programme about them on TV. The idea is, they work for their keep, and their food is only peanuts. And the beauty of it is that when the job is finished you can eat them all up.’

References

- Adank, Patti, Smits, Roel & Van Hout, Roeland. 2004. A Comparison of Vowel Normalisation Procedures for Language Variation Research. *Journal of the Acoustical Society of America*. 116(5):3099-3107.
- Adhikari, Mohamed. 2005. *Not White Enough, Not Black Enough: Racial Identity in the South African Coloured Community*. Athens: Ohio University Press.
- Adhikari, Mohamed. 2009. From Narratives of Miscegenation to Post-Modernist Re-imagining: Towards a Historiography of Coloured Identity in South Africa. In Mohamed Adhikari, Ed. *Burdened by Race: Coloured Identities in Southern Africa*. Cape Town: UCT Press. 1-24.
- Allen, R. W. J., Gasson, J. V. & Vivian, J.C. 1990. Anaesthetic Hazards of the 'passion gap'. *South African Medical Journal*. 78:335-336.
- Anthonissen, Christine. 2013. 'With English the world is more open to you' – Language Shift as a Marker of Social Transformation. *English Today*. 113(29):28-35.
- Anufryk, Volha. 2008. Effects of Language Ability on Phonetic and Phonological Variation of English Intonation by German Speakers. *Proceedings of ISSP*. 233-235.
- Atal, Bishnu S. & Suzanne L. Hanauer. 1971. Speech Analysis and Synthesis by Linear Predictive Coding of the Speech Wave. *Journal of the Acoustical Society of America*. 50:637-655.
- Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.
- Baghai-Ravary, Ladan, Sergio Grau, & Greg Kochanski. 2011. Detecting Gross Alignment Errors in the Spoken British National Corpus. In Proceedings of the workshop *New Tools and Methods for Very-Large-Scale Phonetics Research [VLSP]*. 29-31 January 2011. University of Pennsylvania. Available: <http://arxiv.org/pdf/1101.1684.pdf> [2013 July 19].
- Baker, Adam. 2005. Quantifying Diphthongs: A Statistical Technique for Distinguishing Formant Contours. Proceedings of the conference *New Ways of Analyzing Variation 35 [NWAV]*. 9-12 November 2005. The Ohio State University.

- Bates, Douglas, M. 2005. Fitting Linear Mixed Models in R. *R News*. 5:27-30.
- Bekker, Ian & Georgina Eley. 2007. An Acoustic Analysis of White South African English (WSAfE) Monophthongs. *Southern African Journal and Applied Language Studies*. 25(1):107-114.
- Bekker, Ian. 2009. The Vowels of South African English. PhD thesis. North-West University.
- Bekker, Ian. Forthcoming. The KIT-Split in South African English: A Critical Review. Unpublished manuscript.
- Besten, Michael. 2009. 'We are the original inhabitants of this land': Khoe-San Identity in Post-Apartheid South Africa. In Mohamed Adhikari, Ed. *Burdened by Race: Coloured Identities in Southern Africa*. Cape Town: UCT Press. 134-155.
- Bhorat, Haroon. 2004. The Development Challenge in Post-Apartheid South African Education. In Linda Chisholm, Ed. *Changing Classes: Education and Social Change in Post-Apartheid South Africa*. 31-55.
- Bigi, Brigitte & Daniel Hirst. 2012. Speech Phonetization Alignment and Syllabification (SPPAS): a Tool for the Automatic Analysis of Speech Prosody. In *Proceedings of Speech Prosody*. 22-25 May 2014. Shanghai, China.
- Blishen, Bernard. 1958. The Construction and Use of an Occupational Class Scale. *The Canadian Journal of Economics and Political Science*. 24(4):519-531.
- Boersma, Paul & David Weenink. 2008. Praat: Doing Phonetics by Computer (Version 5.4.26) [Computer program]. Available: <http://www.praat.org/> [2008, February 11].
- Bowerman, Sean. 2004. White South African English: Phonology. In *A Handbook of Varieties of English: Volume I*. Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton, Eds. Berlin: Mouton de Gruyter 931-944.
- Boberg, Charles. 2005. The Canadian Shift in Montreal. *Language variation and change*. 17:133-154.
- Bradley, Robert & Robert Corwyn. 2002. Socioeconomic Status and Child Development. *Annual Review of Psychology*. 53:371-399.
- Branford, William. 1994. English in South Africa. In Robert Burchfield, Ed. *The Cambridge History of the English Language, Volume V: English in Britain*

- and Overseas: Origins and Development*. Cambridge: Cambridge University Press. 430-496.
- Brognaux, Sandrine, Sophie Roekhaut, Thomas Drugman & Richard Beaufort. 2012. Train&align: a New Online Tool for Automatic Phonetic Alignment. In Proceedings of *Spoken Language Technology Workshop [SLT]*. IEEE. 2-5 December 2014. Miami, USA.
- Brown, Justin. 2012. Focusing and Diffusion in ‘Cape Flats English’: a Sociophonetic Study of Three Vowels. MA thesis, University of Cape Town.
- Chambers, Jack. 2009. *Sociolinguistic Theory (Revised Edition)*. West Sussex: Blackwell.
- Chen, Nancy, Wade Shen, Joseph Campbell, & Reva Schwartz. 2009. Large-scale Analysis of Formant Frequency Estimation Variability in Conversational Telephone Speech. In Proceedings of *Interspeech*. 6-10 September 2009. Brighton, UK.
- Chen, Lei, Keelan Evanini & Xie Sun. 2010. Assessment of Non-Native Speech using Vowel Space Characteristics. In Proceedings of *Spoken Language Technology Workshop [SLT]*. IEEE. 2-5 December 2014. Miami, USA.
- Chevalier, Alida. 2011. Social Class Differentiation in South African Indian English: a Sociophonetic Study of Three Vowel Variables. MA dissertation, University of Cape Town.
- Chisholm, Linda & Mohamed Sujee. 2006. Tracking Racial Desegregation in South African Schools. *Journal of Education*. 40:141-159.
- Chisholm, Linda. 2004. Introduction. In Linda Chisholm, Ed. *Changing Classes: Education and Social Change in post-apartheid South Africa*. 1-28.
- Christie, Pam. 1991. *The Right to Learn*. Braamfontein: Ravan Press.
- CMU Pronouncing Dictionary. 2013. Available: <http://www.speech.cs.cmu.edu/cgi-bin/cmudict> [2013, March 20].
- Combrink, Johannes G. H. 1978. Afrikaans: Its Origin and Development. In *Language and Communication Studies in South Africa*. Leonard W. Lanham & Koos P. Prinsloo, Eds. Oxford, Cape Town: Oxford University Press 69-95.
- Das, Rajarshi, Jonathan Izak, Jiahong Yuan & Mark Liberman. 2010. Forced Alignment Under Adverse Conditions. Department of CIS Senior Design Project Report. University of Pennsylvania.

- Da Silva, Arista. 2007. South African English: A Sociolinguistic Analysis of an Emerging Variety. PhD dissertation, University of the Witwatersrand.
- Den Besten, Hans. 1989. From Khoekhoe Foreignertalk via Hottentot Dutch to Afrikaans: the Creation of a Novel Grammar. In *Wheels within Wheels: Papers of the Duisburg Symposium on Pidgin and Creole Languages*. Martin Pütz & Rene Dirven, Eds. Frankfurt: Peter Lang. 207-249.
- Dennis, Tracey. 2008. Changing Sociolinguistic Identities of Young, Middle-Class 'Coloured' People in Post-Apartheid Cape Town. MA thesis, University of Cape Town.
- Deonandan, Raywat, Karen Campbell, Truls Ostbye, Ian Tummos & James Robertson. 2000. A Comparison of Methods for Measuring Socio-Economic Status by Occupation or Postal Area. *Chronic Diseases in Canada*. 21(3):1-6.
- Deumert, Ana. 1999. Variation and Standardisation: The Case of Afrikaans (1880-1922). PhD dissertation, University of Cape Town.
- Department of Education (DOE). 1995a. *White Paper on Education and Training – Notice 196 of 1995: Education and Training in a Democratic South Africa*. Pretoria: Department of Education, Republic of South Africa.
- Department of Education (DOE). 1995a. *Report of the committee to review the organisation, governance and funding of schools*. Pretoria: Department of Education, Republic of South Africa.
- Department of Information (DOI). 1974. *Progress of a People: The Coloureds of South Africa*. Pretoria: Department of Information, Republic of South Africa.
- Du Pré, Roy H. 1994. *Separate but Unequal: The 'Coloured' People of South Africa – A political history*. Johannesburg: Jonathan Ball Publishers.
- Erasmus, Zimitri. 2001. Introduction: Re-imagining Coloured Identities in Post-Apartheid South Africa. In *Coloured by History, Shaped by Place*. Zimitri Erasmus, Ed. Cape Town: Kwela.13-28.
- Evanini, Keelan. 2009. The Permeability of Dialect Boundaries: a Case Study of the Region surrounding Erie, Pennsylvania. PhD dissertation, University of Pennsylvania.
- Evanini, Keelan. 2010. Forced Alignment using P2FA: A Tutorial. Unpublished manuscript. University of Pennsylvania.

- Evanini, Keelan, Stephen Isard & Mark Liberman. 2009. Automatic Formant Extraction for Sociolinguistic Analysis of Large Corpora. In *Proceedings of Interspeech 2009*. 1655-1658.
- Feagin, Crawford. 2002. Entering the Community: Fieldwork. In *The Handbook of Language Variation and Change*. Jack K. Chambers, Peter Trudgill & Nancy Schilling-Estes, Eds. Oxford: Blackwell. 20-39.
- February, Vernon A. 1981. *Mind Your Colour: the 'Coloured' Stereotype in South African Literature*. London: Boston Kegan Paul International.
- Finn, Peter. 2004. Cape Flats English: Phonology. In *A Handbook of Varieties of English: Volume 1*. Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton, Eds. Berlin: Mouton de Gruyter 965-984.
- Fiske, Edward & Helen Ladd. 2004. Balancing Public and Private Resources for Basic Education: School Fees in Post-Apartheid South Africa. *Working Papers Series SAN03-04*.
- Flynn, Nicholas. 2011. *Comparing Vowel Formant Normalisation Procedures*. York Papers in Linguistics Series 4.
- Foulkes, Paul. & Gerhard Docherty. 1999. Urban Voices – overview. In Paul Foulkes and Gerhard Docherty, Eds. *Urban Voices: Accent Studies in the British Isles*. Arnold: London. 1–24.
- Ganzeboom, Harry, Paul de Graaf, Donald Treiman and Jan de Leeuw. 1994. A Standard International Socio-Economic Index of Occupational Status. *Social Science Research*. 21:1-56.
- Gimson, Alfred C. 1989. *An Introduction to the Pronunciation of English*. 4th ed. London, Great Britain: Edward Arnold.
- Goldman, J-P. 2011. EasyAlign: an Automatic Phonetic Alignment Tool under Praat. In *Proceedings of Interspeech 2011*. 3233-3236.
- Gorman, Kyle, Jonathan Howel & Michael Wagner. 2011. Prosodylab-aligner: A Tool for Forced Alignment of Laboratory Speech. *Canadian Acoustics*. 39:192-193.
- Hay, Jennifer & Katie Drager. 2007. Sociophonetics. *Annual Review of Anthropology*. 36:89-104.
- Hopwood, David. 1928. *South African English Pronunciation*. Cape Town: Juta
- Jansen, Jonathan. 2003. How Mergers Shape the Institutional Curriculum. *Education as Change*. 7(2):3-19.

- Kallaway, Peter. 2004. *The History of Education under Apartheid, 1948-1994: The Doors of Learning and Culture shall be Opened*. Cape Town: Pearson Education.
- Kawai, Hisashi & Tomoki Toda. 2004. An Evaluation of Automatic Phone Segmentation for Concatenative Speech Synthesis. In *Proceedings of ICASSP 2004*, Montreal, Canada. 677-680.
- Kies, Ben. 1939. The Policy of Educational Segregation and some of its Effects upon the Coloured People of the Cape. B.Ed thesis, University of Cape Town.
- Kies, Ben. 1945. The Basis of Unity. Address delivered at the 3rd Unity Conference, 4-5 January 1945.
- Kinnes, Irvin. 2000. From Urban Street Gangs to Criminal Empires: the Changing Face of Gangs in the Western Cape. Pretoria: Institute for Security Studies.
- Konopka, Kenneth & Janet Pierrehumbert. 2008. Vowels in Contact: Mexican Heritage English in Chicago. In *Proceedings of SALSA*. 94-104.
- Lobanov, Boris M. 1971. Classification of Russian Vowels Spoken by Different Speakers. *Journal of the Acoustical Society of America*. 49:606-608.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Centre for Applied Linguistics.
- Labov, William. 1972a. *Sociolinguistic Patterns*. 2nd ed. Philadelphia: University of Pennsylvania Press, Inc.
- Labov, William. 1972b. *Language in the Inner City: Studies in Black Vernacular English*. Philadelphia: University of Philadelphia Press.
- Labov, William, Ingrid Rosenfelder & Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear Incrementation, Reversal and Reanalysis. *Language*. 89:30-65.
- Labov, William, Sharon Ash, Charles Boberg. 2006. *The Atlas of North American English*. Berlin, New York: Mouton de Gruyter.
- Lanham, Leonard W. 1967. *The Pronunciation of South African English: a Phonetic-Phonemic Introduction*. Cape Town: A. A. Balkema.
- Lanham, Leonard W. 1978. South African English. In Leonard W. Lanham & Koos P. Prinsloo, Eds. *Language and Communication Studies in South Africa: Current Issues and Directions in Research and Inquiry*. Cape Town: Oxford University Press. 138-166.

- Lanham, Leonard W. 1996. A History of English in South Africa. In Vivian de Klerk, Ed. *Focus on South Africa*. 18-34.
- Lanham, Leonard W. & Carol Ann Macdonald. 1979. *The Standard in South African English and its Social History*. Heidelberg: Groos.
- Lanham, Leonard & Anthony Traill. 1962. *South African English Pronunciation*, Witwatersrand University Press, Johannesburg. Reprinted from *English Studies in Africa* 5(2): 171–206.
- Lass, Roger. 1990. A ‘Standard’ South African Vowel System. In Susan Ramsaran, Ed. *Studies in the Pronunciation of English: A Commemorative Volume in Honour of A.C. Gimson*. London: Routledge. 272-285.
- Lass, Roger. 1995. South African English. In Rajend Mesthrie, Ed. *Language and Social History: Studies in South African Sociolinguistics*. Cape Town: David Phillip. 89-106.
- Lass, Roger. 2002. South African English. In *Language in South Africa*. Rajend Mesthrie, Ed. Cambridge: Cambridge University Press. 104-126.
- Lass, Roger & Susan Wright. 1985. The South African Chain Shift: Order out of chaos? In Roger Eaton, Olga Fischer, Willem Koopman & Frederike van der Leek, Eds. *Papers from the 4th International Conference on English Historical Linguistics*. Amsterdam: John Benjamins. 137-161.
- Lass, Roger & Susan Wright. 1986. Endogeny vs. Contact: ‘Afrikaans influence’ on South African English. *English Worldwide* 7(2):201-224.
- Leibbrandt, Murray, Ingrid Woolard & Christopher Woolard. 2007. Poverty and Inequality Dynamics in South Africa: Post-apartheid Developments in the Light of the Long-Run Legacy. IPC- DRCLAS workshop, Brasilia 11-13 January 2007.
- Lennig, Matthew. 1978. Acoustic Measurement of Linguistic Change: The Modern Paris Vowel System. Doctoral Dissertation, University of Pennsylvania.
- Lewis, Gavin. 1987. *Between the Wire and the Wall*. Cape Town: David Phillip.
- MacKenzie, Clayton G. 1993. New Models of School Management in South Africa: Licence for Change or Loophole for Separatism? *British Journal of Educational Studies*. 41(3):287-301 (Accessed online 27 October 2006 at jstor.org).

- MacKenzie, Laurel & Danielle Turton. 2013. Crossing the Pond: Extending Automatic Alignment Techniques to British English Dialect Data. Paper presented at NWAV 42, Pittsburgh, PA. October 2013.
- Makino, Takehiko & Rika Aoki. 2012. English Read by Japanese Phonetic Corpus: an Interim Report. *Research in Language*. 10(1):79-95.
- Malan, Karen. 1981. An investigation of Non-Standard English Syntax in 12 year old Coloured Children. *The South African Journal of Communication Disorders*. 28:68-80.
- Malan, Karen. 1996. Cape Flats English. In *Focus on South Africa*. V. de Klerk, Ed. Amsterdam: John Benjamins Publishing 125-148.
- Marais, Johannes, S. 1939. *The Cape Coloured People*. London: Longman, Green.
- Markel, John D. & Augustine H. Gray, 1976. *Linear Prediction of Speech*. New York: Springer-Verlag.
- McCormick, Kay. 1989a. Unfiltered Talk: a Challenge to Categories. *York Papers in Linguistics*. 13:203-214.
- McCormick, Kay. 1989b. English and Afrikaans in District Six: a sociolinguistic study. Unpublished PhD thesis, University of Cape Town.
- McCormick, Kay. 1995. Code-Switching, Code-Mixing and Convergence in Cape Town. In Rajend Mesthrie, Ed. *Language and Social History: Studies in South African Sociolinguistics*. Cape Town: David Phillip. 193-208.
- McCormick, Kay. 2002. *Language in Cape Town's District Six*. Oxford: Oxford University Press.
- McCormick, Kay. 2004. Cape Flats English: Morphology and Syntax. In *A Handbook of Varieties of English: Volume II*. Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton, Eds. Berlin: Mouton de Gruyter. 993-1006.
- Mesthrie, Rajend. 1993. English in South Africa. *English Today*. 9(1):27-34.
- Mesthrie, Rajend. 1999. Fifty Ways to Say 'I do': Tracing the Origins of Unstressed Do in Cape Flats English, South Africa. *South African Journal of Linguistics*. 17(1):58-71.
- Mesthrie, Rajend. 2004. Indian South African English: Morphology and Syntax. In *A Handbook of Varieties of English: Volume II*. Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton, Eds. Berlin: Mouton de Gruyter. 974-994.

- Mesthrie, Rajend. 2006. *English in Language Shift: The History, Structure and Sociolinguistics of South African Indian English*. Cambridge University Press.
- Mesthrie, Rajend. 2007. Is Port Elizabeth half-way between Durban and Cape Town? – A Preliminary Account of the English of Coloured and Indian communities in Four South African Cities. *Paper presented at the Linguistic Society of Southern Africa Conference*. University of Potchefstroom, July 2007.
- Mesthrie, Rajend. 2008. ‘Death of the mother tongue’ – is English a Glottophagic Language in South Africa? *English Today*. 24(2):13-19.
- Mesthrie, Rajend. 2010. Socio-phonetics and Social Change: Deracialisation of the GOOSE vowel in South African English. *Journal of Sociolinguistics*. 14(1):3-34.
- Mesthrie, Rajend. 2012a. English in Africa – a Diachronic Typology. In Alex Bergs & Laurel Brinton, Eds. *English Historical Linguistics: An International Handbook Vol II*. Berlin: de Gruyter Mouton. 2092-2106.
- Mesthrie, Rajend. 2012b. Ethnicity, Substrate and Place: The Dynamics of Coloured and Indian English in Five South African Cities in Relation to the Variable (t). *Language Variation and Change*. 24(3):371-395.
- Mesthrie, Rajend. 2012c. Social Change and Changing Accents in South Africa. In Philip Seargeant, & Joan Swann, Eds. *English in the World: History, Diversity, Change*. 316-329). United Kingdom: Routledge.
- Mesthrie, Rajend, Alida Chevalier & Timothy Dunne. 2013. A Study of Variation in the BATH Vowel among White Speakers of South African English in Five Cities. *University of Pennsylvania Working Papers in Linguistics* 19(2):15.
- Millin, Sarah G. 1924. *God's Stepchildren*. London: Constable & Co.
- Mitchell, Alexander G. & Arthur Delbridge. 1965. *The Pronunciation of English in Australia*. Sydney: Angus & Robertson.
- Milroy, Lesley. 1980. *Language and Social Networks*. Oxford: Basil Blackwell.
- Milroy, Lesley & Matthew Gordon. 2003. *Sociolinguistics: Method and Interpretation*. Malden, MA: Blackwell.
- Molteno, Frank. 1987. The 1980 Boycott of Coloured Education in the Cape Peninsula. *British Journal of Sociology of Education*. 8(1):3-24.
- Morreira, Kirsten. 2012. Social change, Class Formation and English: a Study of Young Black South Africans with ‘Model C’ School Backgrounds. PhD thesis, University of Cape Town.

- Naidoo, Jordan. 2005. *Educational Decentralization and School Governance in South Africa: From Policy to Practice*. Paris: International Institute for Educational Planning.
- Nearey, Terrence. 1977. Phonetic Feature Systems for Vowels. Doctoral Dissertation, University of Alberta. Reprinted 1978 by the Indiana University Linguistics Club.
- Nycz, Jennifer, & Paul De Decker. 2005. A New Way of Analyzing Vowels: Comparing Formant Contours using Smoothing Spline ANOVA. Paper presented at NWAV 35, Columbus, OH.
- Pierrehumbert, Janet B., Tessa Bent, Benjamin Munson, Ann R. Bradlow, & J. Michael Bailey. 2004. The Influence of Sexual Orientation on Vowel Production. *Journal of the Acoustical Society of America* 116:1905–1908.
- Ponelis, Fritz. 1994. *The Development of Afrikaans*. Frankfurt: Peter Lang.
- Republic of South Africa (RSA) President's office. 1950. *Population Registration Act No 30 of 1950*. 1950. Cape Town: Government Printer, Republic of South Africa.
- Ponte, Stefano, Simon Roberts & Lance van Stittert. 2007. 'Black Economic Empowerment', Business and the State in South Africa. *Development and Change*. 38(5):933-955.
- Posel, Deborah. 2001. What's in a name? Racial Categorisations under Apartheid and their Afterlife. *Transformation*. 47:50-74.
- R Core Team. 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL <http://www.R-project.org/>.
- Risdal, Megan & Mary Kohn. 2013. Ethnolectal and Generational Differences in Vowel Trajectories: Evidence from African American English and the Recession of the Southern Vowel System. Paper presented at NWAV 42, Pittsburgh, PA. October 2013.
- Roberge, Paul T. 1995. The Formation of Afrikaans. In Rajend Mesthrie, Ed. *Language and Social History: Studies in South African Sociolinguistics*. Cape Town: David Phillip 68-88.
- Rosenberg, Andrew & Julia Hirschberg. 2009. Detecting Pitch Accents at the Word, Syllable and Vowel Level. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the*

- Association for Computational Linguistics, Companion Volume: Short Papers.*
Association for Computational Linguistics.
- Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini & Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite. <http://fave.ling.upenn.edu>.
- Rubin, Mark, Nida Denson, Sue Kilpatrick, Kelly Matthews, Tom Stehlik & David Zyngier. 2014. 'I am Working-Class': Subjective Self-Definition as a Missing Measure of Social Class and Socioeconomic Status in Higher Education Research. *Educational Researcher*. 43:196-200.
- Schiel, Florian & Christoph Draxler. 2004. The Production of Speech Corpora. *Bavarian Archive for Speech Signals*, Tech. Rep.
- Schiffrin, Deborah. 1996. The Transformation of Experience, Identity, and Context. *Towards a Social Science of Language Vol II: Social Interaction and Discourse Structures*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 41-55.
- Schneider, Geoffrey. 2003. Neoliberalism and Economic Justice in South Africa: Revisiting the Debate on Economic Apartheid. *Review of Social Economy*. 61(1):23-50.
- Seekings, Jeremy & Nicoli Nattrass. 2004. Class, Distribution and Redistribution in Post-Apartheid South Africa. *Transformation*. 50:1-30.
- Shirk, Melanie. 1985. Language Contact in a Lower Socio-Economic Suburb of Cape Town: Non-Standard Linguistic Features in White and Coloured Pre-School Children. Unpublished research report, University of Cape Town.
- Smit, Warren. & Kieron Hennessy. 1995. *Taking South African Education out of the Ghetto: An Urban-Planning Perspective*. Cape Town: University of Cape Town Press.
- Soudien, Crain. 1992. Responses to the Education Crisis in South Africa. *International Review of Education*. 38(3):280-285 (Accessed online 14 August 2012 at jstor.org).
- Soudien, Crain. 2004. 'Constituting the Class': an Analysis of the Process of 'Integration' in South African Schools. In Linda Chisholm, Ed. *Changing Classes: Education and Social Change in post-apartheid South Africa*. 89-114.
- Soudien, Crain & Yusef Sayed. 2003. Integrating South African Schools? Some preliminary findings. *IDS Bulletin* 34(1):29-44.

- Soudien, Crain, Peter Kallaway, Michael Cross & Charlotte Fischer. Forthcoming. 'Incindi ye Khala' – A Bitter Harvest: 350 Years of Formal Schooling in South Africa. Unpublished manuscript.
- Southall, Roger. 2004. Political Change and the Black Middle Class in Democratic South Africa. *Canadian Journal of African Studies*. 38(3):521-544.
- Statistics South Africa. 2012. *Census in Brief*. Pretoria: Statistics South Africa.
- Stone, Gerald. 2002. The Lexicon and Sociolinguistic Codes of the Working-Class Afrikaans-Speaking Cape Peninsula Coloured Community. In *Language in South Africa*. Rajend Mesthrie, Ed. Cambridge: Cambridge University Press. 381- 397.
- Tauberer, Joshua & Keelan Evanini. 2009. Intrinsic Vowel Duration and the Post-Vocalic Voicing Effect: Some Evidence from Dialects of North American English. In Proceedings of *Interspeech 2009*.
- Thomas, Erik & Tyler Kendall. 2007. NORM: The Vowel Normalisation and Plotting Suite. [Online Resource: <http://ncslaap.lib.ncsu.edu/tools/norm/>]
- Thomas, Erik. 2011. *Sociophonetics: An Introduction*. Hampshire: Palgrave Macmillan
- Tikly, Leon & Thabo Mabogoane. 1997. Marketisation as a Strategy for Desegregation and Redress: The Case of Historically White Schools in South Africa. *International Review of Education*. 43 (2/3):159-178 (Accessed online 27 October 2006 at jstor.org).
- Traill, Anthony. 2004. The Khoesan Languages. In *Language in South Africa*. Rajend Mesthrie, Ed. Cambridge: Cambridge University Press pp 27-49.
- Trudgill, Peter. 1974. *The Social Stratification of English in Norwich*. London: Cambridge University Press.
- Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.
- Van der Berg, Servaas & Ronelle Burger. Education and Socio-Economic Differentials: A Study of School Performance in the Western Cape. *The South African Journal of Economics*. 71(3):496-524.
- Van der Ross, Richard E. 1986. *The Rise and Decline of Apartheid: A Study of Political Movements among the Coloured People of South Africa, 1880-1985*. Cape Town: Tafelberg.

- Van Rooy, Bertus. 2004. Black South African English: Phonology. In *A Handbook of Varieties of English: Volume 1*. Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton, Eds. Berlin: Mouton de Gruyter 943-954.
- Venter, Al J. 1974. *Coloured: A Profile of Two Million South Africans*. Cape Town: Human and Rousseau.
- Wells, John. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- Watson, Catherine I. & Jonathan Harrington. 1999. Acoustic Evidence for Dynamic Formant Trajectories in Australian English Vowels. *Journal of the Acoustical Society of America*. 106:458-468.
- Wolfram, Walter A. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, DC: Centre for Applied Linguistics.
- Worden, Nigel. 1985. *Slavery in Dutch South Africa*. Cambridge: Cambridge University Press.
- Wissing, Daan. 2004. Black South African English: a New English? Observations from a Phonetic Viewpoint. *World Englishes*. 21(1):129-144.
- Wolpe, Harold. 1988. *Race, Class and the Apartheid State*. London: James Currey.
- Wood, Tahir M. 1987. Perceptions of, and Attitudes Towards Varieties of English in the Cape Peninsula, with Particular Reference to the 'Coloured Community'. MA dissertation, Rhodes University.
- Wu, Chen-huei & Chilin Shih. 2012. A Corpus Study of Native and Non-Native Vowel Quality. Proceedings of *Speech Prosody*. 22-25 May 2014. Shanghai, China.
- Western Cape Government. 2012. Education Management Information System: Find A School. Available: <http://wcedemis.pgwc.gov.za/wced/findaschool.html> [2012, October 4].
- City of Cape Town. 2012. Maps, Policies and Publications. Available: <http://www.capetown.gov.za/en/sdf/Pages/Maps.aspx> [2012, October 4].
- Young, Steve, Gunnar Everman, Mark Gales, Thomas Hain, Dan Kershaw, Gareth Moore, Julian Odell, Dave Ollason, Dan Povey, Valtcho Valtchev & Phil Woodland. 2009. *The HTK Book (For HTK Version 3.4)*. Cambridge University Engineering Department.
- Yuan, Jiahong & Mark Liberman. 2008a. Vowel Acoustic Space in Continuous Speech: An Example of Using Audio Books for Research. Proceedings of the

- conference on “*The First International Workshop on Cataloguing and Encoding of Spoken Language Data*” at CatCod 2008.
- Yuan, Jiahong & Mark Liberman. 2008b. Speaker Identification on the SCOTUS Corpus. Proceedings of *Acoustics* 2008.
- Yuan, Jiahong, & Mark Liberman. 2011. Automatic Detection of ‘g-dropping’ in American English using Forced Alignment. *Automatic Speech Recognition and Understanding (ASRU)*. IEEE Workshop 2011.
- Yuan, Jiahong & Mark Liberman. 2012. Investigating /l/ Variation in American English: A Corpus Approach. *Journal of Speech Sciences*. 1(2): 35-46.
- Yuan, Jiahong. 2012. The Effects of Speaking Rate and Intonation on the Duration of Tones in Mandarin Chinese. Proceedings of *Speech Prosody*. 22-25 May 2014. Shanghai, China.